Investigating the production of an inhibitor of plant β-galactosidase by the plant pathogen *Pseudomonas syringae*

Nattapong Sanguankiattichai

St Catherine's College
University of Oxford

Thesis submitted for the degree of Doctor of Philosophy

February 2021

Abstract

Microbial pathogens are causative agents of various plant diseases. During infection, plants and bacteria participate in a dynamic interaction in the apoplast. Secreted hydrolases are prominent players in plant defence and are therefore targets of pathogen virulence factors. Using *Nicotiana benthamiana* and *Pseudomonas syringae* pv. *tomato* (*Pst*) DC3000 as a model system, we recently discovered a plant β-galactosidase (BGAL1) that facilitates the hydrolytic release of elicitor peptides from bacterial glycosylated flagellin to activate plant defences. The activity of BGAL1 is suppressed by an inhibitor secreted from *Pst*DC3000 during infection. This thesis aims to determine the molecular identity of this inhibitor, characterise its biosynthetic and regulatory pathways and unravel biological roles and evolution of this inhibitor production. Here, I identify the *bim* biosynthetic gene cluster of the BGAL1 inhibitor, which is under the control of the conventional type III secretion system regulators HrpR/S/L and a putative novel virulence regulator HtpX. The *bim* cluster has been acquired via horizontal gene transfer and is widespread among *P. syringae* strains. A search for homologs of *bim* cluster genes also reveals different putative biosynthetic gene clusters of iminosugars in various bacteria. Analytical chemistry analyses suggest that the inhibitor is an iminosugar, an analogue of sugar with N replacing O in the ring, with a putative mass of 151 Da. The production of the inhibitor can also alter galactosylation of plant apoplastic glycoproteins and isolated inhibitor fractions can inhibit other glycosidases such as α-galactosidase, α-glucosidase and β-glucosidase, suggesting that multiple pathways in host plants could be targeted. This work has uncovered a novel small molecule inhibitor secreted by *P. syringae* during infection and highlights the role of glycans and glycosidases in plant-pathogen interactions that remains to be explored.

Acknowledgement

I would like to thank my supervisors Gail Preston and Renier van der Hoorn for their support and guidance throughout my DPhil; Pierre Buscaill and Balakumaran Chandrasekar for pioneering work on this topic; Daniel Krahn for help with analytical chemistry, HPLC and LC-MS; Pedro Bota for help with GC-MS; and members of both the Preston Lab and Plant Chemetics Lab for their support and friendship. I would also like to thank the Interdisciplinary Bioscience Doctoral Training Partnership, UKRI-BBSRC and the Royal Thai Government Scholarship for providing the funding for my DPhil.

Table of Contents

Chapter 1: Introduction 8

1.1 *Pseudomonas syringae* as model plant pathogenic bacteria 9

1.1.1 *P. syringae* biology 9

1.1.2 Toxins and small-molecule virulence factors 12

1.1.3 Virulence regulation 17

1.2 Plant-pathogen interactions and the apoplastic battlefield 21

1.2.1 Plant immune system and manipulation by pathogens 21

1.2.2 Apoplastic hydrolases and inhibitors in plant-pathogen interactions 23

1.3 Glycans, glycosidases and inhibitors in plant-pathogen interactions 24

1.3.1 Polysaccharides 26

1.3.2 Glycoproteins 30

1.3.3 Glycoside conjugates 32

1.4 The discovery of a β-galactosidase inhibitor produced by *P. syringae* during infection and the role of the plant β-galactosidase BGAL1 in immunity 33

1.4.1 Activity-based protein profiling for studying hydrolase activity and inhibition 33

1.4.2 Discovery of β-galactosidase inhibitor and the role of BGAL1 during *P. syringae* infection 35

1.4.3 Plant β-galactosidases 36

1.4.4 Further work on β-galactosidase inhibitor production by *P. syringae* 37

1.5 Iminosugars as inhibitors of glycosidases 38

1.5.1 Iminosugar structure, type and natural occurrence 38

1.5.2 Therapeutic applications of iminosugars 39

1.5.3 Biological roles of iminosugars and applications 42

1.5.4 Biosynthesis of iminosugars 44

1.5.5 Methods for studying iminosugars 48

1.6 Research aims and rationale 49

Chapter 2: The identification of candidate genes involved in BGAL inhibitor production .. 51

2.1 Introduction 51

2.2 Materials and methods 54

2.2.1 Bacterial strains, culture conditions and media 54

2.2.2 BGAL1 inhibitor production *in vitro* 55

2.2.3 Apoplastic fluid extraction 55

2.2.4 FDG assay for β-galactosidase activity 55

2.2.5 Activity-based protein profiling (ABPP) of glycosidase activity 57

2.2.6 Simulation of transposon insertion in the genome 57

2.2.7 Triparental mating conjugation for mini-Tn5 transposon insertion mutagenesis 57

2.2.8 Plating screen for BGAL1 inhibitor mutant colonies 58

2.2.9 The 96 well plate-based screen for BGAL1 inhibitor mutants 59

2.2.10 Inverse PCR 59

2.2.11 Sequencing 61

2.2.12 Mapping transposon insertion sites 61

2.2.13 Simulation of restriction enzyme digest 61

2.2.14 Gene expression profile of candidate genes 62

2.2.15 Orthologs of candidate genes 62

2.3 Results 63

2.3.1 The FDG assay can be used to monitor the β-galactosidase activity of both plant apoplastic fluid BGAL1 and *E coli* LacZ 63

2.3.2 A streamlined workflow was developed for the screen for BGAL1 inhibitor mutants 66

2.3.3 A simulation of random transposon insertion coverage in the genome was used to determine the number of colonies to screen for BGAL1 inhibitor mutants 67
2.3.4 An optimised plating method enabled screening of BGAL1 inhibitor mutant colonies 68
2.3.5 A high throughput FDG assay was used to validate candidate BGAL1 inhibitor mutants 70
2.3.6 Transposon insertion sites in *bim* mutants were identified by inverse PCR followed by sequencing 72
2.3.7 The screen for BGAL1 inhibitor mutants identified candidate genes required for BGAL1 inhibitor production 74
2.3.8 The combination of gene expression profiles and comparative genomics pinpointed important candidate genes 76

2.4 Discussion 80
2.4.1 The screening of enzyme activity and inhibition 80
2.4.2 The identification of transposon insertion sites 82
2.4.3 The candidate genes involved in BGAL1 inhibitor production 83
2.4.4 The saturation screen for identifying gene candidates 86
2.4.5 Conclusion 87

Chapter 3: The characterisation of genes involved in BGAL inhibitor production 88

3.1 Introduction 88

3.2 Materials and methods 91
3.2.1 Bacterial strains, culture conditions and media 91
3.2.2 Generation of gene expression constructs 93
3.2.3 Generation of gene deletion mutants 95
3.2.4 Triparental mating conjugation for transformation of *P. syringae* 99
3.2.5 BGAL inhibitor production *in vitro* 99
3.2.6 FDG assay for β-galactosidase activity 100
3.2.7 RT-PCR 101
3.2.8 Luminescence reporter *in planta* 102
3.2.9 Bacterial growth curve *in vitro* 103
3.2.10 Infection assays 103
3.2.11 Apoplastic fluid extraction 104
3.2.12 Lectin blot 104
3.2.13 GC content mapping 105
3.2.14 Phylogenetic analysis 105
3.2.15 Full-length sequence of PSPTO0834 110

3.3 Results 111
3.3.1 PSPTO0834 and PSPTO0835 from the *bim* gene cluster confer inhibitor production in *P. syringae* and *E. coli* 111
3.3.2 L-aspartate oxidase (NadB) is involved in inhibitor production 114
3.3.3 Other selected *bim* candidate genes were not required for inhibitor production 115
3.3.4 Homologs of the alcohol dehydrogenase PSPTO0834 are present in different putative biosynthetic gene clusters for iminosugars in wide range of bacterial groups 117
3.3.5 The alcohol dehydrogenase GutB1 could not replace the function of PSPTO0834 in the *bim* cluster for BGAL inhibitor production 121
3.3.6 The *bim* cluster likely originated in another bacterial group and was acquired by the ancestor of *Pseudomonas* group through horizontal gene transfer 122
3.3.7 The presence of the *bim* cluster is distributed across major phylogroups of *P. syringae* 123
3.3.8 The *bim* cluster distribution is not limited to strains with modified viosamine in flagellin glycosylation 125
3.3.9 The *bim* cluster was horizontally transferred among *P. syringae* strains 125
3.3.10 The transfer of the *bim* cluster was likely mediated by a site-specific integrase targeting tRNA-lysine site 126
3.3.11 The *bim* cluster has a conserved *hrp* box promoter unique to *P. syringae* 128
3.3.12 Expression of the *bim* cluster and inhibitor production is regulated by type III secretion system regulators 130

3.3.13 The expression of the *bim* cluster and inhibitor production followed T3SS induction in the early hours of infection *in planta* 132
3.3.14 HtpX plays a role in the induction of *hrpL* and *bim* cluster expression 134
3.3.15 Deletion of the *bim* cluster did not affect bacterial growth *in vitro* 136
3.3.16 Deletion of the *bim* cluster did not result in a robust virulence phenotype *in planta* 137
3.3.17 The production of BGAL inhibitor altered the profile of galactose-containing glycoprotein in the plant apoplast during infection 140

3.4 Discussion 142
3.4.1 The biosynthesis of the inhibitor 142
3.4.2 The evolution of inhibitor production 147
3.4.3 The regulation of inhibitor production 153
3.4.4 The role of inhibitor production during infection 155
3.4.5 Conclusion 161

Chapter 4: The characterisation of the BGAL inhibitor 162

4.1 Introduction 162

4.2 Materials and methods 164
4.2.1 Bacterial strains, culture conditions and media 164
4.2.2 *hrpL* promoter activity luminescence reporter assay 165
4.2.3 Apoplastic fluid extraction 165
4.2.4 FDG assay for β-galactosidase activity 166
4.2.5 Relative quantification of inhibitor concentration 167
4.2.6 BGAL inhibitor production *in vitro* 167
4.2.7 Inhibitor purification by cation exchange resin 168
4.2.8 Inhibitor desalting 168
4.2.9 Preparative HPLC (High Performance Liquid Chromatography) 169
4.2.10 LC-MS (Liquid Chromatography – Mass Spectrometry) 170
4.2.11 GC-MS (Gas Chromatography – Mass Spectrometry) 171
4.2.12 Enzymatic assay for different glycosidase activities 172

4.3 Results 174
4.3.1 Inhibition assay with dilution series of inhibitor can be used for relative quantification of inhibitor concentration in the sample 174
4.3.2 Inhibitor production in *P. syringae* increased in type III secretion system inducing media 175
4.3.3 Overexpression of *bim* cluster increased inhibitor production in *P. syringae* 178
4.3.4 *E. coli* strain carrying the *bim* cluster also produced the inhibitor at high level 179
4.3.5 Cation exchange resin can be used to partially purify the inhibitor 181
4.3.6 Elution of the inhibitor from cation exchange resin can be done at high pH 183
4.3.7 The inhibitor is stable in acidic pH but sensitive to basic pH treatment 185
4.3.8 The inhibitor can also be eluted from cation exchange resin using HCl 187
4.3.9 Concentrated HCl treatment can be used to desalt eluate from cation exchange resin 189
4.3.10 Reverse-phase HPLC with C18 column was not suitable for separation of the inhibitor 190
4.3.11 Fmoc derivatisation can be used to modify the inhibitor 192
4.3.12 LC-MS can be used to detect Fmoc-derivatised iminosugars 194
4.3.13 LC-MS can be used to detect FMOC-derivatised inhibitor from *P. syringae* 197
4.3.14 LC-MS can be used to detect FMOC-derivatised inhibitor from *E. coli* expressing the *bim* cluster 200
4.3.15 HILIC HPLC can be used to isolate the inhibitor 202
4.3.16 GC-MS can be used to detect the TMS-derivatised inhibitor from *P. syringae* 205
4.3.17 GC-MS can be used to detect the TMS-derivatised inhibitor from *E. coli* expressing the *bim* cluster 207
4.3.18 Analysis of the putative GC-MS spectrum of the inhibitor 209
4.3.19 The partially purified inhibitor showed inhibitory activity against multiple glycosidases 210

4.4 Discussion 213
4.4.1 Optimising inhibitor production 213

4.4.2 Optimising inhibitor purification 216
4.4.3 Characterising inhibitor specificity 218
4.4.4 Elucidating inhibitor identity 220
4.4.5 Conclusion 224

Chapter 5: General discussion 225

5.1 Identification of genes required for inhibitor production 227

5.2 Biosynthesis of the inhibitor 228

5.3 Evolution of inhibitor production 231

5.4 Regulation of inhibitor production 233

5.5 Biological roles of the inhibitor 234

5.6 Molecular identity of the inhibitor 239

5.7 Conclusion 241

References 243

Appendix 1 260

Appendix 2 261

Chapter 1: Introduction

Plant pathogens pose a serious threat to global food security and economy (Strange and Scott, 2005). With the ever-growing world population, food production has to increase to keep up with demand while, at the same time, a considerable fraction of crops is devastated by plant diseases (Oerke, 2006; Savary *et al.*, 2019). Hence, there is a need to develop measures to combat plant pathogens and prevent or limit their impact on crop production. To achieve this, we need to better understand the biology of plant-pathogen interactions and how adapted microbes cause diseases on plants.

My research focuses on molecular mechanisms of the interaction between plants and bacterial pathogens, using *Pseudomonas syringae* as a model pathogen. During infection, microbial plant pathogens colonise the apoplast, an extracellular space in the plant tissue. This apoplastic interface between plants and bacteria is an important site where various factors are secreted to participate in a dynamic interaction that determines the outcome of the infection. Specifically, this work investigates the recently discovered production of an inhibitor of a plant defence-related β-galactosidase by *P. syringae*. I aim to determine the molecular identity of this inhibitor, characterise its biosynthetic and regulatory pathways and unravel biological roles of this inhibitor production and its evolution. This introduction will cover the biology and virulence mechanisms of *P. syringae*; plant immune responses and manipulations by pathogens; the roles of apoplastic glycans, glycosidases and inhibitors in plant-pathogen interactions; the discovery of BGAL1 and the production its inhibitor; and, finally, iminosugar glycosidase inhibitors.

1.1 *Pseudomonas syringae* as model plant pathogenic bacteria

1.1.1 *P. syringae* biology

Pseudomonas syringae is one of the most important and best studied plant pathogenic bacteria (Mansfield *et al.*, 2012). *P. syringae* is a ubiquitous pathogen that can be found not only in plants or seeds but also various environmental habitats, soil and water, thus the disease transmission can be achieved through various routes such as infected plant materials, insect vectors and environmental aerosols (Arnold and Preston, 2019) (Figure 1.1). *P. syringae* strains can infect a broad range of plants, from crops such as tomato, bean and kiwi, to trees such as horse chestnut, causing different symptoms including water soaking, chlorosis, necrosis and cankers (Arnold and Preston, 2019) (Figure 1.1). As a species complex, the *P. syringae* phylogenetic group is highly diverse and may be classified into multiple closely related species, such as *P. syringae*, *P. savastanoi*, *P. coronafaciens* and *P. amygdali* (Gomila *et al.*, 2017). The diversity of strains within the *P. syringae* species complex may also be classified into different relatively host-specific pathovars and grouped into phylogroups according to phylogenetic relationships (O'Brien, Thakur and Guttman, 2011; Gomila *et al.*, 2017). In a broader taxonomic context, *P. syringae* is part of the gamma-proteobacteria group and belongs to the genus *Pseudomonas*, which also contains human pathogens such as *P. aeruginosa* and soil bacteria such as *P. fluorescens*, some of which are plant pathogens and plant growth promoting bacteria (Silby *et al.*, 2011).

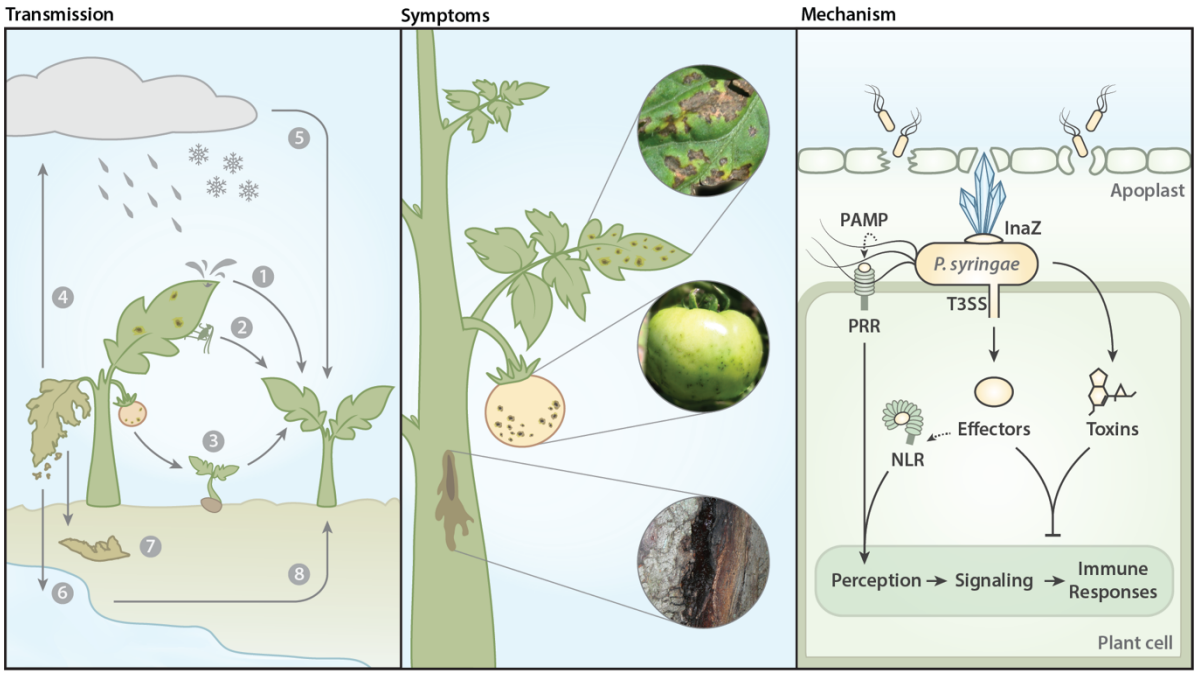


Figure 1.1. Overview of *P. syringae* biology. **Transmission** of *P. syringae* can be through (1) rain splash or aerosols, (2) insect vectors, (3) infected seed, (4-5) precipitates such as rain or snow, (6-8) contaminated water sources or soil from infected plant debris. **Symptoms** of *P. syringae* infection include chlorosis or necrosis on leaves or fruits and cankers or galls in woody tissue. **Mechanism:** *P. syringae* invades plant tissues through natural openings such as stomata and wounds, or frost damage from ice nucleating protein (InaZ) produced by some strains. *P. syringae* colonises the apoplast, an intercellular space in the plant tissue. Pathogen-associated molecular patterns (PAMPs), such as flagellin, can be recognised by plant pattern recognition receptors (PRRs) leading to activation of immune responses. To suppress plant immune response and promote colonization, *P. syringae* uses various virulence factors including effector proteins secreted into plant cells through type III secretion system (T3SS) and small molecules or toxins. Some effectors are recognised by plant nucleotide-binding domain leucine-rich repeat containing proteins (NLRs) leading to activation of immune responses. This figure was previously published in (Arnold and Preston, 2019).

The hemibiotrophic pathogenic lifestyle of *P. syringae* requires multiple adaptations for plant infection from the epiphytic stage through invasion and ultimately endophytic colonisation. For epiphytic colonisation, biofilm formation, stress tolerance and motility are important for their survival (Pfeilmeier, Caly and Malone, 2016; Xin, Kvitko and He, 2018). For invasion of plant tissues, the bacteria can gain access through natural openings, stomata and wounds. Some strains can secrete a phytotoxin coronatine to induce stomata opening, while some can produce an ice nucleation protein that promote frost damage on plants (Figure 1.1) (Pfeilmeier, Caly and Malone, 2016; Xin, Kvitko and He, 2018). Finally, colonisation of the apoplast inside the plant tissue requires manipulation of plant defences and physiology to

create a favourable environment for the pathogen, such as by increasing availability of water and nutrients (Pfeilmeier, Caly and Malone, 2016; Xin, Kvitko and He, 2018). At the late stage of infection, the host plant cells often die leading to necrosis lesions symptoms (Pfeilmeier, Caly and Malone, 2016; Xin, Kvitko and He, 2018). Some *P. syringae* strains can also spread systemically in the host plant (Steele *et al.*, 2010; Donati *et al.*, 2018).

During infection, *P. syringae* uses a variety of virulence factors, effector proteins and small-molecules, to subvert plant defence and promote colonisation (Ichinose, Taguchi and Mukaihara, 2013; Xin and He, 2013) (Figure 1.1). A major virulence mechanism is the type III secretion system, a needle-like protein assembly that delivers effector proteins directly into plant cells, where they localise to various subcellular compartments to modulate plant defence or physiology to promote colonisation. The effector targets can be involved in pathogen perception and defence signalling or cellular trafficking and secretion (Deslandes and Rivas, 2012; Lindeberg, Cunnac and Collmer, 2012). Another important class of virulence factors are small-molecule toxins, hormones such as indole acetic acid (IAA), and hormone-mimicking molecules such as coronatine that are secreted to modulate plant hormone signalling, metabolic enzymes or cause damage symptoms (Carol L. Bender, Alarcón-Chaidez and Gross, 1999).

The repertoire of effectors and toxins often shows polymorphisms and variable distribution among strains of *P. syringae* suggesting a dynamic evolution of this highly diverse species complex coevolving with a broad range of host plants (Hwang *et al.*, 2005; Baltrus *et al.*, 2011; Lindeberg, Cunnac and Collmer, 2012; Baltrus, McCann and Guttman, 2017b; Dillon *et al.*, 2017). Many effectors encoded in a strain also show a degree of functional redundancy (Kvitko *et al.*, 2009). The set of effectors possessed by a strain can reflect the selection for

gain and loss as a result of interactions between the bacterial effectors, their targets in plants and recognition by plants (Dillon *et al.*, 2019). In addition, many virulence factors, including the type III secretion system, effector proteins and toxin biosynthetic gene clusters, are encoded in pathogenicity islands that can be acquired through horizontal gene transfer to contribute to the evolution of the pathogen (Arnold, Pitman and Jackson, 2003).

In this work, I use *P. syringae* pv. *tomato* (*Pst*) DC3000 strain as a model of plant pathogenic bacteria. *Pst*DC3000 has been extensively studied and widely used in plant-pathogen interactions research (Preston, 2000; Xin and He, 2013). *Pst*DC3000 is a causative agent of bacterial speck disease in tomato and can also infect the model plant *Arabidopsis thaliana* (Preston, 2000). Infection of the model plant *Nicotiana benthamiana* requires a mutant strain of *Pst*DC3000 lacking the effector HopQ1, which is able to cause disease since the presence of HopQ1 would otherwise be recognised by the plant and trigger the hypersensitive response (Wei *et al.*, 2007). The genome of *Pst*DC3000 has also been sequenced (Buell *et al.*, 2003). Various virulence factors have been identified including the type III secretion system and at least 28 effectors as well as small molecules such as the phytotoxin coronatine and phytohormone IAA (Preston, 2000; Xin and He, 2013).

1.1.2 Toxins and small-molecule virulence factors

Different *P. syringae* strains can produce a different range of small-molecules or toxins as virulence factors during infection. These compounds belong to various structural classes, such as polyketides, β-lactams, lipopeptides and non-ribosomal peptides, and can act via various mechanisms, including regulating and suppressing plant defence hormone signalling, disrupting plant cell membranes and inhibiting plant metabolic enzymes (Bender, Alarcón-

Chaidez and Gross, 1999; Gross and Loper, 2009). Here compounds that are found in the *P. syringae* group as a whole are discussed, as not all compounds are produced by all strains.

Plant hormones and analogues

Plant hormone signaling can be targeted by *P. syringae* during infection by the production of several compunds such as Indole-3-acetic acid (IAA), a form of the phytohormone auxin, and coronatine, a structural mimic of jasmonyl-isoleucine, an active form of the phytohormone jasmonate. The actions of both compounds can vary but also converge on the suppression of plant defence mediated by another plant hormone, salicylic acid.

IAA is synthesised by *P. syringae* from a tryptophan precursor that goes through an indole-3-acetaldehyde intermediate which is then converted to IAA by an aldehyde dehydrogenase AldA (McClerklin *et al.*, 2017). IAA production functions in bacterial virulence because auxin signaling suppresses salicylic acid-mediated defence response in the plant (McClerklin *et al.*, 2017; Djami-Tchatchou *et al.*, 2020). In addition, IAA is also able to modulate expression of virulence-related gene expression in *P. syringae,* a mechanism proposed to coordinate gene expression at different infection stages. The suppression of the type III secretion system (T3SS) by IAA was proposed to occur at a late stage of infection when the bacterial population is already established so that T3SS-related genes are not required and IAA has accumulated to a higher level to have an effect (Djami-Tchatchou *et al.*, 2020).

Coronatine is a toxin composed of two parts, coronafacic acid (CFA) and coronamic acid (CMA), which are synthesised separately using enzymes encoded by two separate operons and then joined together by a coronafacate ligase enzyme (Geng *et al.*, 2014). CFA is derived from a polyketide synthesis (PKS) pathway, while CMA is derived from isoleucine through a

non-ribosomal peptide synthase (NRPS) (Geng *et al.*, 2014). Coronatine has been widely studied for its various roles in suppression of plant defence. For example, as a structural analogue of jasmonyl-isoleucine, coronatine can exploit jasmonic acid signaling to antagonise salicylic acid signaling which is a key mediator of plant defence response (Geng *et al.*, 2014). Another notable role of coronatine is the induction of stomatal opening to allow bacteria access into the plant tissue (Geng *et al.*, 2014).

Two other phytohormones, ethylene and cytokinin, are also produced by *P. syringae* strains to promote virulence during infection (Weingart *et al.*, 2001; Añorga *et al.*, 2020). Cytokinin production also leads to tumorigenesis in host plants (Añorga *et al.*, 2020).

Lipopeptides

Another class of toxins produced by *P. syringae* are lipopeptides, which include syringopeptin, syringomycin, syringotoxin, syringostatin and pseudomycin. Lipopeptides are composed of a hydrophobic fatty acid chain linked to a hydrophilic cyclic chain of amino acids, resulting in an amphiphilic property that allows them to interact with cell membranes to form pores and cause cell disruption (Bender, Alarcón-Chaidez and Gross, 1999; Gross and Loper, 2009). The structure of lipopeptides can be diverse but the peptide parts are generally synthesised by the operon-encoded non-ribosomal peptide synthase (NRPS), which is enzyme complex organised in modules that sequentially function for amino acid elongation or modification. Fatty acid tails are then transferred to the peptide to make lipopeptides (Bender, Alarcón-Chaidez and Gross, 1999; Gross and Loper, 2009).

A search for non-ribosomal peptide synthase gene clusters also identified the production of another type of linear lipopeptides, such as syringafactin, which is produced during infection.

These lipopeptides can function as surfactants to promote bacterial swarming motility and as humectants to increase water availability in the leaf microenvironments (Berti *et al.*, 2007; Hernandez and Lindow, 2019).

Antimetabolite toxins

Antimetabolite toxins, including phaseolotoxin, tabtoxin and mangotoxin, are a group of small peptide-derived toxins that inhibit enzymes involved in plant amino acid biosynthesis (Arrebola *et al.*, 2011).

Phaseolotoxin is a tripeptide of ornithine, alanine and homoarginine, linked to an inorganic sulfodiaminophosphinyl group, which can be cleaved to release the active octicidine group (Gross and Loper, 2009). This compound inhibits the ornithine carbamoyltransferase enzyme in the host plant required for arginine biosynthesis (Arrebola *et al.*, 2011). The toxin produces a symptom in the plant as a halo of chlorotic zone surrounding a necrotic infection site (Arrebola *et al.*, 2011). The biosynthesis of phaseolotoxin is mediated by the *pht* gene cluster (Gross and Loper, 2009).

Tabtoxin is produced as a conjugate of tabtoxine-β-lactam (TBL) and threonine, which can be hydrolysed by an aminopeptidase to release the the active TBL group that inhibits glutamine synthase enzyme in the host plant. This inhibition blocks the ammonia assimilation pathway leading to toxic ammonia accumulation that disrupts the thylakoid membrane and leads to chlorosis (Arrebola *et al.*, 2011). Biosynthesis of tabtoxin is mediated by the *tab/tbl* gene cluster encoding enzymes that takes an intermediate of lysine biosynthesis pathway as a precursor for production of TBL, which is then conjugated to threonine (Gross and Loper, 2009).

Mangotoxin inhibits the ornithine N-acetyl transferase enzyme in the host plant, which is required for arginine and ornithine biosynthesis. The toxin produces a symptom in the plant as necrotic lesions (Arrebola *et al.*, 2011). The structure of mangotoxin is unknown but is proposed to be an oligopeptide of a non-ribosomal origin. The biosynthesis of mango toxin is mediated by the *mbo* operon (Carrión *et al.*, 2012).

Proteasome inhibitor

Syringolin is a toxin that inhibits the plant proteasome required for salicylic acid-mediated immune signaling pathway. This immune suppression can promote colonisation of the bacteria from wounding sites and counter the stomatal closure defence response (Schellenberg, Ramel and Dudler, 2010; Misas-Villamil *et al.*, 2013). Syringolin structure is a hybrid of cyclic peptide-polyketide synthesised by a combination of non-ribosomal peptide synthase (NRPS) and polyketide synthase (PKS) pathways, encoded by the *syl* operon (Gross and Loper, 2009).

PhevamineA

PhevamineA is a small-molecule virulence factor that suppresses plant immune responses including callose deposition and cell death and impairs plant immune signal potentiation by polyamines and arginine (O'Neill *et al.*, 2018). PhevamineA structure contains a dipeptide and a polyamine. The biosynthetic gene cluster of Phevamine A is the *hsv* operon, which consists of three genes. The first gene encodes an amidino transferase that transfer the amidino group on a polyamine precursor spermidine. This step is followed by sequential condensation with two amino acids, valine and phenylalanine, catalysed by two ATP-grasp type enzymes (O'Neill *et al.*, 2018).

1.1.3 Virulence regulation

The regulation of virulence mechanisms in *P. syringae* involves a network of multiple regulatory systems (Xie, Shao and Deng, 2019; Xie *et al.*, 2020), summarised in Figure 1.2. The type III secretion system (T3SS) is the major pathogenicity and virulence factors of *P. syringae* that are induced during infection. The main genes involved in this system are found in a *hrp* pathogenicity island (Alfano *et al.*, 2000). At the core of T3SS regulation are the HrpR and HrpS enhancer binding proteins, which are encoded by the same operon and function together as a complex to activate the expression of *hrpL*, an alternative sigma factor (Hutcheson *et al.*, 2001). The expression of *hrpL* also requires the sigma factor RpoN (sigma54) (Hutcheson *et al.*, 2001). In turn, HrpL activates the expression of a set of genes having a HrpL binding site (*hrp* box) in their promoter, including type III secretion system components and effector proteins as well as additional uncharacterised genes (Ferreira *et al.*, 2006; Lindeberg *et al.*, 2006; Vencato *et al.*, 2006; Lam *et al.*, 2014; Mucyn *et al.*, 2014).

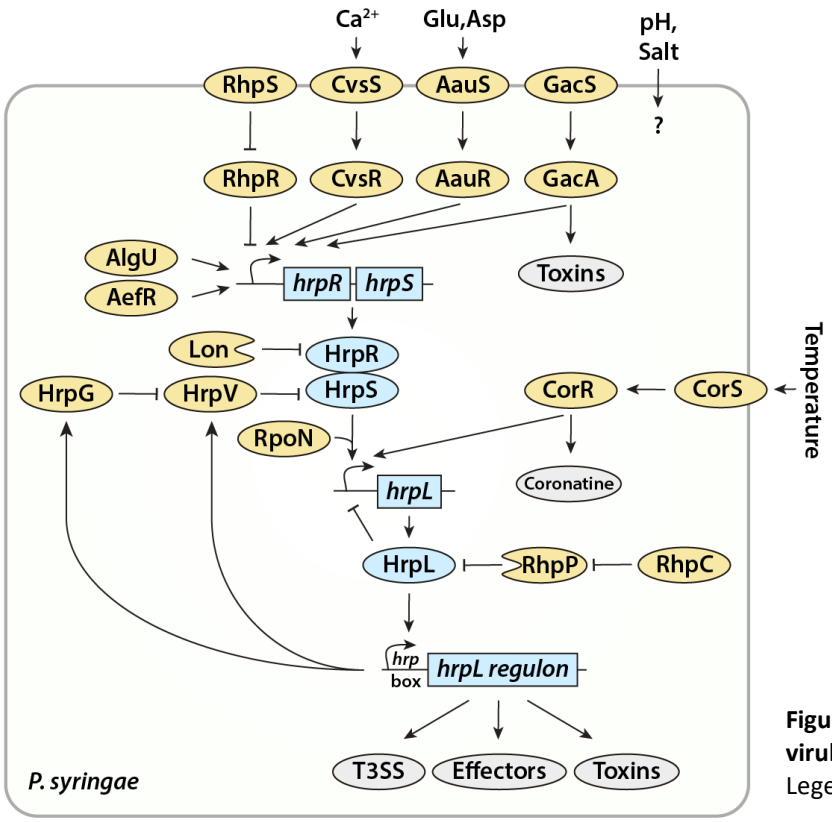


Figure 1.2. Regulation of virulence genes in *P. syringae*. Legend on the next page.

Figure 1.2. **Regulation of virulence genes in *P. syringae***. As key virulence regulators, **HrpR and HrpS** enhancer binding proteins, encoded by the same operon, function together as a complex to activate the expression of *hrpL*, an alternative sigma factor. The expression of *hrpL* also requires the sigma factor **RpoN** (sigma 54) (Hutcheson *et al.*, 2001). In turn, **HrpL** activates the expression of the *hrpL* regulon, a set of genes with HrpL binding site (*hrp* box) in their promoter, including type III secretion system (T3SS), type III-secreted effector proteins and toxin biosynthesis (Vencato *et al.*, 2006). The **RhpR/S** two-component system: RhpR represses *hrpR/S* operon. Upon virulence activation, RhpS dephosphorylates RhpR causing its release from the *hrpR/S* promoter and derepress its expression (Deng *et al.*, 2014). The **GacA/S** two-component system is a regulator of various virulence factors including toxins. GacA can also activate expression of *hrpR/S* (Chatterjee *et al.*, 2003). The **CorR/S** two-component system, the main regulator of coronatine toxin production. CorR can also activate expression of *hrpL* (Sreedharan *et al.*, 2006). The CorR/S is also activated at low temperature (pathovar specific) (Smirnova *et al.*, 2002). Some external signals are known to activate two components systems that can modulate the T3SS expression. For example, amino acids glutamate (Glu) and aspartate (Asp) can activate the **AauS/R** system (Yan *et al.*, 2020), while calcium ion can activate the **CvsS/R** system (Fishman *et al.*, 2017), which in turns upregulate HrpR/S. The *hrp* system is also subjected to negative feedback regulation. For example, **HrpL** can negatively regulate its own gene expression by binding to the promoter of *hrpJ*, an adjacent gene that shares the same upstream promoter region with *hrpL* (Waite *et al.*, 2017). **HrpV**, a protein encoded from the same *hrp* pathogenicity island, functions as a negative regulator by binding to HrpS to inhibit activation of *hrpL* (Preston *et al.*, 1998). In turn, **HrpG** can bind to HrpV and suppress its negative regulation action (Jovanovic *et al.*, 2011). As negative regulators of T3SS, **Lon** protease functions through degradation of HrpR (Bretz *et al.*, 2002), while **RhpP** protease functions through degradation of HrpL (Li *et al.*, 2019). **RhpC** interacts with RhpP and directs it to the periplasm, suppressing its negative impact on T3SS. The regulation of T3SS also crosstalks with regulators of other pathways. For example, **AlgU**, an extracytoplasmic function sigma factor controlling expression of genes involved in alginate production and osmotic and oxidative stress response, can also upregulate expression of *hrpR*/S (Markel *et al.*, 2016). **AefR**, which control quorum sensing signal production, is also involved in the activation of *hrpR/S* expression (Deng *et al.*, 2009). pH and salt can also impact the regulation but the mechanism is unclear (Mole *et al.*, 2007).

Several two-component systems act upstream of the core HrpR/S/L system to play a role in T3SS regulation. These two-component systems are encoded by two genes in the same operon and consist of a membrane-bound sensor histidine kinase that perceives a signal and phosphorylates a response regulator that can bind DNA and modulate target gene expression (Capra and Laub, 2012). The major regulator of HrpR/S is the RhpR/S system. RhpR is a response regulator that binds to a specific site upstream of *hrpR/S* operon and represses its expression. Upon virulence activation, RhpS, a membrane-bound sensor kinase, dephosphorylates RhpR causing its release from the promoter and derepression of expression (Deng *et al.*, 2014). Another important system is the GacA/S system, a regulator of various virulence factors including the biosynthesis of various toxins and secondary metabolites and secretion of extracellular polysaccharides and enzymes (Chatterjee *et al.*, 2003; Mole *et al.*,

2007). GacS is a sensor kinase, while GacA is a response regulator that modulates expression of the regulated genes including *hrpR/S* and *hrpL* (Heeb and Haas, 2001; Ortiz-Martín *et al.*, 2010). The CorR/S system, the main regulator of coronatine toxin production, also influences T3SS expression in a reciprocal manner. The regulator CorR can activate expression of *hrpL* while HrpL can also activate *corR* expression (Sreedharan *et al.*, 2006).

Although the exact signals that many of these sensors and regulators perceive are unknown, the T3SS can be activated by a combination of factors such as presence of sucrose or fructose, low pH, low osmolarity, and lack of complex carbon and nitrogen sources, which are believed to reflect the conditions in the plant apoplast (Mole *et al.*, 2007). On the other hand, other factors, such as glucose and NaCl, can repress T3SS expression. In addition, some external signals are known to activate two components systems that can modulate the T3SS expression. For example, amino acids glutamate and aspartate can activate the AauS/R system (Yan *et al.*, 2020), while calcium ion can activate the CvsS/R system (Fishman *et al.*, 2017), which in turns upregulate HrpR/S. In addition, artificial minimal media designed to mimic these conditions are also useful tools for inducing virulence gene expression *in vitro* (Bronstein *et al.*, 2008; Kim *et al.*, 2009).

The *hrp* system is also subjected to negative feedback regulation. For example, HrpL can negatively regulate its own gene expression by binding to the promoter of *hrpJ*, an adjacent gene that shares the same upstream promoter region with *hrpL* (Waite *et al.*, 2017). In addition, HrpV, a protein encoded from the same *hrp* pathogenicity island, functions as a negative regulator by binding to HrpS to inhibit activation of *hrpL* (Preston *et al.*, 1998; Jovanovic *et al.*, 2011) In turn, HrpG can bind to HrpV and suppress its negative regulation action (Jovanovic *et al.*, 2011). The negative feedback control of the *hrp* system was proposed

to be a mechanism to fine tune the expression level of T3SS and effector proteins to minimise elicitation of plant defence (Waite *et al.*, 2017).

Proteases are also involved in the regulation of T3SS through proteolytic cleavage of key regulatory proteins. Lon protease functions as a negative regulator of T3SS through degradation of HrpR (Bretz *et al.*, 2002), while another protease, RhpP, acts as a negative regulator of T3SS through degradation of HrpL (Li *et al.*, 2019). RhpP is encoded by the same operon as RhpC, a cytoplasmic protein that interacts with RhpP and direct it to the periplasm, suppressing its negative impact on T3SS (Li *et al.*, 2019).

The regulation of T3SS also crosstalks with regulators of other pathways. For example, AlgU, an extracytoplasmic function (ECF) sigma factor controlling expression of genes involved in alginate production and osmotic and oxidative stress response, can also upregulate expression of *hrpR*/S (Markel *et al.*, 2016). Another regulator, AefR, which controls quorum sensing signal production, can also activate *hrpR/S* expression (Deng *et al.*, 2009).

Apart from the T3SS, production of various small-molecule virulence factors or toxins are also regulated by shared regulators. For example, HrpL also regulates the biosynthetic genes for coronatine (Sreedharan *et al.*, 2006), phevamineA (O'Neill *et al.*, 2018) and IAA (Fouts *et al.*, 2002). The GacA/S system controls production of syringomycin, syringolin, tabtoxin, and mangotoxin (Heeb and Haas, 2001; Carrión *et al.*, 2014). Mangotoxin production is additionally regulated by a product of the MgoA operon (Carrión *et al.*, 2014). The CorR/S system, together with CorP, which is encoded by an adjacent gene, controls the expression of biosynthetic genes for coronatine production (Smirnova *et al.*, 2002). In addition, production of many toxins, such as coronatine and phaseolotoxin, are temperature-responsive, induced at low temperature (18-20°C) but suppressed at high temperature (28°C) (Ullrich *et al.*, 1995;

Aguilera *et al.*, 2017). The temperature responsive control of coronatine is mediated by the CorR/S/P system (Smirnova *et al.*, 2002). It was proposed that the CorR response regulator binds DNA to activate regulated genes in a temperature dependent manner or that the CorS membrane-bound sensor kinase changes its conformation in a temperature dependent manner (Smirnova *et al.*, 2002).

1.2 Plant-pathogen interactions and the apoplastic battlefield

1.2.1 Plant immune system and manipulation by pathogens

During infection, microbial plant pathogens colonise the apoplast, an intercellular compartment in the plant tissue. This interface between plant and pathogen is an important site where defence responses occur and the factors produced from both sides come into play (Doehlemann and Hemetsberger, 2013). In this interaction, plants have evolved immune responses against pathogens and pathogens have evolved to manipulate plant processes for infection (Jones and Dangl, 2006; Dodds and Rathjen, 2010). Plant resistance mechanisms against pathogens can be through various pathogen recognition mechanisms as well as loss of susceptibility exploited by pathogens to promote infection (Kourelis and van der Hoorn, 2018).

Upon infection, plants can perceive the presence of invading pathogens by recognising conserved pathogen-associated molecular patterns (PAMPs), such as bacterial flagellin and fungal chitin, using a range of cognate pattern recognition receptors (PRRs) on the cell surface leading to a signalling cascade that activates PAMP-triggered immunity (PTI) (Saijo, Loo and Yasuda, 2017; Tang, Wang and Zhou, 2017). The plant defence responses include generation of reactive oxygen species (ROS burst) (Lamb and Dixon, 1997), cell wall reinforcement and

callose deposition (Bacete *et al.*, 2018), and secretion of pathogenesis-related (PR) proteins (van Loon, Rep and Pieterse, 2006), phytoalexin metabolites (Ahuja, Kissen and Bones, 2012) and antimicrobial peptides (Campos *et al.*, 2018). To subvert plant defence and promote colonisation, pathogens secrete a diverse range of virulence factors, including effector proteins and small-molecule toxins, to modulate different processes in both inside the host plant cells (eg. plant defence signaling and secretion) and in the apoplast (eg. plant hydrolase inhibition) (Toruño, Stergiopoulos and Coaker, 2016). As a counter measure, plants also evolved various nucleotide-binding leucine-rich repeat receptors (NLRs) to recognise specific pathogen effectors and mount effector-triggered immunity (ETI), a more robust defence response that usually leads to the hypersensitive response (HR), a programmed cell death that limits pathogens from spreading further (Cui, Tsuda and Parker, 2015). The outcome of the infection is thus determined by dynamic interactions between various factors produced by both plants and pathogens (Figure 1.1).

In addition to manipulating players of plant defences by various effectors, pathogens have also evolved ways to avoid eliciting plant immune responses in the first place by evading recognition. For example, bacteria can prevent their flagellin from being detected through various means such as disrupting recognition by plant receptors with amino acid polymorphisms in the immunogenic epitope, masking the epitope with post-translational modification (eg. glycosylation), degrading free flagellin with secreted protease or downregulating production of flagellin during infection (Abramovitch, Anderson and Martin, 2006; Trda *et al.*, 2015). Another example way of avoiding effector-triggered immunity is the excision of a genomic island encoding a recognised effector upon exposure to plant hypersensitive response (Pitman *et al.*, 2005). As an important determinant of plant-pathogen

interaction, pathogens have adopted various strategies to overcome immunity in host plants during infection by both countering and evading plant defence.

1.2.2 Apoplastic hydrolases and inhibitors in plant-pathogen interactions

In the apoplast, hydrolases are prominent players used by both plants and pathogens during infection. Consequently, inhibitors of these hydrolases also emerged as an important strategy used by both to subvert each other's responses in the evolutionary arms race (Misas-Villamil and van der Hoorn, 2008; Jashni *et al.*, 2015). For plant defence, hydrolases function in various processes ranging from direct pathogen attack to participating in pathogen perception, signalling and plant immune priming (Van Der Hoorn and Jones, 2004; Thomas and van der Hoorn, 2018). Many hydrolases induced during infections are proteases, or glycosidases such as chitinases and glucanases (van Loon, Rep and Pieterse, 2006). Likewise, plant pathogens also secrete a range of hydrolases such as glycosidases to degrade plant cell wall components or proteases to degrade plant defence proteins (Misas-Villamil and van der Hoorn, 2008; Jashni *et al.*, 2015).

There are several well-studied cases of interactions between plant secreted immune proteases and inhibitor proteins secreted by pathogens. For example, plant papain-like cysteine proteases, including tomato C14, Rcr3 and Pip1, have important functions in plant immunity (Misas-Villamil, van der Hoorn and Doehlemann, 2016) and have become common targets for inhibition by effectors from various pathogens, including EpiC1 and EpiC2b from *Phytophthora infestans* (Tian *et al.*, 2007) , Pit2 from *Ustilago maydis* (Mueller *et al.*, 2013) and Cip1 from *Pseudomonas syringae* (Shindo *et al.*, 2016). In addition, *P. infestans* can suppress plant immune proteases using an effector Avrblb2 which prevents secretion of C14 into the apoplast (Bozkurt *et al.*, 2011). Furthermore, tomato plants can also use the

interaction between plant protease Rcr3 and fungal protease inhibitor Avr2 as a detection mechanism for immune response activation, using the Cf2 receptor-like protein that recognises the Avr2-Rcr3 complex (Rooney *et al.*, 2005). In this system, Rcr3 acts as a decoy for perception of fungal Avr2, which also inhibits another plant immune protease Pip1, a paralog of Rcr3 (Shabab *et al.*, 2008). Another prominent group are subtilases, a large family of proteases (Schaller *et al.*, 2018) that includes members such as SBT3.3 with a role in immune priming (Ramírez *et al.*, 2013) and P69B with a role in activating immune protease RCR3 (Paulus *et al.*, 2020) and cleaving a pathogen-secreted effector protein to release elicitor peptides that activate immune response (Wang *et al.*, 2021). To subvert plant defence, *Phytophthora infestans* secretes Epi1 and Epi10 inhibitors to inhibit plant subtilases (Tian *et al.*, 2004; Tian, Benedetti and Kamoun, 2005). The interplay between hydrolases and inhibitors is, therefore, a crucial part of the infection and another hotspot in plant-pathogen interactions.

1.3 Glycans, glycosidases and inhibitors in plant-pathogen interactions

Glycans are ubiquitously found on a diverse range of extracellular components of the cell. In the apoplast, glycosidases that are active on these glycans and inhibitors that modulate these enzyme activities play important roles in plant-pathogen interactions. Major glycan players involved in this interaction include polysaccharides, glycolipids, glycoproteins, and glycoside conjugates of metabolites, each of which are discussed below. The interactions discussed are summarised in Figure 1.3.

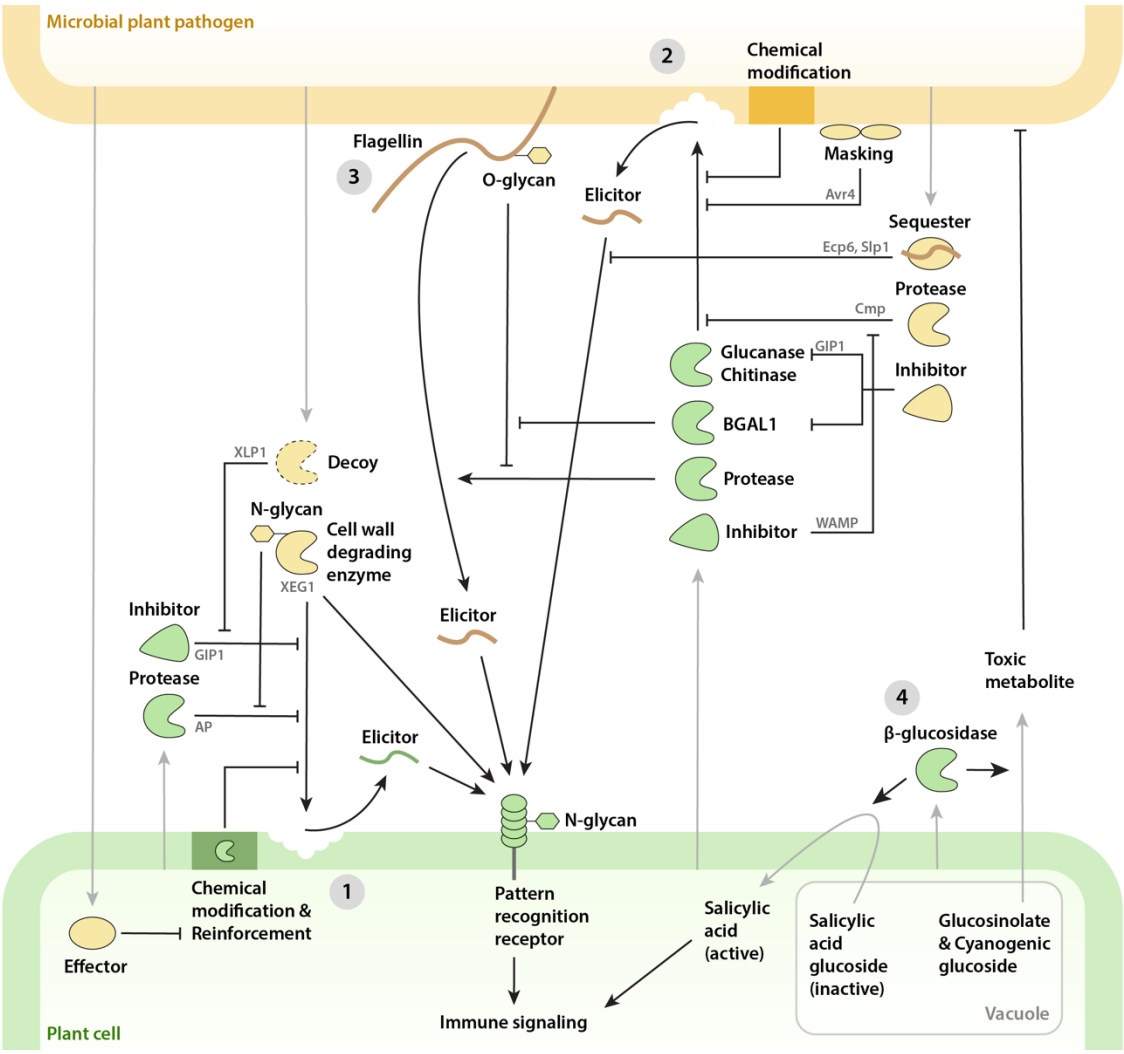


Figure 1.3. **Glycans, glycosidases and inhibitors in plant-pathogen interactions. (1)** Glycosidases are secreted by pathogens to function as cell wall degrading enzymes (Kubicek, Starr and Glass, 2014). To counter these glycosidases, plants can secrete proteinaceous inhibitors to inhibit their activity (Ma *et al.*, 2017) or protease to degrade the glycosidases (Xia *et al.*, 2020). Pathogens can also counter these responses by secreting a decoy to sequester the plant inhibitor (Ma *et al.*, 2017) or use protein N-glycosylation to protect from plant protease (Xia *et al.*, 2020). Plant cell wall polysaccharides are also chemically modified to be more resistant to hydrolysis by glycosidases. Cell wall reinforcement by additional polysaccharide deposition also occurs during defence responses and can be inhibited by some effector proteins from pathogens (Voigt, 2014). In addition, plants can detect invasion by recognising oligosaccharide fragments released from cell wall degradation (Hou *et al.*, 2019) as well as some glycosidases secreted by the pathogen (Ma *et al.*, 2015), using pattern recognition receptors that lead to immune signaling. N-glycosylation of plant pattern recognition receptors is essential for their function (Häweker *et al.*, 2010). **(2)** Plants secrete several glycosidases such as chitinase and glucanase targeting microbial polysaccharides to cause damage and release elicitors that can induce immune response (van Loon, Rep and Pieterse, 2006). Pathogens can counter this attack (Rovenich, Zuccaro and Thomma, 2016) by chemical modification resistant to plant glycosidase, secreting proteins that can bind to the polysaccharide and mask it from plant glycosidase, secreting proteins that can sequester the released elicitors, secreting inhibitors to inhibit plant glycosidases or secreting proteases to cleave plant glycosidase (Naumann, Wicklow and Price, 2011). Plants can also secrete an inhibitor to prevent this protease from cleaving plant glycosidase (Slavokhotova *et al.*, 2014). **(3)** Elicitor peptides from bacterial flagellin can be released by proteolytic cleavage. O-glycosylation of flagellin contributes to protection of the flagellin protein. Plant β-galactosidase BGAL1 facilitates the release of elicitor peptides from flagellin that is glycosylated with modified viosamine. BGAL1 is also inhibited by an inhibitor from bacteria. (Buscaill *et al.*, 2019) **(4)** Plant apoplastic β-glucosidase can cleave the inactive glucoside conjugate of salicylic acid hormone to release ana active form used in immune signaling (Ding and Ding, 2020). β-glucosidase cleavage also activates glucosinolates and cyanogenic glucosides to release toxic metabolites to attack pathogens (Morant *et al.*, 2008).

1.3.1 Polysaccharides

Polysaccharides form parts of the cell wall and extracellular matrix of both plants and microbial pathogens. As discussed below, there are general patterns of strategies used by both plants and pathogens to manipulate the interactions involving the extracellular polysaccharides. The structural and protective roles of polysaccharides make them major targets for attack by glycosidases to cause damage to the cell. The abundance of polysaccharides also makes them targets for pathogen nutrition. To counter glycosidase actions, common strategies are secretion of glycosidase inhibitors or chemical modifications to escape from glycosidase activities. As a counter measure, proteases can be deployed to degrade these secreted proteins. Finally, the cleavage by glycosidases releases oligosaccharide fragments which can be recognised by plant pattern recognition receptors to trigger immune response. The recognition is also expanded to the proteins that are involved.

Plant polysaccharides

The plant cell wall is composed of a complex mixture of polysaccharides, which includes cellulose microfibrils cross-linked with hemicelluloses and embedded in a pectin matrix. Cellulose is a homopolymer of β-1,4-linked glucan. Hemicelluloses are a group of heteropolysaccharides that contains a backbone of β-1,4-linked sugars, including glucose, xylose and mannose, and side chains of differently linked sugars. The most abundant hemicellulose is xyloglucan which has a β-1,4-linked glucan backbone with side chains of variable sequentially linked xylose, galactose and fucose. Another abundant hemicellulose is xylan which has a β-1,4-linked xylose backbone with side chains of arabinose and glucuronic acid. Pectins are polysaccharides that contain galacturonic acid, which can be on its own in homogalacturonan or as a polymer of alternating rhamnose and galacturonic acid in

rhamnogalacturonan. Side chains on rhamnogalacturonan may contain galactan, arabinan or arabinogalactan. (Vorwerk, Somerville and Somerville, 2004; Anderson and Kieber, 2020).

The plant cell wall is dynamically remodeled, for example during growth and development, and its turnover is mediated by a range of glycosidases that are active on different types of polysaccharide components (Minic and Jouanin, 2006; Minic, 2008). In some cases, plant cell wall remodeling enzymes are also susceptibility factors exploited by pathogens. For example, endoglucanase and polygalacturonase responsible for cell wall loosening during fruit ripening are induced by necrotrophic pathogens (Bellincampi, Cervone and Lionetti, 2014).

The plant cell wall also plays an important role during infection (Vorwerk, Somerville and Somerville, 2004; Malinovsky, Fangel and Willats, 2014). One notable plant defence response is the reinforcement of the cell wall with deposition of callose, a β-1,3-linked glucan (Voigt, 2014). This callose deposition is also dynamically modulated by actions of callose synthase and β-glucanase (Wu *et al.*, 2018). Furthermore, the cell wall can also be reinforced by deposition of phenolics such as lignin and crosslinking of glycoproteins such as hydroxyproline-rich proteins (Hématy, Cherk and Somerville, 2009). To counter plant defences, pathogen effector proteins, such as AvrPto from *P. syringae,* can suppress the expression of hydroxyproline-rich proteins and callose deposition, via inhibition of immune kinase signaling (Hauck, Thilmony and He, 2003; Xiang *et al.*, 2008). In addition, chemical modifications of cell wall polysaccharides also contribute to their strength. For example, methylesterification of pectin makes it more resistant to cleavage by cell wall modifying enzymes and the degree of this modification during infection is dynamically controlled by plant pectin methylesterase and the endogenous inhibitor protein PMEI (Malinovsky, Fangel and Willats, 2014; Bacete *et al.*, 2018).

Glycosidases are secreted by pathogens to function as cell wall degrading enzymes, including β-glucosidase and endoglucanase targeting cellulose; β-xylosidase, xylanase and xyloglucanase targeting hemicellulose; and polygalacturonase and pectin lyase targeting pectin (Kubicek, Starr and Glass, 2014). To counter these glycosidases, plants can secrete a range of proteinaceous inhibitors to inhibit their activity, including polygalacturonase-inhibiting protein (PGIP), pectin lyase-inhibiting protein (PNLIP), xylanase-inhibiting protein (XIP) and xyloglucan endoglucanase-inhibiting protein (XEGIP) (Juge, 2006). Furthermore, plants have also evolved ways to detect pathogen invasion and cell wall degradation through receptors that recognise released oligosaccharide fragments as damage-associated molecular patterns (DAMPs) such as oligogalacturonan, cellooligomers and xyloglucan (Hou *et al.*, 2019). Another interesting example of the arms race between plant and pathogen glycosidase is the effector XEG1, a xyloglucanase secreted by *Phytophthora sojae*. Soybean host plants have evolved three ways to counter this effector, which are recognition of XEG1 by a cell surface receptor (Ma *et al.*, 2015), inhibition of XEG1 by GIP1 inhibitor protein (Ma *et al.*, 2017) and degradation of XEG1 by a secreted protease (Xia *et al.*, 2020). In turn, the pathogen has also evolved two additional ways to counter plant defence, which are the sequestration of plant GIP1 inhibitor by a paralogous protein to XEG1 that bind the inhibitor more tightly (Ma *et al.*, 2017) and N-glycosylation of XEG1 that protects it from plant protease (Xia *et al.*, 2020).

Pathogen polysaccharides

Various types of polysaccharides are found in the cell wall and extracellular matrix of different microbial plant pathogens. The fungal cell wall consists of mainly chitin and β-glucan (Geoghegan, Steinberg and Gurr, 2017). The oomycete cell wall contains mainly β-glucan (Raaymakers and Van Den Ackerveken, 2016). The bacterial cell wall contains peptidoglycan,

which is composed of glycan chains of N-acetylglucosamine and N-acetylmuramic acid that are crosslinked by short peptides (Dworkin, 2014). Bacteria also produce exopolysaccharides (EPS) to form a matrix around the cell and biofilm (Heredia-Ponce *et al.*, 2020). *P. syringae* mainly produces alginate (Yu *et al.*, 1999), levan (Laue *et al.*, 2006), cellulose (Arrebola *et al.*, 2015) and Psl-like polysaccharide (Heredia-Ponce *et al.*, 2020). The polysaccharides matrix can have a role in protection of the bacteria and attachment to surface, contributing to epiphytic fitness (Yu *et al.*, 1999). EPS can also suppress plant defences by chelating calcium ions required for plant defence signaling (Aslam *et al.*, 2008). In addition, the outer membranes of Gram-negative bacteria also contain lipopolysaccharides (LPS), consisting of lipid A and chains of conserved glycans and variable O-antigen glycans (Laue *et al.*, 2006). LPS can act as a protective shield against passage of antimicrobials and enhance tolerance (Erbs and Newman, 2012). The long chain O-antigen glycan of LPS in *Xylella fastidiosa* was also shown to delay its recognition by plants and defence responses (Rapicavoli *et al.*, 2018).

As a defence against invading pathogens, plants secrete a range of hydrolases to digest pathogen cell wall components, including glucanase and chitinase that are active against filamentous pathogens or lysozyme-like hydrolase and some chitinases that are active against bacterial peptidoglycan (van Loon, Rep and Pieterse, 2006; Liu *et al.*, 2014). In addition to inflicting damage, plants can also recognise pathogen-associated molecular patterns (PAMPs) released from fragments of microbial cell wall components, such as peptidoglycan, chitin, β-glucan and LPS, using a range of cognate pattern recognition receptors (PRRs) to activate signaling of defence responses (Newman *et al.*, 2013).

To avoid eliciting plant defences, pathogens have also evolved various mechanisms to evade the action of plant glycosidases or receptor recognition (Rovenich, Zuccaro and Thomma,

2016). For example, pathogens can directly counter plant hydrolases with using inhibitor proteins, such as chitinase inhibitor and GIP1 glucanase inhibitor (Rovenich, Zuccaro and Thomma, 2016). Plant chitinases are also cleaved by fungal protease such as fungalysin (Cmp, chitin modifying protein), which is, in turn, inhibited by a plant inhibitor peptide WAMP (Naumann, Wicklow and Price, 2011; Slavokhotova *et al.*, 2014; Jashni *et al.*, 2015). Next, pathogens can also prevent plant chitinase from accessing the pathogen cell wall chitin by masking with the effector protein Avr4 from *Cladosporium fulvum*, which binds to chitin polymers (Rovenich, Zuccaro and Thomma, 2016). Chemical modification from chitin to chitosan is more resistant to chitinase and less recognised (Rovenich, Zuccaro and Thomma, 2016). Finally, secreted LysM effector proteins, such as Ecp6 from *Cladosporium fulvum* and Slp1 from *Magnaporthe oryzae*, can sequester the released chitin fragments to prevent recognition by plant receptors (Rovenich, Zuccaro and Thomma, 2016).

1.3.2 Glycoproteins

A diverse range of glycans can be found on glycoproteins either at asparagine residues (N-linked) or serine, threonine or hydroxyproline residues (O-linked) (Strasser, 2016).

N-linked glycosylation is very common in proteins that are secreted or on the plasma membrane (Strasser, 2016). N-glycosylation can be involved in protein functions in different ways such as folding, stability, protection or binding interactions (Varki, 2017; Lin *et al.*, 2020). In plants, pattern recognition receptors responsible for detection of various elicitors from pathogens have been shown to be N-glycosylated, which is essential for their defence signaling functions (Häweker *et al.*, 2010). For pathogens, the secreted effector protein Slp1 also contains N-glycosylation, which was found to be required for its stability and chitin binding activity (Chen *et al.*, 2014). Glycosylation can also protect the glycoprotein from

protease attacks. For example, N-glycan on XEG1 effector protein protects it from degradation by a host aspartic protease (AP) (Xia *et al.*, 2020). Likewise, O-glycan on flagellin of *P. syringae* protects it from cleavage by host proteases that release an immunogenic elicitor peptide (Buscaill *et al.*, 2019). Flagellin glycosylation is also important for the stability of the flagellin polymer and motility function and the lack of glycosylation can reduce virulence (Taguchi *et al.*, 2009, 2010). Interestingly, plants can counter flagellin glycosylation with the secreted β-galactosidase BGAL1, which facilitates the release of immunogenic peptides from glycosylated flagellin carrying a terminal modified viosamine. It is also hypothesised that some *P. syringae* contains polymorphic glycosylation without the modified viosamine residue to evade the action of BGAL1 (Buscaill *et al.*, 2019). Some *P. syringae* also produce an inhibitor of BGAL1, which is a focus of this thesis and will be discussed in more detail in the next section.

Regarding O-linked glycoproteins, the hydroxyproline rich glycoprotein (HRGP) family is an abundant group of plant cell wall O-glycoproteins with members such as highly glycosylated arabinogalactan proteins (AGPs), moderately glycosylated extensin proteins and other hydroxyproline-rich proteins (Malinovsky, Fangel and Willats, 2014). The diverse members of proteins in this group have been implicated in various processes such as development, growth and signaling (Seifert and Roberts, 2007; Ellis *et al.*, 2010; Malinovsky, Fangel and Willats, 2014). In plant defence, HRGP can be involved in crosslinking and strengthening of the cell wall and is found in cell wall papillae formed in response to pathogens (Deepak *et al.*, 2010; Underwood, 2012). HRGP expression in Arabidopsis was also found to be suppressed by *P. syringae* AvrPto effector (Hauck, Thilmony and He, 2003).

1.3.3 Glycoside conjugates

Many plant defence-related secondary metabolites, or phytoanticipins, are produced and stored as inactive glycoside conjugates, usually with a β-glucosyl group to render them more water soluble for storage and protect the plant from their toxicity (Piasecka, Jedrzejczak-Rey and Bednarek, 2015). These metabolites can be activated by a cleavage of the glucosyl group catalysed by β-glucosidases (Morant *et al.*, 2008). Different compartmentalisations of the inactive glycosides, usually stored in the vacuole, and the glucosidase, usually stored in the apoplast or plastids, prevents their premature encounter and ensures that activation occurs when needed, for example, upon cell disruption during infection or insect foraging (Morant *et al.*, 2008). Examples of such phytoanticipins are glucosinolates, which are produced by Brassicaceae plants. The cleavage of glucosinolates by a specific group of β-glucosidases, called β-thioglucoside glucohydrolases or myrosinase, releases unstable intermediates that decompose to various toxic compounds with anti-microbial and anti-insect activities, such as isothiocyanate (Piasecka, Jedrzejczak-Rey and Bednarek, 2015; Tierens *et al.*, 2001). In addition to the direct toxicity of the products, myrosinase overexpression was also found to increase stomatal defense to *P. syringae*, possibly due to altered hormone signaling pathways (Zhang *et al.*, 2019). Another group of phytoanticipins are cyanogenic glucosides which are widespread in the plant kingdom. These glucosides are activated by β-glucosidases whose cleavage yield unstable intermediates that decompose to release HCN, toxic to insects (Piasecka, Jedrzejczak-Rey and Bednarek, 2015).

Apart from secondary metabolites, plant hormones are also subject to modification as glucoside conjugates. For example, salicylic acid, an important hormone in the signaling for defence activation and systemic resistance, can be chemically modified to modulate the level

of its active form and storage (Ding and Ding, 2020). Conjugation of salicylic acid with a glucosyl group results in an inactive form, which can be stored in the vacuole and later activated by β-glucosidase (Ding and Ding, 2020). β-glucosidase in the apoplast was also implicated in this activation (Maruri-López *et al.*, 2019). In addition, salicylic acid and its conjugate are also modulated by pathogens. For instance, *Agrobacterium tumefaciens* produces a glycosidase SghA that can release the active salicylic acid from its glucoside conjugate (Wang *et al.*, 2019). Another example is abscisic acid, a hormone important for stress responses, which can also be stored as an inactive glucoside conjugate and is activated by a β-glucosidase (Dietz *et al.*, 2000; Lee *et al.*, 2006). Therefore, the action of glycosidase can also contribute to the fine tuning of hormone homeostasis and plant responses.

1.4 The discovery of a β-galactosidase inhibitor produced by *P. syringae* during infection and the role of the plant β-galactosidase BGAL1 in immunity

1.4.1 Activity-based protein profiling for studying hydrolase activity and inhibition

Since various hydrolases and inhibitors play important roles in plant-pathogen interactions, it is of high interest to study the dynamics of their activities during infection. Activity-based protein profiling (ABPP) is a useful tool for studying a wide range of enzyme activities and inhibition. The ABPP approach uses small molecule probes that can react with the active site of enzymes and form a covalent bond leading to irreversible labelling of active enzymes. The specificity and range of enzyme targets are determined by the structure of the reactive portion of the probes, while the probes also carry reporter groups, such as fluorophores or affinity tags, that facilitate subsequent detection of enzyme activity on protein gels or

purification and identification of labelled enzymes (Morimoto and Van Der Hoorn, 2016) (Figure 1.4A). For glycosidase activities, a broad-range glycosidase probe was developed and used to display a range of glycosidase enzymes in plants (Chandrasekar *et al.*, 2014).

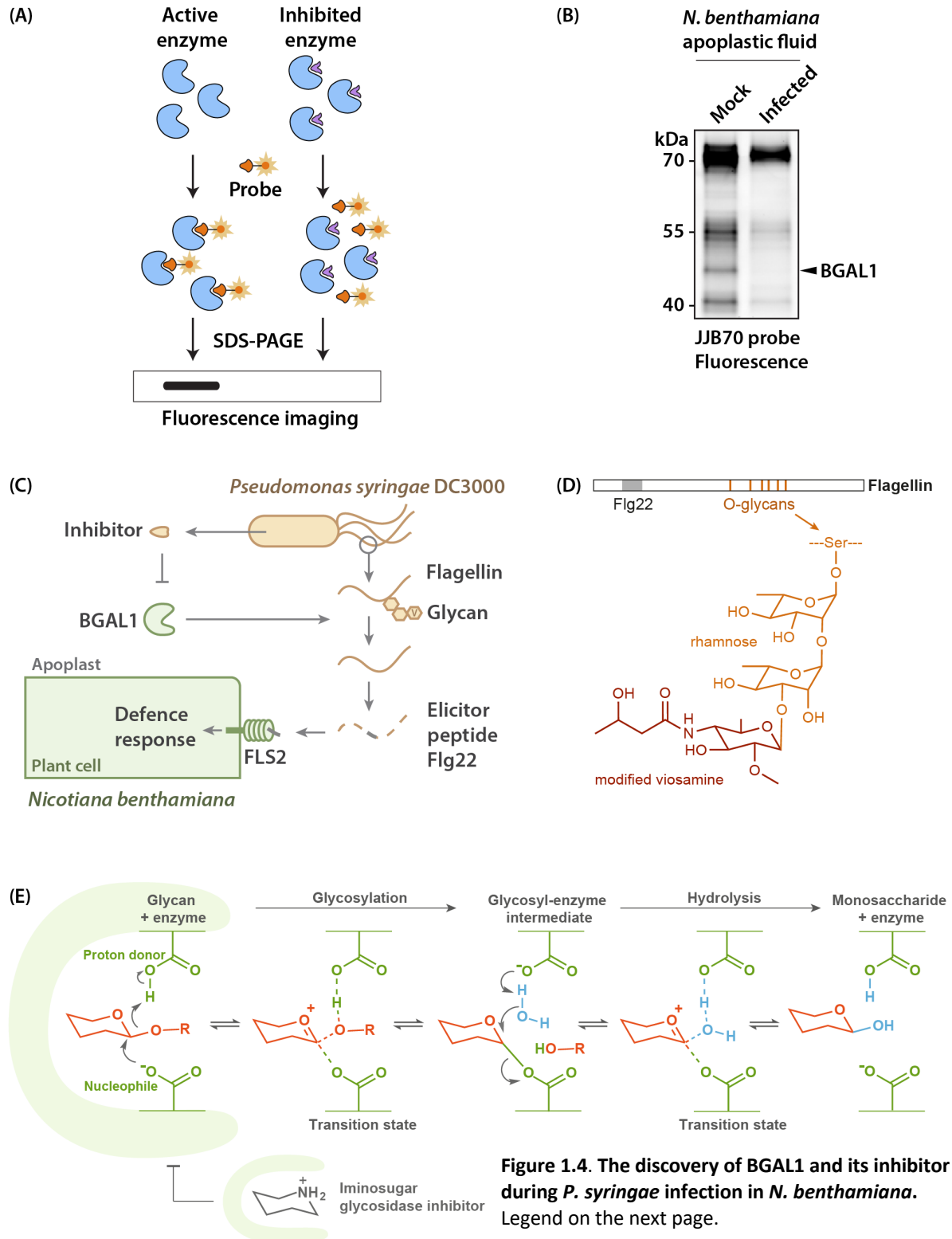


Figure 1.4. The discovery of BGAL1 and its inhibitor during *P. syringae* infection in *N. benthamiana*. Legend on the next page.

Figure 1.4. **The discovery of BGAL1 and its inhibitor during *P. syringae* infection in *N. benthamiana*. (A)** The concept of activity-based protein profiling (ABPP). The ABPP approach uses a small molecule probe that reacts with the active site of enzymes to form a covalent irreversible labelling of active enzymes. The probe also contains a functional group, such as fluorescent tag, allowing the labelled protein activity and its inhibition to be detected after the protein sample is resolved by protein gel electrophoresis (SDS-PAGE). **(B)** ABPP of leaf apoplastic fluid from *N. benthamiana* plant infected with *P. syringae* DC3000 using a glycosidase probe JJB70. The activity of plant β-galactosidase BGAL1 is suppressed during infection compared to mock control. **(C)** Model of the action of BGAL1. BGAL1 facilitates the proteolytic release of elicitor peptides from flagellin protein glycosylated with modified viosamine. The elicitor peptide Flg22 is perceived by the plant receptor FLS2 and triggers defence responses. *P. syringae* can secrete an inhibitor to inhibit BGAL1, possibly to suppress BGAL1-mediated induction of plant defence. **(D)** Schematic diagram of flagellin protein showing the location of the immunogenic peptide Flg22 and six O-glycosylation sites. O-glycans on flagellin in *P. syringae* DC3000 contains rhamnose residues and a terminal modified viosamine residue. **(E)** The catalytic mechanism of glycosidases and its inhibition by iminosugars. Retaining glycosidases, such as BGAL1, catalyse the hydrolysis of a glycosidic bond with the net retention of the configuration of the released monosaccharide. This hydrolysis proceeds through the classical Koshland retaining mechanism, which is a two-step double displacement mechanism. The first step is the formation of a covalent glycosyl-enzyme intermediate by the actions of two side chains of acidic amino acid residues, aspartate or glutamate, in the enzyme active site: one residue acts as a nucleophile attacking the anomeric carbon of the terminal monosaccharide on a glycan substrate (R) to form a covalent bond, while the other residue acts as a proton donor protonating the oxygen as the bond cleaves. The second step is the hydrolysis of the glycosyl-enzyme intermediate by water, releasing the cleaved terminal monosaccharide and regenerating the enzyme active site. Both of these catalytic steps pass through an oxycarbenium ion-like transition state. Iminosugars, a class of glycosidase inhibitors, are structural analogues of sugars with a nitrogen replacing the oxygen in the ring. This nitrogen can be protonated at the physiological pH and carry a positive charge to mimic the transition states during glycosidase catalysis, allowing iminosugars to bind to the active site and act as inhibitors. For clarity, structures of sugars and iminosugars are simplified to show only the ring.

1.4.2 Discovery of β-galactosidase inhibitor and the role of BGAL1 during *P. syringae* infection

Using this ABPP approach, the activity of plant BGAL1, a dominant leaf apoplastic β-galactosidase belonging to the GH35 family, was found to be suppressed during infection of *Nicotiana benthamiana* plants by *Pseudomonas syringae* pv. *tomato* DC3000 (Buscaill *et al.*, 2019) (Figure 1.4B). Further analysis showed that BGAL1 plays a role in plant immunity by facilitating the hydrolytic release of immunogenic elicitor peptides from bacterial flagellin that is glycosylated with modified viosamine (Buscaill *et al.*, 2019) (Figure 1.4C, D). A *bgal1* knockout mutant *N. benthamiana* plant showed higher susceptibility to spray infection of *P. syringae* strains that carry a modified viosamine on flagellin glycan, such as *P. syringae* pv. *tomato* DC3000 and pv. *tabaci* 6605. The incubation of modified-viosamine-containing

flagellin with apoplastic fluid from leaves expressing BGAL1 caused an increase in the release of elicitor peptides, which induced plant immune response as detected by release of reactive oxygen species (ROS burst) (Buscaill *et al.*, 2019).

1.4.3 Plant β-galactosidases

β-galactosidase is an enzyme that catalyses the hydrolysis of a glycosidic bond of a terminal non-reducing β-galactosyl residue from a β-galactoside (Smith and Gross, 2000). This enzyme is distributed across a wide range of phylogenetic groups from prokaryotes to eukaryotes and classified in multiple families of glycoside hydrolases (GH) including GH1, 2, 3, 35, 42, 50 and 59 (Henrissat and Grenoble, 1991). Family GH2, for example, contains *Escherichia coli* LacZ, a well-known β-galactosidase enzyme in molecular biology (Juers, Matthews and Huber, 2012). In plants, β-galactosidases are predominantly found in the GH35 family, an expanded multigene family with many members showing different expression patterns in different tissues and developmental stages, implying diversified roles of these enzymes (Ahn *et al.*, 2007; Tanthanuch *et al.*, 2008; Chandrasekar and van der Hoorn, 2016). Members of the plant GH35 β-galactosidase have activities on various glycosidic linkages of terminal galactose residues (β-1,3; β-1,4 and β-1,6) and substrates including polysaccharides such as galactan, arabinogalactan, xyloglucan, pectin and glycoproteins (Dwevedi and Kayastha, 2010). The catalysis requires two conserved glutamate residues in the active site pocket and follows a classical Koshland retaining mechanism where the hydrolysis product retains the configuration at the anomeric carbon. (Davies and Henrissat, 1995) (Figure 1.4E).

Most of the plant β-galactosidases are localised to the cell wall and have been studied mainly for their roles as cell wall modifying enzymes during development and fruit ripening (Dwevedi and Kayastha, 2010; Chandrasekar and van der Hoorn, 2016). For example, Arabidopsis

MUM2 β-galactosidase is involved in modification of seed coat mucilage pectin, necessary for its hydration property and expansion (Dean *et al.*, 2007). Arabidopsis BGAL10 was shown to be the main β-galactosidase in xyloglucan processing and a *bgal10* mutant displayed an unusual xyloglucan subunit structure and growth defects such as shorter siliques and sepals (Iglesias *et al.*, 2006; Sampedro *et al.*, 2012). Another major function of β-galactosidase is in cell wall modification during fruit ripening. The loss of galactose residues from cell wall polysaccharides by β-galactosidase cleavage is important for loosening of the cell wall and fruit softening (Smith and Gross, 2000; Brummell and Harpster, 2001; Giovannoni, 2004). In addition to cell wall polysaccharides processing, the interactions of β-galactosidases with glycoproteins have also been investigated. A β-galactosidase from radish was found to cleave galactose residues from glycans of arabinogalactan proteins and may contribute to the glycoprotein turnover (Kotake, Dina and Konishi, 2005). BGAL1 in *N. benthamiana* was found to cleave terminal galactose residues from both N- and O-linked glycans and, accordingly, the knockout of *bgal1* increased the level of glycoprotein production with the desired glycosylation (Kriechbaum *et al.*, 2020).

1.4.4 Further work on β-galactosidase inhibitor production by *P. syringae*

Regarding BGAL1 inhibition in the apoplast during *P. syringae* infection, further experiments showed that the observed inhibition was caused by an inhibitor secreted by the bacteria. The inhibitor could also be produced in an *in vitro* culture of *P. syringae* in mannitol glutamate medium, which induces the expression of the type III secretion system (Chandrasekar, 2018; Buscaill *et al.*, 2019). Further characterisation of the inhibitor revealed that the inhibitor is heat-stable (tolerated heating at 95°C), polar (partitioned in the hydrophilic aqueous fraction during methanol/chloroform extraction), basic (bound to cation exchange resin) and has a

small molecular weight (passed through a 3 kDa cut off filter) (Chandrasekar, 2018; Buscaill *et al.*, 2019). These preliminary data together with further investigation in this thesis suggest that the inhibitor is likely to be an iminosugar, which is discussed in the next section.

Additional insights into the production of the BGAL inhibitor by *P. syringae* came from a screen for mutants defective for BGAL inhibitor production, which revealed several candidate genes required for inhibitor production (Chandrasekar, 2018). The screen used a recombinant *P. syringae* strain expressing LacZ as a reporter of β-galactosidase activity, allowing BGAL inhibitor production by bacterial colonies to be distinguished when plated together with X-gal substrate, which yields a blue product upon cleavage. The darker blue color developed from higher enzyme activity in colonies with no inhibitor production allowed isolation of the candidate mutants from agar plates. This method is also used in this thesis. In addition, activity-based protein profiling with a broad-range glycosidase activity probe (Chandrasekar *et al.*, 2014) was also used to test BGAL activity and inhibitor production by the bacteria. Sequencing of the isolated BGAL inhibitor mutants has identified significant leads such as *HrpR*, a major regulator of type III secretion system. However, the screen was not saturated and the biosynthetic genes could not be identified. The inhibitor molecular identity, biosynthetic pathway, its regulation and evolution are still largely unknown.

1.5 Iminosugars as inhibitors of glycosidases

1.5.1 Iminosugar structure, type and natural occurrence

Iminosugars constitute an important group of small-molecule inhibitors of glycosidases (Stütz, 1998). They can be classified as polyhydroxylated alkaloids that have similar structures to monosaccharides but with a nitrogen atom replacing the oxygen atom in the ring structure of

the molecule (Watson *et al.*, 2001). This ring nitrogen atom can be protonated at physiological pH to carry a positive charge that, together with the monosaccharide analogue structure, mimic the characteristics of the transition state of the glycoside hydrolysis catalysed by glycosidases, allowing iminosugars to bind to the enzyme active site and act as inhibitors (Lillelund *et al.*, 2002; Gloster and Davies, 2010) (Figure 1.4E). The structure of the ring backbone and arrangement of hydroxyl groups can vary and dictate the specificity of an iminosugar to different glycosidases (Asano *et al.*, 2000). Iminosugar structures can be classified into five main groups: pyrrolidine, piperidine, pyrrolizidine, indolizidine and nor-tropane (Asano *et al.*, 2000). Many iminosugars have been identified from natural sources, mainly from plants, bacteria and some fungi (Watson *et al.*, 2001). In bacteria, mostly the monocyclic types, pyrrolidine and piperidine, are found. These natural products have also inspired a diverse range of synthetic iminosugars modified for different chemical properties or inhibition specificities (Harit and Ramesh, 2016). Examples of iminosugars, their glycosidase inhibition range and natural sources are shown in Table 1.1.

1.5.2 Therapeutic applications of iminosugars

Iminosugars and their derivatives have been extensively studied in the medical and pharmaceutical fields because of their potential in therapeutic applications, leading to three approved iminosugar-based drugs already on the market (Winchester, 2009; Horne *et al.*, 2011; Nash *et al.*, 2011). Various mechanisms of action of iminosugars targeting different enzymes have been investigated. For example, Miglitol (N-hydroxyethyl-DNJ), a drug used for treatment of type-2 diabetes, inhibits digestive α-glucosidase in the intestine to reduce the release of glucose available for absorption and limit increases in blood glucose level (Scott and Spencer, 2000). Another important application of iminosugars is treatment of lysosomal

Table 1.1. Examples of iminosugars, their reported glycosidase inhibition activities and examples of natural sources. (Asano et al., 2000; Watson et al., 2001)

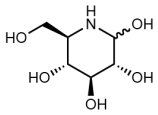
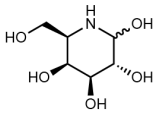
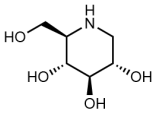
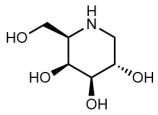
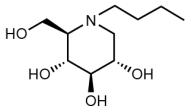
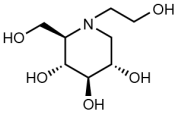
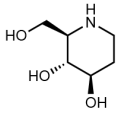
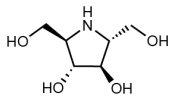
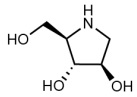
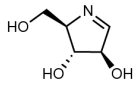
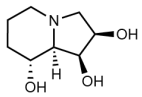
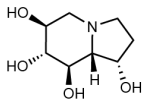
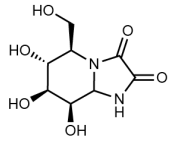
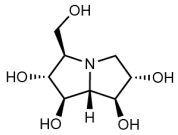
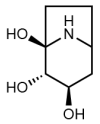
Iminosugar	Inhibited enzyme	Natural source
- Piperidine type (6 membered ring)		
nojirimycin (NJ) 	α-glucosidase, β-glucosidase, β-galactosidase, β-N-acetyl-glucosaminidase	*Streptomyces* (Bacteria)
galactonojirimycin (galactostatin) 	β-galactosidase	*Streptomyces* (Bacteria)
1-deoxynojirimycin (DNJ) 	α-glucosidase, β-glucosidase, β-galactosidase, α-mannosidase, α-fucosidase	*Streptomyces, Bacillus* (Bacteria) *Morus, Hyacinthus, Endospermum* (Plant)
1-deoxygalactonojirimycin (DGJ) 	α-galactosidase, β-galactosidase	Synthetic
N-butyl-DNJ (miglustat) 	ceramide glucosyltransferase	Synthetic
N-hydroxyethyl-DNJ (miglitol) 	α-glucosidase	Synthetic
fagomine 	α-glucosidase	*Morus, Fagopyrum, Xanthocercis* (Plant)

Table 1.1 (Continued). Examples of iminosugars, their reported glycosidase inhibition activities and examples of natural sources. (Asano et al., 2000; Watson et al., 2001)

Iminosugar	Inhibited enzyme	Natural source
- Pyrrolidine type (5 membered ring)		
DMDP (2,5-dideoxy-2,5-imino-D-mannitol) 	α-glucosidase, β-glucosidase, β-galactosidase, β-mannosidase β-xylosidase	*Streptomyces* (Bacteria) *Campanula, Hyacinthus, Endospermum* (Plant)
DAB-1 (1,4-dideoxy-1,4-imino-D-arabinitol) 	α-glucosidase, β-glucosidase, α-mannosidase α-arabinosidase	*Chitinophaga* (Bacteria) *Angylocalyx, Morus, Hyacinthoides, Eugenia* (Plant)
nectrisine 	α-glucosidase, α-mannosidase	*Nectria, Thelonectria* (Fungi)
- Indolizidine type (6 and 5 membered ring)		
swainsonine 	α-mannosidase	*Swainsona, Astralagus, Ipomea* (Plant) *Rhizoctonia, Metarhizium* (Fungi)
castanospermine 	α-glucosidase, β-glucosidase, β-xylosidase	*Castanospermum, Alexa* (Plant)
kifunensine 	α-mannosidase	*Kitasatosporia* (Bacteria)
- Pyrrolizidine type (5 and 5 membered ring)		
casuarine 	α-glucosidase	*Casuaria, Eugenia* (Plant)
- Nortropane type (6 and 5 membered ring)		
calystegine A3 	β-glucosidase, α-galactosidase, β-galactosidase	*Calystegia, Ipomea, Solanum* (Plant)

storage diseases, genetic disorders caused by defective mutants of lysosomal hydrolases that impair the turnover of cellular substrates leading to accumulation of toxic substrates (Parenti, 2009). The first drug, Miglustat (N-butyl-DNJ), inhibits glucosylceramide synthase to reduce the accumulation of glucosylceramide that is not metabolised due to the defective β-glucocerebrosidase in Gaucher disease (Ficicioglu, 2008). Another drug, Migalastat (DGJ), targets a defective α-galactosidase in Fabry disease and acts as a pharmacological chaperone by binding to the enzyme's active site to promote folding and stability (McCafferty and Scott, 2019). In addition, iminosugars, such as DNJ and its derivatives, have also been investigated for inhibition of α-glucosidase I and II in the endoplasmic reticulum, since the cleavage of terminal glucose residues of N-linked-glycan on nascent glycoproteins, catalysed by α-glucosidase I and II, are important for their subsequent interaction with protein folding machineries and the inhibition by iminosugar leads to misfolding and degradation of the proteins (Alonzi *et al.*, 2017). This interference with glycoprotein production can be used as antivirals against broad spectrum of enveloped viruses that require host glycoprotein synthesis machineries (Alonzi *et al.*, 2017). The modulation of glycoprotein synthesis by iminosugars, such as swainsonine and kifunensine, targeting mannosidase in the golgi body have also been investigated for treatments of cancer (Wrodnigg, Steiner and Ueberbacher, 2008; Hossain and Andreana, 2019).

1.5.3 Biological roles of iminosugars and applications

Unlike therapeutic applications, the roles of iminosugars in other organisms are less well studied. In plants, various types of iminosugars can be found in many plant species. This production of iminosugars can act as a defence against insect herbivores by inhibiting their digestive glycosidases (Stütz, 1998). This was also proposed to be a possible target for pest

control agents (Singh *et al.*, 2014; Ya'kobovitz, Butters and Cohen, 2016). However, the biological roles of iminosugar production by microorganisms, including various bacteria, are still unknown. There is also no report of iminosugar production by *Pseudomonas* or involvement in plant-pathogen interactions.

Iminosugars can also be used as tools to study biological processes in other organisms. For example, a library of iminosugars was used to screen for root growth inhibition in Arabidopsis and identified the role of glucosylceramide synthase in root growth inhibition (Rugen *et al.*, 2018). In other studies, iminosugars, DNJ and its derivatives, were used to probe the roles of carbohydrate active enzymes such as α-glucosidase involved in starch degradation during cereal seed germination and growth (Stanley *et al.*, 2011; Andriotis, Rejzek, Barclay, *et al.*, 2016; Andriotis, Rejzek, Rugen, *et al.*, 2016). Other iminosugars were also used to inhibit glycosidases active on cell wall polysaccharides (Nagahashi *et al.*, 1990; Andriotis, Rejzek, Barclay, *et al.*, 2016). In addition, treatments of radish seedlings with iminosugars inhibiting α-glucosidase (DNJ, castanospermine) or α-mannosidase (swainsonine) involved in N-glycan processing in the ER or golgi body, were found to alter N-glycans on glycoproteins (Mega, 2004). The inhibition of seed germination by iminosugars was also proposed as a possible control for root parasitic weeds (Wakabayashi *et al.*, 2015). In bacteria, iminosugars were investigated for effects on bacterial biofilm formation. DNJ could inhibit biofilm formation in *Streptococcus mutans* while a range of synthetic iminosugars was also found to inhibit biofilm formation in *Pseudomonas aeruginosa* (Strus *et al.*, 2016; Yoo *et al.*, 2019). The use of iminosugars as chemical intervention tools to inhibit a class of glycosidases also has an advantage as it can overcome gene redundancy or overlapping enzyme specificities, which are complications in loss-of-function mutagenesis studies. However, the possibility of

multiple effects from an iminosugar inhibiting multiple enzyme targets with similar affinity can also complicate the analysis of resulting phenotypes (Andriotis, Rejzek, Rugen, *et al.*, 2016; Rugen *et al.*, 2018).

1.5.4 Biosynthesis of iminosugars

Biosynthetic pathways for iminosugars have been characterised in a few cases (Fernando *et al.*, 2014). In microbes, biosynthetic gene clusters were identified and pathways were proposed for DNJ, nectrisine and DAB-1 (1,4-dideoxy-1,4-imino-arabinitol). In plants, some pathways have been proposed but biosynthetic enzymes were not identified. The diversity of iminosugar biosynthesis remains to be further explored for many organisms.

First, the biosynthetic gene cluster of DNJ was identified in the bacterium *Bacillus amyloquifaciens* (Clark, Johnson and Horenstein, 2011) and also cloned from *Bacillus subtilis* (Kang *et al.*, 2011). These homologous gene clusters contain three biosynthetic genes encoding for an aminotransferase (*gabT1*), a phosphatase (*yktc1*) and an alcohol dehydrogenase (*gutB1*). Based on these genes and stable isotope feeding studies, the biosynthetic pathway of DNJ was also proposed to start from a transfer of an amine group to fructose-6-phosphate by GabT1, followed by a removal of phosphate group by Yktc1, then oxidation of a hydroxyl group to a keto group by GutB1 leading to the formation of an iminosugar ring of mannojirimycin. Subsequently, to yield DNJ, unknown enzymes catalysing another epimerisation and reduction step are required to convert the stereochemistry of hydroxyl group and remove the C1 position hydroxyl group, respectively (Clark, Johnson and Horenstein, 2011) (Figure 1.5A). In plants, DNJ biosynthetic genes are unknown but, based on stable isotope labelling studies, the biosynthesis is proposed to use a different pathway, starting from imination of an aldehyde group on glucose, followed by the reduction of the

imine into amine group and the reduction of a hydroxyl group to a keto group, which together lead to the formation of an iminosugar ring of nojirimycin that requires another reduction step to form DNJ (Shibano *et al.*, 2004) (Figure 1.5B).

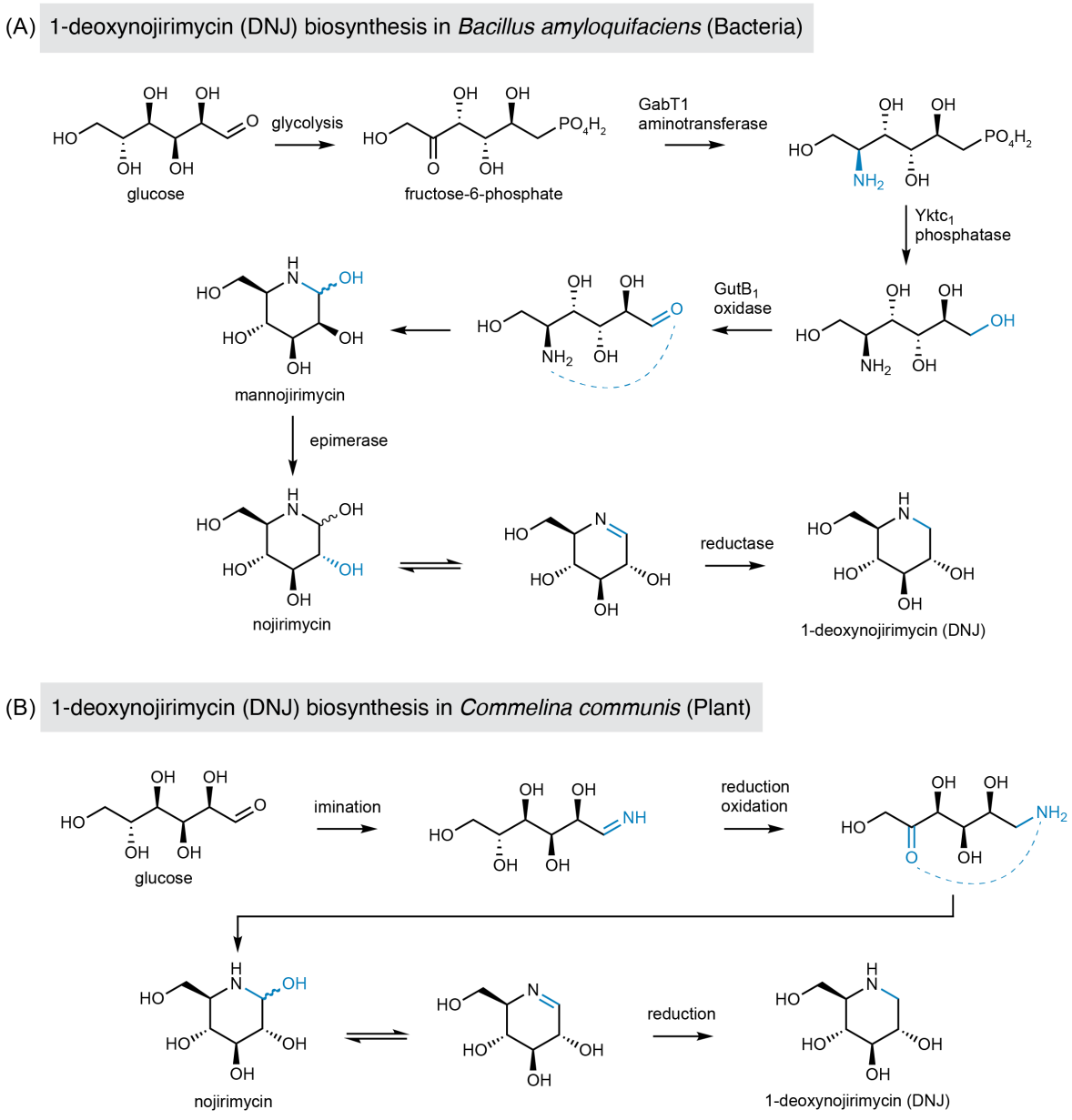
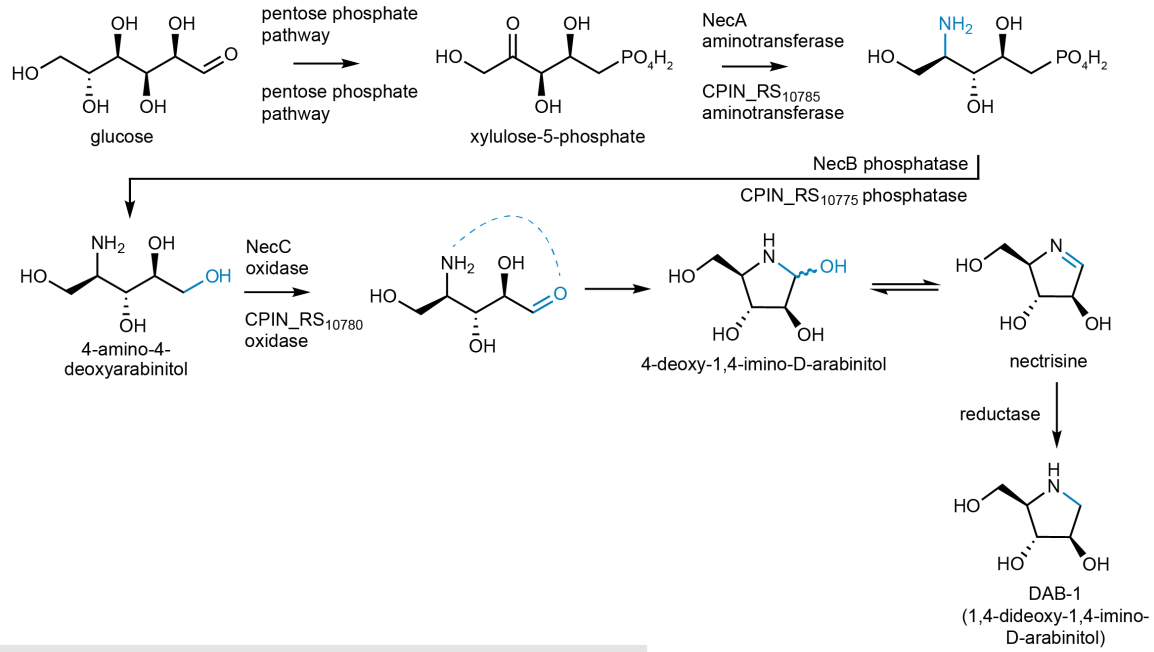


Figure 1.5. Biosynthesis of iminosugars. Chemical changes to the structure at each step are shown in blue. **(A)** 1-deoxynojirimycin (DNJ) in *Bacillus amyloquifaciens* (Clark, Johnson and Horenstein, 2011), **(B)** 1-deoxynojirimycin (DNJ) in *Commelina communis* (Shibano *et al.*, 2004). (continued on next page)

(C) Nectrisine biosynthesis in *Thelonectria discophora* (Fungi) (Top)
& DAB-1 biosynthesis in *Chitinophaga pinensis* (Bacteria) (Bottom)



(D) Swainsonine biosynthesis in *Metarhizium robertsii* (Fungi)

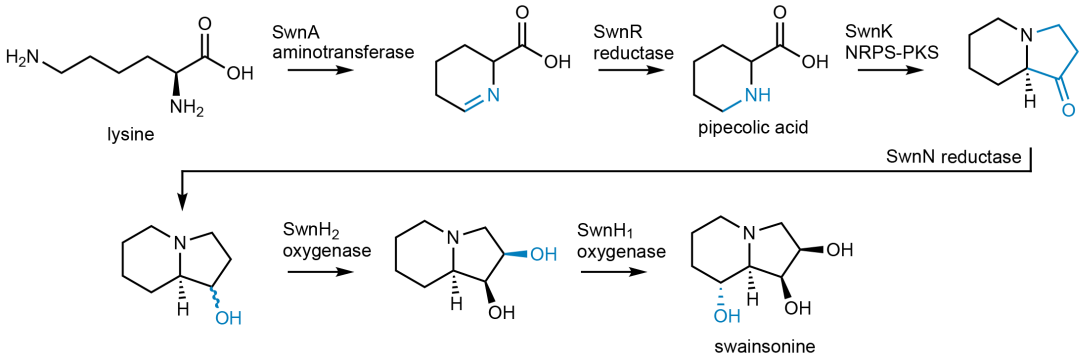


Figure 1.5 (continued). Biosynthesis of iminosugars. Chemical changes to the structure at each step are shown in blue. **(C)** Nectrisine in *Thelonectria discophora* (top of arrows) (Miyauchi *et al.*, 2016) and DAB-1 in *Chitinophaga pinensis* (bottom of arrows) (Nuñez and Horenstein, 2019) and **(D)** Swainsonine in *Metarhizium robertsii* (Fernando et al., 2014).

Second, the biosynthetic gene cluster of nectrisine, characterised in the fungus *Thelonectria discophora*, contains three biosynthetic genes encoding for an aminotransferase (*necA*), phosphotransferase (*necB*) and an oxidase (*necC*) (Miyauchi *et al.*, 2016). The proposed biosynthetic pathway starts with the transfer of an amine group onto D-xylulose-5-phosphate by NecA, followed by the removal of phosphate group by NecB, then the oxidation of a hydroxyl group to a keto group by NecC leading to the formation of an iminosugar ring intermediate, which can spontaneously dehydrate to form nectrisine (Miyauchi *et al.*, 2016) (Figure 1.5C).

Third, the biosynthetic gene cluster of DAB-1, characterised in the bacterium *Chitinophaga pinensis*, also contains three biosynthetic genes encoding an aminotransferase, phosphatase and alcohol dehydrogenase (Nuñez and Horenstein, 2019). The proposed biosynthetic pathway of DAB-1 follows the nectrisine pathway but a final reduction step, catalysed by an unknown enzyme, converts nectrisine into DAB-1 (Nuñez and Horenstein, 2019) (Figure 1.5C).

Overall, the three biosynthetic gene clusters identified share the same set of three enzyme functions an alcohol dehydrogenase, aminotransferase and phosphatase. In DNJ and DAB-1 biosynthesis, the alcohol dehydrogenases and aminotransferases are homologous enzymes from the same family, while the phosphatases are from different families. Interestingly, the enzymes found in fungal nectrisine biosynthesis are not homologous to the bacterial DNJ or DAB-1 biosynthesis, possibly suggesting convergent evolution of iminosugar production.

In contrast to monocylic iminosugars, the biosynthesis of bicyclic iminosugars such as swainsonine is proposed to start from a pipecolic acid precursor, which is derived from lysine. The second part of the ring can be formed by the action of a hybrid non-ribosomal peptide-

polyketide synthase and then additional hydroxyl groups are formed by oxygenase enzymes (Fernando *et al.*, 2014; Luo *et al.*, 2020) (Figure 1.5D).

1.5.5 Methods for studying iminosugars

Many iminosugars have been isolated and characterised from natural products and a range of methods are available and common for iminosugar studies (Molyneux, 1993; Molyneux *et al.*, 2002; Amézqueta and Torres, 2016). First, extraction of iminosugars from plant materials or microbial cultures can normally be done by solvent extraction, such as with water, methanol or ethanol, as iminosugars partition in the hydrophilic fractions due to their highly polar polyhydroxylated alkaloid nature (Molyneux, 1993; Amézqueta and Torres, 2016). Then, purification of iminosugars from the extracts is usually done using cation exchange resins since iminosugars can be protonated in acidified extracts and carry a positive charge to bind to the negatively charged group on cation exchange resins. This allows other non-positively charged compounds to be washed away and iminosugars can subsequently be eluted using alkaline solutions such as NH_4OH (Molyneux, 1993; Amézqueta and Torres, 2016). Furthermore, separation of iminosugars can also be done with liquid chromatography and high-performance liquid chromatography (HPLC). Due to the highly polar nature of iminosugars, hydrophilic interaction liquid chromatography (HILIC) or ion exchange chromatography are normally used instead of the conventional reverse-phase chromatography, which does not efficiently separate highly hydrophilic polar compounds (Rodríguez-Sánchez, Quintanilla-López, *et al.*, 2014a; Amézqueta and Torres, 2016). Finally, for detection, chromatographic methods including LC-MS (liquid chromatography-mass spectrometry) and GC-MS (gas chromatography-mass spectrometry) are commonly used. Since iminosugars are UV transparent, chemical derivatisation to tag the molecule with a UV

absorbing group, such as Fmoc, is required for common liquid chromatography detection (Amézqueta and Torres, 2016). In addition, chemical derivatisation can also be used to modify chemical properties of iminosugars to suit analytical techniques. For example, derivatisation of with Fmoc adds a hydrophobic group to allow separation by reverse-phased liquid chromatography, while derivatisation with trimethylsilane (TMS) leads to a volatile derivative required for gas chromatography analysis (Molyneux *et al.*, 2002; Rodríguez-Sánchez *et al.*, 2011). Finally, the inhibitory activities of iminosugars in the isolated fractions are normally detected using enzymatic assays with chromogenic or fluorogenic substrates of carbohydrate conjugates, which generate detectable signals upon cleavage by corresponding glycosidases (Molyneux *et al.*, 2002).

1.6 Research aims and rationale

I hypothesised that *P. syringae* produces the inhibitor, which may be an iminosugar, to promote host colonisation by inhibiting BGAL or other glycosidases during infection. The main aim of this work is to investigate the recently discovered production of BGAL inhibitor by *P. syringae*. Specifically, **I aim to determine the molecular identity of this inhibitor, characterise its biosynthetic and regulatory pathways and unravel the evolution and biological roles of this inhibitor production** (Figure 1.6). First, revealing the molecular identity of this inhibitor could lead to a discovery of a new natural product. The knowledge of this molecule could potentially be of pharmaceutical use based on its inhibitory activity. Next, elucidating the biosynthetic pathway of this inhibitor will not only contribute to the knowledge of metabolism and biosynthesis of natural products in bacteria but may also find enzyme activities or pathways with useful biotechnological applications. In addition, characterisation of the regulation of this system will also increase fundamental understanding of bacterial virulence

mechanisms and how they are controlled. Furthermore, unravelling the evolution of this inhibitor production will provide more insights into the complex evolution of the pathogen and its virulence mechanisms. Exploring the evolutionary context of the inhibitor production may also lead to identification of similar systems in other bacteria. Finally, understanding the biological roles of the inhibitor and its glycosidase targets can lead to discoveries of novel mechanisms of bacterial virulence or plant defence. Increased knowledge of the interactions between plants and pathogens can also facilitate the improvement of crop resistance and ways to combat plant diseases.

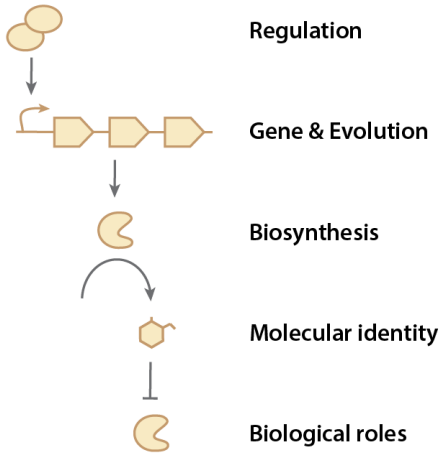


Figure 1.6. **Research aims.** This work investigates the recently discovered production of the BGAL inhibitor by *P. syringae*. I aim to determine the molecular identity of this inhibitor, characterise its biosynthetic and regulatory pathways and unravel the evolution and biological roles of inhibitor production.

In the following chapters of this thesis, chapter 2 will discuss the screen for BGAL inhibitor mutants to identify candidate genes required for inhibitor production. Next, chapter 3 will discuss further characterisation of the candidate genes, which led to the identification of the biosynthetic gene cluster and regulators. These findings also allowed me to explore the evolution and potential biological roles of the inhibitor. Finally, chapter 4 will discuss the approaches I have taken towards the characterisation of the inhibitor, including production, purification and detection of the inhibitor.

Chapter 2: The identification of candidate genes involved in BGAL inhibitor production

2.1 Introduction

During infection, microbial plant pathogens colonise the apoplast, an extracellular space in the plant tissue. This interface between plants and bacteria is an important site where various factors are secreted to participate in a dynamic interaction that determines the outcome of the infection. Apoplastic hydrolases, such as proteases and glycosidases, are prominent players in plant defence and are, therefore, targets of manipulations by multiple inhibitors from pathogens. In a previous study on the interaction between *N. benthamiana* plant and *P. syringae* DC3000 pathogen, a plant apoplastic β-galactosidase BGAL1 was identified to function in plant defence by facilitating the release of an immunogenic elicitor peptide from bacterial flagellin that is glycosylated with a modified viosamine (Buscaill *et al.*, 2019). Furthermore, *P. syringae* DC3000 was found to produce an inhibitor of plant BGAL1 potentially as a virulence mechanism (Buscaill *et al.*, 2019). However, not much is known about this inhibitor and how it is produced by the bacteria. To learn more about this BGAL1 inhibitor production in *P. syringae*, I set out to identify genes that are involved in the process. The knowledge of genes that are required for the production of the inhibitor can provide insights into its biosynthetic and regulatory pathways as well as the evolution of this virulence mechanism.

Previously, candidate genes involved in BGAL1 inhibitor production was identified by a forward genetic screen for BGAL1 inhibitor mutants (*bim*) from a random transposon insertion mutant library of *P. syringae* DC3000 (Chandrasekar, 2018). The screen used a

recombinant *P. syringae* strain expressing LacZ as a reporter of β-galactosidase activity and MG medium, which was found to induce BGAL1 inhibitor production. Inhibitor production by different bacterial colonies could be distinguished when plated together with X-gal substrate that yields blue product upon cleavage by β-galactosidase. The darker blue color developed from higher enzyme activity in colonies with no inhibitor production allowed isolation of the candidate *bim* mutants from agar plates. Subsequently, transposon insertion sites in these mutants were determined to identify the disrupted genes that were candidate genes required for inhibitor production. This approach identified significant candidate genes such as *hrpR*, a major regulator of type III secretion system. However, the screen was still incomplete and did not achieve saturation since only few genes were identified more than once and biosynthetic genes were not identified. Therefore, I aimed to refine the method and conduct a comprehensive screen to identify more candidate genes involved in BGAL1 inhibitor production.

As part of the screen, activity-based protein profiling (ABPP) with a broad-range glycosidase activity probe JJB70 was also used to monitor BGAL1 activity and inhibitor production by the bacteria (Chandrasekar *et al.*, 2014). The probe contains a fluorophore and a reactive group that reacts to the active site of glycosidases forming a covalent bond. After incubation of the sample with the probe, proteins are separated with polyacrylamide gel electrophoresis and visualised by fluorescence scanning. The intensity of fluorescent signal of a band indicates the amount of active protein. Despite being a powerful discovery tool, ABPP application in large scale can be difficult due to cost and availability of the probe and limited throughput. Therefore, I also need to establish a complementary method to quantify BGAL1 activity with high throughput and sensitivity. Synthetic monosaccharide-conjugated substrates can be

useful tools for monitoring activities of glycosidases. For example, FDG (fluorescein-di-β-D-galactopyranoside), a synthetic substrate of galactose conjugated with a fluorophore fluorescein, can be used to monitor β-galactosidase activity since its hydrolysis catalysed by the enzyme yields a fluorescent product that can be measured with sensitive fluorescence measurement and in high-throughput 96-well plate format using a plate reader. I will also explore the use of this method in this work.

This chapter describes the development of the optimised screen for BGAL1 inhibitor mutants and the approaches I took to identify candidate genes involved in BGAL1 inhibitor production in *P. syringae*.

2.2 Materials and methods

2.2.1 Bacterial strains, culture conditions and media

Bacterial strains used in this study are listed in Table 2.1. Cultures were incubated at 28°C for *P. syringae* and 37°C for *E. coli*, with shaking at 220 rpm for liquid cultures. Media used for bacterial cultures were made as follows:

- LB medium: 10g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl.
- Mannitol glutamate (MG) medium (Bronstein *et al.*, 2008): 10 g/L mannitol, 2 g/L L-glutamic acid, 0.5 g/L KH_2PO_4, 0.2 g/L NaCl, 0.2 g/L $MgSO_4.7H_2O$, adjusted to pH 5.5.
- Solid media for plating contain 1.5% (w/v) agar.
- Antibiotics were generally used at the concentrations of 25 µg/ml rifampicin, 12.5 µg/ml gentamicin, 50 µg/ml kanamycin, unless stated otherwise in a specific protocol.

Table 2.1. Bacterial strains used in this study

Strain	Description	Antibiotics	Reference
P. syringae DC3000 Δ*hopQ1* (HQ)	*P. syringae* pv. *tomato* DC3000 lacking the effector *hopQ1* allowing it to cause disease in the model plant *N. benthamiana*	Rifampicin	(Wei *et al.*, 2007)
P. syringae DC3000 Δ*hopQ1 lacZ* (HQ *lacZ*)	*P. syringae* pv. *tomato* DC3000 Δ*hopQ1* carrying pHRP308 plasmid expressing lacZ	Rifampicin, Gentamicin	(Parales and Harwood, 1993; Chandrasekar, 2018)
E. coli mini-Tn5	Donor strain carrying mini-Tn5-Km2 plasmid with promoter less GFP:lux cassette which can serve as a reporter if inserted in the genome in the correct orientation downstream of a promoter; DH5α λpir background	Kanamycin	(De Lorenzo *et al.*, 1990; Fones *et al.*, 2010)
E. coli pRK2013	Helper strain carrying pRK2013 plasmid for conjugal transfer; DH5α λpir background	Kanamycin	(Figurski and Helinski, 1979)
*bim*4	*P. syringae* DC3000 Δ*hopQ1 lacZ* with mini-Tn5 insertion in *hrpR*; BGAL1 inhibitor defective mutant	Rifampicin, Gentamicin, Kanamycin	(Chandrasekar, 2018)
*bim*5	*P. syringae* DC3000 Δ*hopQ1 lacZ* with mini-Tn5 insertion in *gabT1*; BGAL1 inhibitor defective mutant	Rifampicin, Gentamicin, Kanamycin	(Chandrasekar, 2018)

non-*bim*	*P. syringae* DC3000 Δ*hopQ1 lacZ* with mini-Tn5 insertion in *hopI1*; BGAL1 inhibitor producing mutant	Rifampicin, Gentamicin, Kanamycin	This study

2.2.2 BGAL1 inhibitor production *in vitro*

Bacteria were grown overnight in LB medium. The cells were collected by centrifuging at 4,000 x g for 5 minutes then washed once with sterile water and centrifuged again. The cell pellet was then resuspended in MG medium to the OD_{600} of 0.5 and grown overnight. The supernatant containing the inhibitor was collected after centrifuging the bacterial culture and discard the cell pellet. For the strains expressing *lacZ*, the supernatant was heated at 95°C for 5 minutes to inactivate the enzyme to prevent interference in subsequent enzymatic assays.

2.2.3 Apoplastic fluid extraction

N. benthamiana leaves (~4 weeks old) were rolled into a 50 ml syringe filled with ice-cold water. The leaves were infiltrated by pulling the plunger to apply negative pressure drawing air out of the leaf tissue then pushing the plunger to apply positive pressure forcing water into the leaf tissue, while keeping the syringe tip sealed. The infiltrated leaves were blotted dry and rolled into a 20 ml syringe barrel placed inside a 50 ml conical centrifuge tube. Apoplastic fluid was collected from the leaves by centrifugation at 900 x g for 15 minutes at 4°C. The extracted apoplastic fluid was used directly or stored at -20°C.

2.2.4 FDG assay for β-galactosidase activity

Substrate preparation: A stock FDG solution of 2 mM was made by dissolving FDG (Fluorescein di-β-D-galactopyranoside) (Marker Gene Technologies, Eugene, USA, M0250) in 1:1:8 ethanol:DMSO:H_2O (first dissolved in 1:1 ethanol:DMSO, then diluted in the rest of H_2O). Aliquots were stored at -20°C. The stock was diluted to 2 µM in water just before use.

Enzyme preparation: A stock LacZ solution of 1 U/µl was made by dissolving lyophilised β-galactosidase enzyme (Sigma, St. Louis, USA, G6008) in suspension buffer (50% (v/v) glycerol, 5mM Tris, 5 mM $MgCl_2$, 0.5 mM DTT, 0.5 mM β-mercaptoethanol, pH 7.4). Aliquots were stored at -20°C. The stock was diluted to 1 mU/µl in water just before use.

Apoplastic fluid BGAL assay: The assay was set up in 96 well plate with the total reaction volume of 100 µl per well with 0.2 µM FDG and 50 mM MES pH 5. This typically consists of 40 µl apoplastic fluid sample, 10 µl FDG (2 µM), 10 µl MES buffer pH 5 (500 mM) and 40 µl H_2O. For inhibition assays, 40 µl inhibitor sample was used instead of H_2O.

LacZ assay: The assay was set up in 96 well plate with the total reaction volume of 100 µl per well with 60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 0.2 µM FDG and 0.1 mU/µl LacZ. This typically consists of 20 µl 5x Z buffer (300 mM Na_2HPO_4, 200 mM NaH_2PO_4, 50 mM KCl, 5 mM $MgSO_4$, pH 7), 10 µl FDG (2 µM), 10 µl LacZ (1 mU/µl) and 60 µl H_2O. For inhibition assays, 50 µl inhibitor sample was used instead of H_2O.

β-galactosidase activity in the sample was monitored by measuring fluorescence signal from fluorescein, a product of FDG hydrolysis catalysed by β-galactosidase, every minute at 25°C using the plate reader Infinite M200 (Tecan, Mannedorf, Switzerland) at an excitation wavelength of 485 nm and emission wavelength of 535 nm. The rate of change of fluorescence over time was calculated to determine β-galactosidase activity. For comparisons of inhibition, the enzymatic activity was expressed as percentage of the activity compared to no inhibitor control (100%). Note that fluorescence intensity of fluorescein is pH dependent. The same amount of FDG at pH 5 generates less fluorescence than at pH 7.

2.2.5 Activity-based protein profiling (ABPP) of glycosidase activity

Apoplastic fluid samples were incubated in a total reaction volume of 50 µl with 50 mM MES pH 5.0 and 2 µM JJB70, a fluorescent probe targeting active glycosidase, in the dark for 1 hour at room temperature (~22°C). For inhibition assays, equal volumes of apoplastic fluid and inhibitor samples were mixed. The labelling reaction was stopped by adding gel loading buffer (200 mM Tris-HCl (pH6.8), 400 mM DTT, 8% (w/v) SDS, 0.4% (w/v) bromophenol blue, 40% (v/v) glycerol) and heating at 95°C for 5 min. The labelled proteins were separated by polyacrylamide gel electrophoresis and visualised using a Typhoon imager (Cytiva, Marlborough, USA) with Cy2 filters (488 nm excitation and 520 nm emission).

2.2.6 Simulation of transposon insertion in the genome

The genome sequence of *P. syringae* pv. *tomato* DC3000 was downloaded from the Pseudomonas Genome Database (Winsor *et al.*, 2016). Transposon insertions were simulated with the assumption that insertions occur at random sites in the genome. Random integers within the range of the length of the genome were generated and used as insertion locations, which were then mapped to coding sequence of genes in the genome allowing the number of genes hit by the insertions to be calculated. The number of colonies screened were represented by the number of random integers generated.

2.2.7 Triparental mating conjugation for mini-Tn5 transposon insertion mutagenesis

Overnight bacterial cultures in LB medium were collected by centrifuging at 4,000 x g for 5 minutes, washed twice with LB medium (prewarmed to 50°C) and resuspended to the final OD_{600} of 0.5. A mixture of 700 µl recipient strain (*P. syringae* DC3000 Δ*hopQ1 lacZ*), 300 µl

donor strain (*E. coli* mini-Tn5) and 100 µl helper strain (*E. coli* pRK2013) was centrifuged together at 4,000 x g for 5 minutes to form a pellet which was resuspended in 50 µl LB medium and then spotted on to an LB agar plate without antibiotics. The spot was left to dry and the plates were incubated at 28°C for 6 hours. Multiple conjugations were made to obtain independent sets of mutants and obtain the total number of colonies to screen.

2.2.8 Plating screen for BGAL1 inhibitor mutant colonies

The conjugation mixture spot was scraped from the plate, resuspended in 1 ml of MG medium and centrifuged at 4,000 x g for 5 minutes. The cell pellet was then resuspended in 1 ml MG medium supplemented with 50 µM iron citrate and 12.5 µg/ml kanamycin and incubated at 28°C with shaking for 3 hours to induce BGAL1 inhibitor production and inactivate untransformed cells. The cells were then collected by centrifugation at 4,000 x g for 5 minutes, washed with MG medium and resuspended in 1 ml MG medium. This step was done gently to avoid damaging cells, which could lead to LacZ leakage out of the cell and increase background on the plate. The suspension was then plated on MG agar supplemented with 50 µM iron citrate, 12.5 µg/ml each of rifampicin, gentamicin and kanamycin, 0.1 mg/ml X-gal and 0.1 mM IPTG. The pH of the medium was adjusted to 5.5 to reduce the activity of LacZ that was present extracellularly, which can contribute to background. Cell density was optimised, by adjusting the initial bacterial cultures used for conjugation to the OD600 of 0.5 and resuspending the resulting bacterial conjugation mixture in 1 ml, since too high a density can also give too much background. Around 300 µl of bacterial suspension was spread on a 120x120 mm square Petri dish, forming ~ 700 colonies. The plates were incubated at 28°C for 2 days and then left at room temperature (~22°C) for 2 days for colonies to grow to sufficient size and develop differential blue colour without giving too deep a blue colour and high

background. The candidate colonies of BGAL1 inhibitor mutants were identified by their deeper blue colour as a result of higher activity of LacZ cleaving X-gal in the absence of inhibitor production.

2.2.9 The 96 well plate-based screen for BGAL1 inhibitor mutants

Candidate BGAL1 inhibitor mutant (*bim*) colonies picked from the plating screen were inoculated into each well of a 96 well plate containing 150 µl MG medium supplemented with 50 µM iron citrate and 12.5 µg/ml kanamycin and incubated at 28°C with shaking overnight. Mutants that are able to produce the inhibitor (*non-bim,* insertion in *hopI1*) and unable to produce the inhibitor (*bim4,* insertion in *hrpR*), were also included in each plate as controls to allow comparison with other colonies in the screen. The plate was then centrifuged at 4,000 x g for 5 minutes to collect the supernatant, which was transferred to a new 96 well plate and heated on a water bath at 90°C for 5 minutes to inactivate LacZ produced by the bacterial culture. To determine whether the inhibitor was produced, the supernatant was used for FDG assay with LacZ as described earlier by adding the assay reagents into the wells. β-galactosidase activity was measured for each well and used for k-means clustering to classify the candidate mutants to *bim* or non-*bim* mutants. Low activity indicates inhibitor production from non-*bim* mutants while high activity indicates no inhibitor production from *bim* mutants.

2.2.10 Inverse PCR

Regular protocol: Genomic DNA was extracted from an overnight bacterial culture using Wizard genomic DNA purification kit (Promega, Madison, USA) and quantified using a Nanodrop spectrophotometer (Thermo Scientific, Waltham, USA). The genomic DNA was then digested with Fast Digest PaeI (SphI isoschizomer) (Thermo Scientific, Waltham, USA)

using 1 µg DNA and 1 µl enzyme in the final reaction volume of 20 µl incubated at 37°C for 1 hour and heat inactivated at 65°C for 20 minutes. The digested genomic DNA was then self-ligated with T4 DNA ligase (Thermo Scientific, Waltham, USA) using 10 µl (500 ng) digested DNA and 1 µl (5 U) enzyme in the final reaction volume of 50 µl incubated at 16°C for 18 hours and heat inactivated at 65°C for 20 minutes (final concentration of DNA ~10 ng/µl allows self-ligation of DNA fragments to form circular products for subsequent PCR). The ligated DNA was then used as a template for PCR with Q5 high-fidelity DNA polymerase (New England Biolabs, Ipswich, USA) using 2.5 µl DNA and primers oNS22_Tn5_R_R3 and oNS23_Tn5_R_F4 (Table 2) in the final reaction volume of 25 µl with GC enhancer. PCR cycling conditions were 95°C for 3 min, followed by 35 cycles of [95°C for 15 sec, 72°C for 6 min (annealing + extension)], and 72°C for 5 min.

Direct protocol for high-throughput screening: Bacterial cells (small amount in pipette tip from touching a colony or cell pellet or 0.5 – 1 µl liquid culture) were resuspended in 10 µl of water and heated at 95°C for 5 minutes to lyse cells and release DNA. The lysate was then restriction digested by adding 1 µl 10X Fast Digest buffer and 1 µl Fast Digest PaeI (1:10 diluted in 1X Fast Digest buffer) (Thermo Scientific, Waltham, USA) followed by an incubation at 37°C for 1 hour and heat inactivation at 65°C for 20 minutes. The digested mixture was then self-ligated by adding 0.5 µl ATP (10mM) and 1 µl T4 DNA ligase (1:10 diluted in 1X T4 DNA ligase buffer) (Thermo Scientific, Waltham, USA) followed by an incubation at 16°C for 18 hour and heat inactivation at 65°C for 20 minutes. The ligated DNA was then used as a template for PCR with Q5 high-fidelity DNA polymerase (New England Biolabs, Ipswich, USA) with GC enhancer using 2.5 µl DNA template and primers oNS22_Tn5_R_R3 and oNS23_Tn5_R_F4 (Table 2) in the final reaction volume of 25 µl. PCR cycling conditions were 95°C for 3 min,

followed by 35 cycles of [95°C for 15 sec, 72°C for 6 min (annealing + extension)], and 72°C for 5 min.

2.2.11 Sequencing

Inverse PCR products were treated with Illustra ExoProStar (Cytiva, Marlborough, USA) and Sanger sequenced (Eurofins, Luxembourg) using primer oNS15_Tn5_R_seq2 (Table 2.2).

Table 2.2. Primers used in this study

Primer	Sequence (5'-3')	Use
oNS15_Tn5_R_seq2	TTTATTCCTCCTTGTTAACCCG	sequencing inverse PCR products to determineTn5 insertion site
oNS22_Tn5_R_R3	GGAACAGTAGACGAAATTATTGCTTCCATGAAGC TCTTCCAGTCTGATGTCATGCCAT	inverse PCR amplification
oNS23_Tn5_R_F4	TTTATTCCTCCTTGTTAACCCGGGTACCGAGCTCG AATTCGCGCGGCCGCACTTGTGTATAAGAGTCAG	inverse PCR amplification

2.2.12 Mapping transposon insertion sites

Sequencing results were checked for quality and trimmed to obtain only the sequence right next to the transposon boundary, which corresponds to the insertion site in the genome. To identify the position of the insertion site, these sequences were BLAST searched against the *P. syringae* pv. tomato DC3000 genome using the online BLASTN tool in The Pseudomonas Genome Database (pseudomonas.com) (Winsor *et al.*, 2016).

2.2.13 Simulation of restriction enzyme digest

The genome sequence of *P. syringae* pv. tomato DC3000 was downloaded from the Pseudomonas Genome Database (Winsor *et al.*, 2016). Restriction enzymes for inverse PCR need to be selected from enzymes that cut at a suitable frequency in the genome to provide fragments with the lengths that are suitable for inverse PCR amplification and accurate mapping of insertion site from sequencing. The restriction site also has to be present at least

once in the transposon itself to enable the known sequence in the transposon to flank the genome insertion site once the fragments are self-ligated, allowing primers binding and inverse PCR amplification. The simulation of digestion was done by splitting the genome sequence at the sites that match with the restriction sequence of each enzyme. The length of the split fragments was calculated. Fragments larger than 20,000 bp were classified as too big for efficient PCR amplification and fragments smaller than 200 bp were classified as too small to distinguish specific band on the gel and to map sequence of insertion site with high confidence.

2.2.14 Gene expression profile of candidate genes

Gene expression profiling data were obtained from (Lovelace, Smith and Kvitko, 2018).

2.2.15 Orthologs of candidate genes

Proteomes predicted from each bacterial strain genome sequence were downloaded from the Pseudomonas Genome Database (Winsor *et al.*, 2016) or NCBI. Orthologs and orthogroups were predicted using Orthofinder (Emms and Kelly, 2019).

2.3 Results

2.3.1 The FDG assay can be used to monitor the β-galactosidase activity of both plant apoplastic fluid BGAL1 and *E coli* LacZ

Since this work focuses on the inhibitor of a plant β-galactosidase (BGAL1), it is important to establish a tool to monitor this enzyme activity and its inhibition. I first started to establish a method to quantify BGAL1 activity with high throughput and sensitivity. Here, I chose an assay with FDG (fluorescein-di-β-D-galactopyranoside), a synthetic substrate whose cleavage by β-galactosidase yields a fluorescent product that allows enzymatic activity to be monitored with sensitive fluorescence measurement and in high-throughput 96-well plate format. The assay was also optimised to use minimal amount of enzyme and substrate to yield a robust signal while also having a maximum sensitivity for detecting inhibition.

To demonstrate the ability of the FDG assay to detect BGAL1 activity and inhibition, leaf apoplastic fluid from *N. benthamiana* plant was used as a source of BGAL1 and mixed with supernatants collected from MG minimal media cultures of different *P. syringae* DC3000 Δ*hopQ*1 derivatives as sources of inhibitor. Two strains of BGAL1 inhibitor mutants (*bim*4 and *bim*5) were used as negative controls while inhibitor producers, including wild type (HQ), wild type expressing *lacZ* (HQ *lacZ*) and a mutant with a transposon insertion in an unrelated gene (non-*bim*), were used as positive controls for inhibitor production.

Since, in the previous work (Chandrasekar, 2018), the activity and inhibition of BGAL1 was monitored using activity-based protein profiling (ABPP) with a glycosidase specific probe JJB70, I also compared the ABPP method with the FDG assay. As expected for the previously established ABPP, the activity and inhibition of BGAL1 could be detected in the samples

labelled with JJB70 probe. The fluorescent signal from labelled active BGAL1 was reduced upon mixing with supernatant from inhibitor-producing bacteria (HQ, HQ *lacZ*, non-*bim*) compared to non-producers (*bim*4, *bim*5) and blank media control (Figure 2.1A). The water control showed higher activity, likely because the conditions, such as pH or salts, of the reaction mixture was different to samples with minimal media.

Similarly, using the same set of samples, the FDG assay also showed fluorescence signal produced from BGAL1 activity cleaving the substrate, which was inhibited by bacteria-derived inhibitor (Figure 2.1B). FDG hydrolysis correlates well with the ABPP readout, confirming its validity. In addition, FDG hydrolysis was largely absent in apoplastic fluid from the *bgal1* knockout plants (Buscaill *et al.*, 2019), indicating that the assay readout was specific to the activity of BGAL1, which is the predominant β-galactosidase in the apoplastic fluid (Figure 2.1C).

As another standard for the FDG assay, a widely used β-galactosidase enzyme from *E. coli* (LacZ) was also used. Similar to plant BGAL1, FDG hydrolysis by LacZ could be detected and inhibited by samples containing the BGAL1 inhibitor (Figure 2.1D). This provides a useful tool for inhibitor assays since commercially available LacZ can be easily obtained in standardised amount for assays, whereas obtaining BGAL1 from plant apoplastic fluids can be limited by the small volume of isolation from leaves and can have large variability among plants.

Overall, these results demonstrate that the FDG assay can be used to monitor β-galactosidase activity of both *N. benthamiana* BGAL1 and *E coli* LacZ as well as their inhibition by the inhibitor from *P. syringae.* This assay is a useful tool for further works on the BGAL1 inhibitor mutant screen.

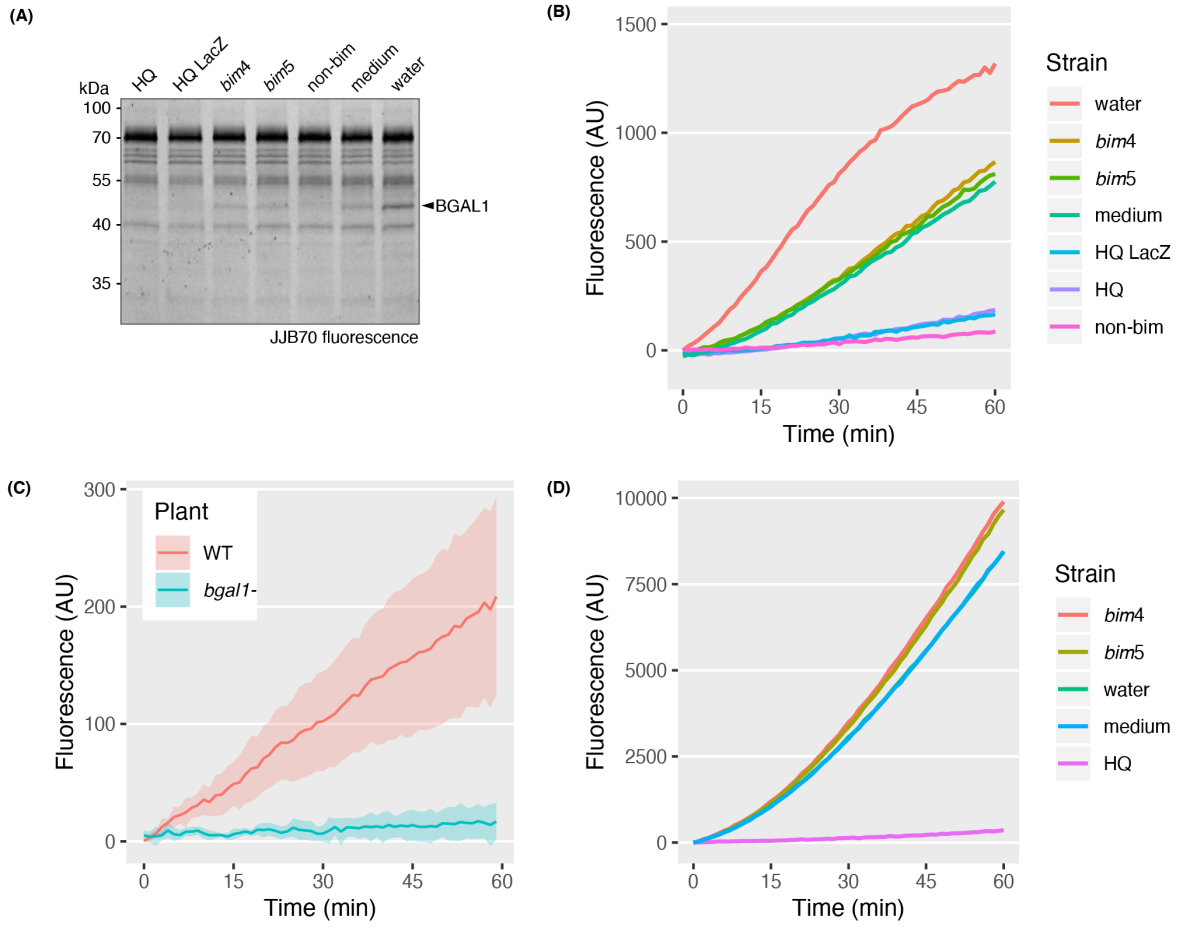


Figure 2.1. **The FDG assay can be used to monitor B-galactosidase activity of both plant apoplastic fluid BGAL1 and *E coli* LacZ and their inhibition by the inhibitor from *P. syringae*. (A) Activity based protein profiling of apoplastic fluid.** Apoplastic fluid from *N. benthamiana* leaves was incubated with supernatant from cultures of different bacterial strains together with fluorescent JJB70 probe in MES buffer pH5. The mixture was then run on a SDS-PAGE gel and fluorescent signal was detected using fluorescent scanning at 488 nm excitation and 520 nm emission. *P. syringae* DC3000 Δ*hopQ*1 (HQ) and its derivative expressing lacZ (HQ lacZ) can produce the inhibitor. The transposon insertion mutants *bim*4 and *bim*5 do not produce the inhibitor, while non-*bim* mutant can still produce the inhibitor. **(B) FDG assay on plant apoplastic fluid.** Apoplastic fluid was incubated with supernatant from cultures of different bacterial strains (same samples as A) together with FDG substrate in MES buffer pH5. Fluorescence of the released fluorogenic product was measured over time indicating enzymatic activity. **(C) FDG assay on plant apoplastic fluid.** Apoplastic fluids from wildtype (WT) and BGAL1 knockout (*bgal1*-) plants were incubated with FDG substrate in MES buffer pH5. Fluorescence of the released fluorogenic product was measured over time indicating enzymatic activity. Shaded area represents standard deviation (N = 6). **(D) FDG assay on *E coli* LacZ.** *E coli* LacZ was incubated with supernatant from cultures of different bacterial strains together with FDG substrate. Fluorescence of the released fluorogenic product was measured over time indicating enzymatic activity.

2.3.2 A streamlined workflow was developed for the screen for BGAL1 inhibitor mutants

In order to identify genes that are involved in BGAL1 inhibitor production in *P. syringae*, I developed a forward genetic screen to identify BGAL1 inhibitor mutants (*bim*). The overview of the optimised screen is outline below (Figure 2.2). First, a mutant library was generated using a mini-Tn5 transposon that inserts into relatively random positions in the *P. syringae* genome, some of which could disrupt genes that are required for the inhibitor production, resulting in *bim* mutants. To aid the identification of *bim* mutants from the library, this screen used a combination of recombinant *P. syringae* constitutively expressing β-galactosidase (*lacZ*) and a chromogenic substrate X-gal in the medium, whose cleavage by LacZ yields a blue product. By plating the bacteria, colonies formed by *bim* mutants unable to produce the inhibitor to suppress the activity of LacZ, could be identified by the deeper blue colour of highly accumulated X-gal cleavage product (Figure 2). Then, the selected candidate colonies were grown in liquid culture and their supernatants were tested for inhibitory activity to confirm the inability to produce inhibitor.

The confirmed *bim* mutants were used for transposon insertion site determination with inverse PCR. Briefly, the bacterial genomic DNA was digested with a restriction enzyme to produce small fragments that were self-ligated to produce circular products, some of which contain the transposon-inserted genomic region. Next, inverse PCR uses primers that bind the transposon and amplify in the outward direction to the flanking region. The gene disrupted in each transposon mutant was then identified by sequencing the inverse PCR product and BLAST against the bacterial genome. This screen yielded candidate genes required for inhibitor production both in biosynthetic and regulatory pathways.

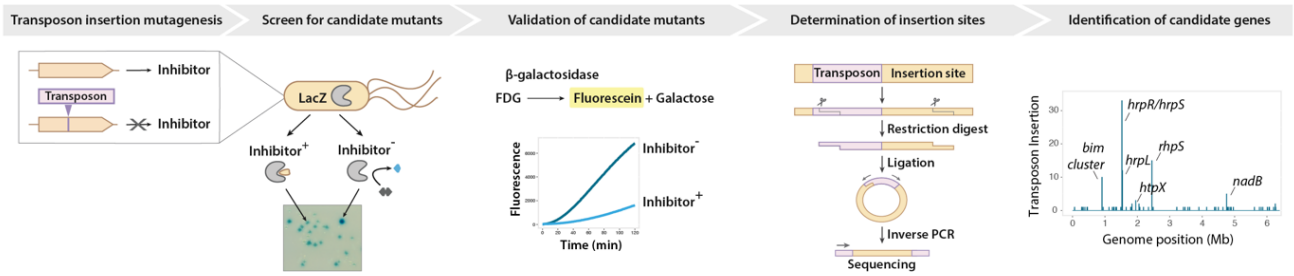


Figure 2.2. Overview of the screen for BGAL1 inhibitor mutants in *P. syringae.* A library of mutants was generated by random transposon insertion mutagenesis using mini-Tn5 transposon. *P. syringae* DC3000 strain expressing lacZ was used for the screen to be able to monitor B-galactosidase activity and inhibitor production. When plated on mannitol glutamate medium supplemented with X-gal, colonies of the desired BGAL1 inhibitor mutants cannot produce the inhibitor to suppress activity of LacZ resulting in more X-gal substrate cleavage and darker blue color from the accumulated product. The candidate colonies were then cultured in 96 well plate and further validated using FDG assay to confirm the absence inhibitor production in the supernatant. The putative BGAL1 inhibitor mutants were then used for inverse PCR and sequencing to determine transposon insertion sites, which were then mapped in the genome to identify candidate genes involved in BGAL1 inhibitor production.

2.3.3 A simulation of random transposon insertion coverage in the genome was used to determine the number of colonies to screen for BGAL1 inhibitor mutants

In a forward genetic approach, a sufficient number of mutant bacterial colonies have to be screened in order to ensure that saturation or near saturation is achieved so that the majority of the genes in the genomes are disrupted by the transposons and the resulting phenotypes are screened. To estimate the number of colonies to be screened, transposon insertions were simulated with the simplified assumption that insertions occur only once at random sites in the genome in each colony. The percentage of genes in the genome being hit by the transposon at least one, five or ten times was calculated given different number of colonies screened (Figure 2.3). The percentage of the genes covered increased sharply as the number of colonies screened increased from the lower end but started to level off as the number of colonies reached a higher level. This suggested that after a certain threshold a higher number of colonies would provide only marginal increase in gene coverage and that threshold could be the effective number. A screen of 150,000 colonies likely provides sufficient degree of

saturation as more than 99.5% of the genes are targeted at least once with the majority being targeted multiple times. This provided the target number of colonies to be used in the screen.

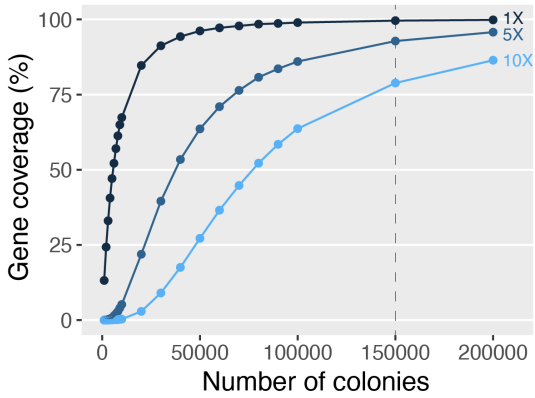


Figure 2.3. A simulation of random transposon insertion coverage in the genome was used to determine the number of colonies from the mutant library screened for BGAL1 inhibitor mutants. The simplified model assumes each colony has one insertion of the transposon in the genome occurring randomly. The simulated insertion sites were then mapped to the coding sequence of genes in the genome. Gene coverage is calculated as the percentage of genes in the genome that are hit at least one, five or ten times provided the number of colonies screened. A total of around 150,000 colonies was predicted give a coverage of more than 99.5% and was used in the screen.

2.3.4 An optimised plating method enabled screening of BGAL1 inhibitor mutant colonies

The next important step to establish in the screen was the detection of the β-galactosidase inhibitor production phenotype as I need to distinguish the *bim* mutants of interest that are unable to produce the inhibitor from the library of transposon insertion mutants. I used a *P. syringae* strain constitutively expressing β-galactosidase (*lacZ* in a plasmid) and a chromogenic substrate X-gal plated in the MG medium, which induces inhibitor production. As β-galactosidase cleaves X-gal and yields a blue coloured product, the degree of blue colour was used as a proxy for enzymatic activity and LacZ inhibition in the bacterial colony. The paler blue indicated inhibitor production by the bacteria to suppress enzymatic activity whereas the deeper blue indicated no inhibitor production by the bacteria to suppress the enzymatic activity. In practice, the plating method had to be optimised to clearly show the different inhibitor production phenotypes. As shown in Figure 2.4A, unoptimised plating

resulted in high background colour and bacterial colonies could not be distinguished. I reasoned that the background was from the presence of LacZ from lysed cells converting X-gal to colour the background blue and this may also deplete the substrate that could instead be converted in the bacterial cells. The main optimisation involved lowering the pH of the medium from 7 to 5.5, which is less optimal for LacZ activity, to suppress the extracellular enzyme present in the agar. The cell density was also optimised not to be too dense, and cells were prewashed gently before plating to minimise excessive leakage of LacZ into the medium. In addition, cells were preincubated in MG medium before plating to induce the production of the inhibitor. This could further reduce the background and improve the difference between inhibitor production phenotypes. This optimised plating method allowed the *bim* mutants to be identified (Figure 2.4B). Nevertheless, distinguishing the mutants by the degree of blue colour was challenging as in some cases the difference was not clear. Since the objective was to try to capture as many *bim* mutants as possible, the initial screen was conducted in a way that any colonies suspected of having deeper blue colour were picked. These were subjected to the next step of validation to ensure false positives were then filtered out and confirm real *bim* mutants. In total, around 150,000 colonies were screened and around 900 candidate *bim* mutant colonies were selected.

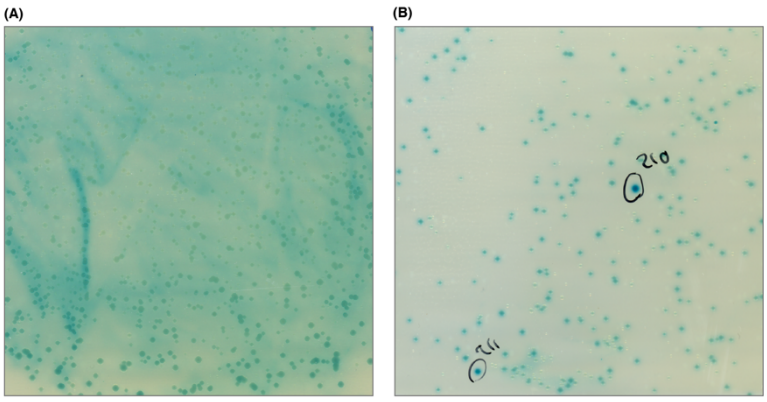


Figure 2.4. An optimised plating method enabled screening of candidate BGAL1 inhibitor mutants. The representative picture of plates from **(A)** unoptimized method and **(B)** optimized method. The candidate *bim* mutant colonies were identified in the circle and labeled.

2.3.5 A high throughput FDG assay was used to validate candidate BGAL1 inhibitor mutants

After the previous screen, candidate *bim* mutants had to be tested against false positives. The high throughput FDG assay with LacZ enzyme was used to further validate the inability of candidate mutants to produce the inhibitor. The selected colonies were inoculated into a 96 well plate and grown overnight in MG medium to induce inhibitor production. The supernatant was then collected, boiled to inactivate any residual enzyme and tested for inhibitor production by FDG assay, monitoring FDG hydrolysis by LacZ overtime using a plate reader. The representative results of one plate are shown in Figure 2.5A. Wells containing the inhibitor producing bacteria showed little or no increase in fluorescence over time as a result of suppressed enzymatic activity, while wells containing *bim* mutants showed faster increase in fluorescence since the inhibitor was not produced. To help distinguish the *bim* and non-*bim* mutants, each plate included three wells of each of the known *bim* mutant (*bim4*) and non-*bim* mutant as controls. The FDG assay fluorescence reading was then used to calculate the rate of β-galactosidase activity, which was then used for classification of strains into two groups of *bim* and non-*bim* mutant by k-means clustering (Figure 2.5B). The mutants that were clustered with the *bim* mutant were selected and subjected to another round of FDG assay validation screen to confirm their *bim* mutant phenotype. The mutants may either have totally loss the ability to produce the inhibitor or partially but drastically reduced ability to produce the inhibitor. Both cases are of interest for identification of genes involved in the inhibitor production. In total, from the testing of around 900 candidate colonies from the initial plating screen, 146 candidate *bim* mutants were confirmed to no longer produce the inhibitor.

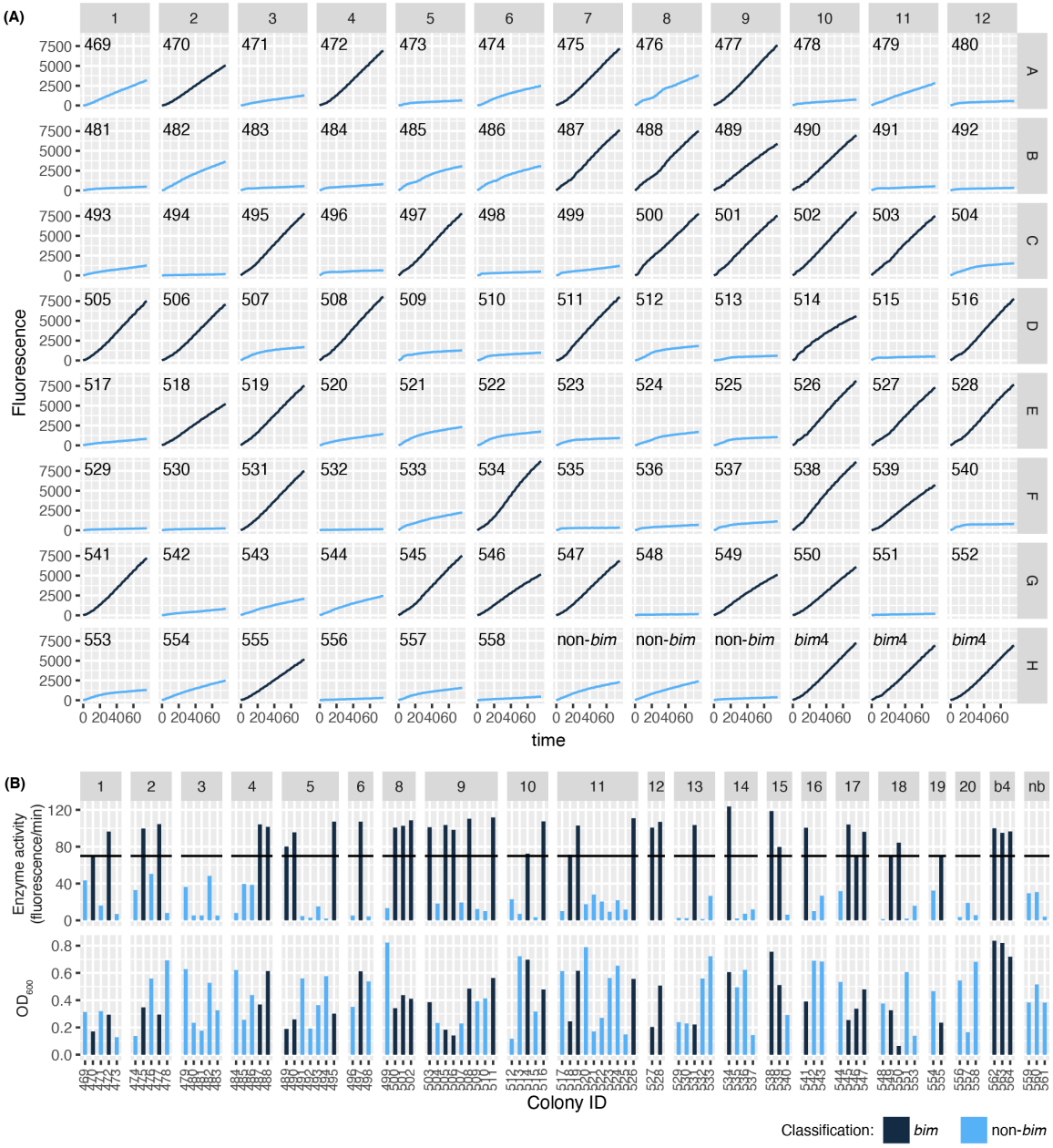


Figure 2.5. A high throughput FDG assay was developed for validation screen of candidate BGAL1 inhibitor mutants. (A) Candidate mutant colonies were numbered and inoculated into each 96 well plate and grown overnight. The culture supernatants were then mixed with the FDG assay reaction mixture and fluorescence signal was measured over time. **(B)** The rate of β-galactosidase activity was calculated for each mutant and was used to classify the BGAL1 inhibitor mutant (*bim*) or non-*bim* by k-means clustering. In every plate, three of the *bim4* (b4) and non-*bim* (nb) controls were used as standards. Non-*bim* mutants still produce the inhibitor resulting in lower enzymatic activity while *bim* mutants cannot produce the inhibitor resulting in higher enzymatic activity.

2.3.6 Transposon insertion sites in *bim* mutants were identified by inverse PCR followed by sequencing

In order to identify genes that are disrupted in *bim* mutants and hence required for inhibitor production, the sequence of transposon insertion sites in the selected *bim* mutants had to be determined. This was achieved by using inverse PCR to amplify the region of the genome adjacent to the transposon insertion site then sequencing the PCR product. Inverse PCR involves digesting the genome into fragments with a restriction enzyme then self-ligating those fragments to produce circular DNA. The fragment that contains the transposon together with the flanking genome regions could then be amplified using primers that bind to the transposon in the outward direction. The resulting PCR product could then be sequenced to determine the insertion site (Figure 2.6A). The adaptation of this concept resulted in the protocol optimised for my screen pipeline as follows.

The first step was to select the restriction enzyme to give fragments that are of optimal length which is not too long for PCR amplification (< 6kb) but long enough to allow unambiguous identification of the genome region (>100 bp). In addition, the restriction site should be present in the transposon so that the fragment containing transposon is included. To select the best restriction enzyme, the genome digest was simulated (Figure 2.6B) to calculate the proportion of fragments that are too big (>12 kb) or too small (<200 bp) (Figure 2.6C). This threshold was twice the length of the set criteria because once the genome region has transposon insertion, the fragment could be digested into up to half the length. EcoRI was not predicted to cut frequently in the genome and thus produces a high proportion of large fragments that cannot be amplified. More frequently cutting enzymes, such as MscI, produced lower proportion of large fragments but fragments can also be too small for

identification. Additionally, the enzyme should produce sticky ends for efficient ligation and be heat inactivatable. The enzyme SphI was chosen based on these described criteria, with ~5% of the genome fragments predicted to be unusable for inverse PCR.

Next, a high throughput format of inverse PCR was developed to allow screening of large number of colonies. This was done in multi well plate format by directly picking bacterial cells and boiling in water to release DNA without a need for purification step. Restriction enzyme was then added to digest the genomic DNA. The digest was heat inactivated before ligase was added to create circular fragments. The concentration of bacteria was also optimised to give sufficiently diluted DNA to promote self-ligation instead of concatenating multiple fragments. The resulting product could then be used as template for PCR. Primers were designed to bind inside the transposon in outward direction to the adjacent genomic region. The primers binding sites were also designed to be adjacent to the border of the transposon to prevent the extra amplification that could arise from erroneous binding and priming of the inverted repeat within the transposon. The PCR products were then sent for Sanger sequencing using a sequencing primer that binds in the PCR primer region but leaves sufficient distance before the insertion site position to allow accurate sequencing read.

This developed inverse PCR protocol was successfully used on different *bim* mutants to obtain PCR products with different sizes from different insertion sites (Figure 2.6D). These PCR fragments were subsequently sequenced to identify transposon insertion sites.

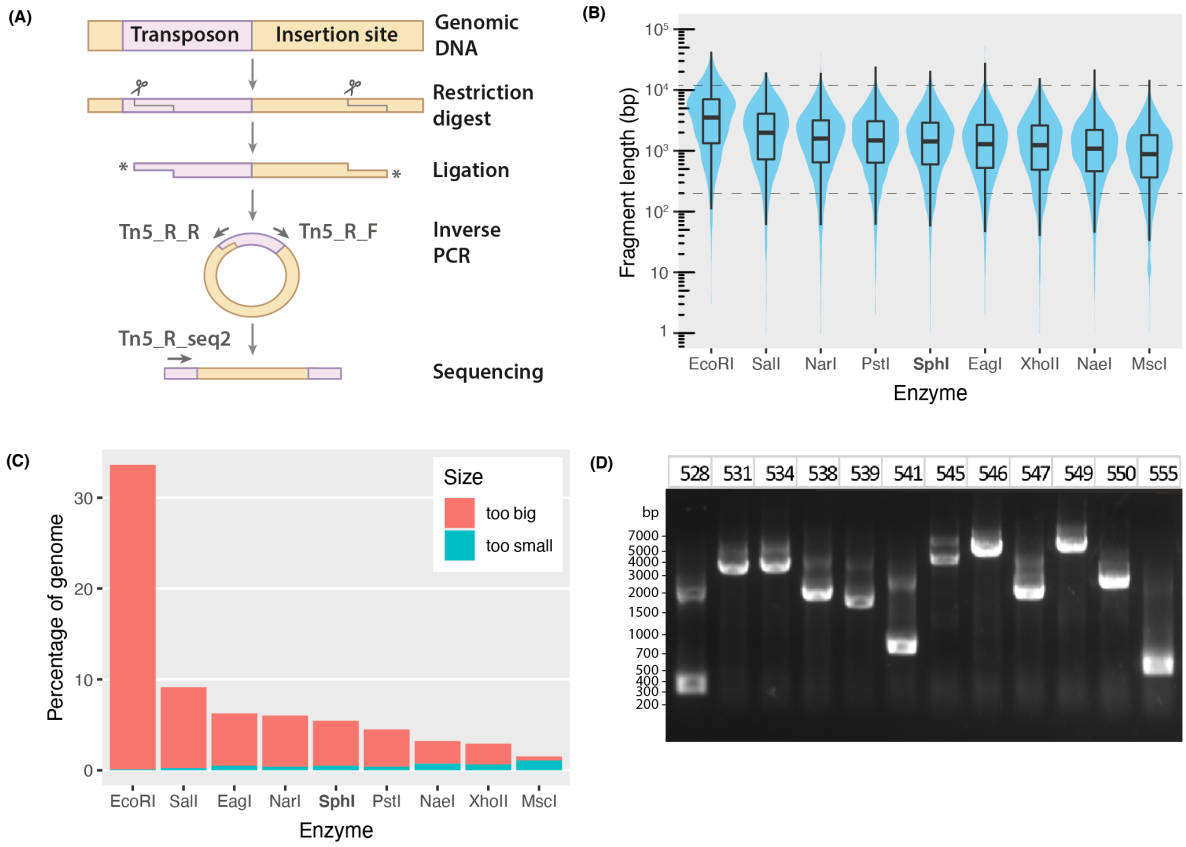


Figure 2.6. Inverse PCR was optimized and used to identify transposon insertion sites in candidate BGAL1 inhibitor mutants. (A) Schematic diagram of inverse PCR protocol. Tn5_R_F and Tn5_R_R are inverse PCR primers. Tn5_R_seq2 is a sequencing primer. **(B)** Distribution of fragment lengths generated when the genome is digested by different restriction enzymes. Box plots show median and interquartile range of the fragment lengths. Violin plots show the distribution of fragment lengths. Dashed lines indicate the threshold of 12 kb and 200 bp. **(C)** Percentage of the genome that is digested by different restriction enzymes into fragments unsuitable for inverse PCR. Fragments larger than 12 kb are classified as too big for PCR amplification and fragments smaller than 200 bp are classified as too small for PCR products to be distinguished on a gel and to give confident mapping of insertion site. **(D)** Example of inverse PCR products separated on agarose gel. Different mutants (numbered) with different insertion sites show different PCR product sizes

2.3.7 The screen for BGAL1 inhibitor mutants identified candidate genes required for BGAL1 inhibitor production

To identify genes that are required for inhibitor production, the sequences of transposon insertion sites determined by inverse PCR of *bim* mutants were used for BLAST searches against the genome so the insertion position could be mapped and the interrupted gene identified. From 146 *bim* mutants sequenced, I identified 135 unique insertion sites in 64 loci (Table 3). When the transposon insertion sites were plotted along the length of the genome,

there were several regions that contained a high number of *bim* mutants, suggesting they corresponded to significant candidate genes (Figure 2.7A). The candidate genes were also plotted for the number of times they were identified with independent insertions in different *bim* mutants since a high number of insertions would indicate the strong candidate genes (Figure 2.7B). Here, I identified strong candidate genes: *hrpR* (22 insertions), *hrpS* (15 insertions), *hrpL* (9 insertions) and *rhpS* (PSPTO2222) (15 insertions), which encode regulators of type III secretion system and a cluster containing PSPTO0834 (3 insertions) and PSPTO0835 (4 insertions), which encode enzymes. Other interesting genes that also appeared multiple times in the screen include *nadB* (5 insertions) and *htpX* (4 insertions).

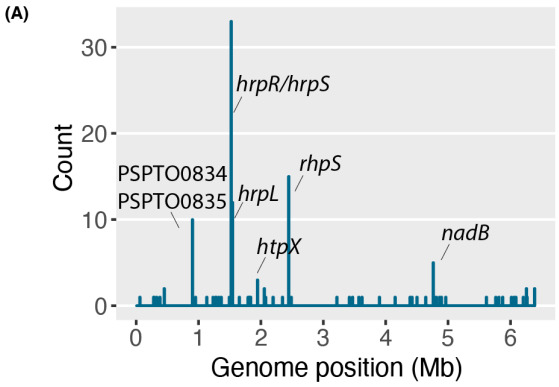


Figure 2.7. Sequencing the transposon insertion sites in BGAL1 inhibitor mutants identified candidate genes involved in inhibitor production. (A) Distribution of the identified insertion site from BGAL1 inhibitor mutant screen across the genome. Histogram binning the count of insertion sites identified in each 10 kb region in the genome. **(B)** Distribution of the amount of times each locus was identified from the screen (the number of independent mutants with transposon insertion hitting the particular locus). Locus was labelled with name followed by PSPTO locus tag number. (continued on next page)

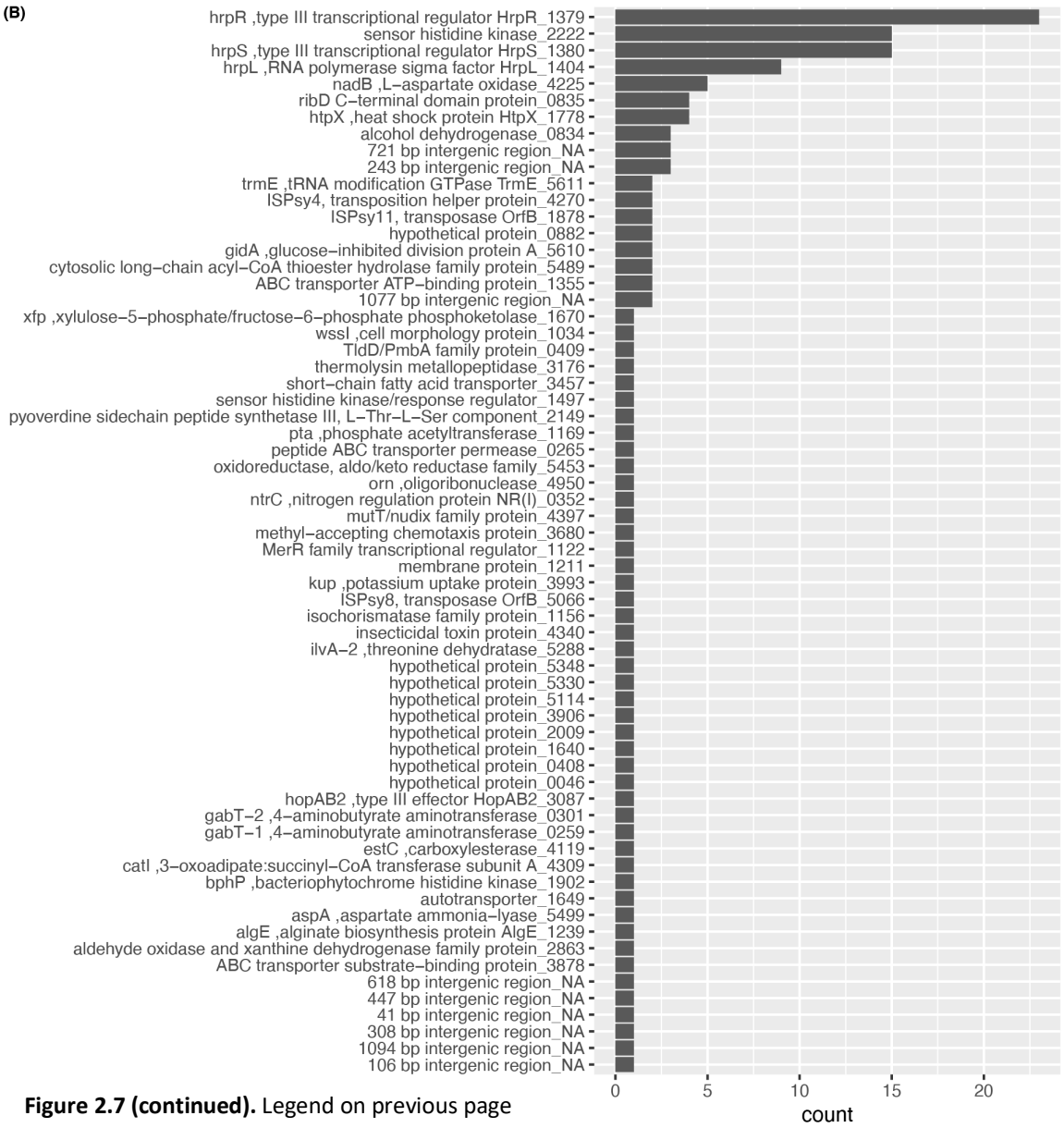


Figure 2.7 (continued). Legend on previous page

2.3.8 The combination of gene expression profiles and comparative genomics pinpointed important candidate genes

To further pinpoint genes that are important for the inhibitor production from the list of candidate genes, I analysed two additional datasets: gene expression during infection and comparative genomics of different *P. syringae* strains.

First, I hypothesised that the expression of genes involved in inhibitor production should be upregulated during infection since the inhibitor is produced during infection. Therefore, I

analysed my list of candidate genes with transcriptomic data (Amelia H Lovelace, Smith and Kvitko, 2018) showing the gene expression profile of *P. syringae* DC3000 during early time points (1, 3 and 5 hours) of infection in *Arabidopsis thaliana* (Figure 2.8). As expected, the expression of Type III secretion system regulators, *hrpR*, *hrpS* and *hrpL*, were induced. PSPTO 0834 and PSPTO 0835 were also highly upregulated supporting the notion that they may function in inhibitor production during infection. Nevertheless, genes that were not upregulated cannot be ruled out, such as *rhpS*, another strong candidate from the list.

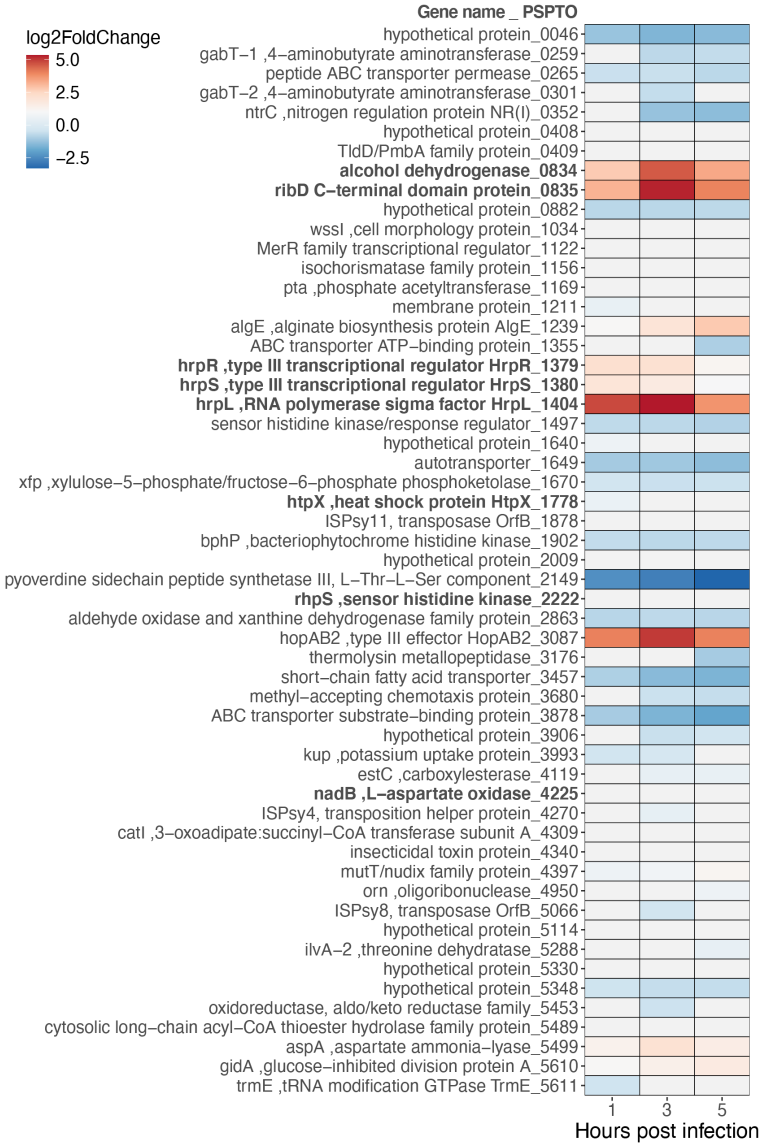


Figure 2.8. Gene expression profile of candidate genes from BGAL1 inhibitor mutant screen during infection. Heat map of log2 fold-change of expression level of genes during early time points of infection (1, 3, 5 hrs) compared to initial bacterial sample before inoculation into plants. Raw data obtained from (Lovelace *et al.*, 2018) All strains are *P. syringae* pathovars, except *Pseudomonas putida* F1*, Pseudomonas fluorescens* SBW25, *Pseudomonas aeruginosa* PAO1 and *E. coli* K12 MG1655.

Next, I hypothesised that genes involved in the inhibitor production should be present in all *P. syringae* strains that can produce the inhibitor but potentially absent (or mostly absent) in the strains that cannot produce the inhibitor. However, they could also be present in non-inhibitor strains if they are also involved in other conserved or housekeeping processes. Therefore, I used a comparative genomics approach to find patterns in the distribution of candidate genes across inhibitor-producing and non-producing strains. A set of *P. syringae* strains, together with other bacterial strains: *P. fluorescens*, *P. putida*, *P. aeruginosa* and *E. coli,* with available genome sequence data were chosen. Their ability to produce BGAL1 inhibitor was tested using FDG assay on supernatants from overnight cultures in MG medium, the same way as the screen above. Using Orthofinder software (Emms and Kelly, 2015), their genomes were compared to find the presence or absence of gene orthologs, which were then mapped on to the list of candidate genes from my mutant screen (Figure 2.9). Again, PSPTO0834 and PSPTO0835 appeared as strong candidates being mostly absent in non-producers and present in inhibitor producers.

The KN91 strain is a notable exception since it can produce the inhibitor but PSPTO0834 is predicted to be absent (Figure 2.9). However, this appears to be due to misannotation of the genome, as the DNA sequence is indeed present in a cluster with PSPTO0835. The proteome database predicted its absence because the gene sequence contains some unassigned bases (N) due to incomplete genome assembly. The candidate genes *hrpR*, *hrpS*, *hrpL* and *rhpS* were also found in inhibitor producing strains. Other notable exceptions were the N6801 and 1302A strains, since they have PSPTO0834 and PSPTO0835 sequences but did not produce the inhibitor in the test (Figure 2.9). This could be because a different regulatory mechanism in these strains preventing the induction of inhibitor production in the conditions tested.

PSPTO0835 from these two strains also contain an amino acid sequence difference, compared to those in other inhibitor producing strains, that could potentially disrupt its function in inhibitor production.

Taken together, these results further support strong candidate genes involved in the inhibitor production: *hrpR*, *hrpS*, *hrpL* and *rhpS*, which encode regulators of type III secretion system and a cluster containing PSPTO0834 and PSPTO0835 which encode putative biosynthetic enzymes.

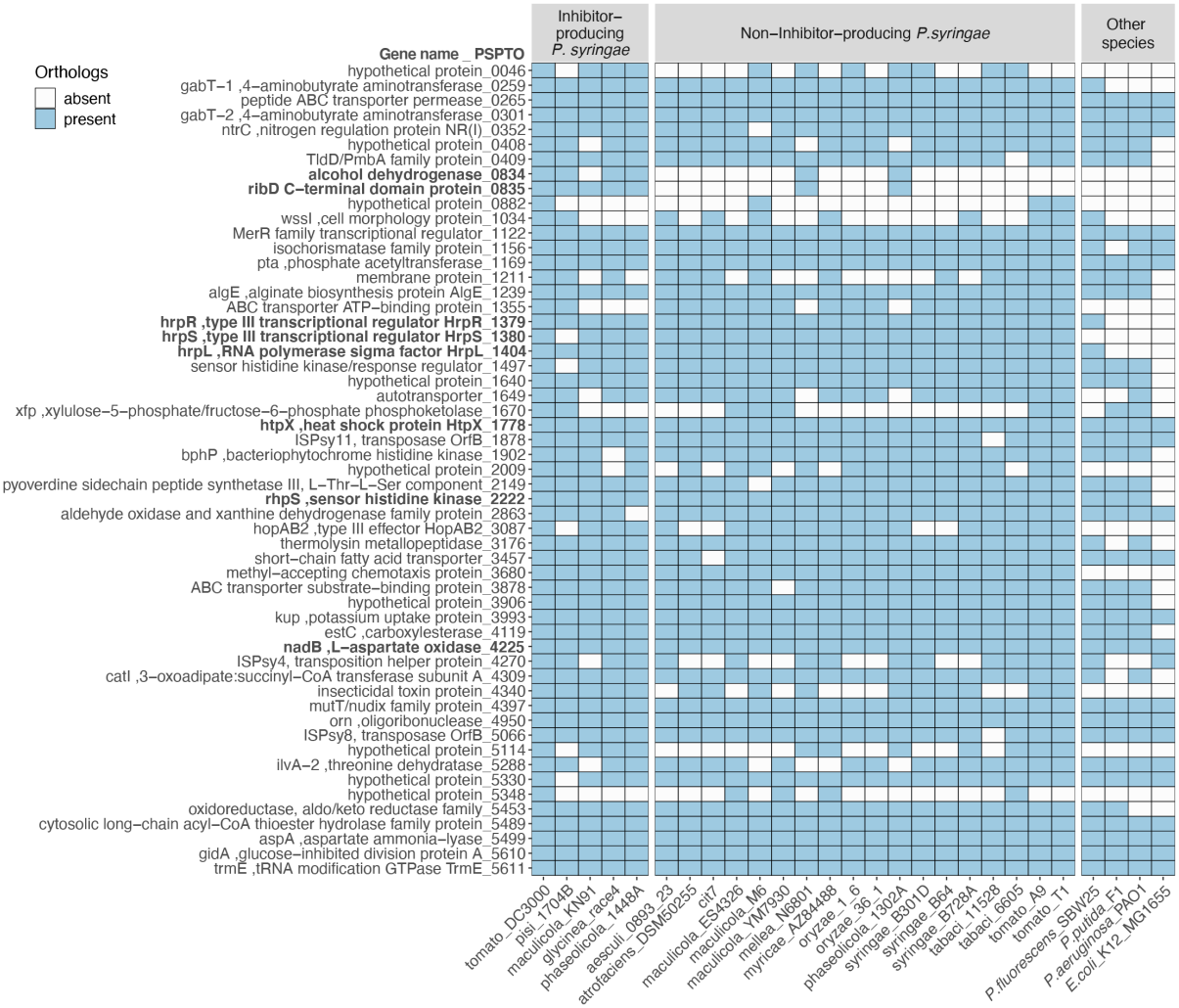


Figure 2.9. Distribution of orthologs of candidate genes from BGAL1 inhibitor mutant screen in different bacterial strains. Heat map showing the presence or absence of orthologs (identified by Orthofinder) of candidate genes from BGAL1 inhibitor mutant screen in the genome of different bacterial strains, classified according to their ability to produce BGAL1 inhibitor. All strains are *P. syringae* pathovars, except *Pseudomonas putida* F1*, Pseudomonas fluorescens* SBW25, *Pseudomonas aeruginosa* PAO1 and *E. coli* K12 MG1655.

2.4 Discussion

In this chapter, I developed a screen for BGAL inhibitor mutants (*bim*) from a transposon insertion mutagenesis library of *P.* syringae, which allowed me to identify candidate genes involved in inhibitor production. There are a number of considerations in the development of this screen, including the methods for monitoring enzymatic activity and inhibition, the determination of transposon insertion sites, the identification of candidate genes and the degree of saturation of the screen.

2.4.1 The screening of enzyme activity and inhibition

My screen was facilitated by the use of synthetic substrates that are specific for the enzyme of interest and yield detectable product upon cleavage by the enzyme. The chromogenic X-gal facilitated high throughput screening of bacterial colonies on plates since it is easy to detect the coloured product without additional instruments and a number of colonies were selected rapidly in the initial screen. This yielded candidates which were then validated by the second round of screening. The fluorogenic substrate FDG then allowed high throughput screening in a multi-well plate format. The fluorescence assay is also sensitive and quantitative allowing validation of the phenotype. If needed, these assays can be repeated in a larger scale to identify more mutants or in other screens such as EMS or UV mutagenesis screen.

In addition to synthetic substrates, commercially available enzymes can also facilitate a screen. In this work, I focused on the inhibitor for plant apoplastic BGAL1. Although the discovery was done with plant apoplastic samples, the variability and amount of samples obtainable from plants can be limiting. In addition, to express the plant enzyme in bacteria

for screening can also be challenging. These problems were overcome by using LacZ, a well-known enzyme with β-galactosidase activity, which is commercially available and easily expressed in bacteria. Although BGAL1 and LacZ belong to different families, GH35 and GH2 respectively, they are both inhibited by the BGAL1 inhibitor and are interchangeable in my screen for inhibitor mutants. This may not always be the case since substrate and inhibitor specificity of enzymes could vary and testing is required. An inhibitor that mimics a substrate is more likely to inhibit a broad range of enzymes with the same substrate specificity because of similarity in the substrate binding groove of the enzyme.

This screening approach can also be applied to other systems that involve enzymatic action or inhibition. A similar application to this work could be the identification of genes involved in other processes by mutant screening. Furthermore, a screening approach similar to this work could be used for other applications, such as metabolic engineering for improved inhibitor production or enzyme engineering for improved substrate specificity, increased enzymatic activity or insensitivity to inhibitors. For instance, a library of *E. coli* containing a random mutations of a biosynthetic enzyme was plated with a colorimetric glycosidase substrate to screen for clones with higher catalytic efficiency that led to enhanced production of a glycosidase inhibitor 1-deoxynojirimycin (DNJ), which resulted in a paler coloured colony from increased inhibition of the enzymatic activity (Jiang *et al.*, 2015). As another example, in a directed evolution experiment, a library of DNA shuffling mutants of LacZ β-galactosidase was screened with a colorimetric fucose-conjugated substrate to select for a change in specificity of the enzyme towards fucosidase activity (Zhang, Dawe and Stemmer, 1997).

2.4.2 The identification of transposon insertion sites

Inverse PCR is a useful approach to determine the unknown sequence adjacent to transposon insertion sites. This is an important step in the screen and several factors are to be considered.

First, since inverse PCR is based on known restriction digest sites and primers binding within the transposon, it is supposed to be reproducible and give the same pattern of band size. This method works with low concentrations of template DNA that promote self-ligation and circularisation. However, in some cases, ligation of multiple fragments occurs resulting in varying lengths of PCR amplicons from one inverse PCR reaction, as observed in some of the *bim* mutants showing multiple bands of PCR products after gel electrophoresis. Nevertheless, this does not always affect the identification of transposon insertion site. Since the sequencing primer binds to the end of the transposon and directs the sequencing read in the outward direction into the insertion site that flanks the transposon, the sequence immediately adjacent to the transposon sequence is always the insertion site. Any sequences that are ligated after that fragment can be discarded. However, careful analysis of the sequencing results can prevent misinterpreting the artificially created hybrid sequence from independent fragments ligation. Another problem from ligation of multiple fragments would be that the template is too long for PCR amplification.

Restriction enzyme choice can also impact the success rate of the inverse PCR. Enzymes should be chosen for having optimal frequency of recognition sites in the genome to yield fragments small enough to allow PCR amplification and long enough to allow accurate mapping to the genome. This is especially important for regions that are repeated multiple times in the genome. Long sequences are needed to give the genomic context. Furthermore,

using one particular restriction enzyme may not be enough to amplify and identify all sequence regions in the genome. Another enzyme could be used to complement the inverse PCR method. In my case, more than 95% of tested *bim* mutants were successfully amplified and identified. For the remaining mutants, alternatives methods such as TAIL PCR using semi-degenerate primers (Liu and Whittier, 1995), or next-generation sequencing using shearing and adapter ligation could be explored (Van Opijnen and Camilli, 2013; Chao *et al.*, 2016).

Primer design was also found to be important in the inverse PCR protocol. The primer pair for PCR amplification binds the transposon in the outward direction pointing into the adjacent genomic region. One challenge of this is that the amplicon includes the inverted repeat sequences present at the border of the transposon, which can erroneously hybridise during PCR amplification and yield an unintended product that interferes with sequencing. Gel extraction of the right band size is one way to overcome this but it is low throughput and efficiency. Instead, this was overcome with a new primer that binds directly to the end of transposon. To allow accurate Sanger sequencing from the insertion site position, the sequencing primer was designed to bind in the inverse PCR primer which was long enough to provide leading space for the beginning of sequencing read. My implementation of inverse PCR allows high throughput workflow which can be adapted to facilitate another similar transposon mutagenesis screen.

2.4.3 The candidate genes involved in BGAL1 inhibitor production

This work has successfully developed a screen and strategy to identify genes involved in BGAL1 inhibitor production. Candidate genes with high confidence appeared multiple times with independent insertions in different *bim* mutants. Comparative genomics and gene

expression data further supported the candidate genes. These genes are further characterised in chapter 3. Nevertheless, the majority of other candidate genes identified from the screen only appeared once in a single *bim* mutant, not indicating strong confidence that they were true positives. These genes could be further validated by gene complementation and independent knockout mutants. Resequencing the genome of the mutants could also be done to identify if there are secondary mutations, such as in the regulators or biosynthetic genes, affecting inhibitor production.

Among the most prominent candidate genes identified in this screen, *hrpR*, *hrpS* and *hrpL* are regulators of type III secretion system, a major virulence factor in *P. syringae* (Tang, Xiao and Zhou, 2006). During infection, enhancer binding proteins HrpR and HrpS act together to activate expression of *hrpL*, a sigma factor that coordinate expression of type III secretion system and effectors which play important roles during infection. Another strong candidate gene, *rhpS* encodes a sensor histidine kinase functioning in a two-component signalling cascade that regulates various genes including *hrpR/hrpS* operon (Tang, Xiao and Zhou, 2006). These results suggest that the inhibitor production is regulated by the signalling cascade that also regulates the type III secretion system. Their co-regulation is consistent with the idea that the inhibitor is a virulence factor produced during infection and in MG media.

A cluster of PSPTO0834 and PSPTO0835 genes was another strong candidate identified in *bim* mutants with multiple insertions both in each gene and its upstream promoter region. Interestingly, this promoter region contains a consensus sequence for HrpL sigma factor binding and is identified as one of the genes regulated by the HrpL regulon (Vencato *et al.*, 2006). Interestingly, a PSPTO0834 mutant was also found to be less virulent in Arabidopsis infection assays (Vencato *et al.*, 2006). The function of this cluster is uncharacterised. PSPTO

0834 is a predicted alcohol dehydrogenase, whereas PSPTO0835 encodes a predicted riboflavin biosynthesis RibD C-terminal domain protein having a putative reductase activity. I hypothesise that these two enzymes are involved in inhibitor biosynthesis.

A study of biosynthesis of 1-deoxynojirimycin (DNJ) (Onose *et al.*, 2013), an iminosugar inhibitor of α-glucosidase, in Bacillus species identified a cluster of three genes encoding biosynthetic enzymes: a GabT1 transaminase, Yktc1 phosphatase and GutB1 oxidoreductase. Interestingly, my BGAL1 inhibitor mutant screen also identified PSPTO0834, the homolog of GutB1. GabT1 and GabT2, two homologs of the transaminase were also identified but not as strong candidates and may act redundantly (Park *et al.*, 2010). However, Yktc1 has no ortholog in *P. syringae*. This suggests that the BGAL1 inhibitor might be a molecule of a similar class to DNJ thus sharing a subset of biosynthetic enzymes.

Another candidate gene *nadB* (L-aspartate oxidase) functions in the NAD biosynthetic pathway, suggesting that NAD cofactor could be involved in the biosynthesis of the inhibitor. However, it is unlikely that cells can survive without NAD, so the *nadB* mutant should still be able to produce NAD via an alternative pathway (Devries *et al.*, 1995). Nevertheless, the mutation of *nadB* could reduce NAD production to the level that can impair inhibitor production. The role of NAD is also supported by the need of the putative oxidoreductase PSPTO0834 and PSPTO0835 in the biosynthesis of the inhibitor.

Another interesting candidate gene is *htpX*. It is unclear how HtpX is involved with inhibitor production. A homolog of *htpX* in *E. coli* encodes a membrane bound protease involved in proteolytic quality control of membrane proteins and response to membrane stress which

may or may not include heat induction (Sakoh, Ito and Akiyama, 2005). It may be involved in part of regulatory pathways for inhibitor production.

2.4.4 The saturation screen for identifying gene candidates

Screen saturation is important for identifying all genes involved in the process of interest during a mutant screen. To determine the target number of mutant colonies to screen or estimate the degree of saturation of a screen, simulation or modelling can be used. In this study, I took the simplistic simulation approach and assumption of purely and uniformly random insertion across the genome. The insertion within coding regions of genes were then scored. This provided useful estimation for my work. However, further considerations could be taken into account. For example, insertion target site bias of mini-Tn5 transposon has been reported, especially towards G (Green *et al.*, 2012). This was indeed observed in my mutants sequencing data with 70% of the insertion site determined start with G. Also, I did not take into account that a transposon insertion could disrupt a gene function or not depending on the position within the gene. Furthermore, insertion in regulatory regions, such as promoters, could disrupt gene function. For instance, I found insertions in the promoter regions of the gene cluster of PSPTO0834 and PSPTO0835. Other approaches have been developed to estimate mutagenesis and screen saturation based on different statistical models and probability distribution (Page and Grossniklaus, 2002; Pollock and Larkin, 2004).

My screen likely reached saturation and identified most possible candidates since multiple mutants were isolated for each strong candidate genes. Still, there are limitations of this forward genetic screen approach using transposon insertion. Genes required for inhibitor production might not be isolated from this approach for various reasons. For example,

essential genes in housekeeping metabolic pathways or genes which are lethal if disrupted would not be identified. In addition, if multiple genes function redundantly in inhibitor production then disrupting one gene is not enough to cause the knockout phenotype. Another possibility is that some genes are missed by not containing a target sequence for transposon insertion or by being too small.

2.4.5 Conclusion

I have successfully developed a screen for *P. syringae* BGAL inhibitor mutants (*bim*) that led to an identification of a number of candidate genes involved in inhibitor production, including PSPTO0834, PSPTO0835 and *nadB* as candidate genes for inhibitor biosynthesis and *hrpR*, *hrpS*, *hrpL*, *rhpS* and *htpX* as candidate genes for regulation of inhibitor production. In the next chapter, these genes will be further characterised to provide more insights into the biosynthetic pathway, regulatory pathway and evolution of the inhibitor production. In addition, I also established the enzymatic assay using a fluorogenic substrate FDG for determination of β-galactosidase activity and inhibition, which will be used in further experiments for inhibitor characterisation in the next chapter.

Chapter 3: The characterisation of genes involved in BGAL inhibitor production

3.1 Introduction

From the previous chapter, I have screened and identified candidate genes involved in BGAL inhibitor production. Consequently, I can validate their roles and investigate different aspects of inhibitor production. First, the biosynthetic pathway of the inhibitor and the enzymes involved are important to know for a new natural product and can have applications in microbial production and metabolic engineering. Next, the evolution of inhibitor production can be explored not only for understandings of inhibitor production in *P. syringae* but also for identification of broader context of inhibitor production in other species with the potential of new natural product discovery. In addition, characterisation of the role of the BGAL inhibitor in plant infection could identify a new virulence mechanism used by *P. syringae* for pathogenesis. Furthermore, the study of the regulation of this inhibitor production could also yield insights into novel regulators of virulence mechanisms in *P. syringae*.

The strong candidate genes required for inhibitor production identified from my screen can be divided into two groups, the first involving biosynthesis of the inhibitor and the second involving regulation of the inhibitor production. For biosynthesis, the candidates PSPTO0834 and PSPTO0835 are from the same gene cluster. PSPTO0834 is a predicted alcohol dehydrogenase, whereas PSPTO0835 is a predicted riboflavin biosynthesis RibD C-terminal domain protein with a reductase activity. Interestingly PSPTO0834 is a homolog of GutB1, which was identified in the biosynthetic gene cluster of an iminosugar glycosidase inhibitor 1-deoxynojirimycin (DNJ) (Clark, Johnson and Horenstein, 2011). The DNJ cluster consists of

three genes encoding enzymes involved in the proposed biosynthetic pathway including GabT1 transaminase, Yktc1 phosphatase and GutB1 oxidoreductase. Another candidate gene identified in my screen, *nadB*, encodes an L-aspartate oxidase enzyme required for the NAD cofactor biosynthesis.

For regulation, the prominent candidates, *rhpS, hrpR*, *hrpS* and *hrpL* are regulators of the type III secretion system, a major pathogenicity factor in *P. syringae* (Xie, Shao and Deng, 2019). During infection, HrpR and HrpS function together as enhancer binding proteins to activate expression of *hrpL*, a sigma factor that initiates expression of pathogenicity and virulence factor genes including the type III secretion system and type III-secreted effectors. *RhpS* is a sensor histidine kinase in a two-component signalling cascade that regulates expression of various genes including the *hrpR/hrpS* operon (Tang, Xiao and Zhou, 2006). Finally, another potential novel regulator, *htpX*, encodes a membrane bound protease with homologs involved in proteolytic quality control of membrane proteins and responses to membrane stress, which may or may not include heat induction (Sakoh, Ito and Akiyama, 2005).

Regarding the role of BGAL1 inhibitor in plant infection, our previous study identified the role of the enzyme BGAL1 in facilitating the release of elicitor peptides from flagellin protein containing glycosylation with modified viosamine. The production of BGAL1 inhibitor by *P. syringae* was hypothesised to target BGAL1 suppression as a mechanism to protect the bacterial strains with the BGAL1 sensitive flagellin glycosylation (Buscaill *et al.*, 2019).

In this chapter, I described the characterization of the biosynthetic gene cluster for the BGAL1 inhibitor and discussed the potential players in the biosynthetic pathway. Next, I investigated the evolution of this gene cluster in *P. syringae* and the homologs in other bacterial species.

Then, I characterised the regulation of inhibitor production in *P. syringae* and identified the regulators involved. Finally, I explored the role of BGAL inhibitor in bacterial virulence during plant infection.

3.2 Materials and methods

3.2.1 Bacterial strains, culture conditions and media

Bacterial strains used in this study are listed in Table 3.1. Cultures were incubated at 28°C for *P. syringae* and 37°C for *E. coli*, with shaking at 220 rpm for liquid cultures. Media used for bacterial cultures were made as follows:

- LB medium: 10g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl.
- Mannitol glutamate (MG) medium (Bronstein *et al.*, 2008): 10 g/l mannitol, 2 g/l L-glutamic acid, 0.5 g/l KH_2PO_4, 0.2 g/l NaCl, 0.2 g/l $MgSO_4.7H_2O$, adjusted to pH 7.
- Solid media for plating contains 1.5% (w/v) agar.
- Antibiotics were generally used at the concentrations of 25 μg/ml rifampicin, 12.5 μg/ml gentamicin, 50 μg/ml kanamycin, 10 μg/ml tetracycline

Table 3.1. Bacterial strains used in this study

Strain	Description	Antibiotics	Reference
HQ	*P. syringae* pv. *tomato* DC3000 lacking the effector *hopQ1* allowing it to cause disease in the model plant *N. benthamiana*	Rifampicin	(Wei *et al.*, 2007)
HQ Δ*bim*	HQ with deletion of the *bim* cluster (PSPTO0834-0838)	Rifampicin	This study
HQ Δ0834	HQ with deletion of PSPTO0834	Rifampicin	This study
HQ Δ0835	HQ with deletion of PSPTO0835	Rifampicin	This study
HQ Δ0836	HQ with deletion of PSPTO0836	Rifampicin	This study
HQ Δ0837	HQ with deletion of PSPTO0837	Rifampicin	This study
HQ Δ0838	HQ with deletion of PSPTO0838	Rifampicin	This study
HQ Δ0836 Δ5452	HQ with deletion of PSPTO0836 and PSPTO5452	Rifampicin	This study
HQ Δ0837 Δ4071	HQ with deletion of PSPTO0837 and PSPTO4071	Rifampicin	This study
HQ Δ*bim* pBBR1MCS2 EV	HQ Δ*bim* carrying pBBR1MCS2 empty vector	Rifampicin, Kanamycin	This study
HQ Δ*bim* pBBR1MCS2 *Pnative*::0834-5	HQ Δ*bim* carrying pBBR1MCS2 with partial *bim* cluster (PSPTO0834-0835) under native promoter	Rifampicin, Kanamycin	This study
HQ Δ*bim* pBBR1MCS2 *Pnative*::0834-8	HQ Δ*bim* carrying pBBR1MCS2 full *bim* cluster (PSPTO0834-0838) under native promoter	Rifampicin, Kanamycin	This study
HQ Δ*hrpA*	HQ with deletion of *hrpA* (PSPTO1381)	Rifampicin	This study
HQ Δ*hrpA* Δ*bim*	HQ Δ*hrpA* with deletion of the *bim* cluster (PSPTO0834-0838)	Rifampicin	This study

*bim*549	HQ with mini-Tn5::GFP::Lux insertion in *nadB* (PSPTO4225), luminescence reporter of gene expression	Rifampicin, Kanamycin	This study
*bim*549 complemented	*bim*549 carrying pBBR1MCS3 with *nadB* (PSPTO4225) driven by *Lac* promoter	Rifampicin, Kanamycin, Tetracycline	This study
*bim*97	HQ with mini-Tn5::GFP::Lux insertion in isochorismatase (*isch*) (PSPTO1156)	Rifampicin, Kanamycin	This study
*bim*97 complemented	*bim*97 carrying pBBR1MCS3 with isochorismatase (*isch*) (PSPTO1156) driven by *Lac* promoter	Rifampicin, Kanamycin, Tetracycline	This study
*bim*154	HQ with mini-Tn5::GFP::Lux insertion in aspartate ammonia lyase (*aal*) (PSPTO5499)	Rifampicin, Kanamycin	This study
*bim*154 complemented	*bim*154 carrying pBBR1MCS3 with aspartate ammonia lyase (*aal*) (PSPTO5499) driven by *Lac* promoter	Rifampicin, Kanamycin, Tetracycline	This study
*bim*502	HQ with mini-Tn5::GFP::Lux insertion in threonine dehydratase (*tdh*) (PSPTO5288)	Rifampicin, Kanamycin	This study
*bim*502 complemented	*bim*502 carrying pBBR1MCS3 with threonine dehydratase (*tdh*) (PSPTO5288) driven by *Lac* promoter	Rifampicin, Kanamycin, Tetracycline	This study
*bim*796	HQ with mini-Tn5::GFP::Lux insertion in *gabT2* (PSPTO0301)	Rifampicin, Kanamycin	This study
*bim*796 complemented	*bim*796 carrying pBBR1MCS3 with *gabT2* (PSPTO0301) driven by *Lac* promoter	Rifampicin, Kanamycin, Tetracycline	This study
*bim*77	HQ with mini-Tn5::GFP::Lux insertion in *htpX* (PSPTO1778)	Rifampicin, Kanamycin	This study
*bim*77 complemented	*bim*77 carrying pBBR1MCS3 with *htpX* (PSPTO1778) driven by *Lac* promoter	Rifampicin, Kanamycin, Tetracycline	This study
*bim*210	HQ with mini-Tn5::GFP::Lux insertion in *htpX* (PSPTO1778)	Rifampicin, Kanamycin	This study
*bim*322	HQ with mini-Tn5::GFP::Lux insertion in *rhpS* (PSPTO2222), luminescence reporter of gene expression	Rifampicin, Kanamycin	This study
*bim*244	HQ with mini-Tn5::GFP::Lux insertion in *hrpR* (PSPTO1379)	Rifampicin, Kanamycin	This study
*bim*442	HQ with mini-Tn5::GFP::Lux insertion in *hrpR* (PSPTO1379), luminescence reporter of gene expression	Rifampicin, Kanamycin	This study
*bim*516	HQ with mini-Tn5::GFP::Lux insertion in *hrpS* (PSPTO1380)	Rifampicin, Kanamycin	This study
*bim*436	HQ with mini-Tn5::GFP::Lux insertion in *hrpS* (PSPTO1380), luminescence reporter of gene expression	Rifampicin, Kanamycin	This study
*bim*331	HQ with mini-Tn5::GFP::Lux insertion in *hrpL* (PSPTO1404)	Rifampicin, Kanamycin	This study
*bim*511	HQ with mini-Tn5::GFP::Lux insertion in *hrpL* (PSPTO1404), luminescence reporter of gene expression	Rifampicin, Kanamycin	This study
*bim*744	HQ with mini-Tn5::GFP::Lux insertion in PSPTO0835, luminescence reporter of gene expression	Rifampicin, Kanamycin	This study
*bim*259	HQ with mini-Tn5::GFP::Lux insertion in *avrptoB* (PSPTO3087), luminescence reporter of gene expression	Rifampicin, Kanamycin	This study

non-*bim*	HQ with mini-Tn5::GFP::Lux insertion in *hopI1* (PSPTO4776)	Rifampicin, Kanamycin	This study
DC3000 WT	*P. syringae* pv. *tomato* DC3000	Rifampicin	Martin Buck Lab
DC3000 Δ*hrpR*	*P. syringae* pv. *tomato* DC3000 lacking *hrpR*	Rifampicin	Martin Buck Lab
DC3000 Δ*hrpS*	*P. syringae* pv. *tomato* DC3000 lacking *hrpS*	Rifampicin	Martin Buck Lab
DC3000 Δ*hrpL*	*P. syringae* pv. *tomato* DC3000 lacking *hrpL*	Rifampicin	Martin Buck Lab
DC3000 Δ*hrpA*	*P. syringae* pv. *tomato* DC3000 lacking *hrpA*	Rifampicin	(E. Roine *et al.*, 1997)
DC3000 Δ*hrcC*	*P. syringae* pv. *tomato* DC3000 lacking *hrcC*	Rifampicin	(E. Roine *et al.*, 1997)
DC3000 Δ*gabT1,2,3*	*P. syringae* pv. *tomato* DC3000 lacking *gabT1*, *2, 3*	Rifampicin	Allan Collmer Lab
E. coli pBBR1MCS3 EV	*E. coli* TOP10 carrying pBBR1MCS3 empty vector	Tetracycline	This study
E. coli pBBR1MCS3 *Plac*::0834-5	*E. coli* TOP10 carrying pBBR1MCS3 with partial *bim* cluster (PSPTO0834-0835) driven by *Lac* promoter	Tetracycline	This study
E. coli pBBR1MCS3 *Plac*::0834-8	*E. coli* TOP10 carrying pBBR1MCS3 with full *bim* cluster (PSPTO0834-0838) driven by *Lac* promoter	Tetracycline	This study
E. coli pBBR1MCS3 *Plac*::0834-5^{mut}	*E. coli* TOP10 carrying pBBR1MCS3 with catalytically inactive mutant *bim* cluster genes (PSPTO0834-0835) driven by *Lac* promoter	Tetracycline	This study
E. coli pBBR1MCS3 *Plac*::gutB1-0835	*E. coli* TOP10 carrying pBBR1MCS3 with gutB1 and PSPTO0835 driven by *Lac* promoter	Tetracycline	This study
E. coli pBBR1MCS2	*E. coli* carrying the broad host range pBBR1MCS2 plasmid used for cloning and gene expression	Kanamycin	(Obranić, Babić and Maravić-Vlahoviček, 2013)
E. coli pBBR1MCS3	*E. coli* carrying the broad host range pBBR1MCS3 plasmid used for cloning and gene expression	Tetracycline	(Obranić, Babić and Maravić-Vlahoviček, 2013)
E. coli pK18mobsacB	*E. coli* carrying the pK18mobsacB plasmid used for generation of deletion mutants	Kanamycin	(Schäfer *et al.*, 1994)
E. coli pRK2013	Helper strain carrying pRK2013 plasmid for conjugal transfer; DH5α λpir background	Kanamycin	(Figurski and Helinski, 1979)

3.2.2 Generation of gene expression constructs

Fragments of the coding sequence of the gene including the preceding ribosome binding site were PCR amplified from genomic DNA using Q5 high fidelity DNA polymerase (New England Biolabs, Ipswich, USA) with primer pairs F and R (Table 3.2). For site-directed mutagenesis to

generate catalytically inactive mutants, the primers mut_R and mut_F (Table 3.2), which contains the desired point mutations sequence, were used together with the R or F primers, respectively, to amplify the gene in two separate fragments. For *gutB1*, the gene was synthesised (Eurofins, Luxembourg). The broad host range pBBR1MCS plasmids were used as vectors (Obranić, Babić and Maravić-Vlahoviček, 2013). pBBR1MCS3 was used for gene expression under control of constitutive *lac* promoter (*Plac*) present in the plasmid. pBBR1MCS2 was used for gene expression under native promoter of the gene (PCR fragment amplified from genomic DNA contains both the native promoter region and coding sequence of the gene). The plasmids were linearised by restriction digest with Fast Digest SmaI (for pBBR1MCS3) or SmaI and NsiI (for pBBR1MCS2, *Plac* removed) (Thermo Scientific, Waltham, USA). The gene fragments were assembled into the restriction digested plasmid backbone by Gibson assembly (Gibson *et al.*, 2009). The assembled product was heat-shock transformed into *E. coli* TOP10 competent cells, plated on LB agar plates with 10 µg/ml tetracycline (for pBBR1MCS3) or 50 µg/ml kanamycin (for pBBR1MCS2) and incubated at 37°C overnight. Colonies were then picked and checked for correct assembly of the plasmid by colony PCR using primer pairs oNS48_pBBR1MCS3_GA_seq_F and oNS49_pBBR1MCS3_GA_seq_R (for pBBR1MCS3) or oNS67_pBBR1MCS3_GA_seq_F2 and oNS49_pBBR1MCS3_GA_seq_R (for pBBR1MCS2) (Table 3.2). Positive colonies were grown overnight in LB medium with appropriate antibiotics and used for plasmid extraction. The plasmid was validated by Sanger sequencing using the same primers used for colony PCR. The validated plasmid constructs were transformed into *P. syringae* by triparental mating conjugation described below.

For catalytically inactive mutants, PSPTO0834 was mutated at two residues (H59A and E60A) predicted to be involved in zinc cofactor binding. PSPTO0835 was mutated at two residues

(T54A and D58A) predicted to be involved in NADP cofactor binding. These residues were chosen since they are conserved residues also found in homologs from *E. coli* with functional annotations in Uniprot database (for PSPTO0834: Threonine dehydrogenase, TDH, P07913 and for PSPTO0835: Riboflavin biosynthesis protein, RibD, P25539) (Bateman *et al.*, 2021).

3.2.3 Generation of gene deletion mutants

Targeted gene deletion in the genome of *P. syringae* was generated by the two-step allelic exchange method (Kvitko and Collmer, 2011; Hmelo *et al.*, 2015) .

Plasmid construct generation: Fragments of around 1 kb flanking each side of the target gene were PCR amplified from genomic DNA using Q5 high fidelity DNA polymerase (New England Biolabs, Ipswich, USA) with primer pairs L_F, L_R (left flanking region) and R_F, R_R (right flanking region) (Table 3.2). The fragments were designed such that the end of the left and right flanking fragments contains the start and stop codon of the target gene, respectively, plus a few adjacent codons in frame. The end result of the deletion would be a drastically truncated gene product encoding a few amino acids at the beginning and end of the protein. The two fragments were assembled by Gibson assembly (Gibson *et al.*, 2009) into the plasmid backbone that was PCR amplified from the suicide vector pK18mobsacB (Schäfer *et al.*, 1994) using Q5 high fidelity DNA polymerase (New England Biolabs, Ipswich, USA) with primers oNS34_pK18mobsacB_GA_F and oNS35_pK18mobsacB_GA_R (Table 3.2). The assembled product was heat-shock transformed into *E. coli* TOP10 competent cells, plated on LB agar with 50 µg/ml kanamycin and incubated at 37°C overnight. Colonies were then picked and checked for correctly assembled plasmid by colony PCR using primers oNS50_pK18mobsacB_GA_seq_F and oNS51_pK18mobsacB_GA_seq_R (Table 3.2). Positive

colonies were grown overnight in LB medium with 50 µg/ml kanamycin and used for plasmid extraction. The plasmid was validated by Sanger sequencing using the same primers used for colony PCR

Mutant strain generation: Plasmid constructs were transformed into *P. syringae* by triparental mating conjugation. The conjugation mixture was plated on LB agar with 25 µg/ml nitrofurantoin and 50 µg/ml kanamycin incubated at 28°C for 2 days. This step selects for *P. syringae* merodiploids from the first cross over between one of the flanking fragments in the suicide pK18mobsacB vector and the homologous region in the genome, resulting in the integration of the construct into the genome. Colonies were picked and grown overnight at 28°C in liquid LB medium containing 25 µg/ml nitrofurantoin and plated onto LB agar plates containing 10% (w/v) sucrose and incubated at 28°C for 2 days. This step counter-selects the *sacB* marker on the integrated construct and selects for the product of the second crossover removing the construct from the genome, resulting in a gene deletion or wild type revertant. Colonies were checked for deletion by colony PCR using primers L_F and R_R (Table 3.2), to confirm the intended truncated size of the predicted mutant gene product.

Table 3.2. Primers used in this study

Primer	Sequence (5'-3')	Use
- Expression construct		
oNS48_pBBR1MCS3_GA_seq_F	GAACGGGTTGGCATGGATTGTAG	colony PCR and sequencing for checking inserts in pBBR1MCS3
oNS67_pBBR1MCS2_GA_seq_F2	CAAGGCGACAAGGTGCTGATG	colony PCR and sequencing for checking inserts in pBBR1MCS2
oNS49_pBBR1MCS3_GA_seq_R	CTGCAAGGCGATTAAGTTGGGTA	colony PCR and sequencing for checking inserts in pBBR1MCS3,2
oNS206_0834_5'full_F	GATAAGCTTGATATCGAATTCCTGCAGCCCC CTGGTAGAAATGGAGTGATTATCAACTTATG	amplify *bim* cluster with ribosome binding site and full 5' sequence
oNS68_0834_prom_F	TGCTTAATGAATTACAACAGTTTTTATGCAC CGTTCCATTCAAACCCTTCGC	amplify *bim* cluster with native promoter
oNS62_0835_R2	CCGCTCTAGAACTAGTGGATCCCCCCTATCG CCTGCATCAGTACGGG	amplify partial *bim* cluster until 0835

oNS63_0838_R2	CCGCTCTAGAACTAGTGGATCCCCCTCAAGGAGGAGGAATCACAGCC	amplify whole *bim* cluster until 0838
oNS180_0834_mut_R	GATAATAATGCCAACTGCAGCTGCTCCGCGTATAATGCCCACCTC	introduce point mutations for catalytically inactive mutant 0834
oNS181_0834_mut_F	AGCAGCTGCAGTTGGCATTATTATCGATGTAGGTAA	introduce point mutations for catalytically inactive mutant 0834
oNS182_0835_mut_R	TGTGGATCAGCAGTAACAATTGCATTGCGCCCTACCATGATTGC	introduce point mutations for catalytically inactive mutant 0835
oNS183_0835_mut_F	AATGCAATTGTTACTGCTGATCCACAATTGACCAATCGCTATGA	introduce point mutations for catalytically inactive mutant 0835
oNS283_0301_gabT2_F	GATAAGCTTGATATCGAATTCCTGCAGCCCGAGGAATACATGAGCCAGACCAACG	amplify gene *gabT2* (PSPTO0301)
oNS284_0301_gabT2_R	CCGCTCTAGAACTAGTGGATCCCCCCGCACGTTTTTCACGCTATTTCAGC	amplify gene *gabT2* (PSPTO0301)
oNS285_4225_nadB_F	GATAAGCTTGATATCGAATTCCTGCAGCCCGGAGCGCACCTATGAGCCAACATTTTCAGCACGAC	amplify gene *nadB* (PSPTO4225)
oNS286_4225_nadB_R	CCGCTCTAGAACTAGTGGATCCCCCGCGTGCGACTGAAGTTCAGC	amplify gene *nadB* (PSPTO4225)
oNS287_2222_F	GATAAGCTTGATATCGAATTCCTGCAGCCCGGAGCTAGGTTGATTCGCAGGTTC	amplify gene *rhpS* (PSPTO2222)
oNS288_2222_R	CCGCTCTAGAACTAGTGGATCCCCCTCAAAGGCGCGGCAGGTC	amplify gene *rhpS* (PSPTO2222)
oNS289_1778_htpX_F	GATAAGCTTGATATCGAATTCCTGCAGCCCGGAGAACGTTTCGACCATGATGC	amplify gene *htpX* (PSPTO1778)
oNS290_1778_htpX_R	CCGCTCTAGAACTAGTGGATCCCCCTGAATCAACCGCGGCGG	amplify gene *htpX* (PSPTO1778)
oNS291_1156_isochr_F	GATAAGCTTGATATCGAATTCCTGCAGCCCGGAGGTGGTTTGATGACCGC	amplify gene isochorismatase (PSPTO1156)
oNS292_1156_isochr_R	CCGCTCTAGAACTAGTGGATCCCCCTCACAAAGGTTGCCCGAAAGG	amplify gene isochorismatase (PSPTO1156)
oNS293_5288_thrdehyd_F	GATAAGCTTGATATCGAATTCCTGCAGCCCGGAGGTGGTTTGTTGCCCGCTTCGTCAGTGAT	amplify gene threonine dehydratase (PSPTO5288)
oNS294_5288_thrdehyd_R	CCGCTCTAGAACTAGTGGATCCCCCTCAGCCCAAAAACAGCCGATACG	amplify gene threonine dehydratase (PSPTO5288)
oNS295_5499_aspammly_F	GATAAGCTTGATATCGAATTCCTGCAGCCCAGGTCACACCTGCATTATGCCGT	amplify gene aspartate ammonia lyase (PSPTO5499)
oNS296_5499_aspammly_R	CCGCTCTAGAACTAGTGGATCCCCCGTCATTCCCTGTCAGGCCTTCA	amplify gene aspartate ammonia lyase (PSPTO5499)
- Deletion construct		
oNS34_pK18mobsacb_GA_F	GCTACTGGGCTATCTGGACAAGGGAA	amplify pK18mobsacB plasmid backbone
oNS35_pK18mobsacb_GA_R	CGGTATCAGCTCACTCAAAGGCGGTA	amplify pK18mobsacB plasmid backbone
oNS50_pK18mobsacb_GA_seq_F	TTGCTCACATGTTCTTTCCTGCG	colony PCR and sequencing for checking inserts in pK18mobsacB
oNS51_pK18mobsacb_GA_seq_R	TTCGCTTGCTGTCCATAAAACCG	colony PCR and sequencing for checking inserts in pK18mobsacB
oNS42_0834-8_del_L_F	TACCGCCTTTGAGTGAGCTGATACCGTAAGTACATCAGCCGCCACC	amplify left flanking region of PSPTO0834
oNS43_0834-8_del_L_R	TCAAGGAGGAGGAATCACAGCCGCCGTCAGGAGCAAGCAT	amplify left flanking region of PSPTO0834
oNS44_0834-8_del_R_F	ATGCTTGCTCCTGACGGCGGCTGTGATTCCTCCTCCTTGA	amplify right flanking region of PSPTO0838
oNS45_0834-8_del_R_R	TTCCCTTGTCCAGATAGCCCAGTAGCGCTCACAACGGACAACTGGA	amplify right flanking region of PSPTO0838

oNS70_0834_del_L_R	GGTATTAAAGATAAGCGAGATTTGGTGTGA AGCCGTCAGGAGCAAGCAT	amplify left flanking region of PSPTO0834
oNS71_0834_del_R_F	ATGCTTGCTCCTGACGGCTTCACACCAAATC TCGCTTATCTTTAATACC	amplify right flanking region of PSPTO0834
oNS72_0834_del_R_R	TTCCCTTGTCCAGATAGCCCAGTAGCGTGGA GACGAAGAGGCTACGG	amplify right flanking region of PSPTO0834
oNS73_0835_del_L_F	TACCGCCTTTGAGTGAGCTGATACCGATGCT TGCTCCTGACGGC	amplify left flanking region of PSPTO0835
oNS74_0835_del_L_R	CCTATCGCCTGCATCAGTACGGATCAATTGA TATCTGACTGAATACAGTAACTTTCAT	amplify left flanking region of PSPTO0835
oNS75_0835_del_R_F	ATGAAAGTTACTGTATTCAGTCAGATATCAA TTGATCCGTACTGATGCAGGCGATAGG	amplify right flanking region of PSPTO0835
oNS76_0835_del_R_R	TTCCCTTGTCCAGATAGCCCAGTAGCTGGTC ATTGCCCAATCCGAAAAC	amplify right flanking region of PSPTO0835
oNS77_0836_del_L_F	TACCGCCTTTGAGTGAGCTGATACCGGTCGC GGTGGGAGAATTGTACTT	amplify left flanking region of PSPTO0836
oNS78_0836_del_L_R	CTATTTTTCTGCAAACCATTGCTTGATGTCTT CCTGCTCCATATAAAGGGGCTC	amplify left flanking region of PSPTO0836
oNS79_0836_del_R_F	GAGCCCCTTTATATGGAGCAGGAAGACATC AAGCAATGGTTTGCAGAAAAATAG	amplify right flanking region of PSPTO0836
oNS80_0836_del_R_R	TTCCCTTGTCCAGATAGCCCAGTAGCCAGAA AGGAGTAGGAAGGAGAGCAG	amplify right flanking region of PSPTO0836
oNS81_0837_del_L_F	TACCGCCTTTGAGTGAGCTGATACCGGCTTG TGGAAAGGAGGTGCTCTT	amplify left flanking region of PSPTO0837
oNS82_0837_del_L_R	TTAGAAACGGGAGGAGGTATGCGCCATACT ATTTTTCTGCAAACCATTGCTTGA	amplify left flanking region of PSPTO0837
oNS83_0837_del_R_F	TCAAGCAATGGTTTGCAGAAAAATAGTATG GCGCATACCTCCTCCCGTTTCTAA	amplify right flanking region of PSPTO0837
oNS84_0837_del_R_R	TTCCCTTGTCCAGATAGCCCAGTAGCCAGGG CTAGGCGGAAAACTGATT	amplify right flanking region of PSPTO0837
oNS85_0838_del_L_F	TACCGCCTTTGAGTGAGCTGATACCGCCGTA GCCTCTTCGTCTCCAC	amplify left flanking region of PSPTO0838
oNS86_0838_del_L_R	TCAAGGAGGAGGAATCACAGCGCTGAATAG ATTTTTCATTACTCACACGACTCT	amplify left flanking region of PSPTO0838
oNS87_0838_del_R_F	AGAGTCGTGTGAGTAATGAAAAATCTATTCA GCGCTGTGATTCCTCCTCCTTGA	amplify right flanking region of PSPTO0838
oNS88_4071_del_L_F	TACCGCCTTTGAGTGAGCTGATACCGACAAC AAGTACCTGATCTTTGAGGG	amplify left flanking region of PSPTO4071
oNS89_4071_del_L_R	TTGAAGAGCCGCTCATGTCAGATCCATGCTG ACTCCTGTGTGGCT	amplify left flanking region of PSPTO4071
oNS90_4071_del_R_F	AGCCACACAGGAGTCAGCATGGATCTGACA TGAGCGGCTCTTCAA	amplify right flanking region of PSPTO4071
oNS91_4071_del_R_R	TTCCCTTGTCCAGATAGCCCAGTAGCGATAT CGTTTCTTCTGGGCTTCGAG	amplify right flanking region of PSPTO4071
oNS92_5452_del_L_F	TACCGCCTTTGAGTGAGCTGATACCGTGGTG TTGGAGAAATCCTTGATCTG	amplify left flanking region of PSPTO5452
oNS93_5452_del_L_R	TGTCATTCAAACGAGTTCAGGCAGCAGCAG GCCATCCATATCGAAGA	amplify left flanking region of PSPTO5452
oNS94_5452_del_R_F	TCTTCGATATGGATGGCCTGCTGCTGCCTGA ACTCGTTTGAATGACA	amplify right flanking region of PSPTO5452
oNS95_5452_del_R_R	TTCCCTTGTCCAGATAGCCCAGTAGCGTATT CTTGAATCCGCTGGGCA	amplify right flanking region of PSPTO5452
oNS176_hrpA_L_F	TACCGCCTTTGAGTGAGCTGATACCGAGGTT TGAGGATGATCTGGACGAG	amplify left flanking region of *hrpA* (PSPTO1381)
oNS177_hrpA_L_R	TAATTAGTAACTGATACCTTTAGCGTTCGTT GCGACCATGATTTGATGCCC	amplify left flanking region of *hrpA* (PSPTO1381)

oNS178_hrpA_R_F	GGGCATCAAATCATGGTCGCAACGAACGCT AAAGGTATCAGTTACTAATTA	amplify right flanking region of *hrpA* (PSPTO1381)
oNS179_hrpA_R_R	TTCCCTTGTCCAGATAGCCCAGTAGCCAAAC CACGGTCGATGAGTTGA	amplify right flanking region of *hrpA* (PSPTO1381)
- RT-PCR		
oNS126_hrpR_qRT_F	TTCACGAGTTGTCCAGCAGG	RT-PCR of *hrpR* (PSPTO1379)
oNS127_hrpR_qRT_R	ACCACACCGAATAACTGGCT	RT-PCR of *hrpR* (PSPTO1379)
oNS130_hrpL_qRT_F	GGCCGATCAGATTCAGATGCT	RT-PCR of *hrpL* (PSPTO1404)
oNS131_hrpL_qRT_R	TACGCAGGGCTTCAAGAAACA	RT-PCR of *hrpL* (PSPTO1404)
oNS132_0834_qRT_F	ATCCCAACGAATACTGTGGCG	RT-PCR of PSPTO0834
oNS133_0834_qRT_R	TGACACCTGCGATATCCAACC	RT-PCR of PSPTO0834
oNS138_avrPtoB_qRT_F	CATACTCGTAACATGCTCCGC	RT-PCR of *avrptoB* (PSPTO3087)
oNS139_avrPtoB_qRT_R	TCGGGAAATGCTGGCGAAATA	RT-PCR of *avrptoB* (PSPTO3087)
oNS144_rpoD_qRT_F	TCCATCGCCAAGAAGTACACC	RT-PCR of *rpoD* (PSPTO0537)
oNS145_rpoD_qRT_R	ATTCGAACTTGTCCACAGCCT	RT-PCR of *rpoD* (PSPTO0537)

3.2.4 Triparental mating conjugation for transformation of *P. syringae*

Overnight bacterial cultures in LB medium were collected by centrifuging at 4,000 x g for 5 minutes, washed with LB medium (prewarmed to 50°C) and resuspended in half of the original volume. A mixture of 700 µl recipient strain (*P. syringae* strain to be transformed), 300 µl donor strain (*E. coli* carrying desired construct) and 100 µl helper strain (*E. coli* pRK2013) was centrifuged together at 4,000 x g for 5 minutes to form a pellet which was resuspended in 50 µl LB medium and then spotted onto an LB agar plate without antibiotics. The spot was left to dry and incubated at 28°C overnight. The spot was then scraped from the plate, resuspended in LB medium and plated on selective LB agar with appropriate antibiotics for the construct transformed. For selection against *E. coli*, 25 µg/ml nitrofurantoin or 50 µg/ml rifampicin was used.

3.2.5 BGAL inhibitor production *in vitro*

Bacteria were pre-grown overnight in 10 ml of LB medium. The cells were collected by centrifuging at 4,000 x g for 5 minutes then washed once with sterile water and centrifuged

again. The cell pellet was then resuspended in 2 ml of MG medium to an OD600 of 1 and grown overnight. The supernatant containing the inhibitor was collected after centrifuging the bacterial culture and discarding the cell pellet. For the strains expressing *lacZ*, the supernatant was heated at 95°C for 5 minutes to inactivate the enzyme to prevent interference in subsequent enzymatic assays.

3.2.6 FDG assay for β-galactosidase activity

Substrate preparation: A stock FDG solution of 2 mM was made by dissolving FDG (Fluorescein di-β-D-galactopyranoside) (Marker Gene Technologies, Eugene, USA, M0250) in 1:1:8 ethanol:DMSO:H_2O (first dissolved in 1:1 ethanol:DMSO, then diluted in the rest of H_2O). Aliquots were stored at -20°C. The stock was diluted to 2 µM in water just before use.

Enzyme preparation: A stock LacZ solution of 1 U/µl was made by dissolving lyophilised β-galactosidase enzyme (Sigma, St. Louis, USA, G6008) in suspension buffer (50% (v/v) glycerol, 5mM Tris, 5 mM $MgCl_2$, 0.5 mM DTT, 0.5 mM β-mercaptoethanol, pH 7.4). Aliquots were stored at -20°C. The stock was diluted to 1 mU/µl in water just before use.

Apoplastic fluid BGAL assay: The assay was set up in 96 well plate with the total reaction volume of 100 µl per well with 0.2 µM FDG and 50 mM MES pH 5. This typically consists of 40 µl apoplastic fluid sample, 10 µl FDG (2 µM), 10 µl MES buffer pH 5 (500 mM) and 40 µl H_2O. For inhibition assays, 40 µl inhibitor sample was used instead of H_2O.

LacZ assay: The assay was set up in 96 well plate with the total reaction volume of 100 µl per well with 60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 0.2 µM FDG and 0.1 mU/µl LacZ. This typically consists of 20 µl 5x Z buffer (300 mM Na_2HPO_4, 200 mM NaH_2PO_4,

50 mM KCl, 5 mM $MgSO_4$, pH 7), 10 µl FDG (2 µM), 10 µl LacZ (1 mU/µl) and 60 µl H_2O. For inhibition assays, 50 µl inhibitor sample was used instead of H_2O.

β-galactosidase activity in the sample was monitored by measuring fluorescence signal from fluorescein, a product of FDG hydrolysis catalysed by β-galactosidase, every minute at 25°C using the plate reader Infinite M200 (Tecan, Mannedorf, Switzerland) at an excitation wavelength of 485 nm and emission wavelength of 535 nm. The rate of change of fluorescence over time was calculated to determine β-galactosidase activity. For comparisons of inhibition, the enzymatic activity was expressed as percentage of the activity compared to no inhibitor control (100%). Note that fluorescence intensity of fluorescein is pH dependent. The same amount of FDG at pH 5 generates less fluorescence than at pH 7.

3.2.7 RT-PCR

P. syringae strains were grown on LB agar plate containing appropiate antibiotics at 28°C for 2 days. The bacterial cells were collected, washed with sterile water and resuspended at an OD600 of 1 in 1 ml of MG medium containing 50 µM iron citrate. The bacterial culture was incubated at 28°C with shaking for 6 hrs to induce type III secretion system expression before bacterial cells were collected from 1 ml of the culture centrifuged at 12,000 x g for 5 min and frozen with liquid nitrogen. The cell pellet was then resuspended in 250 µl lysozyme (5 mg/ml) in TE buffer (10 mM Tris-HCL pH 7.4, 1 mM EDTA) and incubated at 25°C for 10 minutes while shaking. Total RNA was extracted from the sample using Monarch Total RNA miniprep kit (New England Biolabs, Ipswich, USA). The concentration of RNA was measured using a Nanodrop spectrophotometer (Thermo Scientific, Waltham, USA). RNA samples were treated with Turbo DNA-free kit (Invitrogen, Carlsbad, USA) to remove genomic DNA. cDNA was

synthesised from 1 µg of RNA using GoScript reverse transcriptase master mix with random primers (Promega, Madison, USA). The resulting cDNA sample was diluted 10 fold and 1 µl was used as a template for PCR (20 µl reaction volume) using GoTaq green master mix (Promega, Madison, USA) with primers listed in Table 3.2. Cycling conditions were 95°C for 2 min followed by 28 cycles of [95°C for 15 sec, 60°C for 15 sec, 72°C for 15 sec]. The PCR product was detected upon separation by 4% (w/v) agarose gel electrophoresis and staining with ethidium bromide. An equivalent sample to cDNA sample without reverse transcriptase enzyme was used as "no RT control" to ensure no amplification signal was detected from genomic DNA contamination. All PCR product sizes are around 100 bp.

3.2.8 Luminescence reporter *in planta*

P. syringae strains with luminescence reporter constructs were grown in LB medium containing 25 µg/ml rifampicin at 28°C with shaking overnight. The bacterial cells were collected by centrifuging at 4,000 x g for 5 minutes, washed with sterile water and resuspended in sterile water at the OD_{600} of 0.2 (~10^8 CFU/ml). Bacterial suspensions were then infiltrated into areas within *N. benthamiana* leaves using a needleless syringe. The leaves were then detached from the plant, placed on a plastic plate with wet tissue paper for humidity and imaged for luminescence using an Imagequant LAS-4000 imager (Cytiva, Marlborough, USA) with 5 minutes exposure every 5 minutes over a period of 4 hours. Before imaging, the leaves were left in the dark for around 15 minutes to reduce background photon emission from the leaf photosynthetic pigments. The luminescence signal was quantified from the integrated density of the signal normalised to the infiltrated area of the leaf. Images were analysed using Fiji software (Schindelin *et al.*, 2012).

3.2.9 Bacterial growth curve *in vitro*

Bacteria were grown overnight in LB medium. The cells were collected by centrifuging at 4,000 x g and washed with sterile water before being inoculated at the OD600 of 0.05 in 150 µl of LB or MG media in each well of 96 well plate. Bacterial growth was monitored with OD600 readings taken every 30 minutes using the plate reader Infinite M200 (Tecan, Mannedorf, Switzerland). Plates were incubated at 28°C with shaking during the course of the experiment.

3.2.10 Infection assays

Infiltration: *P. syringae* strains were grown in LB medium with 25 µg/ml rifampicin at 28°C with shaking overnight. The bacterial cells were collected by centrifuging at 4,000 x g for 5 minutes, washed with sterile water and resuspended in sterile water at the OD_{600} of 0.0002 (~10^5 CFU/ml). Plants (~4-5 weeks old) were infected by infiltrating the bacterial suspension into leaves using a needleless syringe. The infected plants were maintained in normal growth conditions for 2 days. Leaf discs were then collected from infected tissue, surface sterilised with 70% (v/v) ethanol, washed with sterile water and homogenised in 1 ml of sterile water. The suspension was serially diluted and plated onto LB agar plates containing 25 µg/ml rifampicin and incubated at 28°C for 2 days. Colonies were counted to calculate the number of bacterial colony forming units (CFU) present in the infected tissue. The bacterial suspension used for inoculation was also plated to confirm the concentration of bacterial inoculum.

Spray: *P. syringae* strains were grown in LB medium with 25 µg/ml rifampicin at 28°C with shaking overnight. The bacterial cells were collected by centrifuging at 4,000 x g for 5 minutes, washed with sterile water and resuspended in sterile water at the OD_{600} of 0.2 (~10^8 CFU/ml) with 0.04% (v/v) Silwet L-77. Plants (~4-5 weeks old), previously covered using a clear plastic

propagator lid overnight, were infected by spraying the bacterial suspension on both leaf surfaces until runoff. The infected plants were covered in a clear plastic propagator lid and maintained in normal growth conditions for 3 days. Leaf discs were then collected from infected tissue, surface sterilised with 70% (v/v) ethanol, washed with sterile water and homogenised in 1 ml of sterile water. The suspension was serially diluted and plated onto LB agar plates containing 25 µg/ml rifampicin and incubated at 28°C for 2 days. Colonies were counted to calculate the number of bacterial colony forming units present in the infected tissue. The bacterial suspension used for inoculation was also plated to confirm the concentration of bacterial inoculum.

3.2.11 Apoplastic fluid extraction

N. benthamiana leaves were rolled into a 50 ml syringe filled with ice-cold water. The leaves were infiltrated by pulling the plunger to apply negative pressure drawing air out of the leaf tissue then pushing the plunger to apply positive pressure forcing water into the leaf tissue, while keeping the syringe tip sealed. The infiltrated leaves were blotted dry and rolled into a 20 ml syringe barrel placed inside a 50 ml conical centrifuge tube. Apoplastic fluid was collected from the leaves by centrifugation at 900 x g for 15 minutes at 4°C. The extracted apoplastic fluid was used directly or stored at -20°C.

3.2.12 Lectin blot

Apoplastic fluid samples were collected from *N. benthamiana* leaves and proteins in the sample were precipitated by mixing 200 µl apoplastic fluid + 750 µl acetone (pre-chilled at -20°C) and centrifuged for 5 minutes at 12,000 x g. The pellet was then resuspended in 20 µl gel loading buffer (100 mM Tris-Cl (pH 6.8), 200 mM DTT, 4% (w/v) SDS, 0.2% (w/v)

bromophenol blue, 20% (v/v) glycerol) and heated 95°C for 5 minutes. The protein samples were then separated in 12% (w/v) polyacrylamide gel electrophoresis running at 120 V for 20 min then 180 V for 60 min and transferred onto PVDF membrane using Trans-Blot Turbo system (Bio-Rad, Hercules, USA). The membrane was blocked with Carbo-Free blocking solution (Vector Laboratories, Burlingame, USA) containing 0.05% (v/v) Tween-20, incubated with shaking at room temperature for 30 min and then incubated with 5 µg/ml biotinylated RCA I lectin (Ricinus Communis Agglutinin I, RCA120) (Vector Laboratories, Burlingame, USA) in blocking solution shaking at room temperature for 30 min. The membrane was washed 5 times with TBS buffer (150 mM NaCl, 10 mM Tris, pH 7.5) containing 0.05% (v/v) Tween-20 and then incubated with Streptavidin-HRP (Sigma, St. Louis, USA) (1:5000x dilution in blocking solution) shaking at room temperature for 30 min followed by another 5 times of washing with the TBS buffer containing Tween-20. Chemiluminescence was detected using Clarity ECL western blotting substrates (Bio-Rad, Hercules, USA) with an Imagequant LAS-4000 imager (Cytiva, Marlborough, USA). The membrane was finally stained with Coomassie blue for total protein detection. For infected samples, infection was done by infiltration as described above.

3.2.13 GC content mapping

The genome sequence of *P. syringae* pv. tomato DC3000 was analysed for % GC content in sliding windows of 500 bp across the genome. The %GC was then plotted at the 500 bp interval. The 100 kb region around *bim* cluster was displayed. The analysis was done using R.

3.2.14 Phylogenetic analysis

Alcohol dehydrogenase homologs protein tree: The amino acid sequence of the full length PSPTO_0834 from *P. syringae* pv. tomato DC3000 was used as a query for homologs search

by PSI-BLAST (on www.blast.ncbi.nlm.nih.gov)(Altschul *et al.*, 1997) on RefSeq database with default settings and 5 iterations. The resulting 500 proteins identified were used together with threonine dehydrogenase from *E. coli* (UniProt accession P07913) included as an outgroup. MEGA X software (Stecher, Tamura and Kumar, 2020) was used for multiple sequence alignment and phylogenetic tree reconstruction. The amino acid sequences were aligned with MUSCLE and partial deletion with 95% cutoff was used to remove gaps in the alignment. Maximum Likelihood (ML) phylogenetic tree was reconstructed with LG amino acid substitution model with empirical frequencies (+F) and Gamma distributed evolutionary rate heterogeneity among sites with 4 discrete categories (+G) and allowance for invariant sites (+I). Initial tree was generated with neighbor-joining BIONJ for ML heuristic search with nearest-neighbor-interchange method (NNI). Branch support values were calculated with the Bootstrap method with 100 replications. Branches with support values less than 0.5 were collapsed. The phylogenetic tree was visualized using iTOL software (Letunic and Bork, 2019).

Putative gene clusters containing each homolog of the alcohol dehydrogenase in each genome were assigned by identifying a stretch of consecutive genes with the same orientation on the genome (same strand) as the alcohol dehydrogenase genes, both up and downstream until there is a gene with an opposite orientation. Gene model annotation of the genome was acquired from RefSeq database (N. A. O'Leary *et al.*, 2016) and gene family assignment was based on conserved domain database entry (CDD) linked to each gene (Lu *et al.*, 2020). The association with transposable elements was defined as the presence of transposase/integrase/recombinase or conjugative elements within five genes distance from the first and last gene in the cluster of interest. It is worth noting that many available genome

sequences were not complete assemblies so fragmented genomic regions and incompletely annotated genes are common.

P. syringae species tree: The phylogenetic tree was reconstructed using concatenated coding DNA sequences of 4 conserved housekeeping genes (*rpoD* encoding sigma factor 70, *gyrB* encoding DNA gyrase B, *gltA* encoding citrate synthase and *gapA* encoding glyceraldehyde-3-phosphate dehydrogenase), since they were previously shown to provide robust data for inferring evolutionary relationships among *P. syringae* strains (Hwang *et al.*, 2005). Genome sequences were obtained from available genomes and annotations in NCBI database (Agarwala *et al.*, 2016) and pseudomonas.com database (Winsor *et al.*, 2016). *P. syringae* strains used in the analysis were selected to include all strains that have the *bim* gene cluster identified from protein BLASTP of amino acid sequence of the full length PSPTO_0834 from *P. syringae* pv. *tomato* DC3000 (on www.blast.ncbi.nlm.nih.gov) with default settings, 62 type and pathotype strains (Thakur, Weir and Guttman, 2016) and additional strains to represent different phylogroups. Sequences of the housekeeping genes were retrieved with local BLAST+ search against the genomes (Camacho *et al.*, 2009). Strains that have incomplete genome assembly or annotation at the site of the 4 housekeeping genes were discarded. *P. fluorescens* SBW25 was included as an outgroup. Phylogroups were assigned according to (Baltrus, McCann and Guttman, 2017a). MEGA X software (Stecher, Tamura and Kumar, 2020) was used for multiple sequence alignment and phylogenetic tree reconstruction. The DNA sequences were aligned with MUSCLE and partial deletion with 95% cutoff was used to remove gaps in the alignment. Maximum Likelihood (ML) phylogenetic tree was reconstructed with General Time-Reversible (GTR) DNA substitution model with Gamma distributed evolutionary rate heterogeneity among sites with 4 discrete categories (+G) and

allowance for invariant sites (+I). Initial tree was generated with neighbor-joining BIONJ for ML heuristic search with nearest-neighbor-interchange method (NNI). Branch support values were calculated from Bootstrap method with 100 replications. Branches with less than 0.5 support value were collapsed. The phylogenetic tree was visualized using iTOL software (Letunic and Bork, 2019).

For strains that contain the *bim* cluster, the *bim* cluster together with 20 kb of genomic region flanking each side of the cluster was displayed. Gene model annotation of the genome was acquired from the RefSeq database (N. A. O'Leary *et al.*, 2016) and gene family assignment was based on conserved domain database entry (CDD) linked to each gene (Lu *et al.*, 2020). It is also worth noting that many available genome sequences were not complete assemblies so fragmented genomic regions and incompletely annotated genes are common. Only annotations on the fragment of the contig identified for the *bim* cluster was displayed.

The presence of orthologs of VioT (modified viosamine transferase), essential for the transfer of modified viosamine to the glycan on flagellin protein, as characterised in *P. syringae* pv. *tabaci* 6605 (*Pta*6605) (Yamamoto *et al.*, 2011), was identified by reciprocal best hits criteria (Ward and Moreno-Hagelsieb, 2014). The protein sequence of VioT (A3SK_RS0115305 from *Pta*6605) was used as a query for protein BLASTP search against each proteome of each *P. syringae* strain, then the best protein hits from each of the strain was used as a query for another protein BLASTP search against the proteome of *Pta*6605. If VioT was identified as the best hit of that protein, the strain was classified as having an ortholog of VioT. BLAST search was carried out with local BLAST+ (Camacho *et al.*, 2009).

PSPTO_0834 gene tree: The phylogenetic tree was reconstructed using coding DNA sequences of homologs of PSPTO_0834 identified in *P. syringae* strains. Sequences were obtained from available genomes and annotations in NCBI database (Agarwala *et al.*, 2016). *P. syringae* strains used in the analysis were selected to include all strains that have the *bim* gene cluster identified from protein BLASTP of amino acid sequence of the full length PSPTO_0834 from *P. syringae* pv. *tomato* DC3000 (on www.blast.ncbi.nlm.nih.gov) with default settings. A homolog of PSPTO_0834 identified in *P. fluorescens* Pf275 was included as an outgroup. MEGA X software (Stecher, Tamura and Kumar, 2020) was used for multiple sequence alignment and phylogenetic tree reconstruction. The DNA sequences were aligned with MUSCLE and partial deletion with 95% cutoff was used to remove gaps in the alignment. Maximum Likelihood (ML) phylogenetic tree was reconstructed with General Time-Reversible (GTR) DNA substitution model with Gamma distributed evolutionary rate heterogeneity among sites with 4 discrete categories (+G) and allowance for invariant sites (+I). The initial tree was generated with neighbor-joining BIONJ for ML heuristic search with nearest-neighbor-interchange method (NNI). Branch support values were calculated with the Bootstrap method with 100 replications. Branches with support values less than 0.5 were collapsed. The phylogenetic tree was visualized using iTOL software (Letunic and Bork, 2019).

For promoter region and *hrp* box alignment, a 300 bp region upstream of the start codon of the cluster that contains PSPTO_0834 homolog was used for alignment with Clustal Omega (Madeira *et al.*, 2019). Strains that have incomplete assemblies which contain only a fragment of the promoter region in the contig were discarded. The sequence region around the consensus *hrp* box (Lam *et al.*, 2014) was displayed.

3.2.15 Full-length sequence of PSPTO0834

This gene was originally misannotated, resulting in a truncation. The full-length sequence is shown here with the previously annotated portion highlighted in blue.

>PSPTO0834_CDS

ATGAAAGCACTGGGCTTAATGGATAACCAGAGGCTTGAACTTGTGGATTGCAAGGATCCTGTCATG
CTTGCTCCTGACGGCGTAGAAATCGATATCGTGCTATCAGGTATATGCGGAACTGATCTGGCGGTAT
TGTCGGGCCGTGAAGGTGGAGAGGTGGGCATTATACGCGGGCACGAAGCAGTTGGCATTATTATC
GATGTAGGTAAGGATGTAGTACACCTACAAAAAGGGATGCGGGTGGTGGTTGATCCCAACGAATAC
TGTGGCGTTTGCGAACCTTGCCGTCTTGCTAAAACGCACCTATGCAATGGGGGGGGTGAACGCTGGG
TTGGATATCGCAGGTGTCAACAAACATGGAACTTTTGCCGAGCGCTTCGTTACTCGTGAGCGTTTTG
TGTATCAATTGCCAGACGATATGAGCTGGGCAGCTGGTGTGTTGGTTGAGCCTGTTGCCTGCATTCT
GAATAATATAGACCAGGCGTTCATTCGAGCGGGAGAGCGTGTGTTGATCCTAGGGTCTGGCCCTAT
GAGTCTGATTGCGCAGATCGTTCTGCGCTCAATGGGAGTTGACACGCTCGCCACTGATCGAAACACA
CATCGCATACAGTTCGGCCGCTCACAAAGTCTTGATGTTATACATGCCGATGATCTTGAGTTGCAGAT
GCAGCACCAAGAAAAGTTTGATGTTGTTATCGATACTGTCGGTAATCAGATCGATACAGCTTCACGC
TACATCGGTCGCGGTGGGAGAATTGTACTTTTTGGATTTGATAGTGACTATCACTACATGCTGCCTGT
AAAGTACTTCCTGGTTAACGCTATCAGTATTATTTCTGCTGGAGAATACAATCAGCACTTTCCTAGAG
CAATTCGTCTTGTGCAAAAACTTCCTGAGCTAGGGCGGCTGGTAACGCATCGCTACGTACTAGAAAA
TCACTCGGAGGTTTTCGATGCACTTCTGAACGATGCTTCCGCCCCCAATATAAAAAGCGTATTCACAC
CAAATCTCGCTTATCTTTAA

>PSPTO0834_protein

MKALGLMDNQRLELVDCKDPVMLAPDGVEIDIVLSGICGTDLAVLSGREGGEVGIIRGHEAVGIIIDVGK
DVVHLQKGMRVVVDPNEYCGVCEPCRLAKTHLCNGGVNAGLDIAGVNKHGTFAERFVTRERFVYQLPD
DMSWAAGVLVEPVACILNNIDQAFIRAGERVLILGSGPMSLIAQIVLRSMGVDTLATDRNTHRIQFGRSQ
SLDVIHADDLELQMQHQEKFDVVIDTVGNQIDTASRYIGRGGRIVLFGFDSDYHYMLPVKYFLVNAISIIS
AGEYNQHFPRAIRLVQKLPELGRLVTHRYVLENHSEVFDALLNDASAPNIKSVFTPNLAYL

3.3 Results

- The biosynthesis of the inhibitor

3.3.1 PSPTO0834 and PSPTO0835 from the *bim* gene cluster confer inhibitor production in *P. syringae* and *E. coli*

My screen for *P. syringae* BGAL inhibitor mutants (*bim*) identified PSPTO0834 and PSPTO0835 as strong candidate genes required for inhibitor production. These two genes reside in a cluster, named the *bim* cluster, consisting of five genes (Figure 3.1A). The conserved domains identified in each of the gene product (Table 3.3) suggest their possible roles. There are three potential biosynthetic enzymes: an alcohol dehydrogenase (PSPTO0834), a reductase (PSPTO0835) and a phosphatase (PSPTO0836), a protein of unknown function (PSPTO0837) and a transporter (PSPTO0838).

Table 3.3. Genes in the *bim* gene cluster and conserved domains identified in each gene product according to the CDD database (Lu *et al.*, 2020).

Gene	Conserved domain [NCBI CDD database]
PSPTO0834	medium chain reductase/ zinc dependent alcohol dehydrogenase superfamily [cl16912] L-threonine dehydrogenase family [cd08234]
PSPTO0835	dihydrofolate reductase (DHFR) superfamily [cl17279] 5-amino-6-(5-phosphoribosylamino) uracil reductase (RibD) family [PRK05625]
PSPTO0836	haloacid dehalogenase-like hydrolases (HAD) superfamily [cl21460] phosphoglucomutase/ phosphatase family [cd16423]
PSPTO0837	protein of unknown function (DUF1349) superfamily [cl01397]
PSPTO0838	major facilitator (MFS) transporter superfamily [cl28910] L-fucose permease family [cd17333]

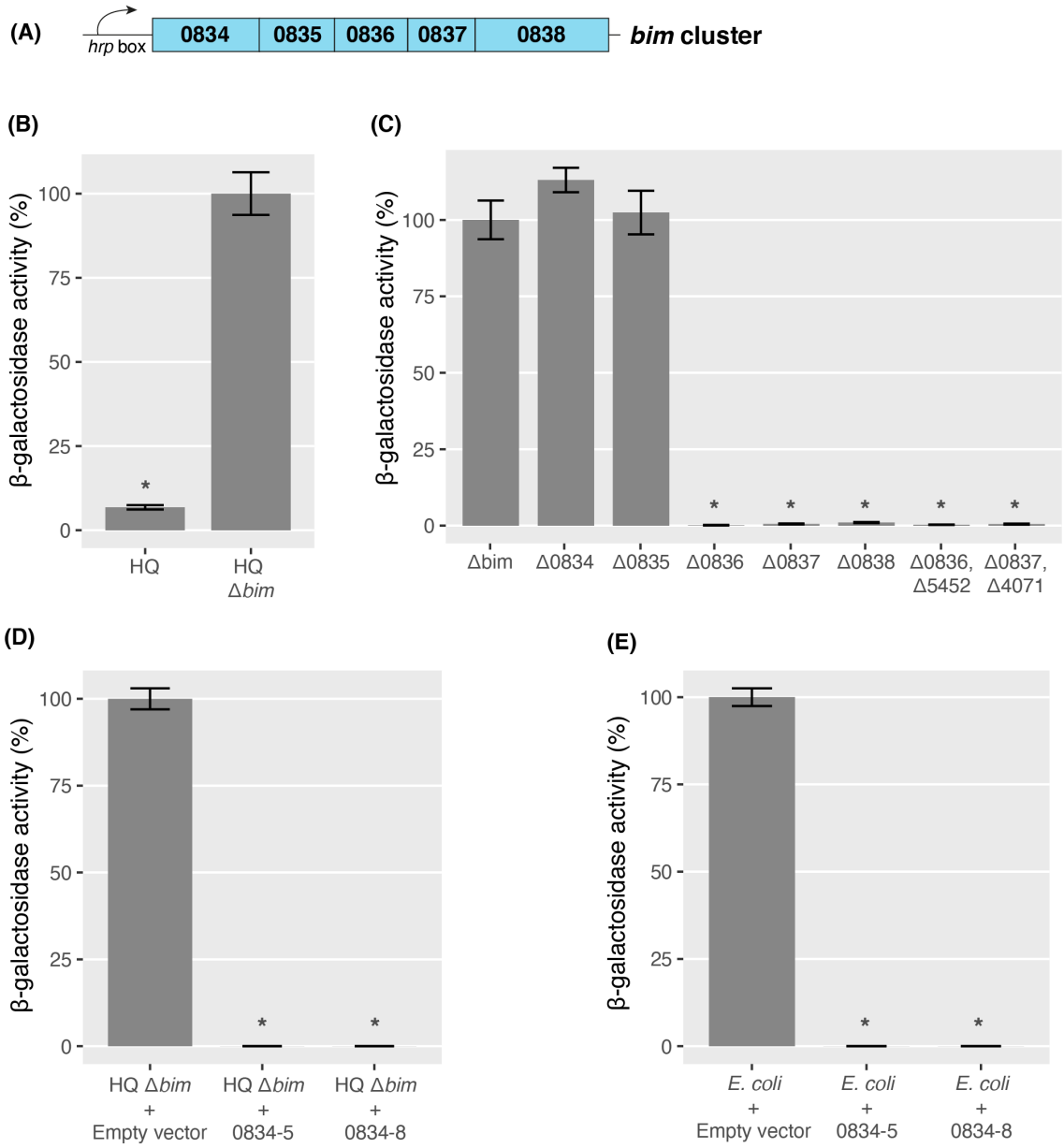


Figure 3.1. PSPTO0834 and PSPTO0835 from the *bim* gene cluster confer inhibitor production in *P. syringae* and *E. coli.* (A) Organization of genes in the *bim* cluster containing five genes, PSPTO0834-8. The curved arrow represents the promoter region that contains the *hrp* box consensus sequence for binding with the sigma factor HrpL, a master regulator of virulence. **(B-E)** Test for inhibitor production. The bacteria were grown in MG medium and supernatant was collected for enzymatic inhibition assay with FDG and LacZ enzyme. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control (HQ Δ*bim* or empty vector). Mean and standard deviation are shown (n=3). Asterisks indicate a statistically significant difference ($P < 0.05$) compared to the no inhibitor control using t-test (B) or ANOVA and Tukey HSD test (C-E).

To confirm that the *bim* cluster was required for inhibitor production, a knockout strain with the whole cluster deleted (HQ Δ*bim*) was generated for comparison with its inhibitor producing background strain (HQ). The bacteria were grown in MG media to induce inhibitor

production then the supernatants of the cultures were tested for their inhibitory activity in enzymatic assay with LacZ enzyme. As expected, the inhibitor production was impaired in the HQ Δ*bim* strain (Figure 3.1B). To further dissect the role of the five genes in the *bim* cluster, knockout mutants with deletions in each individual gene were also tested. Individual deletions of PSPTO0834 and PSPTO0835 impaired inhibitor production while those of PSPTO0836, PSPTO0837 and PSPTO0838 did not (Figure 3.1C). Since redundancy of genes in the genome could mask the phenotype of the individual gene knockout, I also searched for potential homologs of these genes using Orthofinder software (Emms and Kelly, 2019) and generated two additional double knockouts of the identified orthologs: PSPTO5452 (PSPTO0836) and PSPTO4071 (PSPTO0837). However, these double knockouts were still able to produce the inhibitor (Figure 3.1C). These results indicated that only PSPTO0834 and PSPTO0835 are required for the production of the inhibitor in *P. syringae*.

Next, to further validate the role of the *bim* cluster genes, I complemented the HQ Δ*bim* knockout strain with plasmid constructs containing either the partial (PSPTO0834-PSPTO0835) or whole (PSPTO0834-PSPTO0838) *bim* cluster under the native promoter (400 bp region upstream of the start codon of PSPTO0834). Both constructs were able to rescue the inhibitor production phenotype in the HQ Δ*bim* knockout strain (Figure 3.1D). Moreover, when cloned into a plasmid under the constitutive *Lac* promoter, both the partial (PSPTO0834-PSPTO0835) and whole (PSPTO0834-PSPTO0838) *bim* cluster were able to confer inhibitor production in *E. coli* (Figure 3.1E). Together, these results confirmed that *bim* cluster is the biosynthetic gene cluster for the inhibitor and that the combination of PSPTO0834 and PSPTO0835 are sufficient for inhibitor production in the conditions tested.

3.3.2 L-aspartate oxidase (NadB) is involved in inhibitor production

I next investigated the role of *nadB* (PSPTO4225), another important candidate gene identified in the BGAL inhibitor mutant (*bim*) screen. This gene encodes L-aspartate oxidase, an enzyme in the *de novo* biosynthesis pathway of the nicotinamide adenine dinucleotide (NAD) redox cofactor (Kurnasov *et al.*, 2003). NAD (or NADP) could play a role in the biosynthesis of the inhibitor by acting as a cofactor required for the key oxidoreductase enzymes PSPTO0834 and PSPTO0835. To test the involvement of NadB in inhibitor production, the transposon insertion mutant of *nadB* identified from the screen (*bim*549) was tested together with the *bim*549 mutant strain complemented with a plasmid containing *nadB* gene under a constitutive *Lac* promoter. The non-inhibitor producing *bim* cluster knockout strain was used as a negative control. The strains were tested in two growth conditions: the first was growing on LB agar plate then collecting and inoculating the cells in inhibitor-inducing MG medium, the second was growing in liquid LB medium then collecting and inoculating in inhibitor-inducing MG medium. Supernatants from the cultures were then tested for inhibitory activity in enzymatic assay with LacZ enzyme. Some inhibitor production was detected from the *nadB* mutant (*bim*549) in both conditions but was highly impaired in the agar plate grown condition (Figure 3.2A, B). In addition, the complemented *bim*549 strain showed higher degree of inhibitor production than the *bim*549 mutant in both conditions (Figure 3.2A, B). The more impaired inhibitor production from the plate grown cells was also consistent with the way the mutant was selected from the screen, which was based on agar plates. These results suggested that *nadB* has a role in inhibitor production but is not always required. This likely depends on the growth conditions of the cells and their metabolic state.

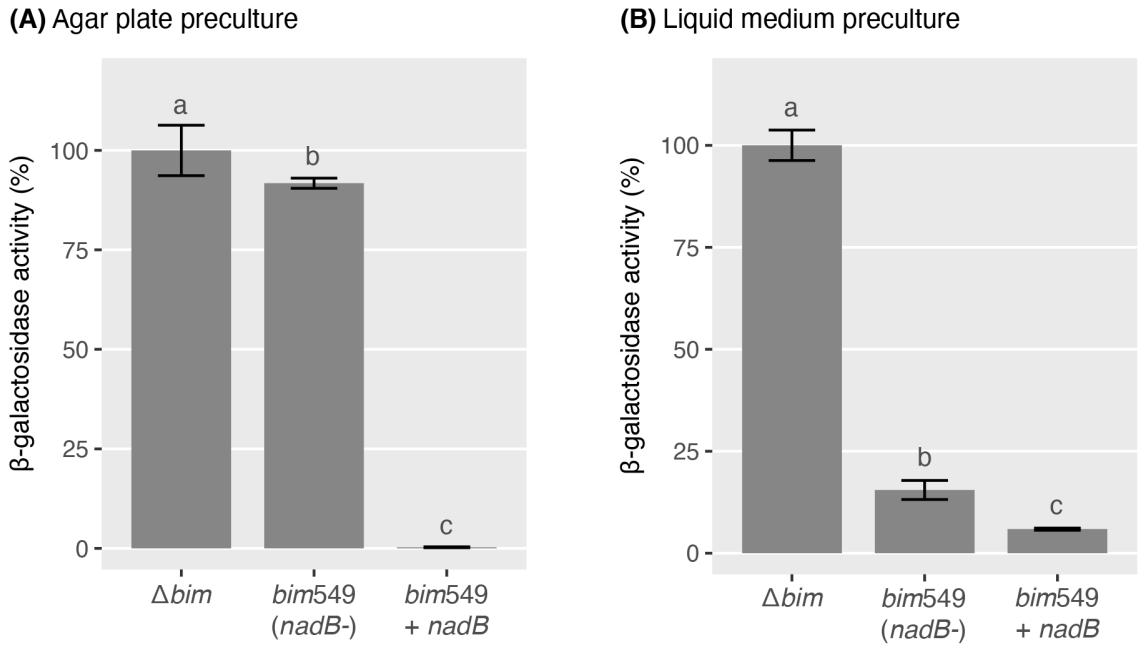


Figure 3.2. L-aspartate oxidase (NadB) is conditionally involved in inhibitor production. The bacteria were collected from a preculture **(A)** on LB agar plate or **(B)** in liquid LB medium, and then grown in MG medium from the starting OD600 of 1 overnight and supernatant was collected for enzymatic inhibition assay with FDG and LacZ enzyme. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control (HQ Δ*bim*). Mean and standard deviation are shown (n=3). Different letters indicate statistically significant difference ($P < 0.05$) from ANOVA and Tukey HSD test. The mutant strain *bim*549 contains a transposon insertion disrupting *nadB* gene and was compared with its corresponding complemented strain with *nadB* in a plasmid.

3.3.3 Other selected *bim* candidate genes were not required for inhibitor production

To identify additional biosynthetic enzymes (in addition to the *bim* cluster) that could be involved in inhibitor production, I looked into other mutants identified in my screen. Candidate mutants containing transposon insertion in genes that encode enzymes present in both *P. syringae* and *E. coli* were selected since they can produce the inhibitor when the *bim* cluster is present. The selection included isochorismatase (*isch*, *bim*97), aspartate-ammonia lyase (*aal*, *bim*154), threonine dehydratase (*tdh*, *bim*502) and 2,4-aminobutyrate aminotransferase (*gabT2*, *bim*796). To test if these candidate genes are required for inhibitor production, each of the mutants were complemented with a plasmid containing each gene under a constitutive *Lac* promoter. The mutants and complemented strains were then tested

for inhibitor production in MG medium. None of the complemented strains could rescue inhibitor production phenotype in the mutants (Figure 3.3A). In addition, to address a potential gene redundancy issue, an independent mutant lacking all three orthologs of gabT2 (gabT1,2,3) was also tested and still showed inhibitor production. These results suggested that none of the tested candidate genes were required for inhibitor production. The inability of these *bim* mutants to produce the inhibitor may be caused by another unidentified mutation within the genome or due to polar effects of the transposon insertion affecting downstream genes.

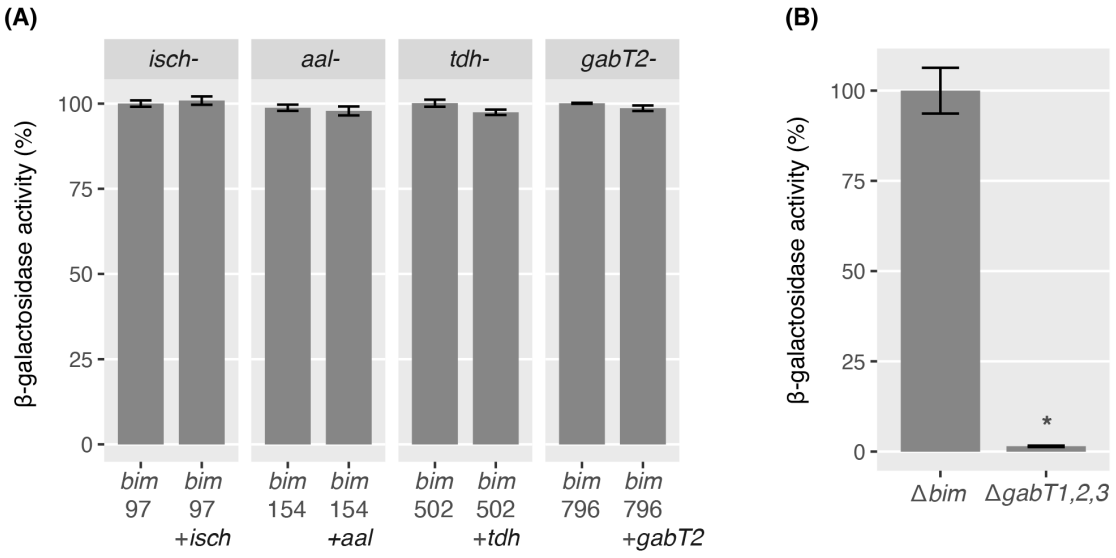


Figure 3.3. Other selected *bim* candidate genes were not required for inhibitor production. The bacteria were grown in MG medium from the OD600 of 1 overnight and supernatant was collected for enzymatic inhibition assay with FDG and LacZ enzyme. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control (HQ Δ*bim*). Mean and standard deviation are shown (n=3). Asterisks indicate a statistically significant difference ($P < 0.05$) compared to the no inhibitor control using t-test. The mutant strains (*bim*97, *bim*154, *bim*502 and *bim*796) contain a transposon insertion disrupting genes (isochorismatase (*isch),* aspartate-ammonia lyase (*aal*), threonine dehydratase (*tdh*) and GABA aminotransferase (gabT2), respectively)). Each strain was compared with their corresponding complemented strains containing the functional gene copy in a plasmid.

- **The evolution of inhibitor production**

3.3.4 Homologs of the alcohol dehydrogenase PSPTO0834 are present in different putative biosynthetic gene clusters for iminosugars in wide range of bacterial groups

The discovery of the *bim* gene cluster as a biosynthetic gene cluster for the BGAL inhibitor allowed me to further explore the evolution of this gene cluster and BGAL inhibitor production. Since a homolog of PSPTO0834, GutB1, has also been identified as an enzyme responsible for the important ring formation step in the biosynthesis of DNJ, an iminosugar-type glycosidase inhibitor (Clark, Johnson and Horenstein, 2011), I reasoned that this alcohol dehydrogenase gene family could be a key marker for biosynthetic gene cluster for iminosugar inhibitors. Therefore, I sought to identify all possible homologs in other bacterial species to potentially explore their evolutionary history. I searched for homologs of PSPTO0834 using psiBLAST and constructed a phylogenetic tree of the identified proteins (Figure 3.4). In addition, I also mapped the possible gene cluster that these homologs belong to as well as the annotations of the neighbouring genes in the cluster (Figure 3.4). Association of the clusters with transposable genetic elements were also identified since many biosynthetic gene clusters can be acquired by horizontal gene transfer between bacterial strains (Figure 3.4).

Overall, there are 8 major clades in the tree (Figure 3.4). Each major clade is represented by a majority of bacterial species of the same genus, which includes *Streptomyces* (clade 1, 2), *Chitinophaga* (clade 3), *Paenibacillus* (clade 4, 5, 7), *Pseudomonas* (clade 6) and *Bacillus* (clade 8). Clade 1 and 2, which are mainly represented by *Streptomyces*, also contain genera in the same Actinobacteria group such as *Frankia*, *Kitasatospora*, *Micromonospora*,

Actinopolyspora, *Actinomadura* and *Phytohabitans*. Some *Streptomyces* can also be found in clade 7. In addition to the well represented clades, *Clostridium* and related genera, such as *Anaerobacterium*, *Anaerocolumna*, *Anaerosacchriphilus* and *Pseudobutyrivibrio*, are found in multiple less represented clades (near clade 1, 7, 8). In addition, there are also less

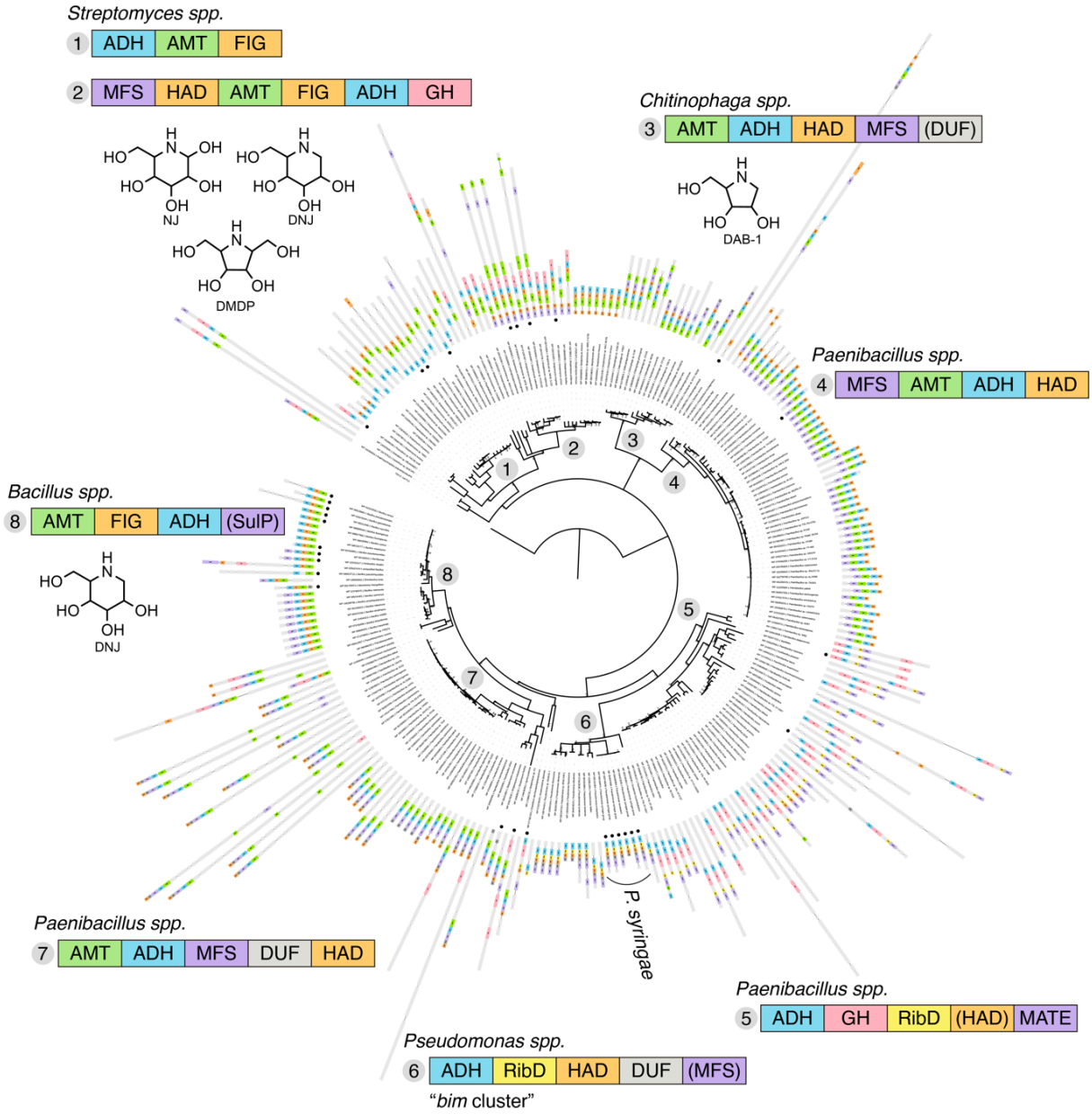


Figure 3.4. Homologs of the alcohol dehydrogenase PSPTO0834 are present in different putative biosynthetic gene clusters for iminosugars in a wide range of bacterial groups. A phylogenetic tree was reconstructed using amino acid sequence of homologs of PSPTO0834 identified by PSI-BLAST. The amino acid sequences were aligned with MUSCLE and partial deletion with 95% cutoff was used to remove gaps in the alignment. Maximum Likelihood (ML) phylogenetic tree was reconstructed. Branch support values were calculated from Bootstrap method with 100 replications. Branches with less than 0.5 branch support value were collapsed. Putative gene

clusters containing each homolog of the alcohol dehydrogenase in each genome were assigned by identifying a stretch of consecutive genes with the same orientation on the genome (same strand) as the alcohol dehydrogenase gene both up and downstream until there is a gene with an opposite orientation. Gene annotation is shown as rectangles. Notable genes of interest are highlighted with different colours. Association with transposable elements (indicated with solid black circle) was defined as the presence of annotated transposase, integrase, recombinase or conjugative elements within five genes distance from the first and last gene in the cluster of interest. It is also worth noting that many available genome sequences were not complete assemblies so fragmented genomic regions and incompletely annotated genes are common. Major clades are labelled with numbers and the dominant genus representing that clade is shown together with the common gene components of the cluster found in that clade. Common genes include homologs of the alcohol dehydrogenase PSPTO0834 (ADH), aminotransferase (AMT), phosphatase (HAD or FIG family), transporter (MFS or MATE or SulP family), reductase (RibD family), glycosidase (GH) and protein of unknown function (DUF1349). Structures of iminosugars identified from strains present in the clades are displayed. For clarity, only a representative subset of strains is shown. The full set of all strains identified is shown in Appendix 1.

Many species identified to have the gene clusters have also been reported to produce iminosugars. For example, nojirimycin was found in *Streptomyces nojiriensis* and *Streptomyces lavendulae* (Inouye *et al.*, 1968); galactonojirimycin was found in *Streptomyces lydicus* (Miyake and Ebata, 1988); 1-deoxynojirimycin (DNJ) was found in *Streptomyces subrutilus* (Hardick *et al.*, 1992), *Streptomyces lavendulae* (Wu *et al.*, 2019), *Bacillus subtilis* (Hardick and Hutchinson, 1993; Kang *et al.*, 2011; Onose *et al.*, 2013)*, Bacillus amyloquifaciens* (Clark, Johnson and Horenstein, 2011; Onose *et al.*, 2013; Seo *et al.*, 2013)*, Bacillus methylotrophicus* (Lee *et al.*, 2018); DAB-1 (1,4-dideoxy-1,4-imino-D-arabinitol) was found in *Chitinophaga pinensis* (Nuñez and Horenstein, 2019); DMDP (2,5-dideoxy-2,5-imino-D-mannitol) was found in *Streptomyces sp.* KSC5791 (Watanabe *et al.*, 1995) and kifunensine was found in *Kitasatospora kifunensis* (Kayakiri *et al.*, 1989). These results support the hypothesis that gene clusters containing alcohol dehydrogenase homologs of PSPTO0834 are markers of biosynthetic gene clusters of iminosugar-type inhibitors.

The gene clusters identified tend to contain a number of recurring genes in different combinations that are characteristic of each clade (Figure 3.4). The recognisable genes include not only all the five genes found in the *bim* cluster but also extra genes, including alcohol dehydrogenase (homolog of PSPTO0834), aminotransferase (aspartate aminotransferase

family), phosphatase (HAD or FIG family), transporter (MFS or MATE or SulP family), reductase (RibD family), glycosidase (amylase family) and protein of unknown function (DUF1349).

For the *bim* gene cluster from *P. syringae*, there are both similarities with other types of clusters and a unique difference. At least one homolog of the last three genes of the *bim* cluster (HAD phosphatase, DUF1349 and MFS transporter) are found in clades 2, 3 and 4, while all three homologs are found together in clade 7 (Figure 3.4). However, the *bim* cluster carries the RibD family reductase (PSPTO0835) but no aminotransferase (AMT), which is unique property in clade 5 and 6 (Figure 3.4). This combination might be unique for production of the *P. syringae* BGAL inhibitor.

Some *Streptomyces* strains from clade 1 and 2 have been reported to produce nojirimycin (NJ) while other *Streptomyces* strains and *Bacillus* strains from clade 8 have been reported to produce deoxynojirimycin (DNJ). The clusters found in these strains share the same set of 3 genes (alcohol dehydrogenase (ADH), aminotransferase (AMT) and FIG family phosphatase). These genes confer the ability to produce nojirimycin, which can then be converted to deoxynojirimycin by an additional reductase step (Clark, Johnson and Horenstein, 2011). This reductase is likely to be present in the genome of the DNJ producing strains but not NJ producing strains resulting in different iminosugars.

Overall, the similar set of genes present among different types of clusters suggests a shared evolutionary history. The different types of gene clusters might have evolved divergently and, likely together with the intrinsic metabolic genes present in each of the strain's genome, become determinants for biosynthesis of different iminosugars.

3.3.5 The alcohol dehydrogenase GutB1 could not replace the function of PSPTO0834 in the *bim* cluster for BGAL inhibitor production

To test the conservation of alcohol dehydrogenase homologs and potentially address the hypothesis of divergence of the cluster genes towards production of different types of iminosugar, I tested if the alcohol dehydrogenase homolog GutB1 from *Bacillus amyloquifaciens* can replace of PSPTO0834 (38% identity of amino acid sequence). *E. coli* strains containing a plasmid with either PSPTO0834-PSPTO0835 or GutB1-PSPTO0835 were tested for BGAL inhibitor production, using the *E. coli* with empty vector plasmid as a negative control. As shown before, the PSPTO0834-PSPTO0835 construct led to a high level of BGAL inhibitor production (Figure 3.5). In contrast, the replacement of PSPTO0834 by GutB1 resulted in only a small fraction of inhibitory activity (Figure 3.5). This suggested that although the two enzymes may share some similar conserved activity, they also have diverged, possibly with regard to their substrate preference, resulting in the inability of GutB1 to completely replace PSPTO0834 function in BGAL inhibitor production.

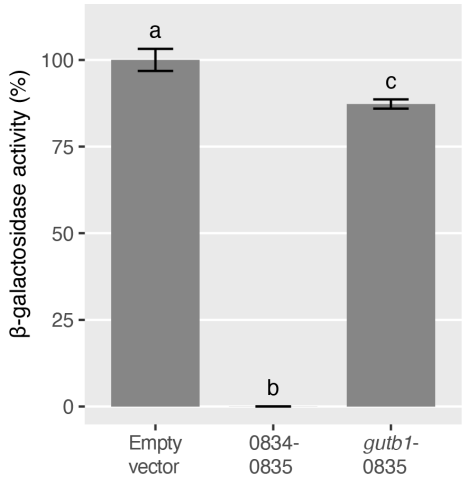


Figure 3.5. The alcohol dehydrogenase homolog GutB1 could not replace PSPTO0834. *E. coli* strains containing a plasmid with either PSPTO0834-PSPTO0835 (from the *bim* cluster) or *gutB1*-PSPTO0835 were tested for BGAL inhibitor production. *E. coli* with empty vector plasmid was used as a negative control. The bacteria were grown in MG medium from the starting OD600 of 1 overnight and supernatant was collected for enzymatic inhibition assay with FDG and LacZ enzyme. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control (empty vector). Mean and standard deviation are shown (n=3). Different letters indicate statistically significant difference ($P < 0.05$) from ANOVA and Tukey HSD test.

3.3.6 The *bim* cluster likely originated in another bacterial group and was acquired by the ancestor of *Pseudomonas* group through horizontal gene transfer

The homologs of PSPTO0834 in the phylogenetic analysis (Figure 3.4) were identified in a fairly distant range of bacterial species from phyla ranging from Actinobacteria such as *Streptomyces*, Firmicutes such as *Bacillus, Paenibacillus* and *Clostridium*, Bacteroidetes such as *Chitinophaga* and Proteobacteria such as *Pseudomonas.* This suggested that the gene cluster could have evolved from horizontal gene transfer across different bacterial groups. In addition, a number of clusters identified, including all from *P. syringae*, are associated with transposable elements (Figure 3.4). Genomic signatures, such as GC content, can be used to infer possible horizontally transferred DNA segments when the genomic GC content differs between the donor and recipient (Ravenhall *et al.*, 2015). We, therefore, calculated the GC content of genomic regions in *P. syringae* DC3000 in sliding windows across the genome. The *bim* cluster showed lower GC content (~45-48%) compared to flanking regions and the average of the whole genome (~58%) (Figure 3.6), supporting the hypothesis of its foreign origin. Interestingly, the GC content of *Paenibacillus* and *Bacillus* strains also fall below 50% (Chen *et al.*, 2007; Luo *et al.*, 2018), as a member of the Firmicutes group with generally low GC content (Bohlin *et al.*, 2017). In addition, *Paenibacillus* and *Bacillus* strains are well represented among species identified for the putative iminosugar gene clusters and contain most of the diversity of genes found among the different types of gene clusters (Figure 3.4). This suggests that the *bim* cluster in *Pseudomonas* clade might have been acquired by horizontal gene transfer from *Paenibacillus* and *Bacillus* or related group. Since all the *Pseudomonas bim* clusters are in the same clade of the phylogenetic tree, the acquisition likely occurred once in *Pseudomonas*. Similarly, the acquisition of the *bim* cluster specifically

in the *P. syringae* group also likely occurred once since they form a clade separate from other *Pseudomonas* species (Figure 3.4).

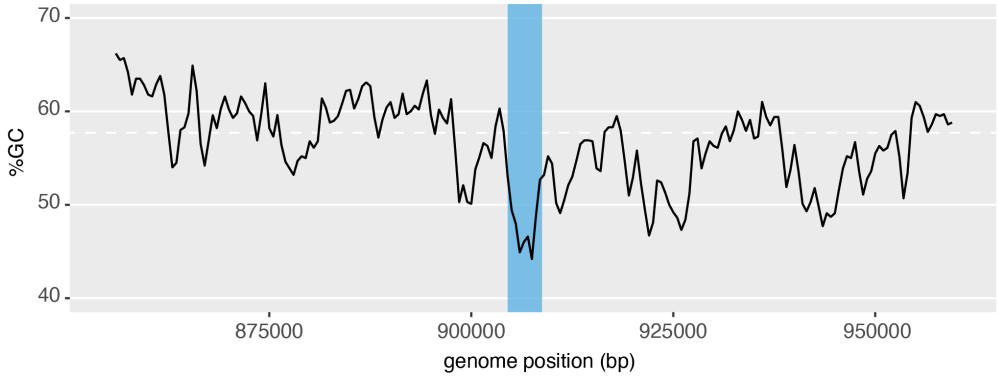


Figure 3.6. The *bim* cluster shows lower GC content than adjacent regions and the average of the whole genome. Genome sequence of *P. syringae* pv. tomato DC3000 was analysed for % GC content in sliding windows of 500 bp across the genome. The %GC was then plotted at the 500 bp interval. Dashed line indicates the average GC content of the whole genome. The 100 kb region around *bim* cluster is displayed. The position of the *bim* cluster is highlighted in blue.

3.3.7 The presence of the *bim* cluster is distributed across major phylogroups of *P. syringae*

I next investigated the distribution of the *bim* cluster in *P. syringae* strains. The protein sequence of PSPTO0834 was used to search for homologs in sequenced *P. syringae* strains in NCBI database to identify the presence of the *bim* cluster. Then, a phylogenetic tree of *P. syringae* strains was constructed from concatenated DNA sequence of four conserved housekeeping genes (*rpoD, gyrB, gltA, gapA*) using strains that were identified to contain the *bim* gene cluster together with other type and pathotype strains (Thakur, Weir and Guttman, 2016) and additional strains to represent different pathovars and phylogroups. The presence of the *bim* cluster in its genomic context were then mapped onto the phylogenetic tree. Strains from all five major phylogroups (1, 2, 3, 4, 5) were found to have the *bim* cluster, although not all members of each phylogroup have the *bim* cluster (Figure 3.7). In contrast, the *bim* cluster was absent in the strains from phylogroups 6 and 7 (Figure 3.7). However, only a few samples were analysed so this observation might not be conclusive. In addition, the

strains with the *bim* cluster cover a range of pathovars isolated from a wide range of host plants but not all strains infecting the same type of host plant have the *bim* cluster. For example, for tomato, the *bim* cluster is present in DC3000 and deltaX but not in T1. Likewise, for bean, the *bim* cluster is present in 1448A but not in B728a (Figure 3.7). Overall, the *bim* cluster is widespread among *P. syringae* strains.

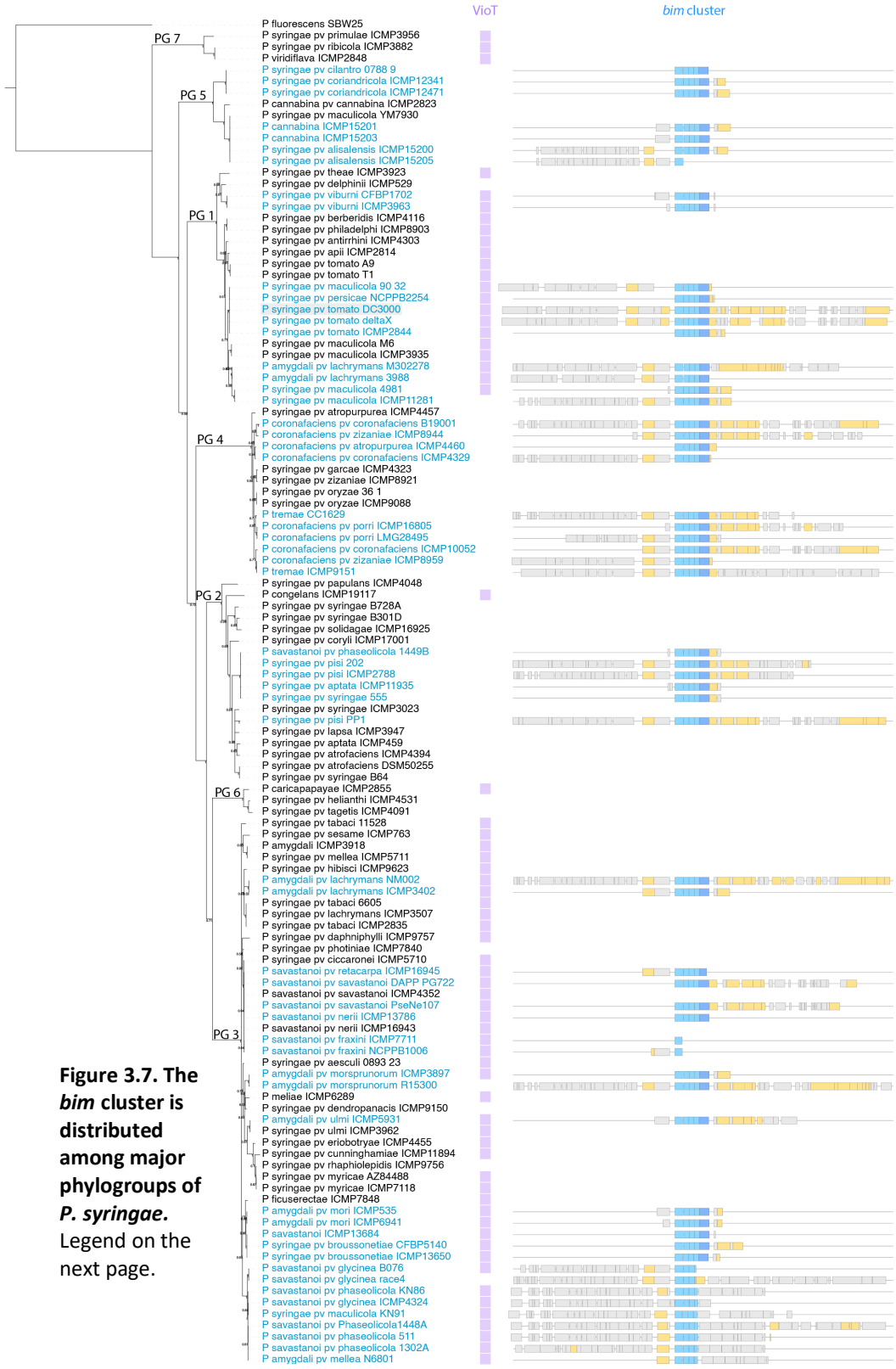


Figure 3.7. The *bim* cluster is distributed among major phylogroups of *P. syringae*. Legend on the next page.

Figure 3.7. The *bim* cluster is distributed among most major phylogroups of *P. syringae.* The phylogenetic tree was reconstructed using concatenated coding DNA sequences of 4 conserved housekeeping genes (*rpoD*, *gyrB*, *gltA* and gapA). *P. fluorescens* SBW25 was included as an outgroup. The DNA sequences were aligned with MUSCLE and partial deletion with 95% cutoff was used to remove gaps in the alignment. Maximum Likelihood (ML) phylogenetic tree was reconstructed. Branch support values were calculated from Bootstrap method with 100 replications. Branches with less than 0.5 branch support value were collapsed. Strains containing the *bim* cluster are labelled in blue. Phylogroup classification is displayed on each branch. VioT solid squares indicate that the strain has a transferase for modified viosamine in flagellin glycosylation. Genomic regions (20 kbp) around the *bim* cluster are displayed. Annotated genes are shown as rectangles. The *bim* cluster genes PSPTO0834-7 are labelled in light blue, PSPTO0838 is labelled in dark blue. Annotated transposase, integrases, recombinases or conjugative elements are labelled in yellow. Incomplete assembly of genome is common as can be seen when the genome region has no gene annotations (rectangles) displayed. For clarity, only a representative subset of strains is shown. The full set of all strains identified is shown in Appendix 2.

3.3.8 The *bim* cluster distribution is not limited to strains with modified viosamine in flagellin glycosylation

Since the plant BGAL1 enzyme was identified to play a role in plant defence by facilitating the release of elicitor peptides from flagellin that is glycosylated with modified viosamine, the production of BGAL inhibitor by the *bim* cluster could be a counter defence mechanism to protect the bacterial strains with the sensitive flagellin glycosylation (Buscaill *et al.*, 2019). To test if there is a correlation between inhibitor production and modified viosamine glycosylation in *P. syringae* strains, I also searched for the presence of orthologs of VioT, a transferase enzyme responsible for the transfer of modified viosamine residue in flagellin glycosylation (Yamamoto *et al.*, 2011). VioT could be found in phylogroups 1, 3 and 7 (Figure 3.7). However, the *bim* cluster for inhibitor production is also present in phylogroups 2, 4 and 5, which do not have VioT. In contrast, phylogroup 7 strains have VioT but no *bim* cluster (Figure 3.7). This suggests that inhibitor production is not limited to the BGAL1-flagellin glycosylation interaction and there are likely additional roles of the inhibitor in the interaction between *P. syringae* and plant during infection.

3.3.9 The *bim* cluster was horizontally transferred among *P. syringae* strains

The scattered pattern of distribution of the *bim* cluster among the strains of *P. syringae* in the phylogenetic tree and the close association of all identified *bim* clusters with transposable elements (Figure 3.7) suggests that the *bim* cluster was horizontally transferred among the *P. syringae* strains. To test this hypothesis, I also constructed a gene tree based on the coding DNA sequence of PSPTO0834 from the *bim* cluster to assess the phylogenetic relationships between the distributed *bim* clusters and compare it to the phylogenetic relationships between the different strains that contain them. As expected for horizontal gene transfer, incongruence between the species tree and gene tree was observed (Figure 3.8). Multiple clades of closely related *bim* clusters are distributed in multiple strains from different groups. These results support horizontal gene transfer of the *bim* cluster between strains of *P. syringae*. It is likely that multiple instances of horizontal gene transfer occurred in ancestors of multiple strains, then the clusters were also vertically inherited to the present strains analysed. This also hints at potentially useful functions of the *bim* cluster that have allowed it to be repeatedly acquired and maintained.

3.3.10 The transfer of the *bim* cluster was likely mediated by a site-specific integrase targeting tRNA-lysine site

From the genomic context of *bim* clusters mapped out in the phylogenetic tree (Figure 3.7), the *bim* clusters seem to be downstream of two conserved genomic regions. This may appear as if they were vertically inherited. However, upon closer inspection, both conserved genomic sites of *bim* cluster are downstream of a tRNA-lysine encoding site. The tRNA-lysine gene is present in only two copies at two sites in the genome, which are conserved across *P. syringae* strains (Figure 3.9). As examples, the genomic sites from three strains, DC3000, 1448A and B728a, are shown. Both tRNA-lysine sites were followed by an integrase or transposase, which

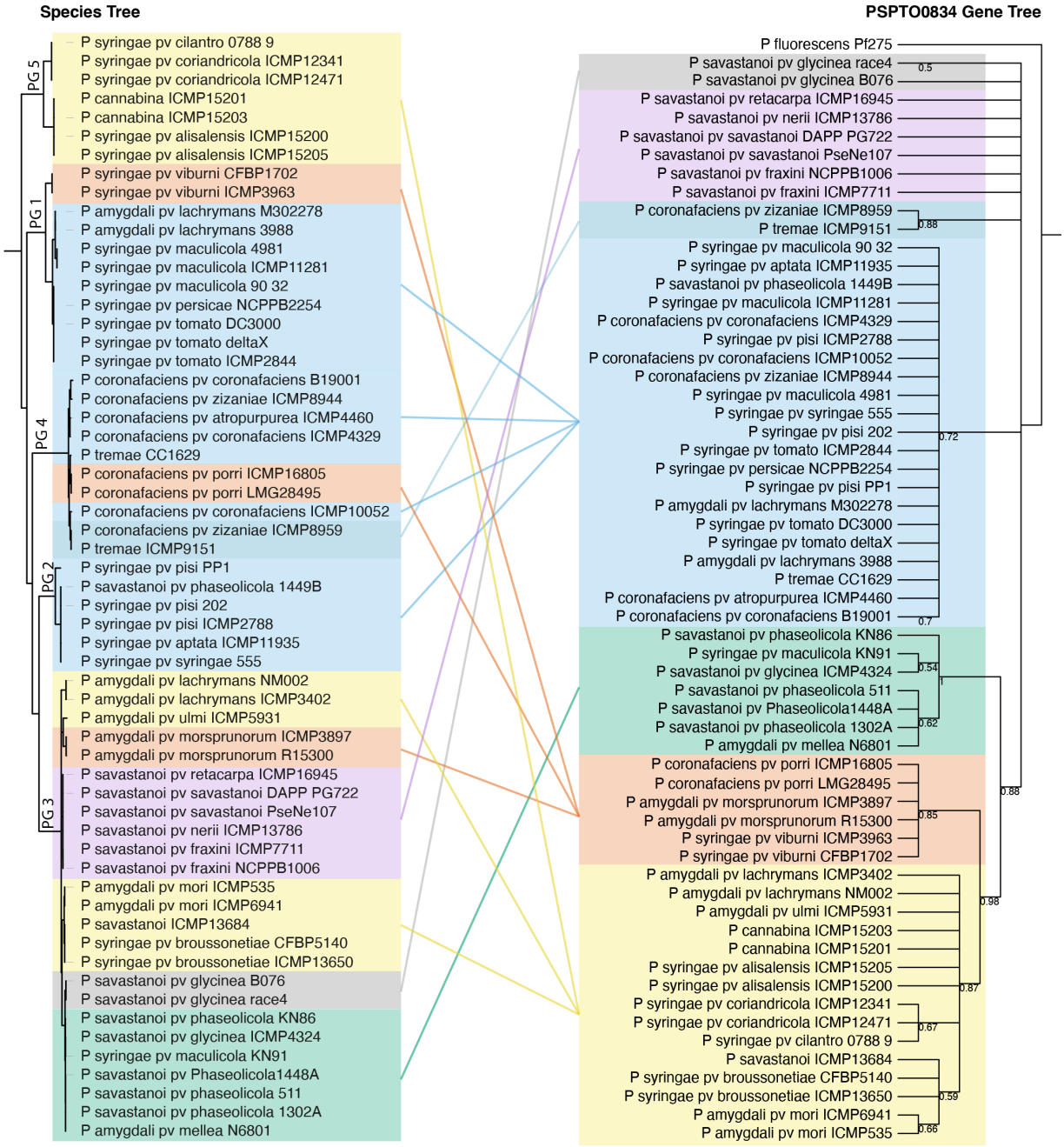


Figure 3.8. The *bim* cluster was horizontally transferred among *P. syringae* strains. Phylogenetic tree of *P. syringae* strains (species tree), is the same tree as Figure 7 but trimmed to show only strains that contain the *bim* cluster for clarity. Phylogenetic tree of *bim* cluster (gene tree) was reconstructed from coding DNA sequences of homologs of PSPTO0834 identified in *bim* clusters from *P. syringae* strains. A homolog from *P. fluorescens* Pf275 was included as an outgroup. The DNA sequences were aligned with MUSCLE and partial deletion with 95% cutoff was used to remove gaps in the alignment. Maximum Likelihood (ML) phylogenetic tree was reconstructed. Branch support values were calculated from Bootstrap method with 100 replications. Branches with less than 0.5 branch support value were collapsed. The corresponding clades in species tree and gene tree are coloured the same and connected with lines.

likely targeted this site. In DC3000 and 1448A, either one of the two tRNA-lysine sites contain integrations of the *bim* cluster, while B728a does not contain *bim* cluster (Figure 3.9). In addition, homologs of these integrases also exist nearby tRNA-lysine sites in other species of *P. aeruginosa* (pseudomonas.com database, (Winsor *et al.*, 2016)). Moreover, most integrases are commonly known to target tRNA genes and many exhibit a specific site preference (Williams, 2002). Together, these findings suggest that the horizontal transfer of the *bim* cluster was mediated by an integrase that targets the tRNA-lysine sites.

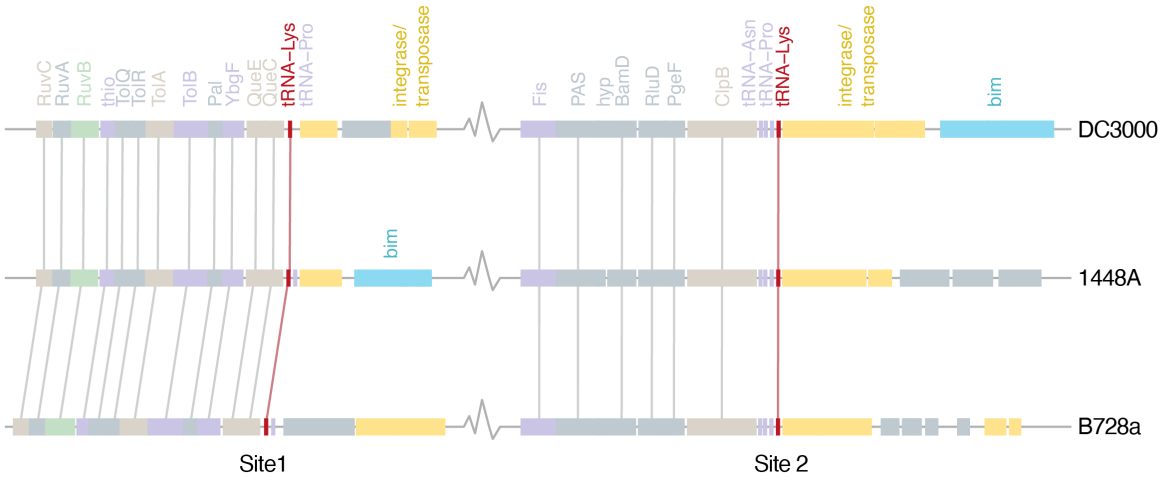


Figure 3.9. The transfer of the *bim* cluster is likely to have been mediated by a site-specific integrase targeting tRNA-lysine sites. Genomic regions around the two conserved tRNA-lysine sites (red) in the genomes of *P. syringae* pv. *tomato* DC3000, pv. *phaseolicola* 1448A and pv. *syringae* B728a. Genes are labeled with names and homologs are connected by lines to show synteny of the region. Annotated integrases or transposases are labeled in yellow. The *bim* cluster is labelled in blue.

3.3.11 The *bim* cluster has a conserved *hrp* box promoter unique to *P. syringae*

The promoter region of the *bim* cluster in DC3000 contains a consensus *hrp* box sequence for binding of HrpL, a major regulator of expression of type III secretion system (T3SS) and virulence factors in *P. syringae* (Lam *et al.*, 2014). This is consistent with the fact that T3SS regulators were identified in my screen for genes required for inhibitor production as well as the production of the inhibitor in T3SS-inducing MG medium. To assess the conservation of this regulation, the promoter regions of *bim* clusters identified from *Pseudomonas* strains

were aligned. All of the *bim* clusters from *P. syringae* strains have the conserved *hrp* box sequence in their promoter regions, while the *bim* cluster found in *P. fluorescens* lacks the *hrp* box in its promoter region (Figure 3.10). In addition, the pattern of sequence variation within the promoter region was also consistent with the phylogenetic relationship of the *bim* cluster gene tree, suggesting they were inherited together (Figure 3.10). These data indicate that the *bim* cluster was transferred to an ancestor of *P. syringae* and acquired the *hrp* box sequence, allowing it to be regulated by the type III secretion system regulator before it was transferred into multiple strains of *P. syringae*. The co-regulation of the *bim* cluster with the major virulence mechanism induced during infection also suggests a possible beneficial role in plant infection which allowed the cluster and this regulation to be selected and maintained.

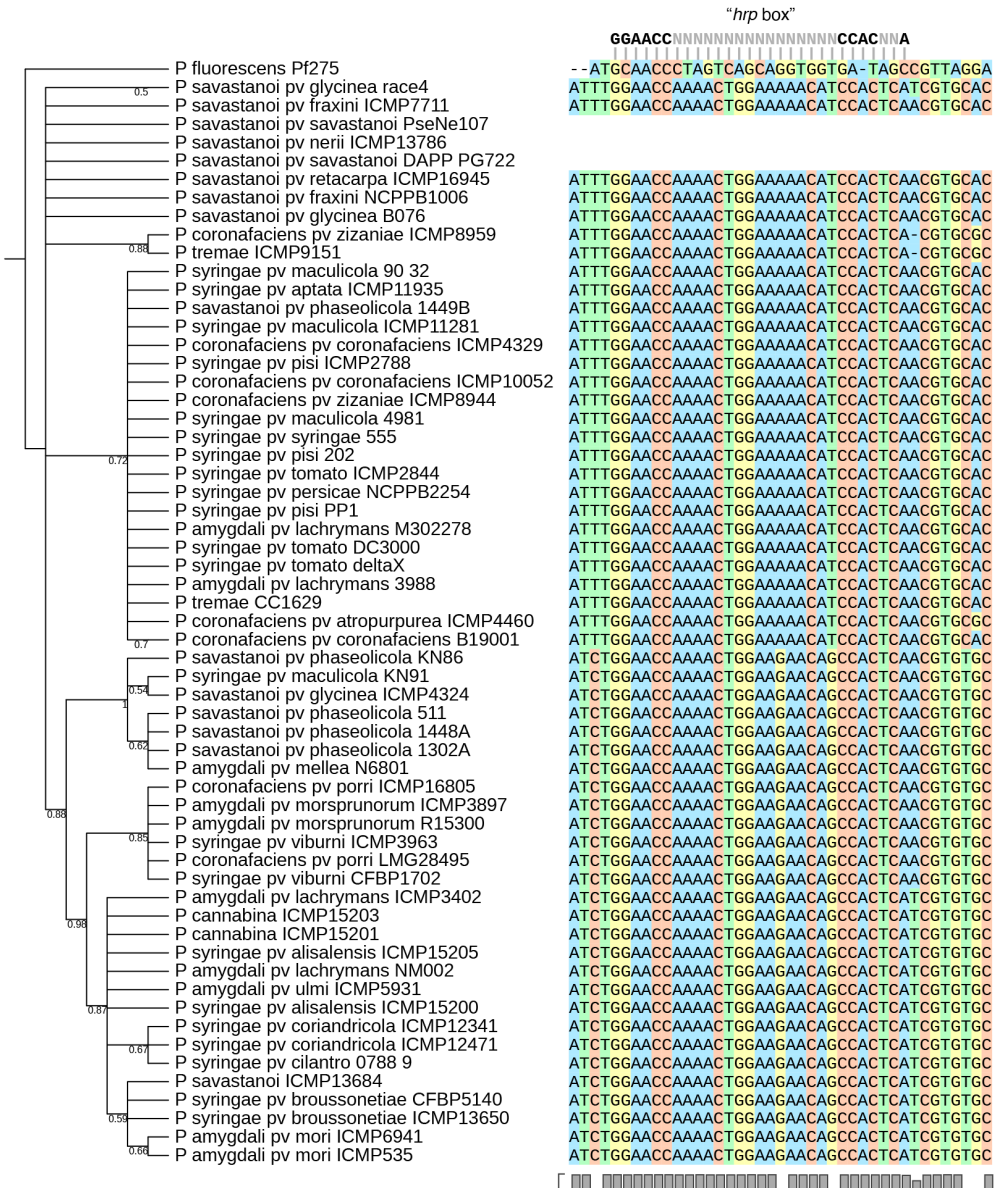


Figure 3.10. **The *bim* cluster has a conserved *hrp* box promoter unique to *P. syringae*.** Phylogenetic tree of the *bim* cluster (same as Figure 8) was reconstructed from coding DNA sequences of homologs of PSPTO0834 identified in the *bim* clusters from *P. syringae* strains. A homolog from *P. fluorescens* Pf275 was included as an outgroup. The DNA sequences were aligned with MUSCLE and partial deletion with 95% cutoff was used to remove gaps in the alignment. Maximum Likelihood (ML) phylogenetic tree was reconstructed. Branch support values were calculated from Bootstrap method with 100 replications. Branches with less than 0.5 branch support value were collapsed. Sequences of promoter regions around 150 bp upstream of the start codon of PSPTO0834 were aligned with Clustal Omega. The strains with incomplete genome assembly around this site were omitted. The consensus *hrp* box sequence is displayed above the alignment and % consensus of each position is displayed below the alignment.

- The regulation of inhibitor production

3.3.12 Expression of the *bim* cluster and inhibitor production is regulated by type III secretion system regulators

My screen for *P. syringae* BGAL inhibitor mutants (*bim*) identified *rhpS*, *hrpR*, *hrpS* and *hrpL* as strong candidate genes required for inhibitor production. These genes all encode regulators of the type III secretion system (T3SS) and virulence factor expression (Xie, Shao and Deng, 2019), which acts in a transcriptional regulatory cascade. To confirm the involvement of the T3SS regulators in the production of the BGAL inhibitor, I tested an independent set of mutant strains lacking each of the major T3SS regulators (*hrpR*, *hrpS*, *hrpL*) and structural genes for the T3SS (*hrpA*, *hrcC*) for inhibitor production compared to the wild type DC3000 strain in T3SS-inducing MG medium. As expected, the mutants of regulators (*hrpR*, *hrpS*, *hrpL*) were unable to produce the inhibitor, while the mutants of T3SS structural genes were still able to produce the inhibitor (Figure 3.11A). These results demonstrate that the inhibitor production is dependent on T3SS regulation but not the T3SS machinery.

Next, I tested if the T3SS regulators act at the transcriptional level by inducing *bim* cluster expression to enable inhibitor production. The non-inhibitor producing mutants isolated from my screen containing transposon insertions in the T3SS regulators (*rhpS*, *hrpR*, *hrpS*, *hrpL*)

were tested along with the inhibitor-producing HQ background strain as well as mutant strains unrelated to inhibitor production carrying a transposon insertion in *hopI1* (non-*bim*) and a random insertion. The strains were grown in MG medium to induce T3SS expression and then total RNA was extracted and subjected to RT-PCR analysis to assess the level of expression of the *bim* cluster by monitoring transcript of PSPTO0834. As expected, all the non-inhibitor producing mutants of T3SS regulators (*rhpS*, *hrpR*, *hrpS*, *hrpL*) showed reduced level of *bim* cluster transcript compared to the inhibitor producing strains (HQ, non-bim, random) (Figure 3.11B), confirming that the regulators were required for induction of *bim* gene expression. In addition, the expression of *avrPtoB*, an effector gene regulated by T3SS, was also monitored and showed the same pattern as the *bim* cluster (Figure 3.11B), consistent with the same coregulation by conventional T3SS regulators.

These results demonstrate that T3SS regulators are required for the induction of the *bim* cluster expression and inhibitor production. This is also consistent with the identification of the conserved *hrp* box in the promoter of *bim* cluster.

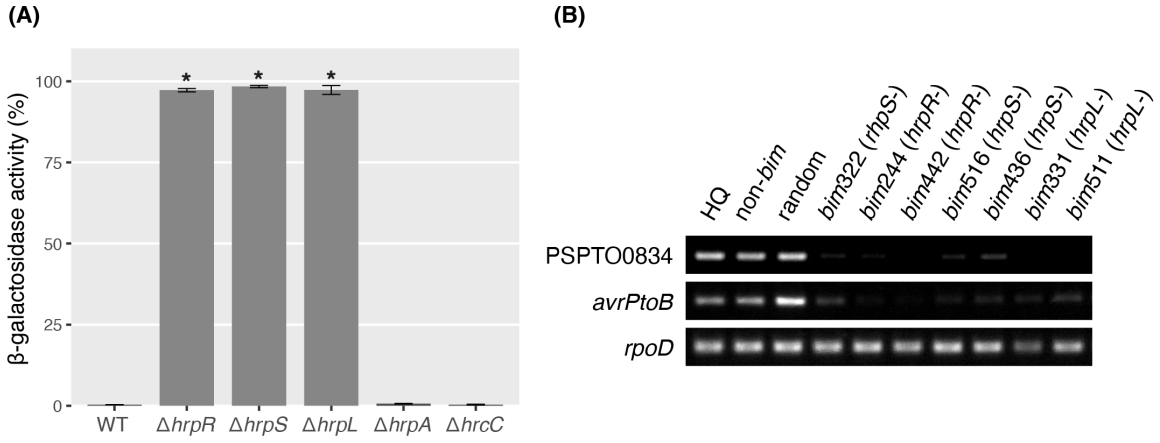


Figure 3.11. Expression of the *bim* cluster and inhibitor production is regulated by type III secretion system regulators. (A) Test for inhibitor production. The bacteria were grown in MG medium from a starting OD600 of 1 overnight and supernatant was collected for enzymatic inhibition assay with FDG and LacZ enzyme. The enzymatic activity was expressed as percentage of the activity compared to maximum activity. Mean and standard deviation are shown (n=3). Asterisks indicate a statistically significant difference ($P < 0.05$) compared to the wild type strain (WT) using ANOVA and Tukey HSD test. WT refers to *P. syringae* DC3000 strain,

Δ*hrpR*/Δ*hrpS*/Δ*hrpL* strains are deleted for the type III secretion system regulators and the Δ*hrpA*/Δ*hrcC* strains are deleted for structural type III secretion machineries. **(B)** RT-PCR. The different bacterial strains were grown from a starting OD600 of 1 in MG medium for 6 hours to induce type III secretion system expression. Total RNA was then extracted and used for RT-PCR using primers specific for transcripts of PSPTO0834, *avrPtoB* (type III secreted effector gene) or *rpoD* (housekeeping gene control). HQ refers to *P. syringae* DC3000 *ΔhopQ1* strain used as a background strain for transposon insertion mutagenesis. The rest of the strains are transposon inserted mutants. Non-*bim* strain containing a transposon insertion in *hopI1* gene, which encodes a secreted effector protein unrelated to inhibitor production, was used as a control. Random strain contains a transposon insertion in a random site which does not affect inhibitor production and was used as another control. The strains *bim*322 contains a transposon insertion disrupting *rhpS, bim*244 and *bim*442 contain a transposon insertion disrupting *hrpR, bim*516 and *bim*436 contain a transposon insertion disrupting *hrpS, bim*331 and *bim*511 contain a transposon insertion disrupting *hrpL.*

3.3.13 The expression of the *bim* cluster and inhibitor production followed T3SS induction in the early hours of infection *in planta*

I next monitored the dynamics of *bim* gene expression during the early hours of infection *in planta*. This was done by detecting luminescence upon infiltration of *N. benthamina* leaves with bacterial strains containing a lux gene cassette in a mini-Tn5 transposon insertion downstream of the promoter of each gene in the T3SS regulatory pathway (*rhpS*, *hrpR*, *hrpS* and *hrpL*), *bim* cluster (PSPTO0835) and T3SS effector *avrPtoB*. The insertion of the lux-containing transposons in these strains also led to gene disruption. I reasoned that the strains could still serve as reporters of gene induction, however, after the initial induction, the pathogenicity or virulence phenotypes were likely to be affected. Nevertheless, the induction pattern observed during the first two hours after infection followed the pattern of the T3SS regulatory cascade where *hrpR/S and hrpL* were induced, which in turn induced the expression of the *bim* cluster and the T3SS effector gene *avrPtoB* (Figure 3.12A, B). However, *hrpL* expression declined after two hours, likely because the disruption of *hrpL* render the strain non-pathogenic and was affected by the plant immune responses. Consistently, BGAL inhibitor production also started to be detectable after induction of *bim* cluster expression (Figure 3.12C). These results confirm the coregulation of the *bim* cluster expression, and inhibitor production, with the conventional T3SS regulation *in planta*.

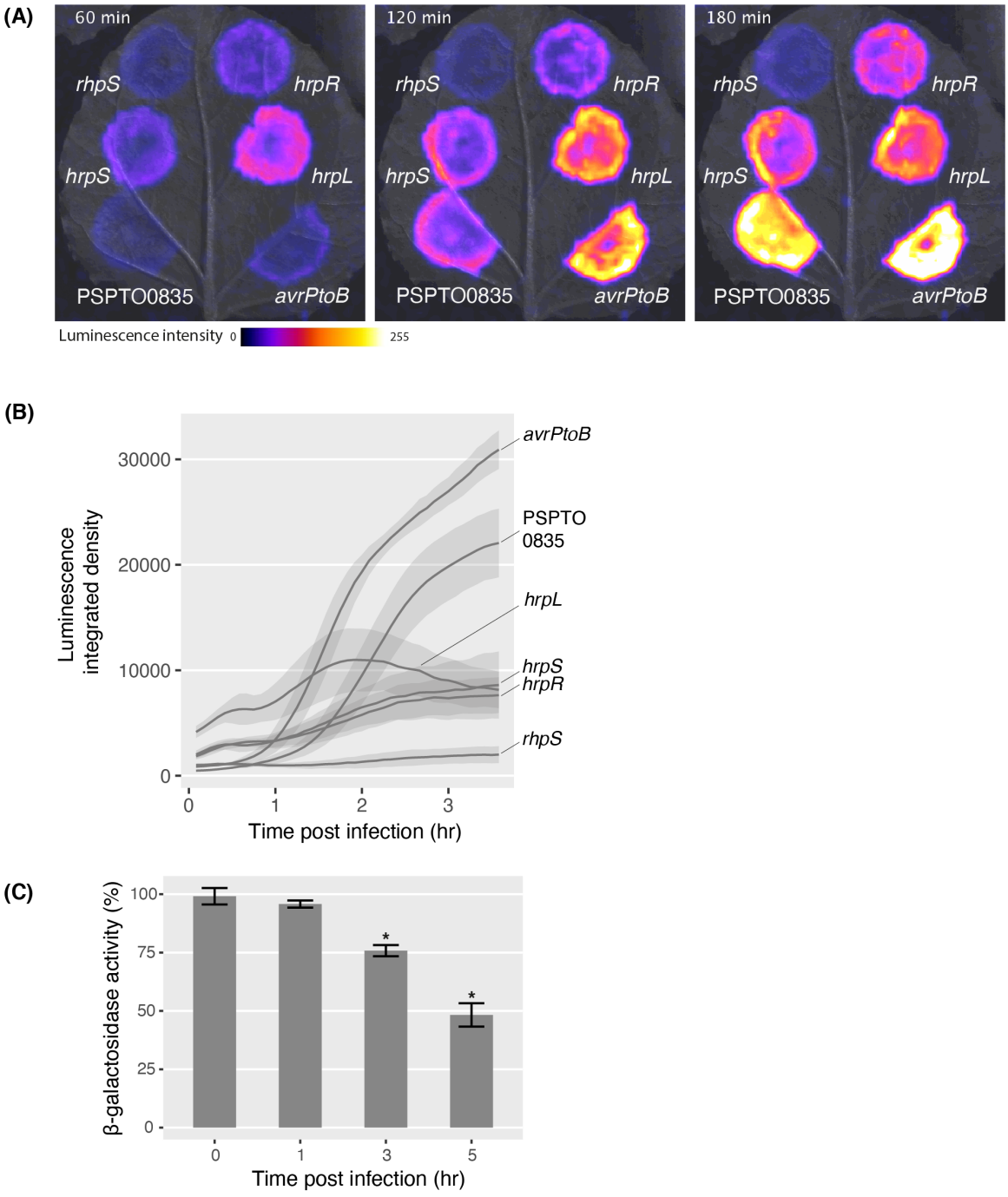


Figure 3.12. The expression of the *bim* cluster and inhibitor production followed T3SS induction in early hours of infection *in planta*. (A) luminescence reporter of gene expression. The *P. syringae* strains used contain luminescence reporter constructs inserted downstream of the promoter of each gene indicated. The bacterial suspension was infiltrated into spots on *N. benthamiana* leaves at OD_{600} of 0.2 (~10^8 CFU/ml). The leaves were then detached from the plant and imaged for luminescence signal over time to monitor gene expression dynamics. **(B)** The luminescence signal from (A) was quantified by integrated density of the signal from the infiltrated area of the leaf. Mean and standard deviation are shown (n=3). **(C)** *P. syringae* HQ strain was infiltrated into on *N. benthamiana* leaves at OD_{600} of 0.2 (~10^8 CFU/ml). Then, apoplastic fluid was extracted at each timepoint and tested for BGAL activity with FDG assay. The enzymatic activity was expressed as percentage of the activity compared to timepoint 0. Mean and standard deviation are shown (n=3). Asterisks indicate a statistically significant difference ($P < 0.05$) compared to timepoint 0 using ANOVA and Tukey HSD test.

3.3.14 HtpX plays a role in the induction of *hrpL* and *bim* cluster expression

In addition to the conventional regulators of the T3SS, *htpX* (PSPTO1778) was also identified in my screen for *P. syringae* BGAL inhibitor mutants (*bim*) as a candidate gene that affects inhibitor production. *htpX* encodes a putative membrane-associated protease, likely having a regulatory function. To confirm the role of HtpX, I tested inhibitor production from the culture of the isolated *htpX* transposon insertion mutant (*bim*77) and complemented the mutant with a plasmid carrying *htpX* under a constitutive *Lac* promoter. Compared to the no inhibitor mutant strain lacking the *bim* cluster, the *htpX* mutant showed only a small degree of inhibition but when the mutant was complemented a high level of inhibition was restored (Figure 3.13). These data indicate that HtpX has a positive role in the production of the inhibitor, although not absolutely required.

To test if the role of HtpX in the inhibitor production is in the transcriptional regulation pathway, I monitored transcript levels of T3SS regulators *hrpR* and *hrpL,* as well as the HrpL-regulated genes *avrPtoB* and the *bim* cluster gene PSPTO0834. Strains were grown in MG medium to induce T3SS expression then total RNA was extracted for RT-PCR. The non-*bim* mutant containing a transposon insertion in *hopI1*, unrelated to inhibitor production was included as a control. Both mutants of *htpX* (*bim*77 and *bim*210) showed reduced expression level of *hrpL*, PSPTO0834 and *avrPtoB* (Figure 3.13B). In addition, the complementation of *htpX* in the *bim*77 mutant restored the level of *hrpL*, PSPTO0834and *avrPtoB* expression (Figure 3.13B), confirming that the observed phenotype was specific to *htpX*. In contrast, the expression level of the upstream regulator *hrpR* was not as drastically affected (Figure 3.13B). These results demonstrate that HtpX isrequired for the induction of T3SS expression and

inhibitor production at the level of *hrpL* expression (Figure 3.13C). HtpX could therefore function in another layer of regulation of T3SS expression.

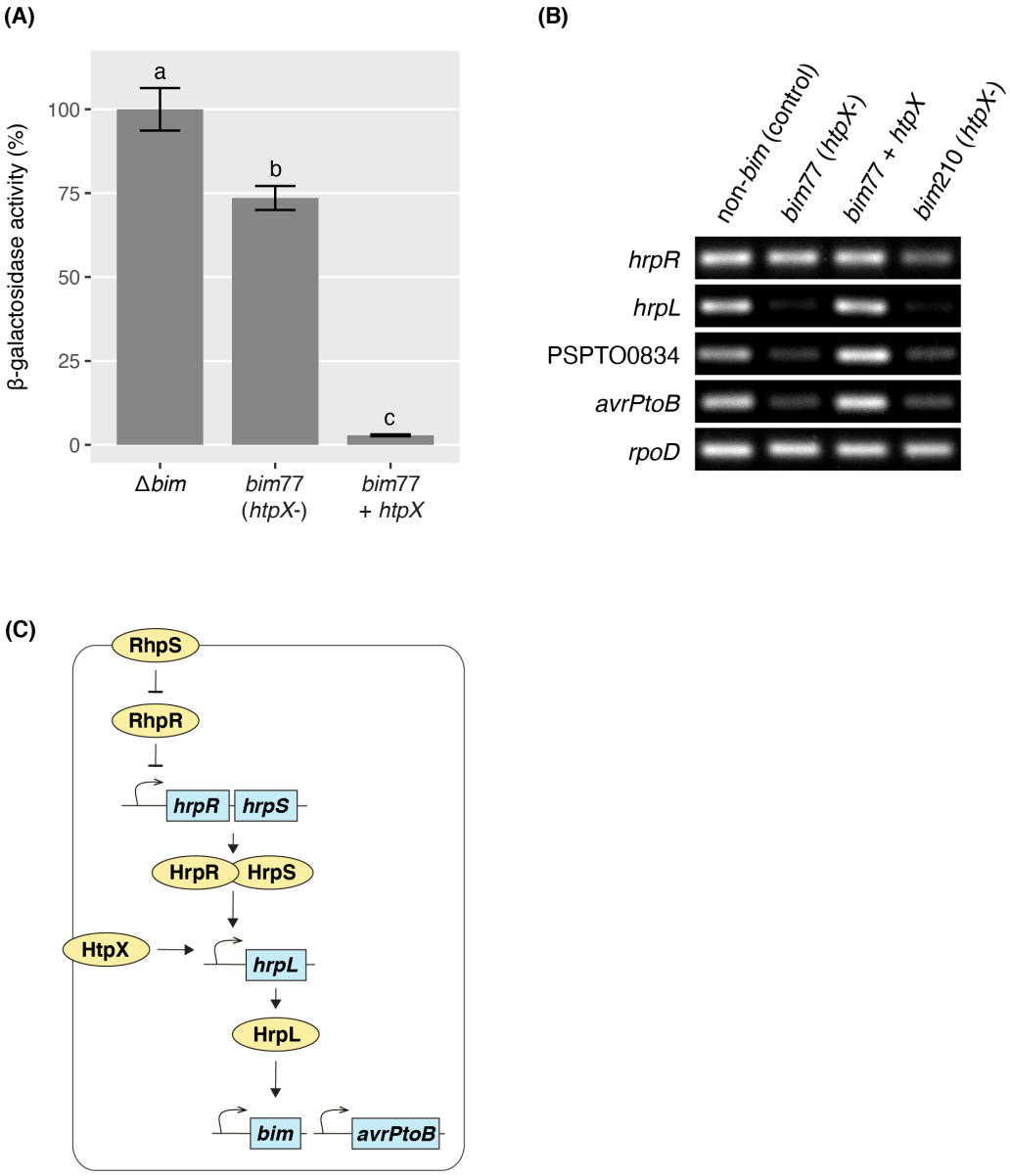


Figure 3.13. HtpX plays a role in the induction of *hrpL* and *bim* cluster. (A) test for inhibitor production. The bacteria were grown in MG medium and supernatant was collected for enzymatic inhibition assay with FDG and LacZ enzyme. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control (HQ Δ*bim*). Mean and standard deviation are shown (n=3). Different letters indicate statistically significant difference ($P < 0.05$) from ANOVA and Tukey HSD test. Comp means complemented strain. **(B)** RT-PCR. The different bacterial strains were grown in MG medium for 6 hours to induce type III secretion system expression. Total RNA was then extracted as used for RT-PCR using primers specific for transcripts of *hrpR*, *hrpL*, PSPTO0834, *avrPtoB* effector or *rpoD* housekeeping control. Two independent transposon insertion mutants disrupting *htpX* were used (*bim*77 and *bim*210). *bim*77 complemented with *htpX* in a plasmid (*bim*77+*htpX*) was also used to confirm the phenotype of *htpX*. Non-*bim* strain containing a transposon insertion in *hopI1* gene, which encodes a secreted effector protein unrelated to inhibitor production, was used as a control. **(C)** Schematic diagram illustrating the main type III expression system regulatory cascade discussed in this work.

- The biological roles of the inhibitor during infection

From previous sections, I have discovered that the *bim* gene cluster is widespread among *P. syringae* strains, being horizontally transferred and acquired multiple times. In addition, the expression of the *bim* cluster is coregulated with the type III secretion system (T3SS), the main virulence system of *P. syringae*. Moreover, the plant BGAL1 enzyme, which is inhibited by the inhibitor, has been characterised for its function in plant defence. These findings suggest that inhibitor production could be beneficial to the bacteria during plant infection. Therefore, I set out to test the role of inhibitor production during infection. The inhibitor-producing HQ strain and the non-producing mutant (HQ Δ*bim*) were used for comparisons.

3.3.15 Deletion of the *bim* cluster did not affect bacterial growth *in vitro*

First, I tested if the deletion of the *bim* cluster affects bacterial growth *in vitro*. Both HQ and HQ Δ*bim* strains were inoculated in LB medium or a T3SS-inducing MG medium and their growths were monitored by measuring the optical density (OD600) over time. Both strains showed no difference in growth in either conditions (Figure 3.14), suggesting that the inhibitor production did not have an important role for normal growth of the bacteria *in vitro*.

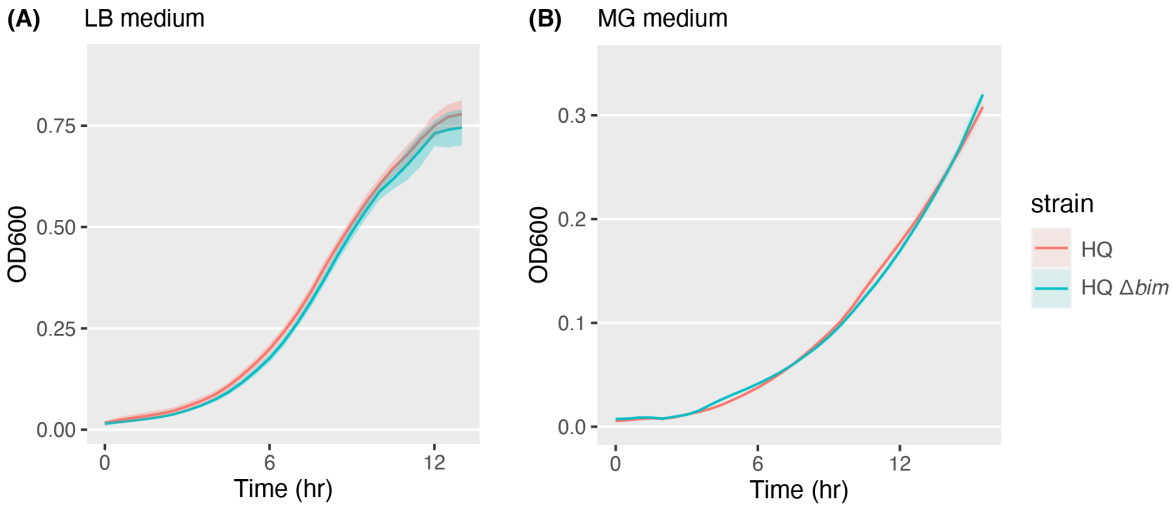


Figure 3.14. Deletion of the *bim* cluster did not affect the bacterial growth *in vitro.* Bacteria (HQ and HQ Δ*bim* strains) were pre-grown overnight in LB medium. The cells were then collected and inoculated at the OD600 of 0.05 in LB **(A)** or MG media **(B)** in a 96 well plate. Bacterial growth was monitored with OD600 readings over time with incubation at 28°C and shaking during the course of the experiment. Mean and standard deviation are shown (n=3).

3.3.16 Deletion of the *bim* cluster did not result in a robust virulence phenotype *in planta*

I next tested if inhibitor production promotes bacterial growth during infection. I infected *N. benthamiana* leaves with the inhibitor-producing HQ strain and the non-inhibitor-producing mutant lacking the *bim* cluster (HQ Δ*bim*) and monitored the number of bacterial colony forming units (CFU) in the plant tissue to detect bacterial growth. I also tested different infection methods. First, no significant effect of *bim* cluster deletion on bacterial growth was observed with the infection by infiltration of bacterial inoculum into leaves (Figure 3.15A). Since the role of inhibitor production could be in the preinvasive stage of infection that would be bypassed by infiltration, I then investigated infection by spraying bacterial inoculum on leaf surface. I observed a significant but small effect of *bim* cluster deletion reducing bacterial growth in two experiments, but this was not reproducible in subsequent experiments (Figure 3.15 B, C). Next, I speculated that the effect of the inhibitor production could be redundant and masked by actions of other effectors produced by the bacteria to suppress plant defence. Therefore, I tested the effect of inhibitor production in a background of *hrpA* knockout strain, which lacks a structural protein of the type III secretion system rendering it unable to inject effector proteins into the plant cell. Again, there was no significant difference between the inhibitor producing and non-producing mutant (Figure 3.15D). Finally, I speculated that the inhibitor production by the bacteria at the time of plant inoculation might not be high and quick enough to show a clear effect since the bacteria was first grown in a rich LB medium, which suppressed inhibitor induction. I thus tried to pre-induce the inhibitor production in

the bacteria by first incubating the bacterial inoculum in the inducing MG medium before spraying on the plant. However, I still did not observe a significant effect on bacterial growth (Figure 3.15E).

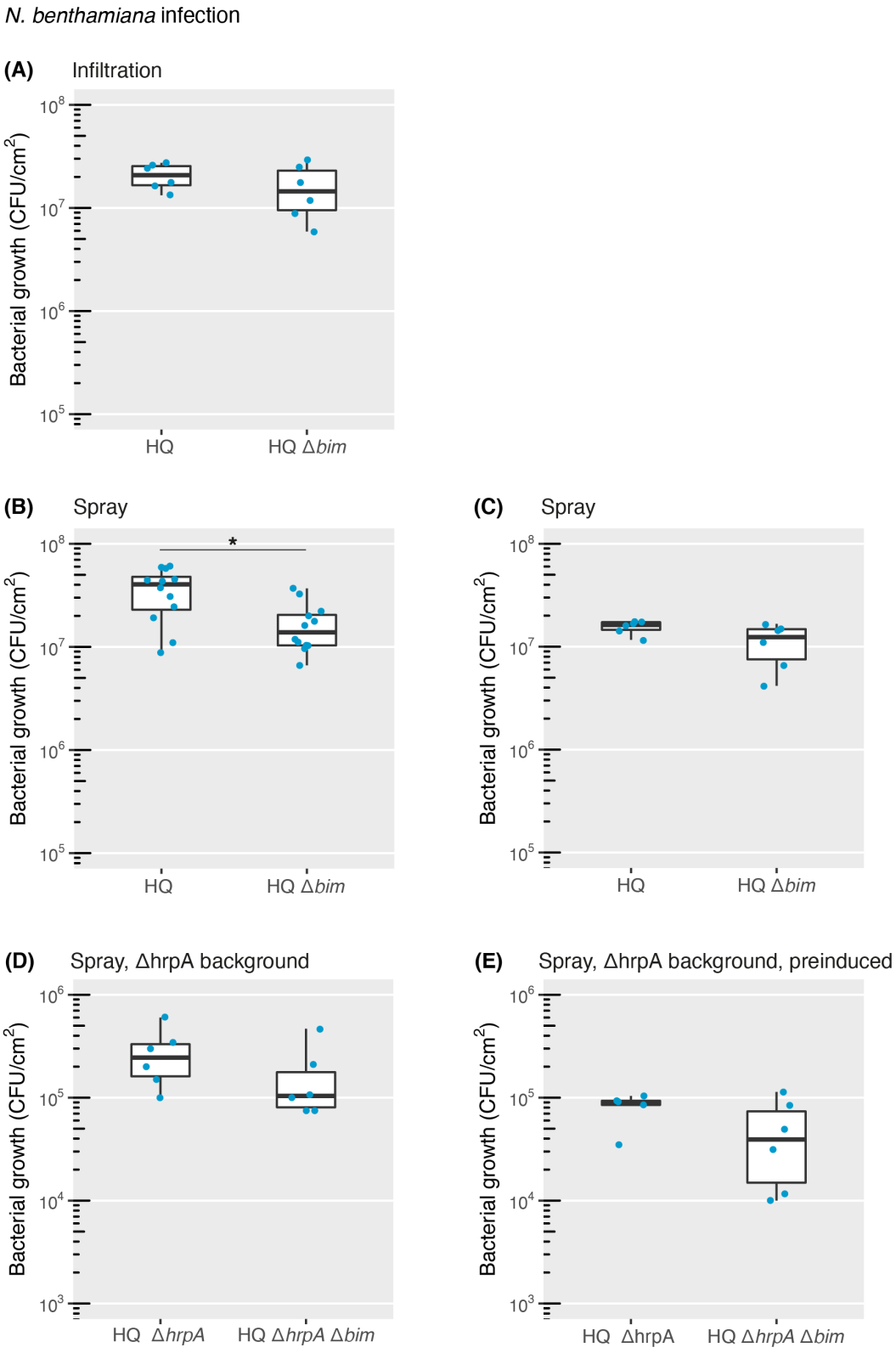


Figure 3.15. Deletion of the *bim* cluster did not result in a robust virulence phenotype *in N. benthamiana* infections. *N. benthamiana* plants were infected with **(A-C)** the inhibitor-producing HQ and non-producing HQ Δbim strains or **(D-E)** inhibitor-producing HQ Δ*hrpA* and non-producing HQ Δ*hrpA* Δ*bim* strains, additionally lacking the type III secretion system. Infection was done by **(A)** infiltration of bacterial suspension into the leaf (~10^5 CFU/ml + 0.04% (v/v) Silwet L-77) and incubating inoculated plants for 2 days or **(B-E)** spraying of bacterial suspension on both leaf surfaces (~10^8 CFU/ml) and incubating inoculated plants for 3 days. Leaf discs were then collected from infected tissue, surface sterilised, homogenised then serially diluted and plated on LB agar. In **(E)** Bacterial inoculum was incubated in MG medium prior to infection to pre-induce inhibitor production. The number of bacterial colony forming units (CFU) present per cm^2 area of the infected leaf tissue was calculated. Asterisk indicates a statistically significant difference ($P < 0.05$) from t-test.

In a previous study on HrpL regulon, PSPTO0834 was identified to be under control of HrpL and its knockout strain was found to have impaired growth compared to the wild type DC3000 strain in *Arabidopsis* (Vencato *et al.*, 2006). Therefore, I also tested my strains with *Arabidopsis* plants by both infiltration and spray infection. However, I did not observe a significant difference between the inhibitor producing and non-producing strains (Figure 3.16).

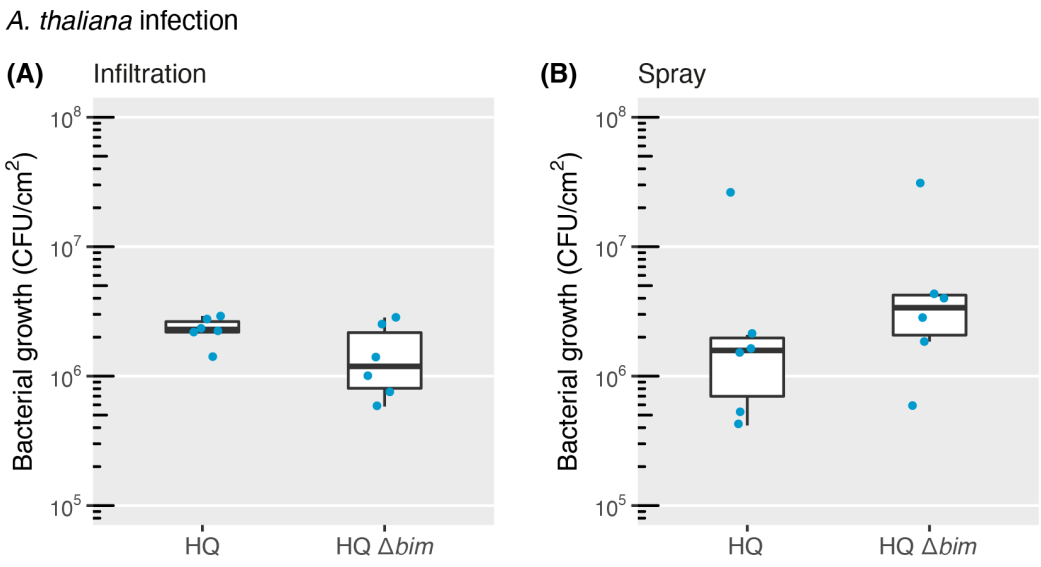


Figure 3.16. Deletion of the *bim* cluster did not result in a robust virulence phenotype in *A. thaliana* infections. *Arabidopsis* plants were infected with the inhibitor-producing HQ or non-producing HQ Δ*bim* strain by **(A)** infiltration of bacterial suspension into the leaf (~10^5 CFU/ml) and incubating inoculated plants for 2 days or **(B)** spraying of bacterial suspension on leaf surface (~10^8 CFU/ml + 0.04% (v/v) Silwet L-77) and incubating inoculated plants for 3 days. Leaf discs were then collected from infected tissue, surface sterilised, homogenised then serially diluted and plated on LB agar. The number of bacterial colony forming units (CFU) present per cm^2 area of the infected leaf tissue was calculated.

Overall, the presence of the *bim* cluster and inhibitor production had no robust effects on bacterial growth during plant infection in the conditions tested. There was a tendency for the non-inhibitor producing HQ Δ*bim* strain to show slightly less growth than the wild type strain *in planta* when inoculated by spraying but the effect was minor and not consistently statistically significant.

3.3.17 The production of BGAL inhibitor altered the profile of galactose-containing glycoproteins in the plant apoplast during infection

Finally, I sought to explore other of BGAL inhibitor production by *P. syringae* during plant infection. Since there could be glycoproteins carrying galactose residues that are targets of BGAL enzyme, I hypothesised that the galactose-containing glycoproteins in the plant apoplast could be altered during infection by the BGAL inhibitor. To test this, I used a biotinylated lectin protein RCAI, which has an affinity to bind galactose residues (Itakura *et al.*, 2007), to probe the profile of glycoproteins in the apoplastic fluid extract from *N. benthamiana* leaves.

First, I determined whether plant BGAL1 has an impact on glycan profiles of apoplastic glycoproteins by comparing wild type and BGAL1 knockout mutant (*bgal1-1*) plants. The apoplastic fluid samples from *bgal1-1* plants showed an accumulation of galactose-positive signals, which were largely absent in wild type plants (Figure 3.17A). The detected glycoprotein signals seemed to have a smeary pattern, suggesting heterogenous glycan modifications with different molecular weights. In addition, glycoprotein signals were excluded from some regions corresponding to protein bands detectable in coomassie staining

(eg. ~70 and 180 kDa), suggesting that galactose residues were not found in all apoplastic proteins.

Next, I tested apoplastic fluid samples from wild type plants infected with either inhibitor-producing HQ or non-producing HQ Δ*bim* strain. The samples from HQ infected leaves showed an accumulation of galactose-positive signals that also increased from 3 to 4 days post infection, mainly at high molecular weight (Figure 3.17B). This accumulation is similar to the *bgal1-1* mutant plant (Figure 3.17A). In contrast, the galactose-positive signal was largely absent in samples from HQ Δ*bim* infected leaves (Figure 3.17B), similar to wild type plants (Figure 3.17A). These data suggest that the suppression of BGAL by the inhibitor during infection with *P. syringae* modified the glycoproteins in the plant apoplast. The identity and function of these galactose-containing glycoproteins as well as the effect of differential glycosylation with or without the galactose residue remains to be tested further.

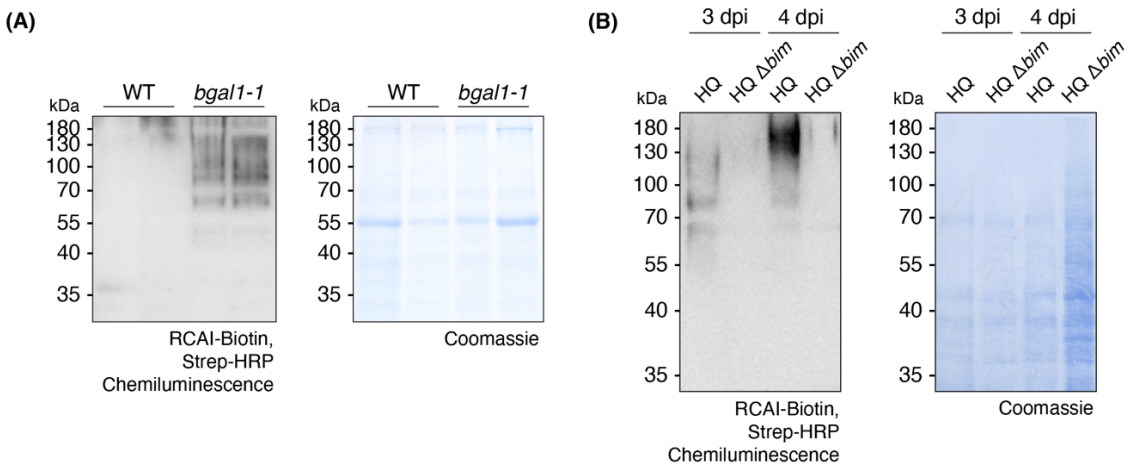


Figure 3.17. The production of BGAL inhibitor altered the profile of galactose-containing glycoproteins in the plant apoplast during infection. The biotinylated galactose-binding lectin protein RCAI was used to probe a protein blot from SDS-PAGE separation of protein extracts in apoplastic fluids of *N. benthamiana* leaves from **(A)** wild type plant (WT) or BGAL1 mutant plant (*bgal1-1*), **(B)** wild type plant infected with either inhibitor-producing HQ or non-producing HQ Δ*bim* strain at 3 or 4 days post infection (dpi). The biotinylated lectin was detected by HRP-conjugated streptavidin and ECL chemiluminescence substrate. Coomassie blue staining of the same blot was used to assess total protein loading. This experiment was done once.

3.4 Discussion

In this chapter, I characterised the biosynthetic gene cluster for the BGAL1 inhibitor and investigated the evolution of this gene cluster in *P. syringae* as well as homologs in other bacterial species. Then, I characterised the regulation of inhibitor production in *P. syringae* and identify the regulators involved. Finally, I explored the potential roles of BGAL inhibitor in bacterial virulence during infection.

3.4.1 The biosynthesis of the inhibitor

In this study, I have identified and confirmed that the *bim* gene cluster is the biosynthetic gene cluster of the BGAL inhibitor. The deletion of the *bim* cluster from *P. syringae* abolished inhibitor production, which could then be complemented with the *bim* cluster. Moreover, heterologous expression of the *bim* cluster in *E. coli* also led to inhibitor production. Further analysis of the five genes in the cluster (*PSPTO0834-8*) showed that the first two genes, *PSPTO0834* and *PSPTO0835* were sufficient to initiate BGAL inhibitor production in both *P. syringae* and *E. coli*. The other three genes (*PSPTO0836-8*) were not required, possibly due to a redundancy with other genes in the genome that can fulfil the same functions. Alternatively, these genes might be essential for inhibitor production at conditions that were not tested in this study, such as different plant environments containing different metabolites. For example, the encoded proteins could act on a different substrate that can also feed into the inhibitor biosynthesis pathway. Likewise, the putative transporter could be required for uptake of substrates used as precursors for inhibitor production. In addition, since these genes are present in different combinations in different versions of homologous gene clusters in other bacterial strains, they could also be essential in the strains where the *bim* cluster

evolved but not in *P. syringae* and *E. coli*. On the other hand, they might have functions unrelated to inhibitor production and their presence in the *bim* cluster is a remnant of the evolutionary history of the cluster.

Previous studies on biosynthesis of iminosugar can help inform the understanding of BGAL inhibitor biosynthesis by the *bim* cluster. The first essential gene in the *bim* cluster, PSPTO0834, encodes an alcohol dehydrogenase in the L-threonine dehydrogenase family, which is homologous to GutB1, an enzyme described for the biosynthesis of 1-deoxynojirimycin (DNJ) (Clark, Johnson and Horenstein, 2011). This enzyme catalyses the key step for iminosugar formation by oxidising the hydroxyl group to a ketone group that then spontaneously reacts with the amine group, forming the N-containing ring of iminosugar (Figure 1.5). This supports my hypothesis of iminosugar nature of the BGAL inhibitor and can explain the essential role of PSPTO0834 in biosynthesis of the inhibitor. The acceptance of a substrate containing both amine and hydroxyl groups is likely the key characteristic of this enzyme family. The second essential gene, PSPTO0835, encodes an uncharacterised reductase that belongs to the dihydrofolate reductase family. Interestingly, in the proposed biosynthetic pathway of DNJ, the final step involves a yet unidentified reductase that is responsible for the reduction of the double bond between C and N in the ring (Figure 1.5). This reaction is reminiscent of the known role of dihydrofolate reductase in the reduction of the double bond in a dihydrofolate ring (Cao *et al.*, 2018), suggesting that PSPTO0835 might function in a similar reduction reaction. Another similarity between genes in the *bim* cluster and the known biosynthetic pathway is the putative phosphatase PSPTO0836. Phosphatase Yktc1 in DNJ biosynthesis catalyses the removal of a phosphate group from the fructose phosphate precursor of the pathway (Figure 1.5). Although they belong to a different family

and no homolog of Yktc1 is found in *P. syringae*, a similar phosphatase step is likely to occur in the biosynthesis of the BGAL inhibitor. Since PSPTO0836 was not essential in the inhibitor production, this step could be catalysed by other phosphatases. Another enzyme identified in DNJ biosynthesis was aminotransferase GabT1, which is responsible for the transfer of the amine group onto the fructose sugar backbone, allowing the subsequent N ring formation (Figure 1.5). Interestingly, I also identified its homolog, GabT2, once in my screen of BGAL inhibitor mutants. However, the inhibitor production was not rescued by complementing the mutant strain with GabT2 and an independent mutant lacking all GabT homologs (1, 2 and 3) was still capable of producing the inhibitor. These data indicated that GabT homologs are not required for BGAL inhibitor production. Other aminotransferases are likely to be involved instead. In addition, a similar set of biosynthetic steps was also proposed for iminosugar DAB-1 (Nuñez and Horenstein, 2019), highlighting this framework of iminosugar biosynthetic pathway. The different set of enzymes used for BGAL inhibitor production also suggests that it is different molecule from other previously characterised inhibitors.

Another gene in the *bim* cluster, PSPTO0837, encodes a domain of unknown function DUF1349. This domain is classified in a concanavalin-like lectin/glucanase superfamily (Pfam database, (El-Gebali *et al.*, 2019)) and may have functions related to carbohydrate binding or hydrolysis. One protein containing this domain was studied in *Saccharomyces cerevisiae* for a role in the regulation of a galactose-regulated metabolic pathway (Choi *et al.*, 2008). Nevertheless, the function of this protein domain is still unknown.

The last gene in the *bim* cluster, PSPTO0838, belongs to the Major Facilitator Superfamily (MFS) of transporters. It is classified in a family of fucose-galactose-glucose:H^+ symporter with members such as fucose permease and glucose/galactose permease (Pao, Paulsen and Saier,

1998). Thus, PSPTO0838 could function as a transporter for sugar uptake. PSPTO0838 was not required for inhibitor production in both *P. syringae* and *E. coli* at the conditions tested so it is unclear how or whether it plays a role in inhibitor production. Furthermore, there is also a clade of *P. syringae* with *bim* clusters that has lost PSPTO0838, suggesting that it was not essential for inhibitor production. Nevertheless, transporter proteins are likely required for the secretion of the inhibitor. Since the inhibitor has a sugar mimicking property, it might be exported by sugar transporters. In addition, there are multiple multidrug transporters characterised for efflux of wide range of cationic metabolites, including MFS and MATE family transporters (Omote *et al.*, 2006; Du *et al.*, 2015). One or many of these transporters might be able to transport the inhibitor.

Apart from the *bim* cluster, *nadB* was required for optimal inhibitor production since its mutant was impaired in inhibitor production and its complementation restored inhibitor production. *nadB* encodes an L-aspartate oxidase functioning in *de nov*o biosynthesis of NAD (Kurnasov *et al.*, 2003). It is likely involved in inhibitor production since at least two key enzymes (PSPTO0834 and PSPTO0835) are oxidoreductases, which are dependent on NAD(P)(H) cofactors. I have not determined which form of the cofactor is required for each enzyme although I could speculate from its homolog. PSPTO0834 is homologous to GutB1 which uses NAD+ as a cofactor for an oxidation reaction (Wu, Arciola and Horenstein, 2013), whereas PSPTO0835 is in the same family as dihydrofolate reductase which uses NADPH as a cofactor in reduction reaction (Cao *et al.*, 2018). Nevertheless, *nadB* was not essential since its mutant could still produce some inhibitor. The fact that the mutant was not lethal also suggested that the biosynthesis of NAD, an essential cofactor for cells, was not completely abolished. This is consistent with the report that there is an alternative pathway for NAD

biosynthesis in *Pseudomonas aeruginosa* (Devries *et al.*, 1995). Interestingly, different growth conditions resulted in different phenotypes of the *nadB* mutant. The mutant previously grown on LB agar was more impaired in inhibitor production than those grown in liquid LB media, suggesting a difference in metabolic state of the cells and regulation of the pathway and availability of NAD. This is consistent with a study with *P. aeruginosa* reporting that surface-attached cells showed lower levels of enzyme bound NADH and overall NAD(H) production than planktonic cells (Perinbam *et al.*, 2020).

Additional unidentified genes are likely required for other steps in the biosynthesis of the inhibitor. My attempts to test other candidates (threonine dehydratase, aspartate ammonia lyase, isochorismatase, and GabT2) from my transposon insertion mutagenesis screen of BGAL inhibitor mutants did not yield extra positive leads since they could not be complemented. These isolated mutants might contain secondary mutations elsewhere in the genome that impact inhibitor production. However, preliminary sequencing analysis of coding sequences of known players (the *bim* gene cluster and T3SS regulators) did not identify any mutations in these mutant strains. Alternatively, additional enzymes in biosynthesis of the inhibitor might be identified from a more targeted approaches using the knowledge from previously characterised biosynthesis of iminosugars. For example, phosphatase and aminotransferase are likely required. A screen with targeted gene deletions could be tested for their effects on inhibitor production but this could suffer from similar limitations as the initial random mutant screen such as mutant lethality and gene redundancy. Conditional/inducible silencing could be used to overcome the former and multiple simultaneous deletions or silencing could be used to overcome the latter. In addition, the choice of culture media might also affect the isolation of mutants as different substrates and

pathways might be used in different media. For instance, auxotrophic mutants might not grow in minimal media. On the other hand, a screen for genes in the inhibitor biosynthetic pathway could be done by testing if their overexpression can increase inhibitor production. However, the enhancing effect of the gene may only be detectable if it is the limiting factor of the pathway.

Once a more complete biosynthetic pathway of the inhibitor is known, its components can be tested further with *in vitro* reconstruction using purified enzymes and substrates. This approach would also overcome complications of gene redundancy when using bacterial cells. However, using the right substrates for each enzyme is still a challenge. Additional information could also be gained from studying each purified enzyme for the characterisation of substrate specificities and kinetics.

3.4.2 The evolution of inhibitor production

The identification of *bim* biosynthetic gene cluster of the BGAL inhibtor allowed me to further explore the evolution of this inhibitor biosynthesis. I hypothesised that the alcohol dehydrogenase (PSPTO0834) is a key distinguishing step in the biosynthesis of iminosugar since it was also found for previously characterised iminosugars (DNJ, DAB-1) (Clark, Johnson and Horenstein, 2011; Nuñez and Horenstein, 2019). In addition, I speculated that its role in the oxidation of hydroxyl to ketone group in a substrate that also contains an amine group, which allows subsequent formation of iminosugar ring, is a unique property, unlike phosphatase or aminotransferase steps which can act on other metabolites. Therefore, I decided to search for all homologs of the alcohol dehydrogenase PSPTO0834 in order to identify candidate gene clusters for iminosugar production and study their evolution in a

broader context than the *bim* cluster alone. Indeed, with this approach, I identified gene clusters in various bacterial groups in combinations with common gene components characteristic of the iminosugar gene cluster discussed previously. Several strains that carry these putative clusters indeed produce iminosugars, supporting my hypothesis that this gene is the determinant of iminosugar production. Interestingly, many bacterial species identified for the putative iminosugar gene clusters are plant-associated or plant-growth-promoting bacteria including *Paenibacillus* species (Berge *et al.*, 2002; Timmusk, Grantcharova and Wagner, 2005; Grady *et al.*, 2016), *Bacillus* species (Allard-Massicotte *et al.*, 2016; Radhakrishnan, Hashem and Abd Allah, 2017), endophytes such as *Kosakonia oryzendophytica* (Li *et al.*, 2016) and the N-fixing symbiont *Frankia* (Sellstedt and Richau, 2013). Iminosugars produced by these strains might have a role in plant colonization or interactions with plant-associated microbiome.

The recurring biosynthetic genes found in the putative clusters include alcohol dehydrogenase, aminotransferase, phosphatase (HAD or FIG family), reductase (RibD family) This set of enzymes follow the framework of the previously characterised biosynthetic pathways of iminosugar (Clark, Johnson and Horenstein, 2011; Nuñez and Horenstein, 2019). The variation in the components of the gene cluster as well as the metabolic potential of the strains carrying them (eg. presence of other enzymes for further modifications of the iminosugar backbone) might determine the type of iminosugar produced. It would be interesting to test these uncharacterised strains to discover novel iminosugars. In addition, although many gene homologs are shared between different types of clusters, some of them probably diverged to have different substrate specificities and produce different iminosugars. For instance, GutB1 from the DNJ biosynthesis was unable to completely replace the function

of its homolog, PSPTO0834, in BGAL inhibitor biosynthesis. With further characterisation, the diversity of these genes products could provide useful resources for metabolic engineering for enhanced production of known iminosugars or novel iminosugars. For example, characterisation of alcohol dehydrogenase homologs in selected *Bacillus* and *Paenibacillus* strains showed that they exhibit different catalytic turnover rate and substrates preference (Wu, Arciola and Horenstein, 2013). In addition, a point mutation was also made in the alcohol dehydogenase to alter the amino specificity so that the enzyme accepts polyol substrates, demonstrating a potential for enzyme engineering (Wu, Arciola and Horenstein, 2013).

Apart from biosynthetic enzymes, transporters (MFS, MATE, SulP family) and glycosidases are also frequently found in putative iminosugar gene clusters. The transporters might mediate uptake of substrates or secretion of iminosugars. As for the glycosidase, since it may be counterintuitive to produce a glycosidase at the same time as a glycosidase inhibitor, the glycosidase that is present in the inhibitor biosynthesis cluster might be resistant to inhibition by the produced iminosugar. Further studies may lead to the discovery of transporters to enhance iminosugar secretion or iminosugar-insensitive glycosidases.

Putative iminosugar biosynthetic gene clusters were identified across different bacterial genera from a broad range of phyla suggesting that the gene cluster could have evolved from horizontal gene transfer between different bacterial groups, as often the case for biosynthetic gene clusters (Ochman *et al.*, 2000). In addition, in many species including some *Bacillus*, *Streptomyces*, *Clostridium* and all of *P. syringae*, the clusters are present nearby transposable elements which could account for their mobilization (Koonin, Makarova and Aravind, 2001).

Regarding the *bim* cluster, its acquisition likely happened once in the ancestor of *Pseudomonas* since all *bim* clusters in *Pseudomonas* species are grouped in the same clade of the phylogenetic tree (HGT1, Figure 3.18). Similarly, the ancestor of *P. syringae* group likely acquired the *bim* cluster once since their *bim* clusters form a clade separate from those in other *Pseudomonas* species (HGT2, Figure 3.18). In addition, the promoter regions of *bim* clusters in *P. syringae* also uniquely contain a conserved *hrp* box sequence, mediating coregulation with the type III secretion system induced during infection in plants (Lam *et al.*, 2014). This *hrp box* was likely acquired after the transfer of *bim* cluster into the ancestor of *P. syringae* and then inherited together with the *bim* cluster (Figure 3.18). Possibly, the expression of the *bim* cluster in plants conferred some selective advantage. The association of transposable elements with the *bim* cluster could mediate its horizontal transfer among different *P. syringae* strains (HGT3, Figure 3.18). After the acquisition, the *bim* cluster was also maintained and vertically inherited in multiple strains.

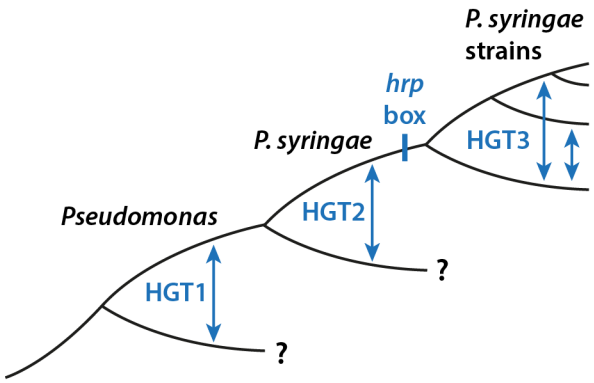


Figure 3.18. Possible evolutionary history of the *bim* cluster. The ancestor of the *bim* cluster was acquired once in the *Pseudomonas* group through a horizontal gene transfer event from an unknown species (HGT1). This ancestor of *bim* cluster was then horizontally transferred between several *Pseudomonas* species, including once into the *P. syringae* group (HGT2). Once in *P. syringae*, the *bim* cluster acquired a *hrp* box sequence in its promoter, allowing its expression to be coregulated with the type III secretion system, the main virulence mechanism induced during infection. Subsequently, the *bim* cluster was distributed among different strains of *P. syringae* through multiple horizontal gene transfer events (HGT3). After the acquisition, the *bim* cluster was also maintained and vertically inherited in multiple strains.

The origin of the *bim* cluster acquisition might be impossible to conclusively identify but might be speculated. Horizontal gene transfer can often be inferred from genomic signature, such as GC content, which can differs between genomes of distant donor and recipient strains (Ravenhall *et al.*, 2015). In *P. syringae* DC3000, the *bim* cluster has lower GC content than the nearby region and the average of the whole genome, consistent with its implicated foreign origin. Interestingly, *Paenibacillus* and related groups have genomic GC contents within the same range, as a member of the Firmicutes group (Chen *et al.*, 2007; Bohlin *et al.*, 2017; Luo *et al.*, 2018). However, GC content alone cannot be used as a definitive prove. In addition, *Paenibacillus* strains are well represented among species identified for the iminosugar gene cluster and contain all of the diversity of gene components found in the *bim* cluster. Thus, the acquisition of the *bim* cluster in *P. syringae* clade could be from *Paenibacillus* or related bacteria. However, there is also a possibility that my phylogenetic analysis was limited by the availability of genome sequences in the database and that the representation of bacterial taxa are biased or incomplete.

The horizontal gene transfer of the *bim* cluster between *P. syringae* strains (HGT3, Figure 3.18) is supported by four findings. First, the *bim* cluster has a scattered distribution in the phylogenetic tree of *P. syringae*. Second, the apparent incongruence between the species tree of *P. syringae* strains and gene tree of the *bim* cluster PSPTO0834 is indicative of horizontal gene transfer. Third, the *bim* cluster can be found at two different genomic locations in different strains. Fourth, the *bim* cluster is always flanked by transposable elements. Furthermore, the phylogenetic distribution pattern of the *bim* cluster suggests that multiple instances of horizontal gene transfer occurred in ancestors of multiple strains, then those clusters were vertically inherited to the present strains analysed. The transfer of the

bim cluster between *P. syringae* strains was likely mediated by a tRNA-lysine site-specific integrase since the *bim* cluster tends to be downstream of either of the two conserved genomic sites encoding tRNA-lysine and integrases that are commonly known to target tRNA genes and many exhibit a specific site preference (Williams, 2002). In addition, the finding that the *bim* cluster was repeatedly acquired and maintained in multiple strains also suggests that inhibitor production is beneficial and promotes its selection. Similarly, horizontal transfer of pathogenicity islands has been accounted for the evolution of a wide range of pathogenicity and virulence-related genes such as effectors, toxins and the type III secretion system (Arnold, Pitman and Jackson, 2003). Interestingly, the pathogenicity island PPHGI-1 in *P. syringae* pv. *phaseolicola* 1302A genome is also found next to a tRNA-lysine site and has an excision ability mediated by a XerC recombinase (Pitman *et al.*, 2005).

The widespread presence of the *bim* cluster across most major *P. syringae* phylogroups with pathovars that infect a wide range of plant species suggests that the inhibitor could act on a glycosidase mechanism that is conserved across multiple plant species. Previously, the plant BGAL1 enzyme was implicated in plant defence with its role in facilitating the release of elicitor peptides from flagellin that is glycosylated with modified viosamine. The production of BGAL inhibitor by the *bim* cluster was attributed to be a counter mechanism to protect the bacterial strains with the sensitive flagellin glycosylation (Buscaill *et al.*, 2019). However, the *bim* cluster is not limited to *P. syringae* strains using modified viosamine in flagellin glycan suggesting that the function of BGAL inhibitor production is not limited to the interaction between BGAL1 and flagellin glycosylation. The inhibitor could target the action of BGAL on additional substrates or other glycosidases.

3.4.3 The regulation of inhibitor production

The expression of the *bim* cluster and inhibitor production was confirmed to be under control of regulators of type III secretion system (T3SS), the major virulence mechanism in *P. syringae*. The mutations disrupting *rhpS*, *hrpR*, *hrpS* and *hrpL* all impaired *bim* cluster expression and inhibitor production. Furthermore, the promoter of the *bim* cluster contains the *hrp* box consensus sequence, a HrpL binding site (Vencato *et al.*, 2006; Mucyn *et al.*, 2014), supporting the direct regulation of *bim* cluster expression by HrpL. The *bim* cluster is therefore part of the HrpL regulon, which is induced during infection under the conventional regulatory cascade of T3SS regulators (Xie, Shao and Deng, 2019). This mode of regulation is also consistent with the dynamics of *bim* gene expression and inhibitor production during early hours of infection. In addition, inhibitor production is also detected in the T3SS-inducing minimal medium such as MG medium (Bronstein *et al.*, 2008). The co-regulation of the *bim* cluster with the T3SS induced during infection also suggests a possible role of the inhibitor in plant infection.

In addition to conventional T3SS regulators, I also identified HtpX as potentially a novel regulator of T3SS expression and inhibitor production. HtpX, is a membrane-bound metalloprotease, with a catalytic domain on the cytoplasmic side, implicated in proteolytic quality control of membrane proteins and membrane stress responses (Sakoh, Ito and Akiyama, 2005). HtpX is heat-inducible in *E. coli*, *Bacillus subtilis* and *Xylella fastidiosa* but not *Streptococcus gordonii* (Sakoh, Ito and Akiyama, 2005; Lin, Huang and Shaw, 2012). In a transcriptomics study on *P. syringae*, the expression of *htpX* was highly induced upon treatment with phenol as a membrane stress, and slightly induced upon treatment with H_2O_2 as an oxidative stress (Sandhu, 2008).

In this study, the mutant of *htpX* was impaired in inhibitor production, which was rescued with the complementation of *htpX*. Further gene expression analysis by RT-PCR showed that HtpX acted at the transcriptional level of *hrpL* and was required for normal expression of HrpL and its downstream regulon such as the *bim* cluster and effectors. Nevertheless, the role of HtpX is likely as a modulator of gene expression but not an on/off induction switch since low levels of inhibitor production and *bim* cluster expression were still detectable in the *htpX* mutant. Interestingly, a mutant of *htpX* was also identified in a transposon mutagenesis screen as one of the mutants defective in the induction of T3SS *in planta* (Xiao, 2005).

HtpX likely functions as a positive regulator of *hrpL* expression, possibly by its proteolytic activity causing degradation of a negative regulator of *hrpL* or activation of a positive regulator of *hrpL*. The targets of HtpX cleavage could be membrane proteins that are associated with regulation of T3SS expression (Xie *et al.*, 2020). There are also examples of proteases regulating T3SS including Lon and RhpP, which function as negative regulators of T3SS by cleaving HrpR and HrpL, respectively (Bretz *et al.*, 2002). It would be interesting to identify the proteolytic substrate of HtpX and the transcription profiling of genes regulated by its action. In addition, it would also be interesting to determine if HtpX function is related to temperature. Unlike other toxins such as phaseolotoxin and coronatine, which are induced at low temperature (18°C) but not at high temperature (28°C) (Ullrich *et al.*, 1995; Aguilera *et al.*, 2017), BGAL inhibitor production occurs at both low and high temperatures. HtpX could function to promote T3SS induction at high temperature.

3.4.4 The role of inhibitor production during infection

My findings on the evolution and regulation of the *bim* cluster indicate that the BGAL inhibitor likely has a role in bacterial colonization of the plant environment. In addition, the previous work on the plant BGAL1 enzyme identified its role in plant defence by facilitating the release of elicitor peptides from flagellin that contains modified viosamine in the glycan. The BGAL1 knockout *N. benthamiana* plant was more susceptible to the bacterial infection (Buscaill *et al.*, 2019). Therefore, I initially focused my hypothesis regarding the role of the BGAL inhibitor with the expectation that the *P. syringae* strain producing BGAL inhibitor should be better protected from BGAL1 and show better growth during infection than the mutant lacking the inhibitor. Interestingly, a previous study that identified PSPTO0834 (from the *bim* cluster) in a regulon under control of HrpL found the PSPTO0834 mutant to be less virulent in *Arabidopsis* plant (Vencato *et al.*, 2006).

However, my inhibitor mutants did not show a consistent and robust difference, with both infiltration and spray infections on both *N. benthamiana* and *A. thaliana* leaves. It might be argued that there seems to be a trend in which the non-inhibitor producing mutant was slightly less virulent, particularly in spray inoculated plants, although the effect was small and not consistently statistically significant. I speculate that this lack of a robust virulent phenotype of inhibitor production could be due to the insufficient level of inhibitor production when inoculating the bacteria or a redundancy of multiple virulence mechanisms, including effector proteins and toxins, suppressing plant defence and masking the effect of the BGAL inhibitor. I attempted to address these by testing a *hrpA* knockout strain unable to secrete effector proteins and preinduction of inhibitor production by incubating the bacterial inoculum in the inducing MG medium before infection. Nevertheless, I still could not observe

a robust difference. Another possibility is that other toxins, such as coronatine that is still produced in the *hrpA* knockout strain and suppresses plant defence, mask the effects of the BGAL inhibitor. This was also shown to be the case in a study on phevamine A, another toxin produced by *P. syringae*, where the phenotype of this toxin production was only observed in the background strain that lacks all effectors and coronatine production (O'Neill *et al.*, 2018). In addition, the phenotype of BGAL inhibitor production may not be as drastic as observed in the BGAL1 knockout plant as inhibitor may not completely inhibit the plant enzyme at all time. Alternatively, the effects of inhibitor might be more pronounce during the later stage of infection as the bacterial population establishes and the inhibitor accumulates at a higher level in the plant tissue.

According to my phylogenetic analysis, BGAL inhibitor production is not restricted to the *P. syringae* strains with modified-viosamine-glycosylated flagellin that is sensitive to plant BGAL, suggesting that additional actions of plant BGAL on different substrates are likely to also be targeted by the inhibitor. Indeed, I found an evidence that the BGAL inhibitor affects the apoplastic protein glycosylation profile, using protein blots probed with a lectin specific for galactose-containing epitope. BGAL inhibitor production by *P. syringae* suppressed BGAL1 activity causing the accumulation galactose on apoplastic glycoproteins. The smeary pattern of the protein signal was likely due to the heterogeneity of complex glycosylation present on the proteins and the high molecular weight observed was likely due to highly glycosylated proteins, such as arabinogalactan proteins (Seifert and Roberts, 2007). In addition, the fact that the lectin galactose binding signal was not observed on all the apoplastic proteins suggests that only a subset of proteins carry glycans with galactose. However, it is likely that this detection approach still missed some glycoproteins that are present in low abundance or

strongly associated to plasma membrane or cell wall, which are not washed out in the apoplastic fluid extract. It would be interesting to identify these glycoproteins and study the effect of the presence and absence of galactose on their glycan.

Apart from the role in plant immunity, the effect of plant BGAL1 on apoplastic glycoproteins was also investigated in the context of recombinant glycoprotein production in *N. benthamiana* (Kriechbaum *et al.*, 2020). BGAL1 can cleave terminal galactose residues from both N- and O-linked glycans. Accordingly, the knockout of BGAL1 increased the level of accumulation of glycoproteins with the desired glycosylation (Kriechbaum *et al.*, 2020). This effect was observed because BGAL1 is the dominant β-galactosidase in the apoplast, although it could not account for all the β-galactosidase activity. Coexpression of the *bim* cluster to produce the BGAL inhibitor from *Agrobacterium* used for agroinfiltration to produce these glycoproteins might give further protection against glycosidases and increase the yield of desired glycoproteins. The use of a small molecule inhibitor could be beneficial since it can simultaneously target multiple enzymes with the same specificity and thus address the potential challenges of redundancy or polyploidy that would require multiple genes to be identified and knocked out.

Another important consideration for the role of BGAL1 is the effects of the presence or absence of galactose on the functions of glycans or glycoproteins. Cleavage of galactose by BGAL1 could have a direct role in removal of galactose (either as a monomer or sequential polymeric chain) or indirect role as an initiator of further hydrolytic processing cascade when a new epitope is exposed after galactose removal. For example, there could be multiple glycosidases in the apoplast that could act sequentially to cleave the different types of monosaccharide residues on a glycan chain. Alternatively, removal of glycans can stop its

protective effect on the glycoprotein against cleavage by proteases, as exemplified by the role of glycosylation on flagellin or a pathogen effector protein against plant defence protease (Buscaill *et al.*, 2019; Xia *et al.*, 2020). Overall, glycosylation can impact various aspects of glycoprotein functions such as structure, stability, activation or binding (Varki, 2017; Lin *et al.*, 2020).

Next, regarding possible targets of BGAL1, galactose can be found in various sites in the plant apoplast, including glycoproteins and cell wall associated polysaccharides. Glycoproteins can be glycosylated either at asparagine residues (N-linked) or serine and threonine or hydroxyproline residues (O-linked) (Strasser, 2016). For N-linked glycosylation, a terminal galactose residue is found in the Lewis a glycan (Strasser, 2016). The biological roles of this glycan are unclear. In *Arabidopsis*, no growth or developmental phenotype was observed in mutant and recombinant lines lacking or overproducing this glycan type (Strasser *et al.*, 2007). Nevertheless, in general, N-linked glycosylation can have multiple effects on glycoproteins so one might speculate possible roles of altered glycan related to the functions of glycoproteins. For example, N-linked glycosylation was shown to be required for pattern recognition receptors in plant immunity (Häweker *et al.*, 2010) and glycosylation may also affect activity of proteases (Goettig, 2016).

For O-linked glycosylation, the addition of single galactose can be found on serine in some proteins (Strasser, 2016). Moreover, galactose is also commonly found in the glycan on hydroxyproline-rich glycoprotein (HRGP) family, an abundant group of cell wall proteins, including highly glycosylated arabinogalactan proteins (AGPs), moderately glycosylated extensins and other hydroxy/proline rich proteins (Malinovsky, Fangel and Willats, 2014). Proteins in this group have been implicated in various processes such as development, growth

and signaling (Seifert and Roberts, 2007; Ellis *et al.*, 2010; Malinovsky, Fangel and Willats, 2014). Interestingly, there are also examples related to plant-microbe interactions. HRGP can be involved in crosslinking and strengthening of the cell wall and are found in cell wall papillae that are formed in response to pathogens (Deepak *et al.*, 2010; Underwood, 2012). HRGP expression was also found to be suppressed by *P. syringae* AvrPto effector in *Arabidopsis* suggesting possible role in extracellular defense (Hauck, Thilmony and He, 2003). Another study reported that an AGP protein in *Arabidopsis* was required for efficient *Agrobacterium* transformation suggesting possible role in either bacterial binding or plant defence signaling (Gaspar *et al.*, 2004).

The plant cell wall consists of a complex mixture of polysaccharides, which include cellulose microfibrils cross-linked with hemicelluloses and embedded in a pectin matrix (Vorwerk, Somerville and Somerville, 2004; Malinovsky, Fangel and Willats, 2014). Hemicellulose xyloglucan contains galactose residues in the side chains that are important for mechanical strength of cell wall and related physiological functions. Abnormal xyloglucan structure depleted for galactose in mutants of *Arabidopsis* β-galactosidase or galactosyltransferase resulted in growth and developmental defects (Peña *et al.*, 2004; Sampedro *et al.*, 2012; Kong *et al.*, 2015). In addition, some pectic polysaccharides such as rhamnogalacturonan also contain galactose or its polymer galactan side chains (Malinovsky, Fangel and Willats, 2014). During infection, enzymatic remodeling of cell wall occurs and the altered composition of the cell wall can affect plant susceptibility and resistance (Vorwerk, Somerville and Somerville, 2004; Underwood, 2012; Bellincampi, Cervone and Lionetti, 2014; Malinovsky, Fangel and Willats, 2014; Bacete *et al.*, 2018). In addition, molecules from the cell wall are released enzymatically to serve as damage-associated molecular patterns (DAMPs) that elicit plant

defence responses (Malinovsky, Fangel and Willats, 2014). The action of a glycosidase inhibitor might interfere with these aspects of plant defence. However, specific roles of galactose in cell wall polysaccharides during infection are not well characterised and hence the possible role of BGAL and its inhibitor are still unclear. Some plant β-galactosidases are implicated in cell wall loosening during fruit ripening so the inhibitor could also modulate modifications of the cell wall.

In addition, there could also be a role for the BGAL inhibitor during infection in other tissues or developmental stages of the plant or with other factors such as interactions with other microbes or insects. These conditions were not tested in my infection assay. For example, multiple co-infecting pathogens can compete and exclude one another or cooperate to promote their colonisation in plant tissues (Abdullah *et al.*, 2017). Production of secondary metabolites have also been studied in complex interactions within a microbiome (Carrión *et al.*, 2019). For competition, the inhibitor might act as an antibiotic or inhibit glycosidases used by another microbe to gain nutrients. For cooperation, the inhibitor might provide protective effect on other microbes, which, together help colonise the plant. Alternatively, in the interaction with insects, iminosugar inhibitors are produced by some plants to function in defence against herbivores by inhibiting their digestive enzymes (Stiitz, 1999). There are also studies exploring the inhibitor produced by fungal endophytes for plant pest resistance (Singh *et al.*, 2014) and the use of iminosugars as plant pest control (Ya'kobovitz, Butters and Cohen, 2016). Similarly, the inhibitor produced by *P. syringae* during infection might function to deter insect herbivores from the infected tissue. On the other hand, some *P. syringae* can also be transmitted by insects such as aphids (Stavrinides, McCloskey and Ochman, 2009), so inhibitor production might have a role inside the insect.

3.4.5 Conclusion

I have identified the *bim* cluster as the biosynthetic gene cluster of the BGAL inhibitor. I found that only PSPTO0834 and PSPTO0835 from the cluster were sufficient to initiate inhibitor production in both *P. syringae* and *E. coli*. In addition, I confirmed a role of NadB, which is likely involved in generation of the NAD cofactors required for oxidoreductases PSPTO0834-5. Overall, the biosynthetic pathway of BGAL inhibitor is likely to be similar to the framework of previously described iminosugars. Next, I investigated the evolution of the *bim* cluster and found putative iminosugar biosynthetic gene clusters homologs in other bacterial species. The *bim* cluster was likely acquired by horizontal gene transfer into *P. syringae* and subsequently transferred among different strains to become widespread. Furthermore, I also characterised the regulation of inhibitor production in *P. syringae* by the conventional type III secretion system (T3SS) regulators (RhpS, HrpR, HrpS, HrpL) and identified HtpX as a new potential regulator of inhibitor production and T3SS. Finally, I explored the role of the BGAL inhibitor in bacterial virulence during plant infection. Although no robust virulence phenotype was observed, I have discussed other possibilities that can be explored in future work and have identified additional potential targets of the inhibitor such as the modification of apoplastic glycoproteins.

Chapter 4: The characterisation of the BGAL inhibitor

4.1 Introduction

From the discovery of BGAL inhibitor production by *P. syringae*, it is of high interest to identify the molecular identity of this inhibitor as it could be a novel natural product. In addition, the optimised methods for production and purification of the inhibitor will also be useful for its characterisation.

Previous study discovered that *P. syringae* could produce the inhibitor in mannitol glutamate (MG) medium, which induces the type III expression system (Buscaill *et al.*, 2019). Further characterisation showed that the inhibitor was a heat stable, polar and basic small molecule (Buscaill *et al.*, 2019). The inhibitor could also be partially purified by cation exchange resin and elution with NH_4OH solution. These properties are consistent with the properties of iminosugars, sugar analogues with an N atom replacing the O atom in the ring (Molyneux, 1993). Although the identity of the inhibitor remains unknown, these preliminary data serve as useful guides for further characterisation of the inhibitor.

In the previous chapter, I identified the *bim* biosynthetic gene cluster of the inhibitor and demonstrated that *bim* gene expression was regulated by type III secretion system regulators. These together provided clues for further work to characterise the inhibitor. Knowledge of the regulation of *bim* gene expression can be used during the optimisation of culture conditions for inhibitor production. In addition, the knowledge of the *bim* cluster can be used for genetic manipulation of bacterial strains, by both deletion and overexpression in *P. syringae* and heterologous expression in *E. coli*. Together, they serve as useful tools for comparative analysis approach for the detection and identification of the inhibitor. In

addition, one of the essential genes in the *bim* cluster, a putative alcohol dehydrogenase PSPTO0834, is a homolog of GutB1, an enzyme described for biosynthesis of an iminosugar DNJ, supporting the hypothesis that BGAL inhibitor is an iminosugar (Clark, Johnson and Horenstein, 2011).

This chapter investigates the production, purification, characterisation and identification of the inhibitor. I seek to optimise the production of the inhibitor both in terms of culture conditions and genetic manipulation in order to produce sufficient amount of the inhibitor for further analysis. Then, I aim to further optimise the purification method to obtain the inhibitor and analytical chemistry methods to detect and characterise the inhibitor in order to ultimately elucidate its molecular identity.

4.2 Materials and methods

4.2.1 Bacterial strains, culture conditions and media

Bacterial strains used in this study are listed in Table 4.1. Cultures were incubated at 28°C for *P. syringae* and 37°C for *E. coli*, with shaking at 220 rpm for liquid cultures. Media used for bacterial cultures were made as follows:

- LB medium: 10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl
- KB medium: 10 g/l peptone, 1.5 g/l KH_2PO_4, 5 mM $MgSO_4$, 1% (v/v) glycerol
- Mannitol glutamate (MG) medium (Bronstein *et al.*, 2008): 10 g/l mannitol, 2 g/l L-glutamic acid, 0.5 g/l KH_2PO_4, 0.2 g/l NaCl, 0.2 g/l $MgSO_4.7H_2O$, pH 5.5 or 7. Where applicable, 50 µM Iron citrate was added.
- M9 medium: 6 g/l Na_2HPO_4, 3 g/l KH_2PO_4, 0.5 g/l NaCl, 1 g/l NH_4Cl, 0.1 mM $CaCl_2$, 2 mM $MgSO_4$, 22 mM sugar (glucose or fructose), pH 5.5 or 7
- HMM medium: 6.6 mM K_2HPO_4, 43.4 mM KH_2PO_4, 3.8 mM $(NH_4)_2SO_4$, 0.85 mM $MgCl_2$, 0.85 mM NaCl, 10 mM fructose, pH 5.5
- Antibiotics were generally used at the concentrations of 25 µg/ml rifampicin, 12.5 µg/ml gentamicin, 25 µg/ml kanamycin, 10 µg/ml tetracycline.

Table 4.1. Bacterial strains used in this study

Strain	Description	Antibiotics	Reference
HQ	*P. syringae* pv. *tomato* DC3000 Δ*hopQ1*, the lack of the effector *hopQ1* allows it to cause disease in the model plant *N. benthamiana*	Rifampicin	(Wei *et al.*, 2007)
HQ Δ*bim*	*P. syringae* DC3000 Δ*hopQ1* lacking the *bim* biosynthetic gene cluster for the inhibitor	Rifampicin	This study Chapter 2
HQ (p_{lac}::*bim*)	*P. syringae* DC3000 Δ*hopQ1* carrying pBBR1MCS3 with partial *bim* biosynthetic gene cluster (PSPTO0834-5) under *lac* promoter	Rifampicin, Tetracycline	This study Chapter 2

HQ (p_{native}::*bim*)	*P. syringae* DC3000 Δ*hopQ1* carrying pBBR1MCS2 with partial *bim* biosynthetic gene cluster (PSPTO0834-5) under native promoter	Rifampicin, Kanamycin	This study Chapter 2
E. coli EV	*E. coli* Top10 carrying pBBR1MCS3 empty vector	Tetracycline	This study Chapter 2
E. coli (p_{lac}::bim^{wt})	*E. coli* Top10 carrying pBBR1MCS3 with partial *bim* biosynthetic gene cluster (PSPTO0834-5) under *lac* promoter	Tetracycline	This study Chapter 2
E. coli (p_{lac}::bim^{mut})	*E. coli* Top10 carrying pBBR1MCS3 with catalytically inactive mutant partial *bim* biosynthetic gene cluster (PSPTO0834-5) under *lac* promoter	Tetracycline	This study Chapter 2
DC3000 pBS63	*P. syringae* pv. *tomato* DC3000 carrying pBS63 plasmid containing p_{hrpL}*::lux* construct as a luminescent reporter of *hrpL* promoter activity	Rifampicin, Kanamycin	(McCraw *et al.*, 2016)

4.2.2 *hrpL* promoter activity luminescence reporter assay

Bacteria were grown overnight in LB medium. The cells were collected by centrifuging at 4,000 x g and washed with sterile water before being inoculated at the OD600 of 0.05 in 150 µl of appropriate media in each well of 96 well plate. OD600 and luminescence readings were taken every 30 minutes using the plate reader Infinite M200 (Tecan, Mannedorf, Switzerland). Plates were incubated at 28°C with shaking during the course of the experiment. The ratio of luminescence to OD600 was used as the normalized activity of the luminescence reporter.

4.2.3 Apoplastic fluid extraction

N. benthamiana leaves (~4 weeks old) were rolled into a 50 ml syringe filled with ice-cold water. The leaves were infiltrated by pulling the plunger to apply negative pressure drawing air out of the leaf tissue then pushing the plunger to apply positive pressure forcing water into the leaf tissue, while keeping the syringe tip sealed. The infiltrated leaves were blotted dry and rolled into a 20 ml syringe barrel placed inside a 50 ml conical centrifuge tube. Apoplastic fluid was collected from the leaves by centrifugation at 900 x g for 15 minutes at 4°C. The extracted apoplastic fluid was used directly or stored at -20°C.

4.2.4 FDG assay for β-galactosidase activity

Substrate preparation: A stock FDG solution of 2 mM was made by dissolving FDG (Fluorescein di-β-D-galactopyranoside) (Marker Gene Technologies, Eugene, USA, M0250) in 1:1:8 ethanol:DMSO:H_2O (first dissolved in 1:1 ethanol:DMSO, then diluted in H_2O). Aliquots were stored at -20°C. The stock was diluted to 2 µM in MilliQ water just before use.

Enzyme preparation: A stock LacZ solution of 1 U/µl was made by dissolving lyophilised β-galactosidase enzyme (Sigma, St. Louis, USA, G6008) in suspension buffer (50% (v/v) glycerol, 5mM Tris, 5 mM $MgCl_2$, 0.5 mM DTT, 0.5 mM β-mercaptoethanol, pH 7.4). Aliquots were stored at -20°C. The stock was diluted to 1 mU/µl in water just before use.

Apoplastic fluid BGAL assay: The assay was set up in 96 well plate with the total reaction volume of 100 µl per well with 0.2 µM FDG and 50 mM MES pH 5. This typically consists of 40 µl apoplastic fluid sample, 10 µl FDG (2 µM), 10 µl MES buffer pH 5 (500 mM) and 40 µl H_2O. For inhibition assays, 40 µl inhibitor sample was used instead of H_2O.

LacZ assay: The assay was set up in 96 well plate with the total reaction volume of 100 µl per well with 60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 0.2 µM FDG and 0.1 mU/µl LacZ. This typically consists of 20 µl 5x Z buffer (300 mM Na_2HPO_4, 200 mM NaH_2PO_4, 50 mM KCl, 5 mM $MgSO_4$, pH 7), 10 µl FDG (2 µM), 10 µl LacZ (1 mU/µl) and 60 µl H_2O. For inhibition assays, 50 µl inhibitor sample was used instead of H_2O.

β-galactosidase activity in the sample was monitored by measuring fluorescence signal from fluorescein, a product of FDG hydrolysis catalysed by β-galactosidase, every minute at 25°C using the plate reader Infinite M200 (Tecan, Mannedorf, Switzerland) at an excitation

wavelength of 485 nm and emission wavelength of 535 nm. The rate of change of fluorescence over time was calculated for β-galactosidase activity. For comparisons of inhibition, the enzymatic activity was expressed as percentage of the activity compared to no inhibitor control (100%). Note that the fluorescence intensity of fluorescein is pH dependent. The same amount of FDG yields less fluorescence at pH 5 than pH 7.

4.2.5 Relative quantification of inhibitor concentration

Inhibition assay was done with LacZ as described above using two-fold dilution series of the inhibitor sample to be quantified and no inhibitor control. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control (100%) and plotted against the inhibitor dilution. The dose response curve was fitted with log-logistic model using drc package in R (Ritz *et al.*, 2015) to calculate the dilution of the inhibitor at IC50 (the concentration of inhibitor required for the enzymatic activity to be inhibited by 50%). The concentration of inhibitor in the sample can then be relatively quantified by comparing the difference in dilution at IC50 of each sample. The more dilutions required to reach IC50, the higher concentration of the inhibitor in the sample.

4.2.6 BGAL inhibitor production *in vitro*

Bacteria were grown overnight in LB medium. The cells were collected by centrifuging at 4,000 x g for 5 minutes then washed once with sterile water and centrifuged again. The cell pellet was then resuspended in appropriate media to the OD600 of 1 and grown overnight at 28°C (both *E. coli* and *P. syringae*) with 220 rpm shaking. The bacterial culture was then centrifuged at 4,000 x g and the supernatant containing the inhibitor was collected.

4.2.7 Inhibitor purification by cation exchange resin

Bacterial culture supernatant was mixed with 50% volume methanol and 20% volume chloroform then centrifuged at 10,000 x g for 5 minutes for phase separation of polar metabolites, macromolecules (e.g. proteins) and non-polar compounds. The hydrophilic water-methanol top fraction containing the inhibitor was collected. Amberlyst 15 hydrogen form cation exchange resin was loaded into a column and washed with 2.5 M HCl (15x bed volume) followed by water (15x bed volume). Supernatant containing the inhibitor were passed through the column to allow the inhibitor to bind to the resin. Then, the resin was washed with 70% (v/v) ethanol (15x bed volume) followed by water (15x bed volume). The inhibitor was eluted with 2.5 M HCl (10x bed volume). Elution with NH_4OH was used at 1 M. The eluate was freeze dried then resuspended in water.

4.2.8 Inhibitor desalting

Inhibitor sample, dissolved in 100 µl water, was mixed with 900 µl concentrated HCl (12 M) and incubated at room temperature for 15 minutes to allow salts to precipitate out of the solution. The mixture was then centrifuged at 12,000 x g for 1 minute to pellet the salts and the supernatant containing the inhibitor was collected. The supernatant was neutralised by diluting to 10 ml with water before passing through Amberlite IRA96 free base anion exchange resin (3 ml bed volume) followed by a wash with 4 ml water. The flowthrough was collected, freeze dried then resuspended in water.

4.2.9 Preparative HPLC (High Performance Liquid Chromatography)

HPLC was carried out using Prominence UFLC HPLC system (Shimadzu, Kyoto, Japan) equipped with LC-20AR semi-preparative pump, CBM-20A system controller, Sil-10AP autosampler, FRC-10A fraction collector, SPD-20A UV detector (220 nm) and DGU-20A3R degassing unit.

RP-HPLC (Reversed-Phase HPLC): RP-HPLC was carried out using SunFire C18 OBD Prep column (100 Å pore size, 5 µm particle size, 10 mm inner diameter, 150 mm length, part no. 186002563) (Waters, Milford, USA). The sample was dissolved in water and filtered with 0.45 µm syringe filter before 100 µl was injected into the HPLC system. The HPLC was carried out in reversed-phase mode using a linear gradient of solvent B (0.1% (v/v) formic acid in acetonitrile) in solvent A (0.1% (v/v) formic acid in water) with a flow rate of 13 ml/min. The gradient was from 2% to 98% B over 60 minutes. Fractions of 13 ml were collected every minute. Fractions were freeze dried then resuspended in 400 µl water.

HILIC HPLC (Hydrophilic Interaction Liquid Chromatography HPLC): HILIC HPLC was carried out using XBridge BEH Amide OBD Prep column (130 Å pore size, 5 µm particle size, 10 mm inner diameter, 250 mm length, 20 ml column volume, part no. 186008262) with XBridge BEH Amide Prep Guard cartridge (130 Å pore size, 5 µm particle size, 10 mm inner diameter, 10 mm length, part no. 186006597) (Waters, Milford, USA). The column was washed with 60% (v/v) acetonitrile (50x column volume) then equilibrated with 95% (v/v) acetonitrile (20x column volume). The sample was dissolved in 50% (v/v) acetonitrile and filtered with 0.2 µm syringe filter before 200 µl was injected into the HPLC system. The HPLC was carried out in HILIC mode using a linear gradient of solvent B (0.1% (v/v) formic acid in acetonitrile) in solvent A (0.1% (v/v) formic acid in water) with a flow rate of 13 ml/min. The gradient started

at 95% B for 5 minutes then from 95% to 50% B over 30 minutes then maintained at 50% for 5 minutes then from 50% to 95% B over 2.5 minutes followed by 2.5 minutes at 95% B. Fractions of 6.5 ml were collected every minute. Fractions were freeze dried then resuspended in 400 µl water.

4.2.10 LC-MS (Liquid Chromatography – Mass Spectrometry)

Sample preparation (Fmoc derivatisation): Samples (100 µl) were mixed with 100 µl potassium borate buffer (0.4 M, pH 9.5) followed by 200 µl Fmoc-Cl (25 mM in acetonitrile) (9-Fluorenylmethoxycarbonyl chloride, Sigma). The mixture was incubated for 30 minutes at room temperature then the reaction was stopped with the addition of 4 µl acetic acid and filtered through 3 kDa cutoff filter (Amicon Ultra, Millipore). For use in enzyme inhibition assays, the reaction was quenched with 0.1 M glycine before addition of acetic acid.

LC-MS: LC-MS analysis was conducted using Micromass LCT Premier Mass Spectrometer (Waters, Milford, USA) equipped with ACE Equivalence 3 C18 column (50 x 2.1 mm) (Avantor, Radnor, USA). Liquid chromatography was carried out in reversed-phase mode using a linear gradient of solvent B (0.1% (v/v) trifluoroacetic acid in acetonitrile) in solvent A (0.1% (v/v) trifluoroacetic acid in water) with a flow rate of 1 ml/min. The gradient was from 3% to 98% B over 7 minutes then from 98% to 3% B over 2 minutes followed by 1 minute at 3% B. UV absorption was detected at the wavelength of 254 nm. Electrospray ionisation was used for mass spectrometry. Data were analysed and visualised using Mnova software (Mestrelab Research)

4.2.11 GC-MS (Gas Chromatography – Mass Spectrometry)

Sample preparation (trimethylsilylation): Samples (100 µl) were freeze-dried and then resuspended in 40 µl pyridine. Derivatisation of analyte samples was done with 60 µl MSTFA (N-Methyl-N-trimethylsilyl trifluoroacetamide, TS48910, Thermo Scientific, Waltham, USA) at 60°C for 1 hr. The mixture was then centrifuged for 5 minutes at 13,000 x g to separate insoluble precipitates and the supernatant was transferred to a glass vial for autosampler loading to GC-MS machine.

GC-MS: GC-MS analysis was conducted using Intuvo 9000 GC system with 5977B single quadrupole MSD system and 7693 autosampler (Agilent, Santa Clara, USA). The GC system was equipped with a guard chip and two Agilent 122-5512UI-INT columns connected sequentially. The sample (0.5 µl) was injected in splitless mode at 230°C. Helium was used as carrier gas and the flow was set constant at 1 ml/min through the first column and 1.2 ml/min through the second column. The oven temperature was initially held at 70°C for 5 min then increased to 330°C at the rate of 7.5°C/min (for *E. coli* samples) or increased to 300°C at the rate of 7.5°C/min then to 330°C at the rate of 50°C/min (for *P. syringae* samples). The transfer line temperature was 250 °C. The sample was ionised by electron impact ionisation at 70 eV. The MS source temperature was 230°C and the MS quadrupole temperature was 150°C. The mass spectrometer was autotuned before running each batch of samples. The acquisition of mass spectra was in scan mode for ions from 50 to 600 mass units with the scan speed of n=2. Solvent delay of 7.5 min was set before acquisition. Data were analysed and visualised using MassHunter qualitative analysis software (Agilent, Santa Clara, USA).

4.2.12 Enzymatic assay for different glycosidase activities

Substrate preparation: Substrates (table 4.2) were dissolved in DMF (Dimethylformamide) at the stock concentration of 10 mg/ml. The substrates were diluted in water to the concentration of 10 µg/ml before use in the assay.

Enzyme preparation: Enzymes (table 4.3) if supplied as lyophilised powder were dissolved in resuspension buffer (50 % (v/v) glycerol, MES 100 mM pH 5.5 or 6.5). Enzymes supplied as ammonium sulphate suspension were used as supplied. The stock enzyme solutions were diluted in water to different concentrations depending on the enzyme before use in the assay.

Apoplastic fluid assay: The assay was set up in a 96 well plate with the total reaction volume of 100 µl per well with 1 µg/ml substrate and 100 mM MES pH 5.5. This typically consists of 70 µl apoplastic fluid sample, 10 µl substrate (10 µg/ml), 10 µl MES pH 5.5 (1 M) and 10 µl H_2O. For inhibition assay, 10 µl inhibitor sample was used in place of H_2O.

Enzyme standard assay: The assay was set up in a 96 well plate with the total reaction volume of 100 µl per well with 1 µg/ml substrate and 100 mM MES pH 5.5 or 6.5. This typically consists of 10 µl enzyme (various concentrations Table 4.3), 10 µl substrate (10 µg/ml), 10 µl MES pH 5.5 or 6.5 (1 M) and 70 µl H_2O. For inhibition assay, 10 µl inhibitor sample was used in place of H_2O.

Glycosidase activity in the sample was monitored by measuring fluorescence signal from 4-MU (4-methylumbelliferone), a product of 4-MU-monosaccharide conjugate cleavage catalysed by each corresponding glycosidase, every minute at 25°C using the plate reader Infinite M200 (Tecan, Mannedorf, Switzerland) at an excitation wavelength of 360 nm and

emission wavelength of 460 nm. The rate of change of fluorescence over time was calculated for glycosidase activity. For comparisons of inhibition, the enzymatic activity was expressed as percentage of the activity compared to no-inhibitor-control (100%). Note that the fluorescence intensity of 4-MU is pH dependent. The same amount yields less fluorescence at pH 5.5 than pH 6.5.

Table 4.2. Substrates used for glycosidase activity assay

Substrate	Sigma Cat. No.
4-Methylumbelliferyl α-D-galactopyranoside	m7633
4-Methylumbelliferyl β-D-galactopyranoside	m1633
4-Methylumbelliferyl α-D-glucopyranoside	m9766
4-Methylumbelliferyl β-D-glucopyranoside	m3633
4-Methylumbelliferyl α-D-mannopyranoside	m3657
4-Methylumbelliferyl β-D-mannopyranoside	m0905
4-Methylumbelliferyl α-L-fucopyranoside	m8527
4-Methylumbelliferyl β-D-fucoside	m5510
4-Methylumbelliferyl-β-D-xylopyranoside	m7008
4-Methylumbelliferyl-β-D-glucuronide hydrate	m9130
4-Methylumbelliferyl-N-acetyl-α-D-glucosaminide	474500
4-Methylumbelliferyl N-acetyl-β-D-glucosaminide	m2133
4-Methylumbelliferyl N-acetyl-β-D-galactosaminide	m9659

Table 4.3. Enzymes used for glycosidase activity assay

Enzyme	Sigma Cat.No.	Assay pH	Concentration (mU/µl)
α-Galactosidase from green coffee beans	g8507	6.5	0.006
β-Galactosidase (LacZ) from *Escherichia coli*	g6008	6.5	2
α-Glucosidase from *Saccharomyces cerevisiae*	g5003	6.5	0.005
β-Glucosidase from almonds	49290	5.5	0.09
α-Mannosidase from *Canavalia ensiformis* (Jack bean)	m7257	5.5	0.07
β-Mannosidase from *Helix pomatia*	m9400	5.5	0.012
α-L-Fucosidase from bovine kidney	f5884	5.5	0.08
β-Glucanase mixture (used for xylosidase) from *Trichoderma longibrachiatum*	g4423	5.5	0.9
β-Glucuronidase from *Escherichia coli*	g7646	6.5	25
β-N-Acetylglucosaminidase from *Canavalia ensiformis* (Jack bean)	a2264	5.5	0.12

4.3 Results

4.3.1 Inhibition assay with dilution series of inhibitor can be used for relative quantification of inhibitor concentration in the sample

For optimisation of inhibitor production and purification, it is important to be able to quantify and compare the amount of the inhibitor in the sample. Since the inhibitor identity is unknown and the absolute concentration of the inhibitor cannot be determined, I established a relative quantification method instead. The concept is based on the IC50 (the concentration of inhibitor required for the enzymatic activity to be inhibited by 50%) which is constant for an inhibitor in an assay and can be used as a reference point. The inhibitor sample can be prepared in a serial dilution series and tested in enzymatic inhibition assay to produce a dose response curve. The concentration of inhibitor in the sample can then be relatively quantified by comparing the difference in dilution factor at IC50 of each sample. The more dilutions required to reach IC50, the higher concentration of the inhibitor in the original sample.

To test this approach, I used a commercially available β-galactosidase inhibitor PETG (2-phenylethyl β-D-thiogalactoside). PETG stock solutions of different concentrations were used in two-fold serial dilutions which were then tested in inhibition assays with the LacZ enzyme. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control (100%) and plotted against the inhibitor dilution. The dose response curve fitted with log-logistic model allowed the dilution factor at IC50 to be determined. The most concentrated inhibitor stock required higher dilution to reach the IC50 point while the less concentrated stock required lower dilution to reach IC50 (Figure 4.1). The difference in dilution factors at IC50 was also consistent with the difference in inhibitor concentrations of

different samples. For example, 1M and 500 mM samples showed the difference of 1 log2 unit of dilution which corresponds to 2-fold difference (Figure 4.1). These results validated the dilution series approach for relative quantification of the inhibitor which was used in further experiments to compare inhibitor concentrations in different samples.

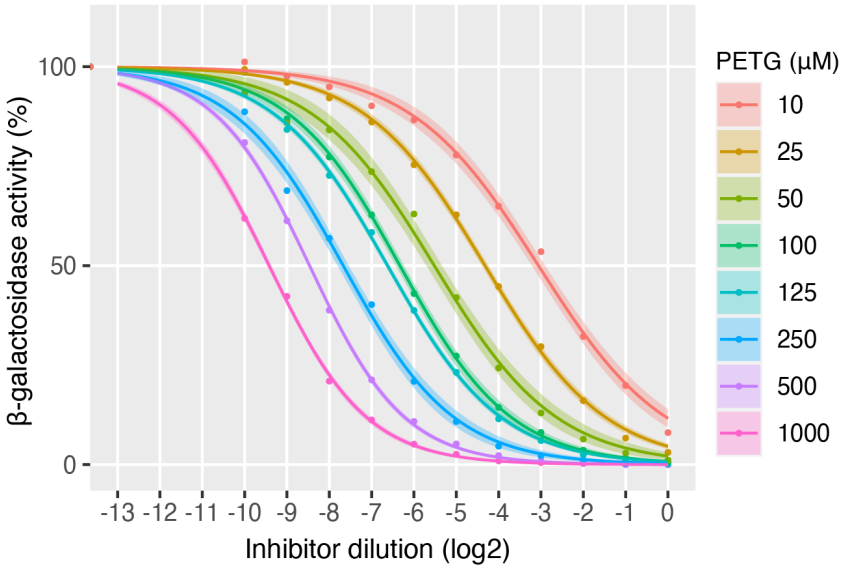


Figure 4.1. Inhibition assay with inhibitor dilution series can be used for relative quantification of inhibitor concentration. β-galactosidase inhibitor PETG at different concentrations was used in two-fold dilution series which were tested in inhibition assay with FDG and LacZ enzyme. The enzymatic activity was expressed as a percentage of the activity compared to no inhibitor control and plotted against the inhibitor dilution (log2 scale). Dose response curves (with 95% confidence interval) were fitted with log-logistic model to estimate the dilution of the sample required to reach IC50 (the inhibitor concentration required for 50% inhibition of enzymatic activity). The fold difference in dilution required to reach IC50 reflects the difference in inhibitor concentrations in the samples.

4.3.2 Inhibitor production in *P. syringae* increased in type III secretion system inducing media

To produce sufficient amount of the inhibitor for further inhibitor characterisation experiments, I needed to maximise the production of the inhibitor from the bacteria. I first investigated different bacterial culture media that could lead to a high amount of inhibitor production. Since the expression of the *bim* biosynthetic gene cluster for the inhibitor is

regulated by the type III secretion system regulator HrpL, a range of media were tested for *hrpL* induction and inhibitor production. To monitor the level of *hrpL* induction, *P. syringae* DC3000 pBS63 strain containing the lux cassette under the control of the *hrpL* promoter (McCraw *et al.*, 2016) was used to report luminescence signal from *hrpL* activity. The bacterial culture was grown with continuous luminescence monitoring and then the supernatant was collected and tested for inhibitory activity in enzymatic assay. In rich media such as LB and KB, there was no induction of *hrpL* (Figure 4.2A) and the inhibitor was not detected (Figure 4.2B). In contrast, other minimal media: MG, M9 and HMM, induced the expression of *hrpL* (Figure 4.2A) and inhibitor production was detected (Figure 4.2B). Since the addition of iron in MG medium was reported to increase the induction of the type III secretion system expression, this was also tested and showed the consistent upregulation of *hrpL* (Figure 4.2A) and higher inhibitor production with iron supplementation (Figure 4.2B). To further compare the amount of inhibitor produced in different inducing media, dilution series of the inhibitor supernatant were tested in inhibition assays. Cultures grown in M9 with fructose pH 5.5 produced the highest level of inhibitor production (Figure 4.2C). Overall, I have tested media that maximise the induction of type III secretion system and inhibitor production. M9 medium with fructose at pH 5.5 can be used for inhibitor production by *P. syringae* in further experiments.

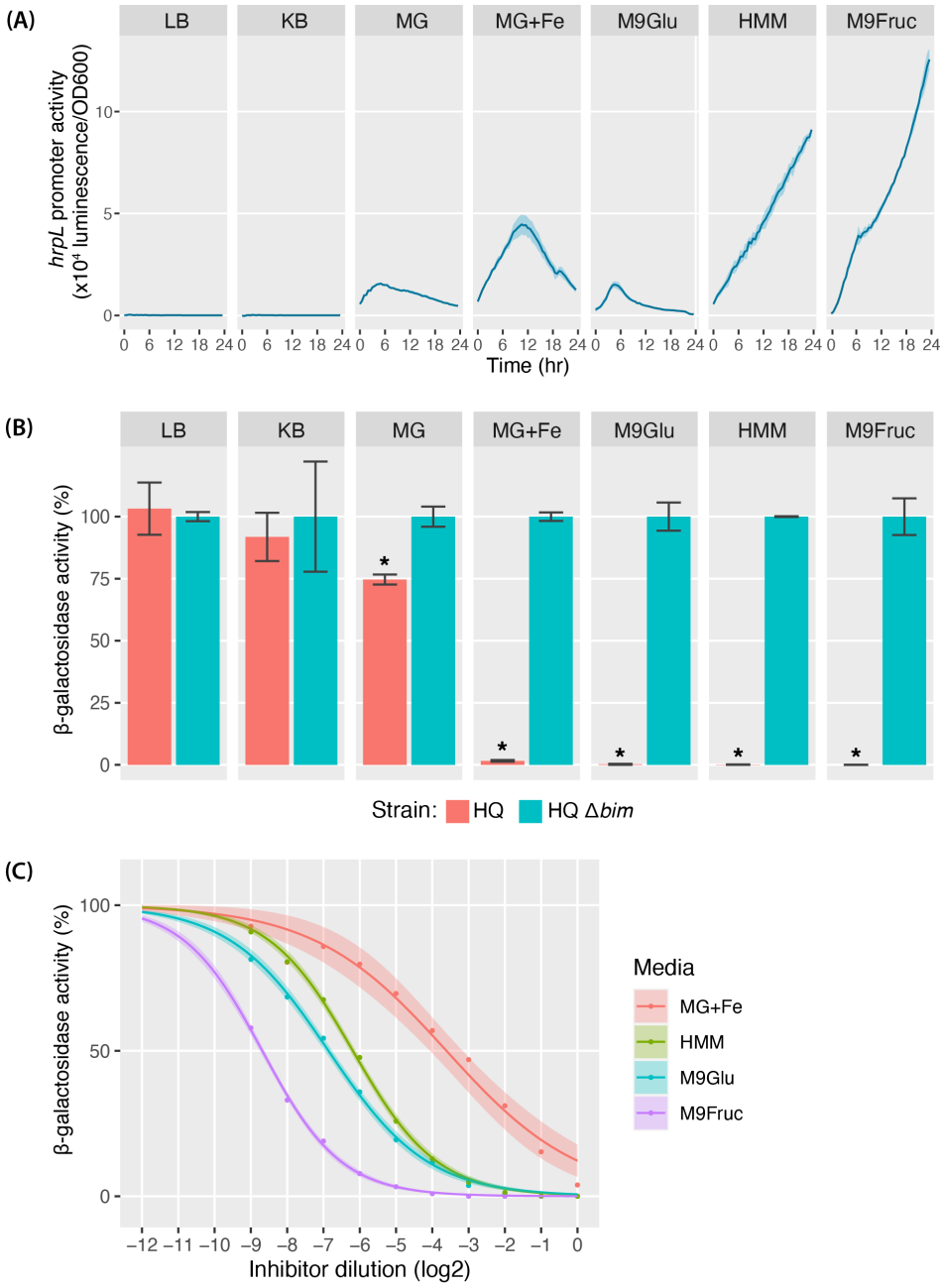


Figure 4.2. Inhibitor production in *P. syringae* increased in type III secretion system inducing media. (A) Type III secretion system induction in different media. *P. syringae* DC3000 pBS63 strain (carrying *hrpL* promoter with lux luminescence reporter construct) was inoculated in different media and measured for luminescence and OD600 over time. The reporter activity was expressed as luminescence signal normalised with OD600. Shaded areas show standard deviation (n=3) **(B) Inhibitor production in different media.** *P. syringae* HQ (inhibitor producing) and HQ Δ*bim* mutant (non-inhibitor-producing) were grown in different media overnight and supernatants were collected for inhibition assay with FDG and LacZ enzyme. Enzyme activity was expressed as percentage of the activity compared to no inhibitor control. Asterisks denote p-value < 0.05 from t-test comparing HQ and HQ Δ*bim* in each condition. Mean and standard deviation are shown (n=3) **(C) Relative quantification of inhibitor production in different media.** Supernatants from *P. syringae* HQ cultures in different media were used in inhibition assay with dilution series to compare level of inhibitor production in each condition. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control and plotted against the inhibitor dilution (log2 scale). Dose response curves (with 95% confidence interval) were fitted with log-logistic model to estimate the dilution of the sample required to reach IC50 (the inhibitor concentration required for 50% inhibition of enzymatic activity). The fold difference in dilution required to reach IC50 reflects the difference in inhibitor concentrations in the samples.

4.3.3 Overexpression of *bim* cluster increased inhibitor production in *P. syringae*

In addition to culture media, genetic manipulation of the strain can also be investigated for the improvement of inhibitor production. Since the *bim* biosynthetic gene cluster responsible for inhibitor production was identified, I tested whether its overexpression can be used to increase inhibitor production in *P. syringae*. The *bim* cluster genes (PSPTO0834-5) demonstrated to be essential for inhibitor production were cloned into a plasmid and transformed into *P. syringae* HQ strain. The *bim* genes were placed either under control of the constitutive *lac* promoter (Obranić, Babić and Maravić-Vlahoviček, 2013) or the native promoter of the *bim* cluster (a 400 bp region upstream of the start codon of PSPTO0834, which includes the *hrp* box for HrpL sigma factor binding). The transformant strains were cultured in M9 fructose medium to induce inhibitor production and the culture supernatant was collected to test for inhibitor in enzymatic assay with dilution series for relative quantification. As expected, both strains overexpressing the *bim* cluster showed a higher level of inhibitor production than the wild type strain (Figure 4.3). The native promoter construct led to higher increase in inhibitor production (~10 fold) than *lac* promoter construct (~3 fold). These results indicate that overexpression of the *bim* cluster can be used to further increase inhibitor production by *P. syringae*.

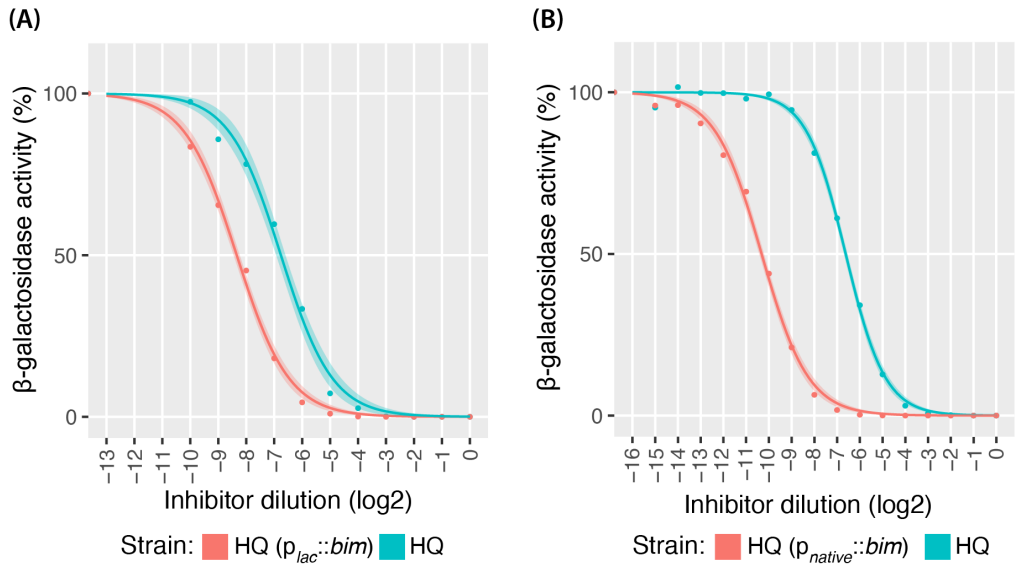


Figure 4.3. Overexpression of *bim* cluster in *P. syringae* increased inhibitor production. Supernatants from cultures of *P. syringae* HQ and HQ transformed with a plasmid carrying *bim* cluster under **(A)** *lac* or **(B)** native promoter were used in inhibition assay with dilution series to compare level of inhibitor production. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control (100%) and plotted against the inhibitor dilution (log2 scale). Dose response curves (with 95% confidence interval) were fitted with log-logistic model to estimate the dilution of the sample required to reach IC50 (the inhibitor concentration required for 50% inhibition of enzymatic activity). The fold difference in dilution required to reach IC50 reflects the difference in inhibitor concentrations in the samples.

4.3.4 *E. coli* strain carrying the *bim* cluster also produced the inhibitor at high level

I next investigated the use of *E. coli* (p*lac*::*bim*^{wt}) strain, carrying a plasmid with the *bim* biosynthetic genes (PSPTO0834-5) for inhibitor production, as an alternative to *P. syringae*. This could be beneficial as inhibitor production under control of the constitutive *lac* promoter is unrestrained by type III secretion system regulation, unlike in *P. syringae*. As a negative control, I included *E. coli* (p*lac*::*bim*^{mut}) strain, carrying catalytically inactive mutant *bim* cluster genes (PSPTO0834-5). In a range of media tested, the wild type *bim* transformant showed detectable inhibitor production at varying levels, while no inhibitor was detected from the mutant *bim* transformant (Figure 4.4A, B). Inhibitor production in LB was the lowest compared to other media, likely due to the peptide-rich nature of the medium with no additional carbon source. M9 based media showed the highest inhibitor production, similar to *P. syringae*. Finally, the level of inhibitor production by *E. coli* is comparable to that of *P. syringae* in M9 fructose medium (Figure 4.4C). These results suggest that *E. coli* can be used as another source of inhibitor production.

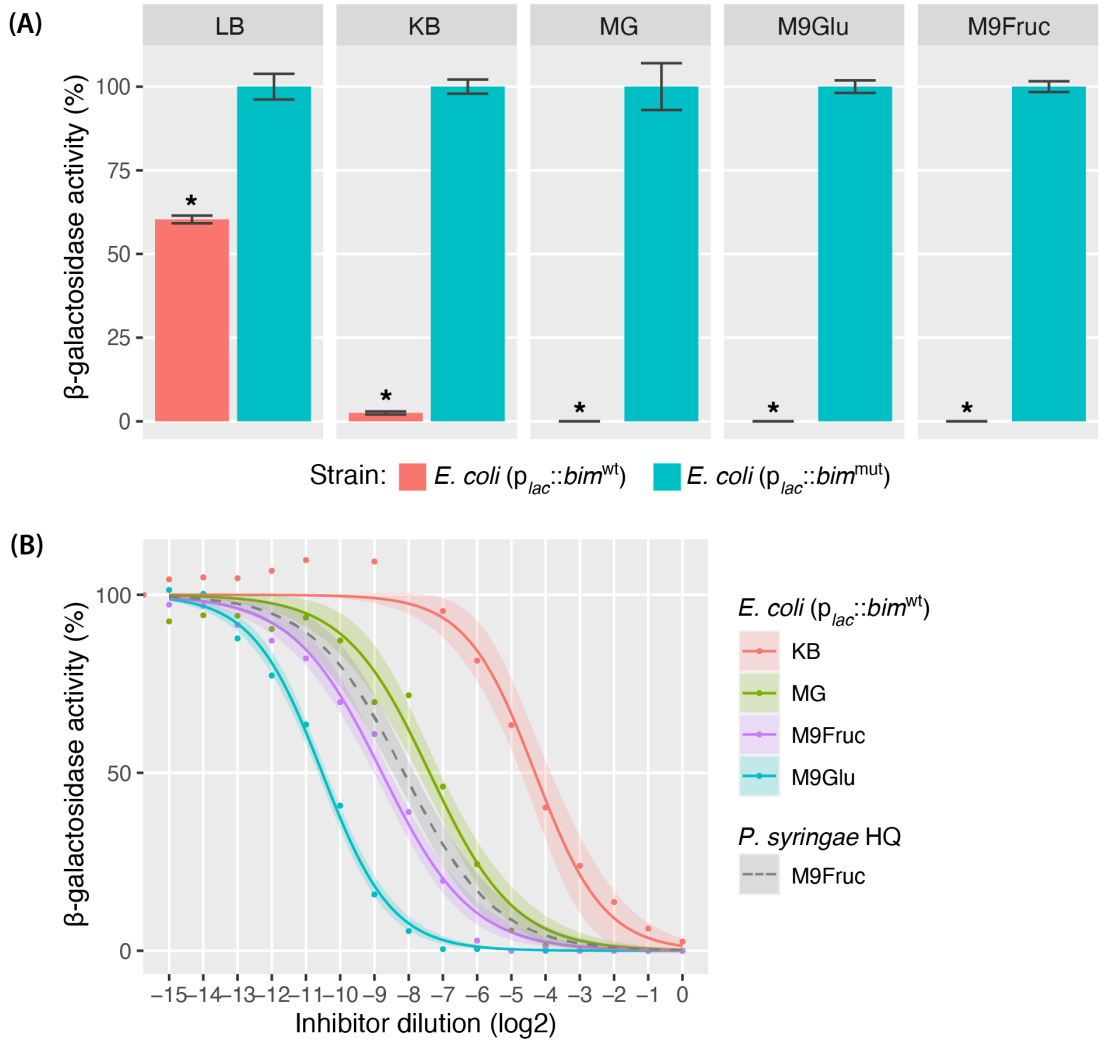


Figure 4.4. *E. coli* carrying *bim* cluster can also produce the inhibitor at high level. (A) Inhibitor production in different media. Inhibitor-producing *E. coli* (p*lac*::*bim*^{wt}) and non-inhibitor-producing *E. coli* (p*lac*::*bim*^{mut}) were grown in different media overnight and supernatants were collected for inhibition assay with FDG and LacZ enzyme. Enzyme activity was expressed as percentage of the activity compared to no inhibitor control. Asterisks denote p-value < 0.05 from t-test comparing wild type and mutant in each condition. Mean and standard deviation are shown (n=3) **(B) Relative quantification of inhibitor production in different media.** Supernatants from cultures of *E. coli* (p*lac*::*bim*^{wt}) and *P. syringae* HQ in different media were used in inhibition assay with dilution series to compare level of inhibitor production in each condition. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control and plotted against the inhibitor dilution (log2 scale). Dose response curves (with 95% confidence interval) were fitted with log-logistic model to estimate the dilution of the sample required to reach IC50 (the inhibitor concentration required for 50% inhibition of enzymatic activity). The fold difference in dilution required to reach IC50 reflects the difference in inhibitor concentrations in the samples.

4.3.5 Cation exchange resin can be used to partially purify the inhibitor

It was previously established that the inhibitor could be isolated with cation exchange resin (Buscaill *et al.*, 2019). Compounds containing amine groups, such as iminosugars, can be protonated to become positively charged and bind to cation exchange resin. Other components in the mixture can be washed away. Then, during elution, the high pH deprotonates the amine group the compound can be eluted from the resin (Amézqueta and Torres, 2016).

I first tested this approach for inhibitor purification. *P. syringae* HQ was cultured in M9 fructose medium to produce the inhibitor and HQ Δ*bim* mutant lacking the inhibitor was used as a negative control. The supernatants were collected and methanol-chloroform extraction to obtain a water-soluble metabolite fraction which was then passed through cation exchange resin to bind the inhibitor. The resin was then washed with water then the inhibitor was eluted with 1M NH_4OH. To assess the inhibitor purification, fractions were collected and tested for inhibitory activity. As expected, the inhibitor was found in the input supernatant from HQ culture, then, once bound to the resin, was largely absent in water wash fraction (Figure 4.5A). Finally, inhibitor was detected in the fraction eluted from the resin with NH_4OH (Figure 4.5A).

To evaluate the composition and assess the purity of the fractions, both the bacterial supernatant input and the cation exchange resin eluate samples were subjected to GC-MS analysis. The input sample showed predominant signals from phosphate and fructose, which were the major components of the culture medium (Figure 4.5B). In contrast, the eluate fraction showed less dominant phosphate and fructose signals but higher enrichment of some

amino acids (Figure 4.5B). This suggests that the purification with cation exchange resin removed enriched amine containing compounds. Nevertheless, I did not identify a distinguishable peak that is present in the wild type but not Δ*bim* mutant sample, which would be indicative of the inhibitor (Fig 5B). Overall, these results suggest that cation exchange resin can be useful for enrichment of the inhibitor.

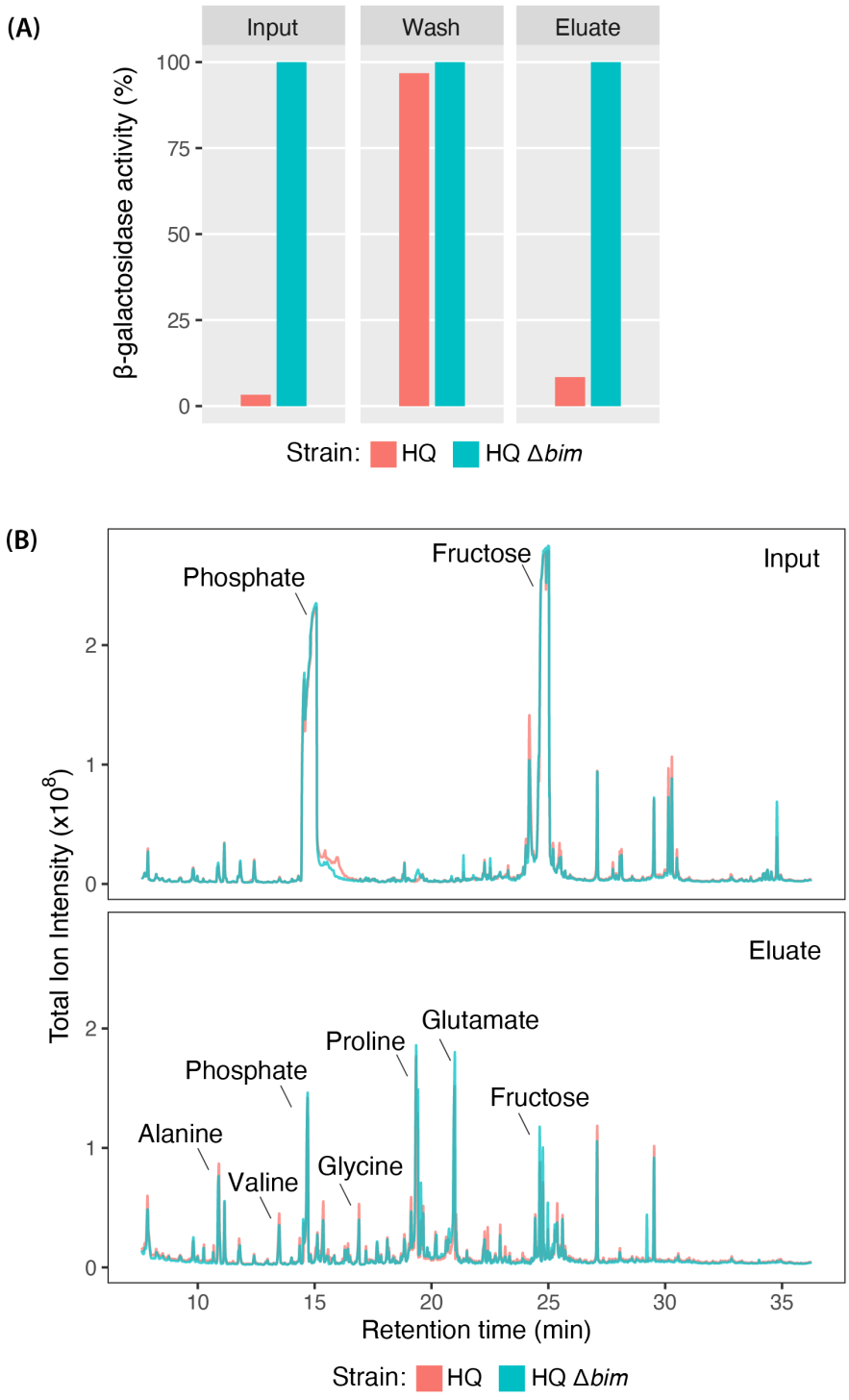


Figure 4.5. The inhibitor could be partially purified from bacterial culture using cation exchange resin. (A) Test of inhibitor in different isolated fractions. Supernatants from cultures of *P. syringae* HQ (inhibitor producing) and HQ Δ*bim* mutant (non-inhibitor-producing) were subjected to methanol-chloroform extraction to obtain a water-soluble metabolite fraction which was then used as input to the cation exchange resin to bind the inhibitor. The resin was then washed with water and the inhibitor was eluted with 1M NH_4OH. Fractions collected from each step were tested in inhibition assay with FDG and LacZ enzyme. Enzyme activity was expressed as percentage of the activity compared to no inhibitor control. One representative purification round is shown. **(B) GC-MS analysis.** The input and eluate fractions from cation exchange purification were dried and derivatised with MSTFA before GC-MS analysis. Total ion chromatogram is shown annotated with dominant detected compounds.

4.3.6 Elution of the inhibitor from cation exchange resin can be done at high pH

Since the eluate from the cation exchange resin with 1M NH_4OH contains other coeluting compounds, I next investigated ways to reduce the presence of coeluting compounds with washing steps. Elution of a compound is dependent on its pKa dissociation constant. Less basic compound with lower pKa can be deprotonated and thus eluted at lower pH. Therefore, differential washing steps with lower pH can remove the weaker basic compounds while retaining the inhibitor for the final elution. For example, washing with pyridine has been reported to remove some amino acids from the resin (Molyneux, 1993).

To test this approach, the inhibitor was bound to cation exchange resin as before and subjected to washing steps with 1M pyridine (pH 9.6) or increasing concentrations of NH_4OH [1 mM (pH 10.1), 10 mM (pH 10.6), 100 mM (pH 11.1), 1 M (pH 11.6)]. The output fractions were then tested for inhibition to determine the inhibitor elution point. The results showed that 1M pyridine could not elute the inhibitor from the resin (Figure 4.6A). Inhibitor elution occurred with 100 mM to 1 M NH_4OH (Figure 4.6B). This indicated that the inhibitor can be eluted at pH higher than 11 and the weaker basic pyridine solution can be used to wash the resin. This also suggests a fairly strong basic property of the inhibitor.

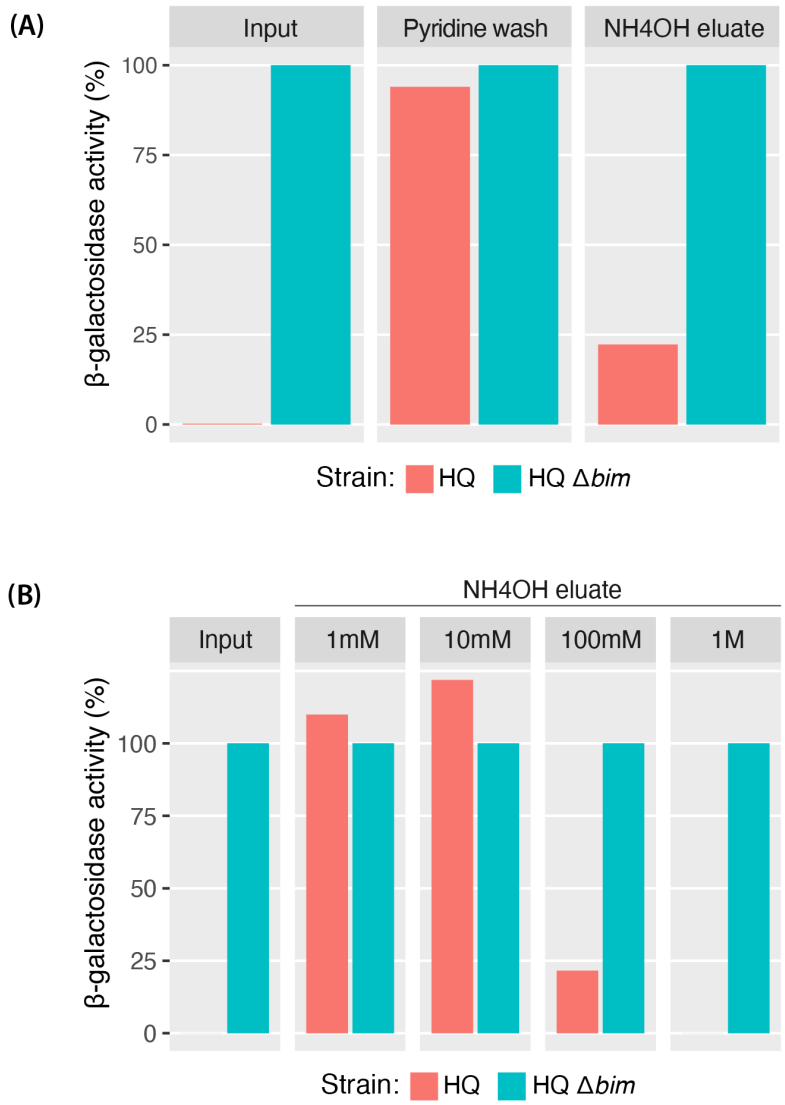


Figure 4.6. Elution of the inhibitor from cation exchange resin can be done with high pH. Supernatants from cultures of *P. syringae* HQ (inhibitor producing) and HQ Δ*bim* mutant (non-inhibitor-producing) were subjected to methanol-chloroform extraction to obtain a water-soluble metabolite fraction which was then used as input to the cation exchange resin to bind the inhibitor. The resin was then washed with water then **(A)** 1 M pyridine or **(B)** increasing concentrations of NH_4OH (1, 10, 100mM) before the final elution of the inhibitor with 1M NH_4OH. Fractions collected from each step were tested for presence of the inhibitor in inhibition assay with FDG and LacZ enzyme. Enzyme activity was expressed as percentage of the activity compared to no inhibitor control. One representative purification round is shown.

4.3.7 The inhibitor is stable in acidic pH but sensitive to basic pH treatment

Since the inhibitor isolation processes tested so far involved high pH treatment, I evaluated whether the inhibitor is stable or sensitive to different pH conditions. The purification protocol can then be optimised for the pH at which the inhibitor is stable.

First, the supernatant of HQ culture containing the inhibitor was subjected to treatments with different bases (1M pyridine, 1M NH_4OH and 1M NaOH) or acids (1M CH_3COOH, 1M HCl) for one hour then neutralised. I next determined whether the inhibitory activity was affected using enzymatic activity assays. The culture with the non-inhibitor-producing HQ Δ*bim* strain was included as a negative control. Treatment with acids, such as CH_3COOH and HCl, did not abolish inhibitory activity (Figure 4.7A). However, treatments with bases, such as pyridine, NH_4OH and NaOH, reduced the inhibitory activity (Figure 4.7A).

Next, I evaluated the stability of the inhibitor at different pH levels by adjusting the pH of the bacterial culture supernatant containing the inhibitor and incubating it overnight before neutralising to pH 7, with either NaOH or HCl, and testing for inhibitory activity. The inhibitory activity was maintained upon incubation at acidic pH (pH2, 4, 5.5) and only slightly reduced at pH 6 (Figure 4.7B). In contrast, the inhibitory activity was reduced upon incubation at pH 7 and drastically reduced upon incubation at basic pH (pH8, 10, 12) (Figure 4.7B). Overall, these results confirmed that the inhibitor is stable at acidic but not basic pH.

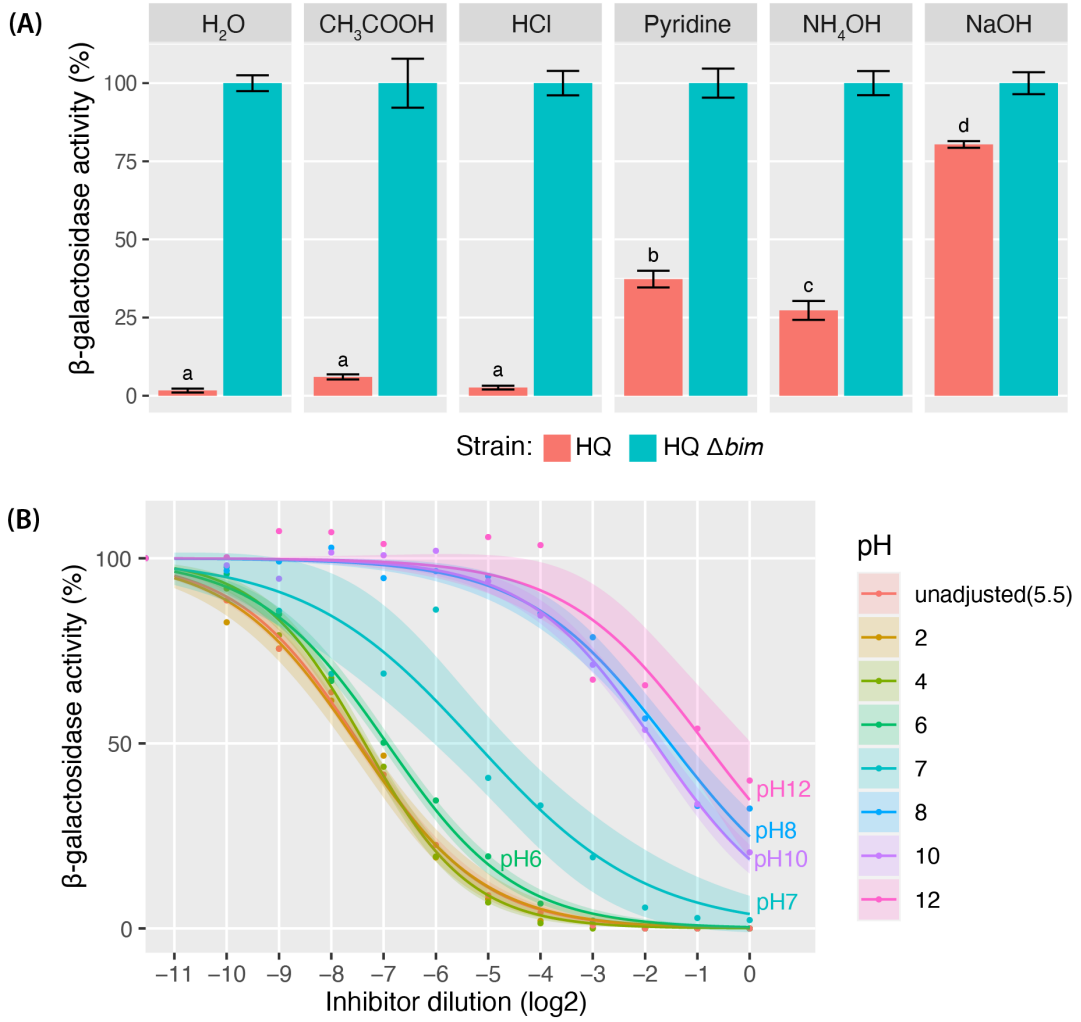


Figure 4.7. The inhibitor is stable in acidic pH but sensitive to basic pH treatment. (A) Effects of acid or base treatment on the inhibitor. Supernatants from cultures of *P. syringae* HQ (inhibitor producing) and HQ Δ*bim* mutant (non-inhibitor-producing) were subjected to treatments of 1 M of different acids or bases for 1 hr before being neutralised. The mixtures were then tested in inhibition assay with FDG and LacZ enzyme. Enzyme activity was expressed as percentage of the activity compared to no inhibitor control. Different letters (a,b,c,d) denote significant difference (p-value < 0.05) from ANOVA followed by Tukey test comparing HQ inhibitor samples of all treatments. Mean and standard deviation are shown (n=3) **(B) effects of pH on the inhibitor.** Supernatants from *P. syringae* HQ culture were adjusted to different pH and incubated for 24 hr before being neutralised to pH 7. The mixtures were then tested in inhibition assay with dilution series to compare level of inhibitor in each condition. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control and plotted against the inhibitor dilution (log2 scale). Dose response curves (with 95% confidence interval) were fitted with log-logistic model to estimate the dilution of the sample required to reach IC50 (the inhibitor concentration required for 50% inhibition of enzymatic activity). The fold difference in dilution required to reach IC50 reflects the difference in inhibitor concentrations in the samples.

4.3.8 The inhibitor can also be eluted from cation exchange resin using HCl

Since the inhibitor is sensitive to basic pH, washing or eluting the cation exchange resin with a base should be avoided. HCl was chosen as an alternative eluent since the inhibitor is stable at acidic pH. HCl provides an excess of H^+ that can displace cations previously bound to the resin, including the inhibitor. I modified the previous cation exchange purification protocol to wash with 70% (v/v) ethanol and water and elute with 1 M HCl. The fractions were then tested for inhibition. The inhibitor was indeed detected in the eluate (Figure 4.8A).

The use of HCl with cation exchange resin not only eluted the inhibitor but also salt cations that binds to the resin (such as Na^+, K^+). The resulting eluate can therefore contain a high amount of salts, which is not desirable because it can interfere with the downstream experiments. One way to minimise this is to use bacterial culture media that contain lower salt concentrations. In addition, the lower amount of cations in the input sample can reduce competition for resin binding sites, resulting in more binding of inhibitor to the resin. Therefore, I used MG medium, which contains lower salt content than M9 medium, for inhibitor production by *E. coli* (p*lac*::*bim*^{wt}). The inhibitor isolation was then carried out following the same protocol as above. I next tested the concentration of the inhibitor by comparing the input bacterial culture supernatant and the output of cation exchange purification in the inhibition assay with dilution series. This showed that around 256-fold concentration of the inhibitor in the eluate compared to the input could be achieved (Figure 4.8B). Since 700 ml of input was loaded onto the resin and the eluate was finally resuspended in 1 ml, the theoretical maximum concentration factor is 700-fold. Therefore, the yield of this isolation was estimated to be ~36%. Overall, these results demonstrate the use of the modified cation exchange resin method for isolating the inhibitor.

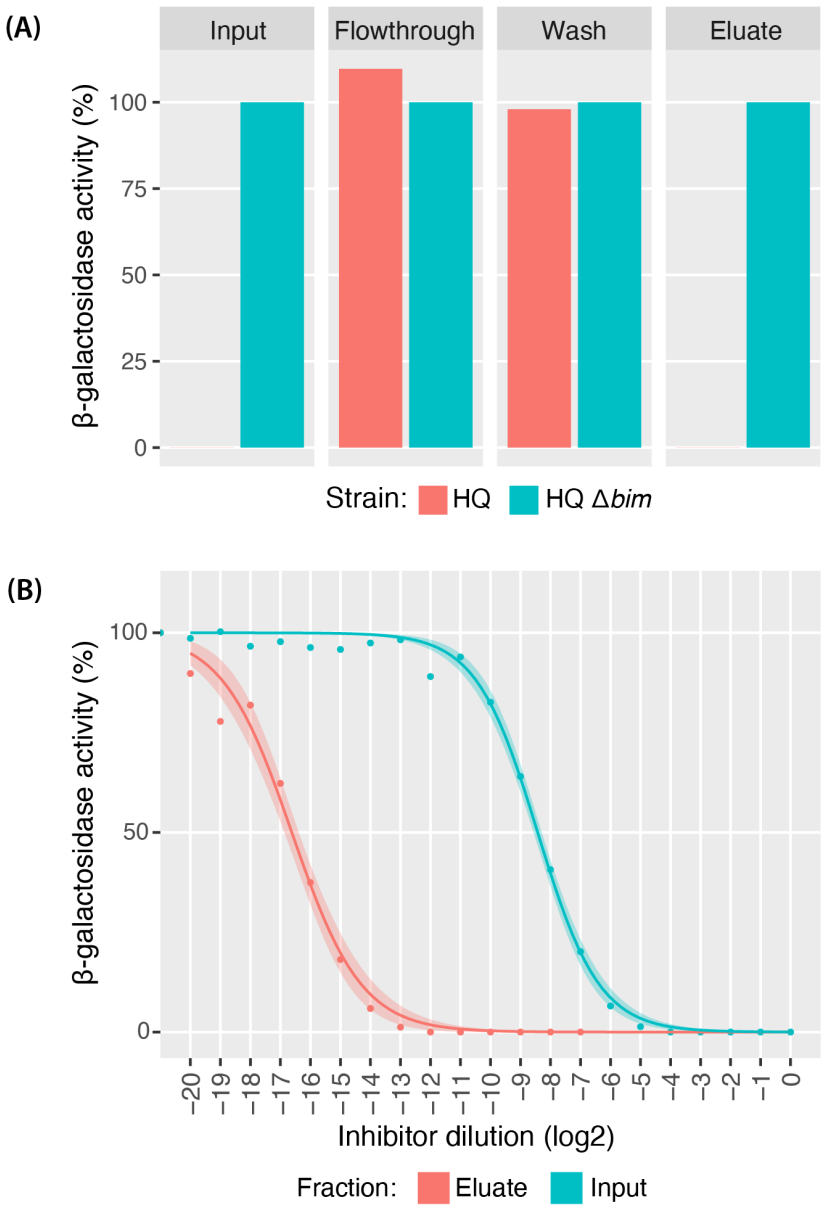


Figure 4.8. The inhibitor can also be eluted from cation exchange resin using HCl. (A) Test of inhibitor in different isolated fractions. Supernatants from cultures of *P. syringae* HQ (inhibitor producing) and HQ Δ*bim* mutant (non-inhibitor-producing) were subjected to methanol-chloroform extraction to obtain a water-soluble metabolite fraction which was then used as input to the cation exchange resin to bind the inhibitor. The resin was then washed with water and 70% (v/v) ethanol before the final elution with 1M HCl. Fractions collected from each step were tested in inhibition assay with FDG and LacZ enzyme. Enzyme activity was expressed as percentage of the activity compared to no inhibitor control. One representative purification round is shown. **(B) Test of enrichment of inhibitor from purification.** The input and final eluate fractions were used in inhibition assay with dilution series to compare level of the inhibitor. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control and plotted against the inhibitor dilution (log2 scale). Dose response curves (with 95% confidence interval) were fitted with log-logistic model to estimate the dilution of the sample required to reach IC50 (the inhibitor concentration required for 50% inhibition of enzymatic activity). The fold difference in dilution required to reach IC50 reflects the difference in inhibitor concentrations in the samples.

4.3.9 Concentrated HCl treatment can be used to desalt eluate from cation exchange resin

Since the eluate from cation exchange resin can contain high amount of salts, which is not desirable, a method for desalting can be beneficial. I tested the use of concentrated HCl (10-12 M) to precipitate cations of salts, such as Na^+ and K^+, from the solution by the common ion effect of saturating amount of Cl^- from HCl causing salts, as NaCl and KCl, to crash out of the solution (Ruud, 2010). The visible salt precipitate can then be separated from the supernatant, which contains the soluble inhibitor. To assess the efficiency of this desalting method, the inhibitor fraction before and after desalting were tested in inhibition assays with dilution series to compare the amount of inhibitor. The desalted fraction recovered around 66% of the input inhibitor (Figure 4.9). This result suggests the possible use of concentrated HCl for desalting the inhibitor fraction.

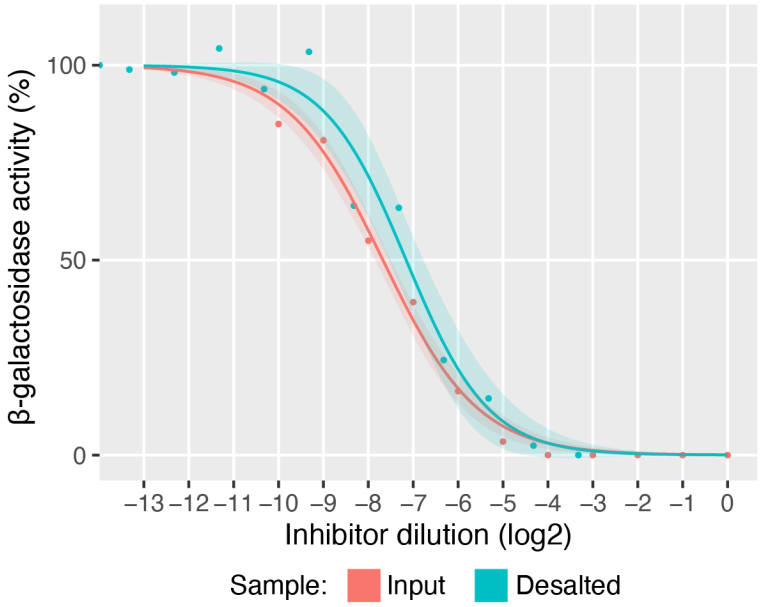


Figure 4.9. The inhibitor from cation exchange resin eluate can be desalted with concentrated HCl treatment. The eluate fraction of cation exchange resin purification of supernatant from *P. syringae* HQ (inhibitor producing) culture was used as an input for concentrated HCl treatment to precipitate salts then the output was neutralised then used in inhibition assay with dilution series to compare level of the inhibitor. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control and plotted against the inhibitor dilution (log2 scale). Dose response curves (with 95% confidence interval) were fitted with log-logistic model to estimate the dilution of the sample required to reach IC50 (the inhibitor concentration required for 50% inhibition of enzymatic activity). The fold difference in dilution required to reach IC50 reflects the difference in inhibitor concentrations in the samples.

4.3.10 Reverse-phase HPLC with C18 column was not suitable for separation of the inhibitor

After cation exchange resin purification of the inhibitor, I could enrich for the inhibitor but there were still other coeluting compounds. The purification method therefore needed to be refined further. I evaluated whether the inhibitor could be isolated using HPLC (high-performance liquid chromatography), a separation method commonly used for compound purification. Since the inhibitor does not absorb UV, only inhibitory activity could be used as an indicator of the inhibitor in a fraction. Therefore, a supernatant from HQ culture was separated with HPLC with the common C18 column in reverse-phase mode and fractions were collected and tested for inhibition. The inhibitor was detected only in eluate fraction 5 and 7 (Figure 4.10A). I next assessed the composition and purity of these fractions with GC-MS analysis. All fractions showed predominant signals from phosphate and fructose which are the major components of the culture medium (Figure 4.10B). Together, these data suggest that the inhibitor is too hydrophilic to interact with the C18 stationary phase in the column leading to insufficient retention and early elution together with other hydrophilic compounds. The reversed-phase HPLC with C18 column was therefore not suitable for purification of the inhibitor. In addition, it was noted that, during fractionation, the glass tube used to collect fraction 5 was rinsed and reused for fraction 7. The tube might not be completely cleaned and the inhibitor detected in fraction 7 was the leftover residue of the inhibitor from fraction 5.

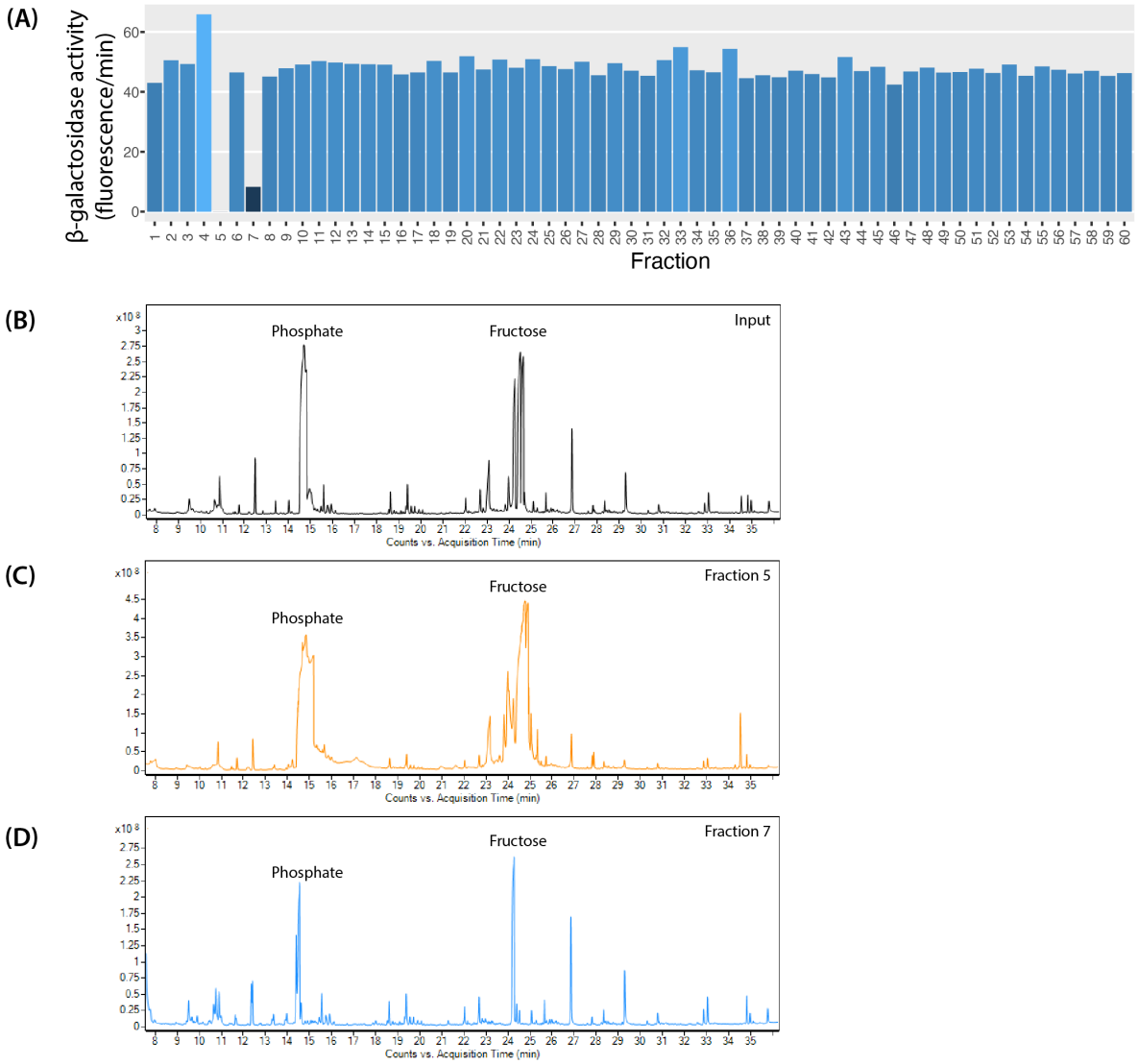


Figure 4.10. Reverse-phase HPLC with C18 column cannot isolate the inhibitor. (A) Test for inhibitor in HPLC isolated fractions. Supernatant from *P. syringae* HQ (inhibitor producing) culture was subjected to methanol-chloroform extraction to obtain a water-soluble metabolite fraction which was then used as input in a preparative reverse-phase HPLC with C18 column. Fractions output from the HPLC run were collected and tested in inhibition assay with FDG and LacZ enzyme. Enzyme activity was expressed as percentage of the activity compared to no inhibitor control. One representative purification round is shown. **(B) GC-MS analysis.** The input and fractions 5 and 7 were dried and derivatised with MSTFA before GC-MS analysis. Total ion chromatogram is shown annotated with dominant detected compounds.

4.3.11 Fmoc derivatisation can be used to modify the inhibitor

Since the inhibitor is a small hydrophilic molecule and non-UV-absorbing, routine liquid chromatography separation methods could not work efficiently. To overcome this limitation, I attempted to modify the inhibitor by chemical derivatisation to attach a group with desirable properties. The selected Fmoc-Cl reagent reacts with an amine group, which was predicted to be present in the inhibitor. Fmoc contains aromatic rings that absorbs UV light and adds a hydrophobic group to the target compound, which would facilitate separation by reverse-phase HPLC (Amézqueta and Torres, 2016) (Figure 4.11A).

To test if the inhibitor can be derivatised this way, the inhibitor fraction was incubated for 30 minutes with or without Fmoc-Cl and borate buffer pH 9.5, and then tested in the inhibition assays to determine the remaining inhibitory activity. The Fmoc treated sample showed two-fold lower inhibition activity than the buffer only control (Figure 4.11B) suggesting that the inhibitor has been derivatised and thus lost inhibitory activity. Since the inhibitor sample did not completely lose its inhibitory activity, this derivatisation was likely to be partial and a proportion of inhibitor molecules still remained underivatised. However, the inhibitory activity was also lower in the buffer only control compared to the untreated sample, which could be due to the inactivation of the inhibitor by the alkaline pH of the buffer (Figure 4.11B). Nevertheless, the results demonstrated Fmoc derivatisation of the inhibitor and confirmed the amine containing nature of the inhibitor.

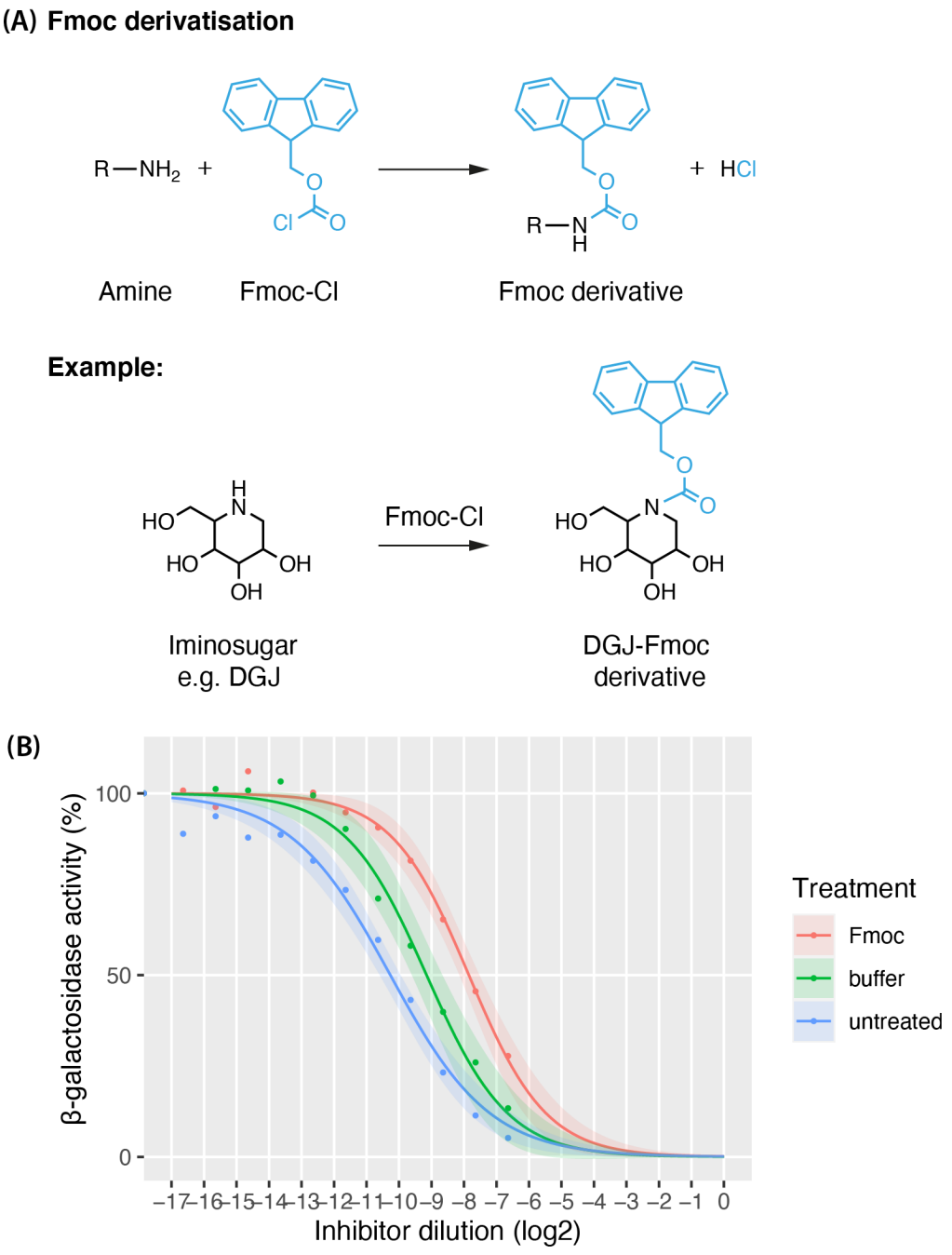


Figure 4.11. Fmoc derivatisation can be used to modify the inhibitor. (A) Schematic diagram of Fmoc derivatisation reaction, **(B)** The partially purified inhibitor from cation exchange resin purification of supernatant from *P. syringae* HQ culture was treated with Fmoc-Cl derivatisation reagent in potassium borate buffer pH 9.5 or only buffer or water and incubated for 30 minutes at room temperature before being quenched with 0.1 M glycine and acetic acid acidification. The mixtures were then tested in inhibition assay with dilution series to compare level of inhibitor in each condition. The enzymatic activity was expressed as percentage of the activity compared to no inhibitor control and plotted against the inhibitor dilution (log2 scale). Dose response curves (with 95% confidence interval) were fitted with log-logistic model to estimate the dilution of the sample required to reach IC50 (the inhibitor concentration required for 50% inhibition of enzymatic activity). The fold difference in dilution required to reach IC50 reflects the difference in inhibitor concentrations in the samples.

4.3.12 LC-MS can be used to detect Fmoc-derivatised iminosugars

After establishing that Fmoc derivatisation can be done on the inhibitor to overcome the challenge of inhibitor hydrophilicity and UV absorbing property, I next investigated whether it can be used to facilitate LC-MS (Liquid Chromatography – Mass Spectrometry) analysis to detect the inhibitor and gain more information on its mass. As a proof of concept, I first tested this approach on a commercially available inhibitor, 1-deoxygalactonojirimycin (DGJ), an iminosugar with similar properties to the inhibitor of interest.

First, a blank run with only Fmoc derivatisation reagents and buffer but no sample was run in LC-MS to assess background. The UV detection showed a prominent peak at 7 min, which corresponds to the mass 179.1 Da detected by mass spectrometry (Figure 4.12A). This signal originated from the 9-methylene-fluorene by product from the Fmoc-Cl reagent. It is worth noting that in the LC-MS system, the sample that has passed through the LC separation step enters the UV detector first before going to the MS detector. This results in a short time lag (around 0.2-0.3 min) from the signal detected by UV light to the signal detected by MS. In this case, the peak, which appeared at 7.02 min in the UV chromatogram, appeared at 7.25 min in the total ion chromatogram (Figure 4.12A). In addition, the mass detected by LC-MS usually corresponds to the predicted compound molecular mass + 1 because an adduct is formed with a H^+ during ionisation. In this case, 9-methylene-fluorene with the mass of 178.08 Da was indeed detected by LC-MS as 179.1 Da (Figure 4.12A).

Next, I tested the DGJ sample derivatised with Fmoc in the LC-MS. The UV peak of DGJ-Fmoc could be detected at 5.7 min (Figure 4.12B). MS analysis of that sample showed a mass of 386.15 Da, which is expected for the DGJ-Fmoc H^+ adduct (163.1 Da [DGJ] + 1 Da [H^+ adduct]

+ 223.2 Da [Fmoc group] - 1 Da [H loss from amine group during bond formation with Fmoc]). To identify the DGJ-Fmoc peak, the chromatogram was extracted for the ion with mass of 386.15 Da, which appeared at 5.9 min (Figure 4.12B). This peak was not detected in the underivatized DGJ sample, confirming that it originated from FMOC derivatised DGJ (Figure 4.12C). This result indicates that Fmoc derivatisation coupled with LC-MS can be used to detect iminosugars.

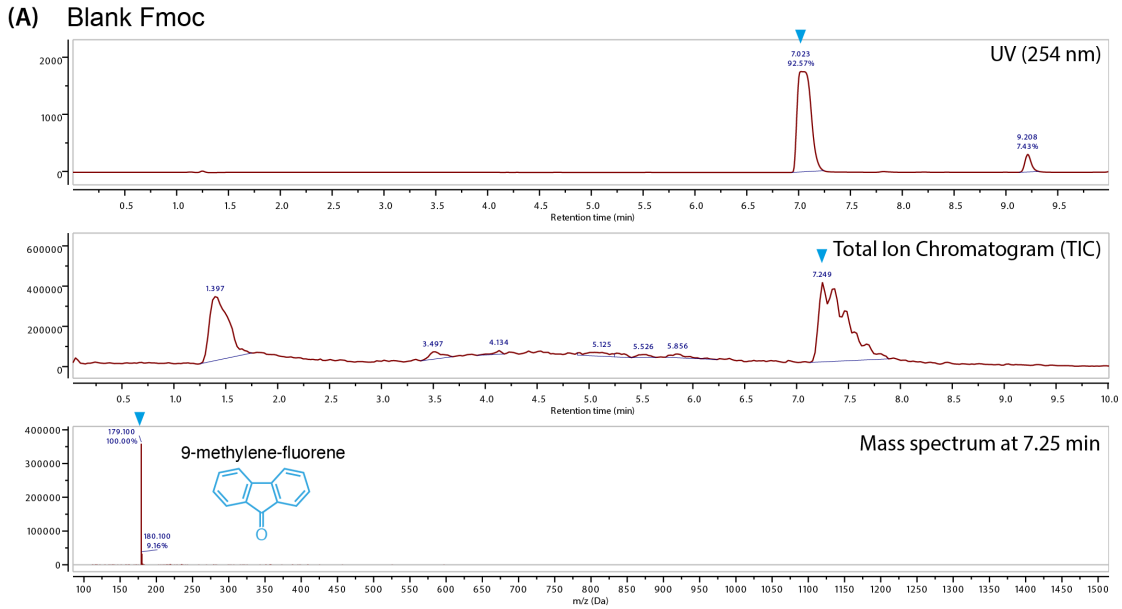


Figure 4.12. LC-MS can be used to detect Fmoc derivatised iminosugar. LC-MS analyses results showing chromatogram of UV absorption at 254 nm, total ion chromatogram, mass spectra of peak of interest and extracted ion chromatogram of mass of interest. Input samples were: **(A)** blank sample with only Fmoc derivatisation reagents, **(B)** iminosugar DGJ (Deoxygalactonojirimycin) with Fmoc derivatisation reagents, **(C)** DGJ without Fmoc derivatisation. (continued on next page)

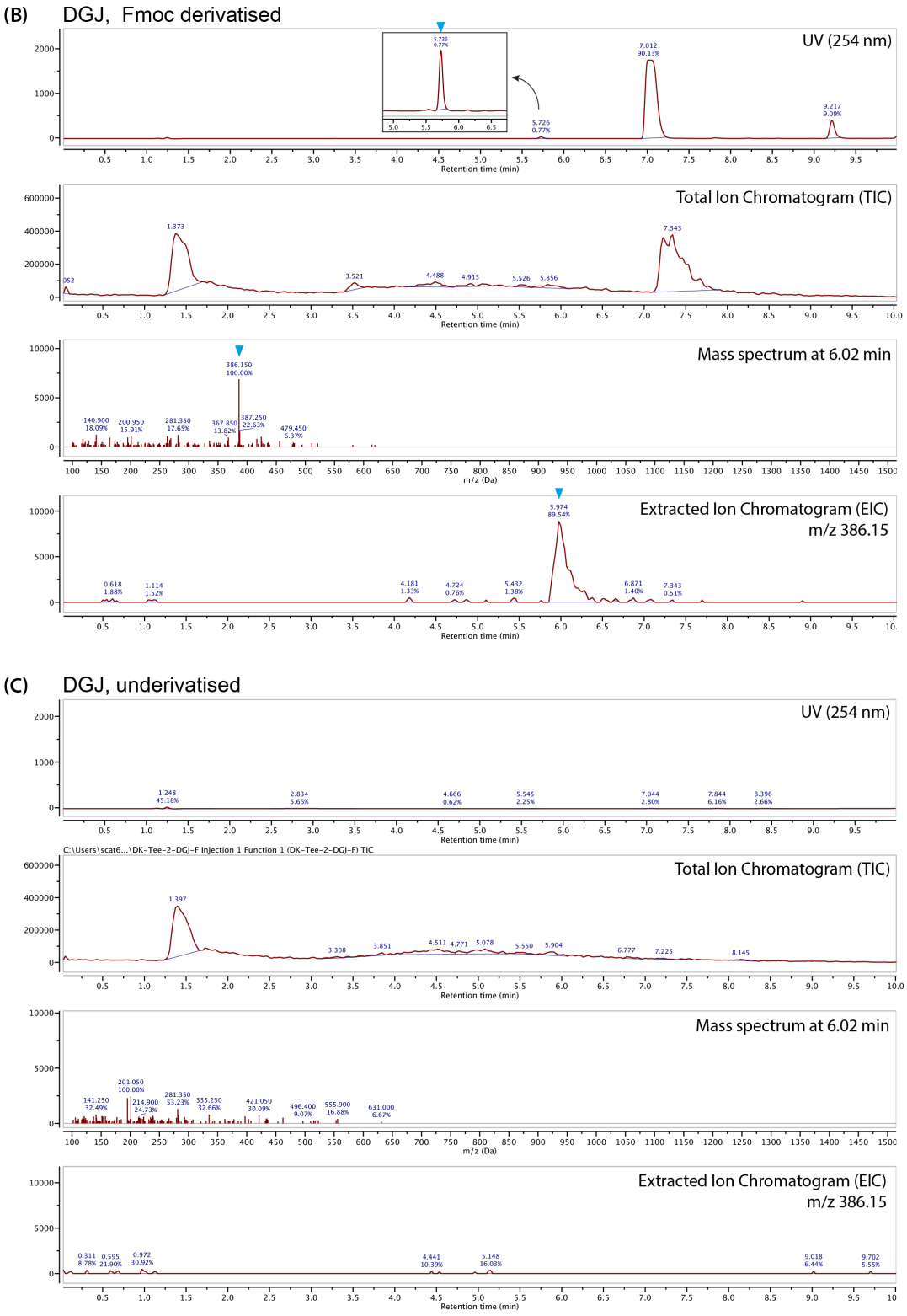


Figure 4.12 (continued). LC-MS can be used to detect Fmoc derivatised iminosugar. LC-MS analyses results showing chromatogram of UV absorption at 254 nm, total ion chromatogram, mass spectra of peak of interest and extracted ion chromatogram of mass of interest. Input samples were: **(A)** blank sample with only Fmoc derivatisation reagents, **(B)** iminosugar DGJ (Deoxygalactonojirimycin) with Fmoc derivatisation reagents, **(C)** DGJ without Fmoc derivatisation. (continued on next page)

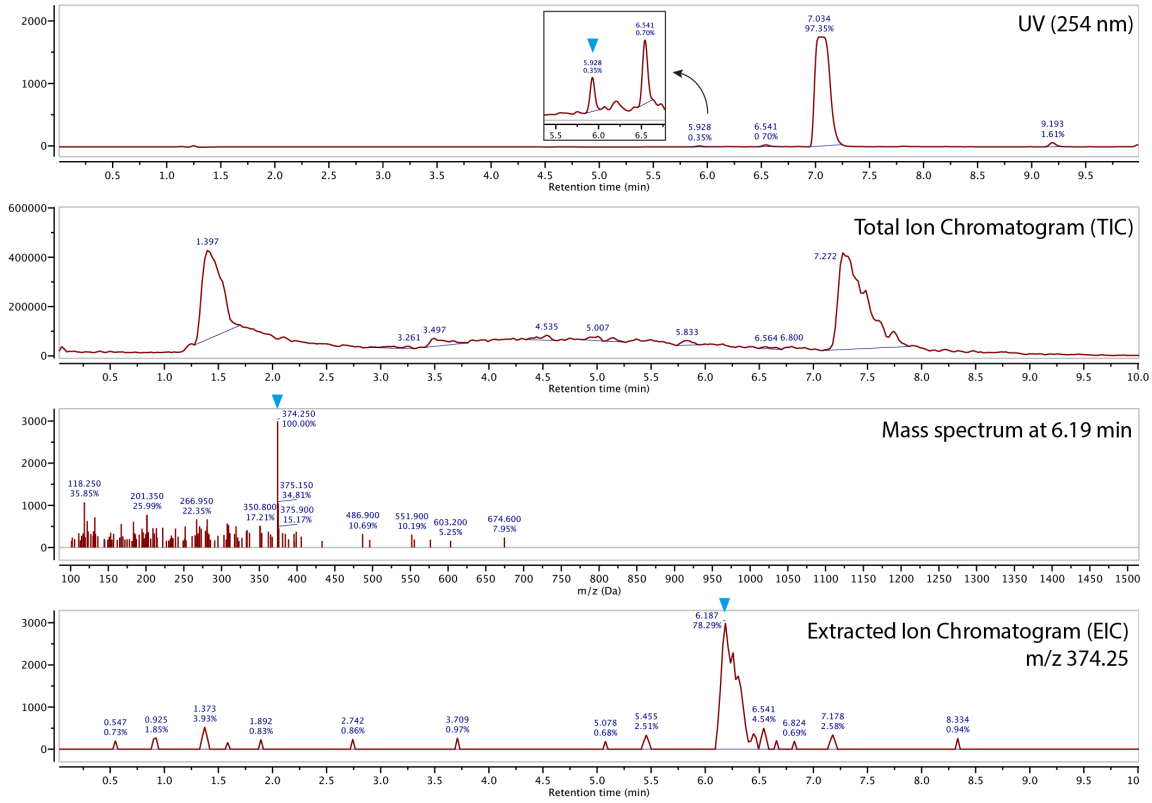
4.3.13 LC-MS can be used to detect FMOC-derivatised inhibitor from *P. syringae*

After establishing that LC-MS can be used to detect Fmoc-derivatised iminosugars, I applied this approach to the inhibitor fraction isolated from *P. syringae* HQ cultures purified on cation exchange resin. Since the identity and mass of the inhibitor was unknown, I could not look for a specific mass in the spectra. Instead, I used a comparative approach comparing the fraction from the inhibitor-producing HQ strain to the non-producing HQ Δ*bim* mutant. A differential peak present in HQ but not HQ Δ*bim* would be indicative of the inhibitor.

In the Fmoc-derivatised HQ inhibitor sample, I could detect a UV peak at 5.92 min with a corresponding mass of 374.25 Da from the extracted ion chromatogram peak at 6.19 min (Figure 4.13A). In contrast, in the Fmoc-derivatised HQ Δ*bim* sample, only a trace of this signal could be seen, not substantially higher than the background noise (Figure 4.13B). In addition, this peak was not detected in both underivatized samples, confirming that it originated from the Fmoc-derivatised inhibitor (Figure 4.13C, D). These results suggested the Fmoc-derivatised inhibitor might have a mass of 374.25 Da.

Figure 4.13 (next page). Fmoc derivatisation and LC-MS analysis identified a putative peak of the inhibitor isolated from *P. syringae.* LC-MS analyses results showing chromatogram of UV absorption at 254 nm, total ion chromatogram, mass spectra of peak of interest and extracted ion chromatogram of mass of interest. Input samples were: **(A)** inhibitor eluate from cation exchange resin purification of *P. syringae* HQ culture, Fmoc derivatised. Putative signals from the inhibitor are labelled with blue arrowheads, **(B)** no inhibitor eluate from cation exchange resin purification of *P. syringae* HQ Δ*bim* culture, Fmoc derivatised, **(C)** inhibitor eluate from cation exchange resin purification of *P. syringae* HQ culture, underivatized, **(D)** no inhibitor eluate from cation exchange resin purification of *P. syringae* HQ Δ*bim* culture, underivatized.

(A) HQ, Fmoc derivatised



(B) HQ Δ*bim*, Fmoc derivatised

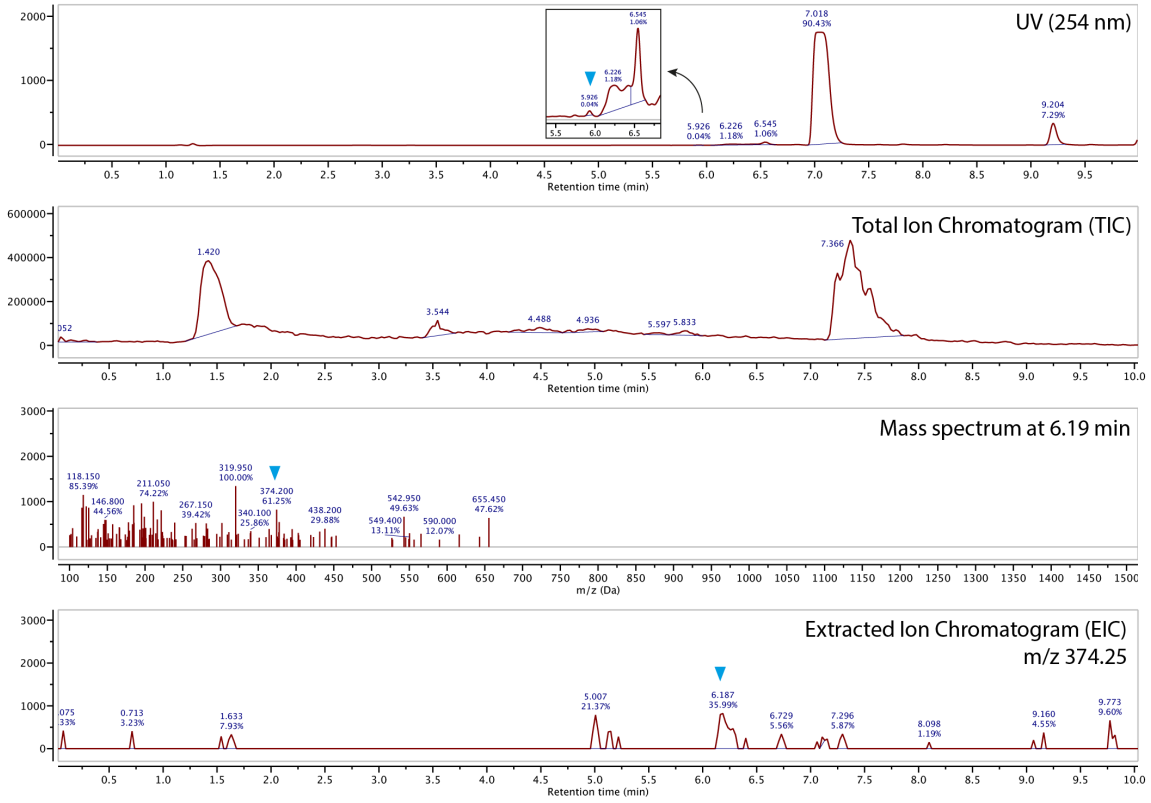
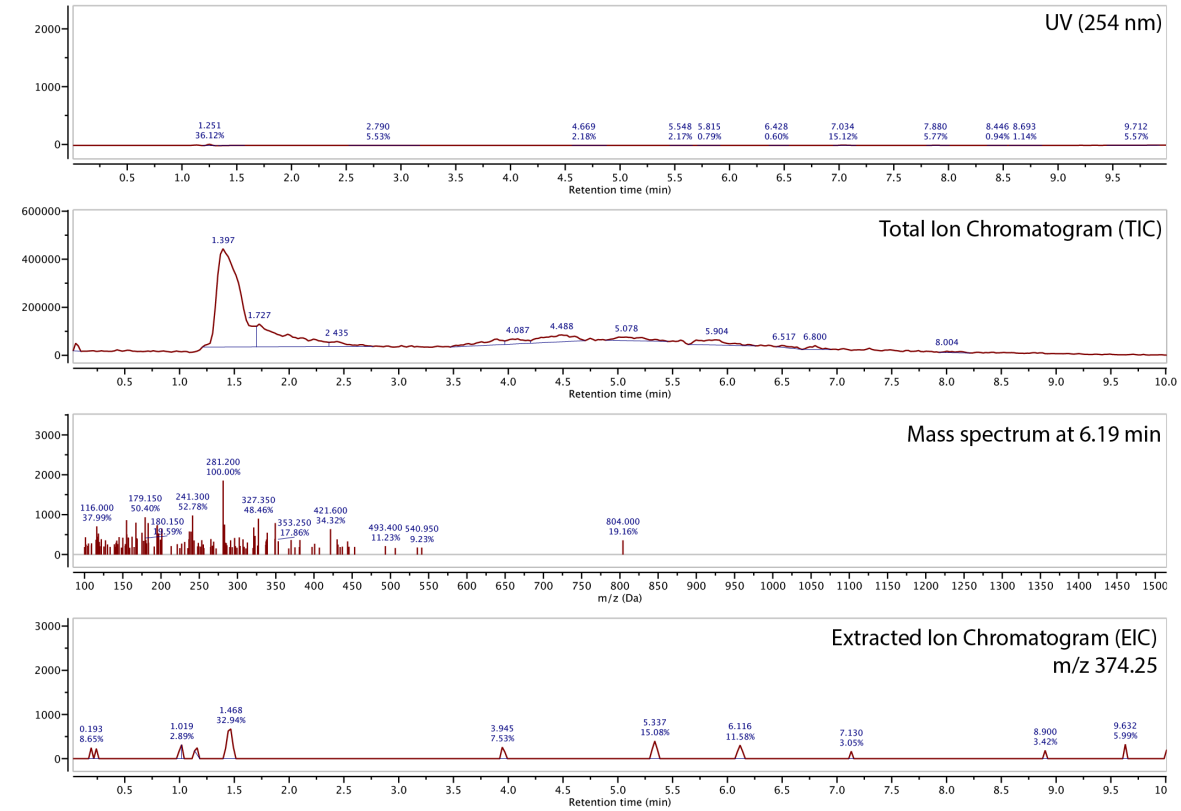


Figure 4.13 (continued). Legend on previous page

(C) HQ, underivatised



(D) HQ Δ*bim*, underivatised

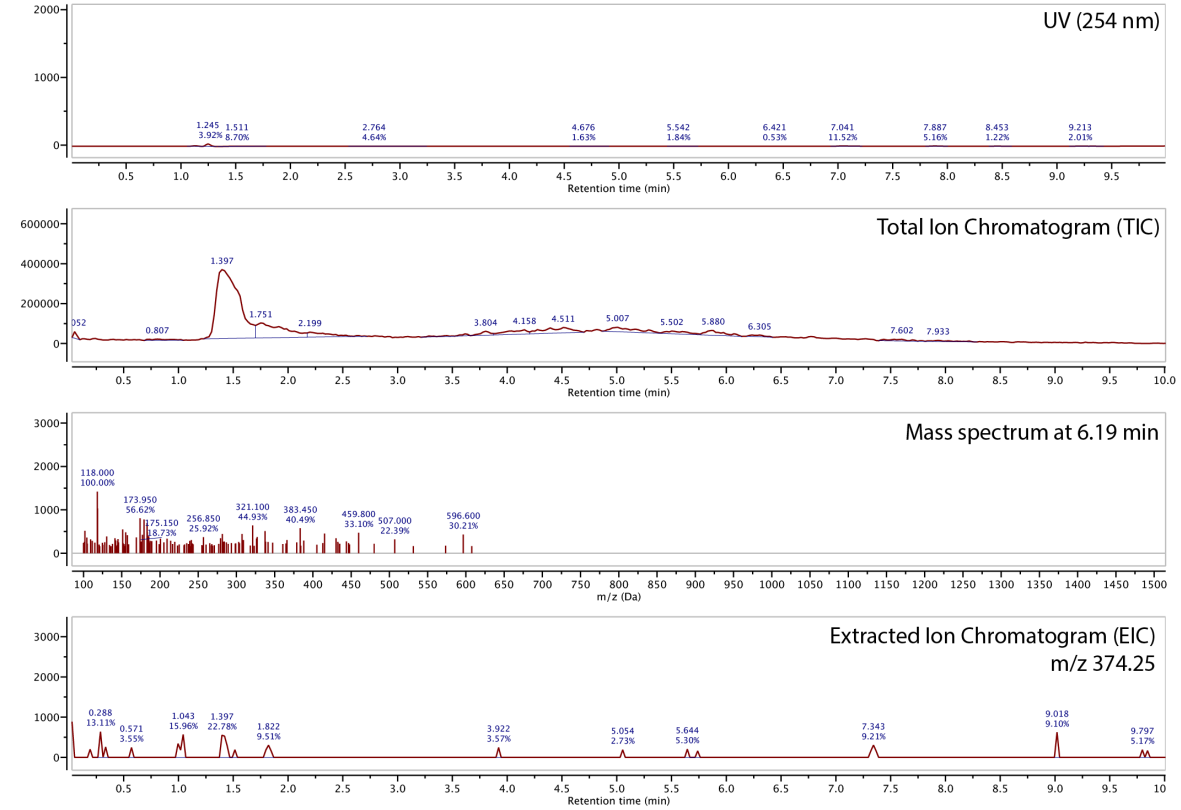


Figure 4.13 (continued). Legend on previous page

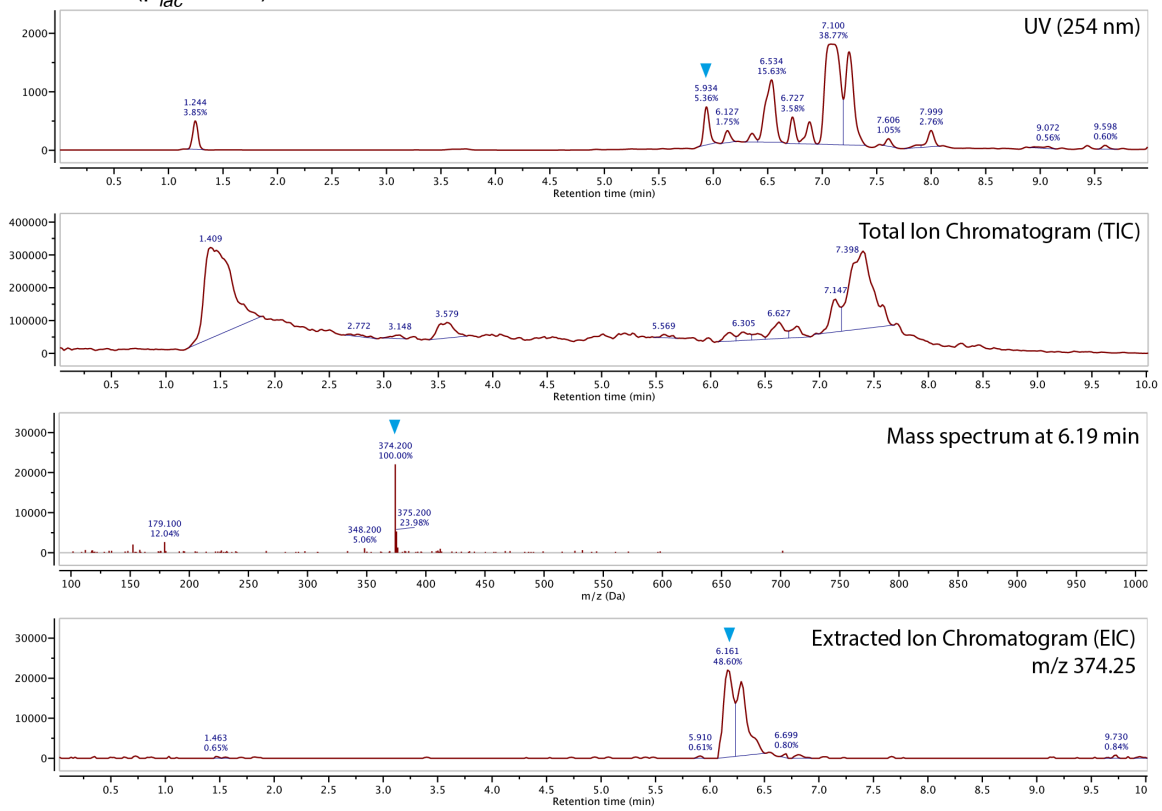
4.3.14 LC-MS can be used to detect FMOC-derivatised inhibitor from *E. coli* expressing the *bim* cluster

To further confirm the finding of the candidate inhibitor peak found in inhibitor samples from *P. syringae*, the same protocol was carried out with samples from the inhibitor-producing *E. coli* (p*lac*::*bim*^{wt}) strain and non-inhibitor-producing *E. coli* (p*lac*::*bim*^{mut}) strain as a negative control. Again, the inhibitor sample derived from *E. coli* (p*lac*::*bim*^{wt}) showed the same UV peak at 5.9 min and the mass of 374.2 Da at around 6.19 min (Figure 4.14A). These signals were also absent in the sample derived from *E. coli* (p*lac*::*bim*^{mut}) strain that produced no inhibitor (Figure 4.14B).

Overall, these results showed that Fmoc derivatisation can be used together with LC-MS to detect the putative inhibitor. The putative mass of Fmoc-derivatised inhibitor was 374.2 Da, suggesting that the putative inhibitor has a mass of ~151 Da (374.2 Da - 1 Da [H^+ adduct] - 223.2 Da [Fmoc group] + 1 Da [H loss from amine group during bond formation with Fmoc]).

Figure 4.14 (next page). Fmoc derivatisation and LC-MS analysis identified a putative peak of the inhibitor isolated from *E. coli* carrying *bim* cluster. LC-MS analyses results showing chromatogram of UV absorption at 254 nm, total ion chromatogram, mass spectra of peak of interest and extracted ion chromatogram of mass of interest. Input samples were: **(A)** inhibitor eluate from cation exchange resin purification of *E. coli* (p*lac*::*bim*^{wt}) culture, Fmoc derivatised. Putative signals from the inhibitor are labelled with blue arrowheads. **(B)** no inhibitor eluate from cation exchange resin purification of *E. coli* (p*lac*::*bim*^{mut}) culture, Fmoc derivatised. Positions equivalent to the putative inhibitor signals are labelled with grey arrowheads.

(A) *E. coli* (p_{lac}::*bim*^{wt}), Fmoc derivatised



(B) *E. coli* (p_{lac}::*bim*^{mut}), Fmoc derivatised

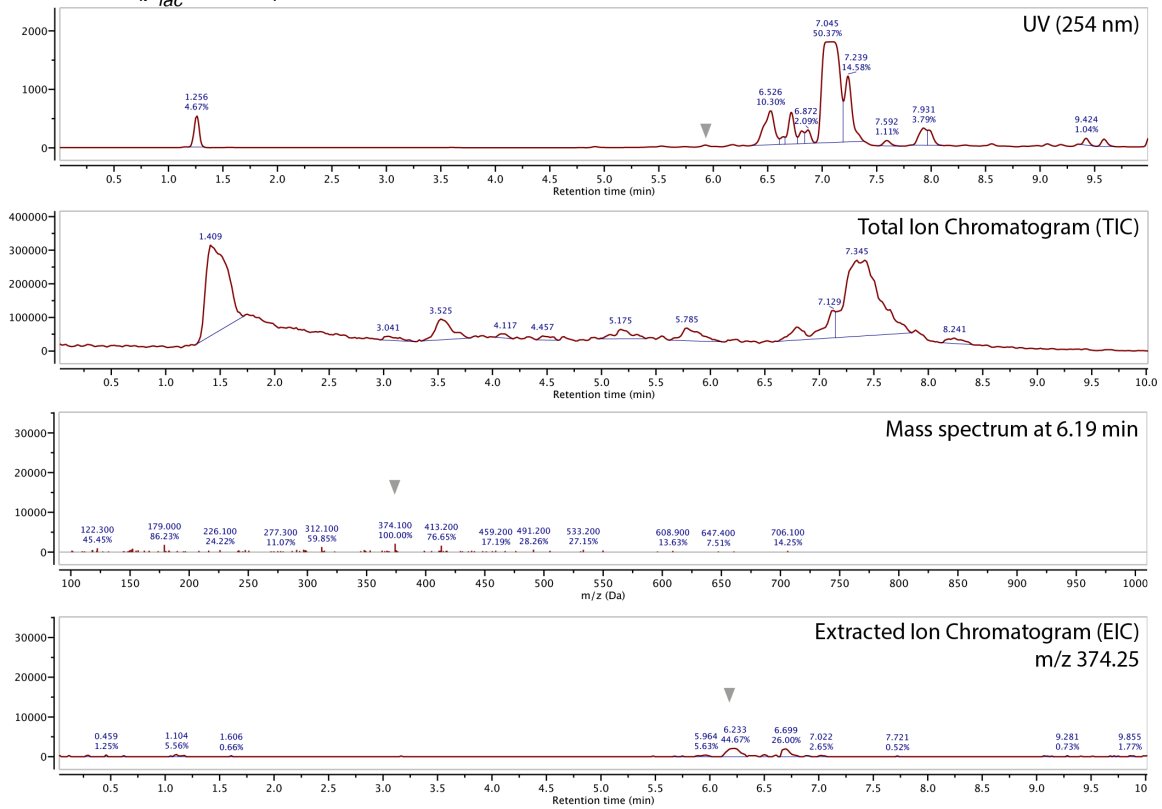


Figure 4.14 (continued). Legend on previous page

4.3.15 HILIC HPLC can be used to isolate the inhibitor

Fmoc derivatisation was useful for inhibitor detection and separation in typical liquid chromatography systems. However, since the inhibitor molecule was chemically modified, detection of native inhibitor by itself was not possible and the use for inhibitor purification purposes was limited. Although removal of the Fmoc group from a derived compound could be achieved by a treatment with bases such as piperidine, it is not suitable for the base-sensitive inhibitor. On the other hand, HILIC (Hydrophilic Interaction Liquid Chromatography) HPLC can be a useful method for separation of highly hydrophilic compounds, such as sugars, since the more polar stationary phase can interact and retain hydrophilic compounds so that separation is achieved (Rodríguez-Sánchez, Quintanilla-López, *et al.*, 2014b). I therefore tested if HILIC HPLC could be used to further purify the inhibitor fraction isolated from cation exchange resin.

I separated the inhibitor isolated from *E. coli* (p*lac*::*bim*^{wt}) strain with HILIC HPLC and collected all fractions. UV detection of the run showed different peaks at different retention time points suggesting that some separation was achieved (Figure 4.15A). To identify inhibitor-containing fractions, all collected fractions were tested in an inhibition assay. Inhibitory activity was found to be highest in fraction 14 and less so in fraction 17 and the least in fraction 12 (Figure 4.15B). The appearance of the inhibitor in fractions in the middle of the run suggested that the inhibitor had some retention in the column. The presence of the inhibitor in multiple fractions might be due to multiple isomers of the inhibitor molecule or sub-optimal HILIC HPLC program, which is can cause retention time drift.

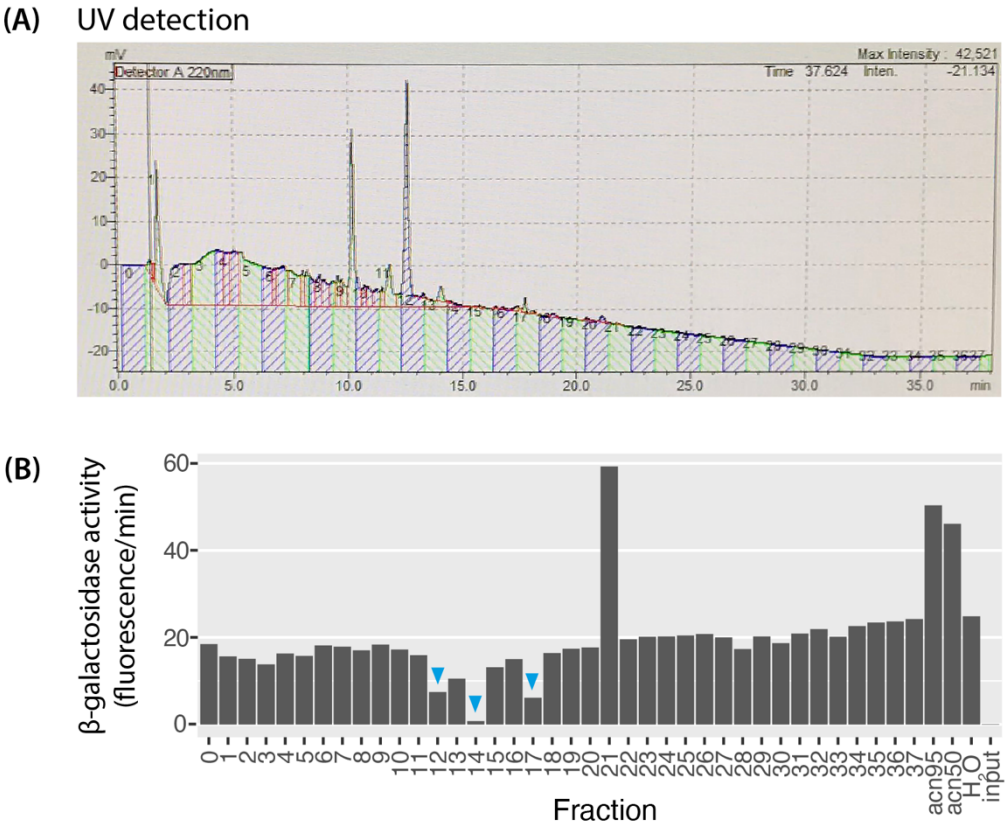


Figure 4.15. HILIC HPLC can be used for separation of the inhibitor. Inhibitor eluate from cation exchange resin purification of *E. coli* (p*lac*::*bim*^{wt}) culture was used as an input in the preparative HPLC run on HILIC mode. **(A)** HPLC output fractions were collected and tested for presence of the inhibitor by inhibition assay with FDG and LacZ enzyme. **(B)** UV absorption chromatogram of the HPLC run. Solid arrowheads indicate detected inhibition.

To test if I could detect the inhibitor signal as before, the inhibitor-containing fractions were subjected to Fmoc derivatisation and LC-MS analysis. The signal from the previously observed putative inhibitor, with the UV peak at 5.9 min and the mass of 374.2 Da peak at 6.2 min, was detected in both fractions 14 and 17 (Figure 4.16A, B). In addition, compared to Fmoc LC-MS analysis of inhibitor obtained from cation exchange resin alone (Figure 4.14), the HILIC output fractions showed significantly lower signals from other peaks, suggesting that further purification was achieved. These results demonstrated that HILIC HPLC can be used for native inhibitor isolation and further confirm the putative inhibitor peak detected from FMOC LC-MS analysis.

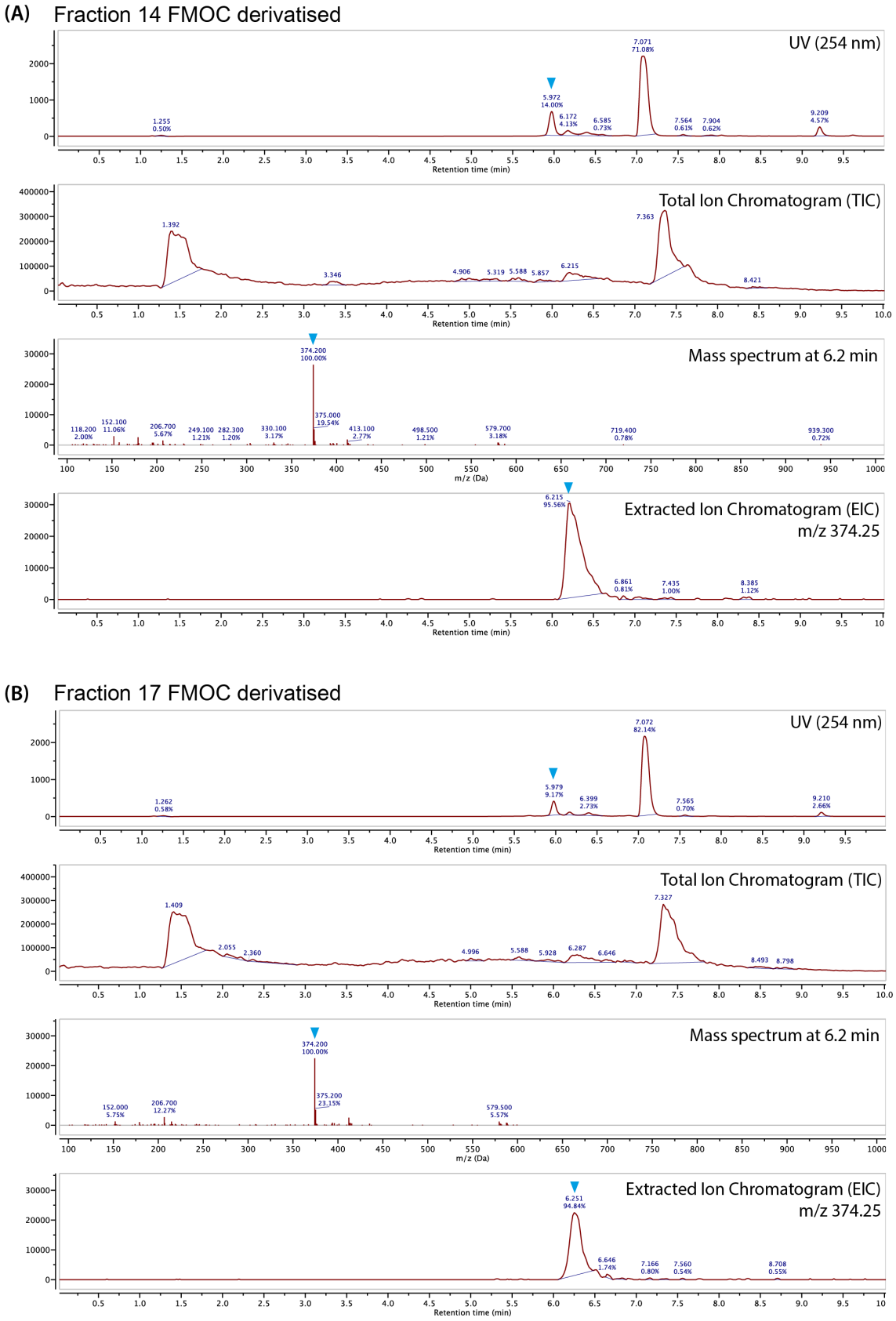


Figure 4.16. Fmoc derivatisation and LC-MS analysis identified a putative peak of the inhibitor from HILIC HPLC isolated fractions. LC-MS analyses results showing chromatogram of UV absorption at 254 nm, total ion chromatogram, mass spectra of peak of interest and extracted ion chromatogram of mass of interest. Input samples were: **(A)** fraction 14 and **(B)** fraction 17 from HILIC HPLC purification of inhibitor eluate from cation exchange resin purification of *E. coli* (p*lac*::*bim*^{wt}) culture. Samples were Fmoc derivatised prior to LC-MS analysis. Putative signals from the inhibitor are labelled with blue arrowheads.

4.3.16 GC-MS can be used to detect the TMS-derivatised inhibitor from *P. syringae*

I next used GC-MS (Gas Chromatography - Mass Spectrometry) analysis as another independent approach to characterise the inhibitor to further support my findings. GC-MS could yield not only information on the molecular mass but additional information from fragmentation patterns. However, analysis of polar compounds by GC-MS requires chemical derivatisation to render the compound more volatile so that it is compatible to gas chromatography. TMS derivatisation (trimethylsilylation), by a treatment with MSTFA, is a common derivatisation that adds trimethylsilyl (TMS) groups to amine or hydroxyl groups of a compound. Thus, TMS derivatisation can be done on iminosugars (Amézqueta and Torres, 2016) (Figure 4.17).

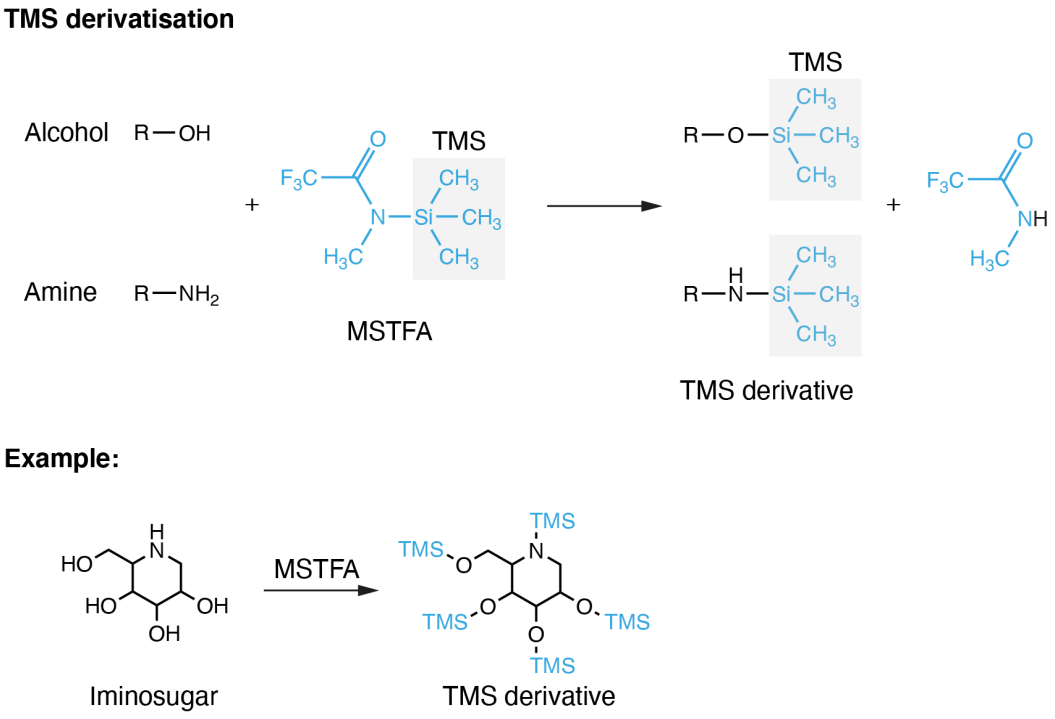


Figure 4.17. Schematic diagram illustrating TMS derivatisation. TMS derivatisation (trimethylsilylation) can be done by a treatment with MSTFA that adds trimethylsilyl (TMS) groups to amine or hydroxyl groups of a compound. Thus, TMS derivatisation can be done on iminosugars

I took the same comparative approach as used for LC-MS analysis and compared cation exchange resin isolated inhibitor fractions from *P. syringae* HQ and HQ Δ*bim* strains. First, a blank sample with only derivatisation reagent and solvent was analysed to establish background signals (Figure 4.18A). Next, a comparison of total ion chromatogram (TIC) signals from HQ and HQ Δ*bim* samples showed a distinguishable peak at 23.1 min in HQ but not HQ Δ*bim* sample (Figure 4.18B, C), suggesting a candidate peak of the inhibitor. The mass spectrum of the candidate inhibitor is shown in Figure 4.18D. This mass spectrum has a 217 Da peak, corresponding to a fragment of three carbon with two TMS groups, which is commonly found in sugar-type compounds (Harvey and Vouros, 2020). For more focussed analysis, I then looked for the extracted chromatogram (EIC) of mass 217 Da. A prominent peak candidate of the inhibitor at 23.1 min was detected in HQ sample but not blank or HQ Δ*bim* samples (Fig 18 B, C). Another prominent peak at 22 min was likely to be the incompletely derivatised form of the inhibitor, since its mass spectrum lacks the larger fragments corresponding to the fully derivatised inhibitor. These results suggested the putative peak and mass spectrum of the inhibitor.

Figure 4.18 (next page). TMS derivatisation and GC-MS analysis identified a putative peak and mass spectra of the inhibitor isolated from *P. syringae.* GC-MS analysis results showing total ion chromatogram and extracted ion chromatogram of m/z 217. Input samples were: **(A)** Blank sample of only TMS derivatisation reagents, **(B)** inhibitor eluate from cation exchange resin purification of *P. syringae* HQ culture, TMS derivatised. Putative signals from the inhibitor are labelled with blue arrowheads, **(C)** no inhibitor eluate from cation exchange resin purification of *P. syringae* HQ Δ*bim* culture, TMS derivatised, **(D)** Mass spectrum of the putative peak of the inhibitor from (B). Solid arrowheads indicate candidate inhibitor peak, while empty arrowheads indicate equivalent positions in no-inhibitor samples.

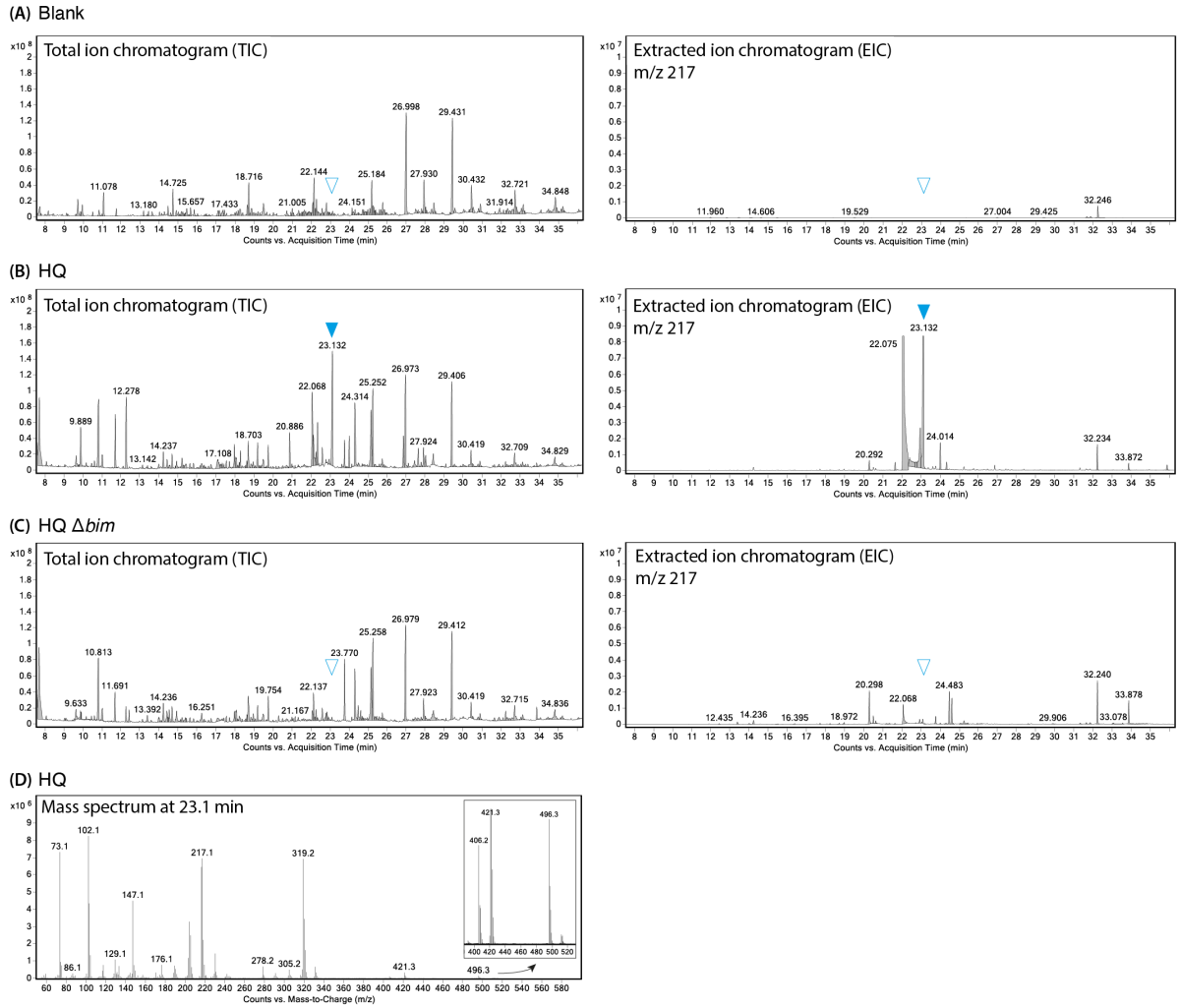


Figure 4.18 (continued). Legend on previous page.

4.3.17 GC-MS can be used to detect the TMS-derivatised inhibitor from *E. coli* expressing the *bim* cluster

To confirm the findings from the *P. syringae* derived inhibitor sample obtained using GC-MS, I next analysed samples derived from inhibitor-producing *E. coli* (p*lac*::*bim*^{wt}) and non-inhibitor-producing *E. coli* (p*lac*::*bim*^{mut}). Again, a distinguishable peak was detected at 22.7 min in *E. coli* (p*lac*::*bim*^{wt}) but not in *E. coli* (p*lac*::*bim*^{mut}) (Figure 4.19B, C). The mass spectrum of this peak (Figure 4.19D) was the same as that detected for the *P. syringae* candidate inhibitor peak (Figure 4.18D), although its retention time was slightly different since the gas chromatography run conditions were different. In addition, the extracted chromatogram (EIC)

of mass 217 also showed a prominent candidate inhibitor peak in the sample from *E. coli* with the wild-type *bim* cluster but not in *E. coli* with the mutant *bim* cluster (Figure 4.19B, C). This result provided further support that the peak and mass spectrum detected is from the candidate inhibitor.

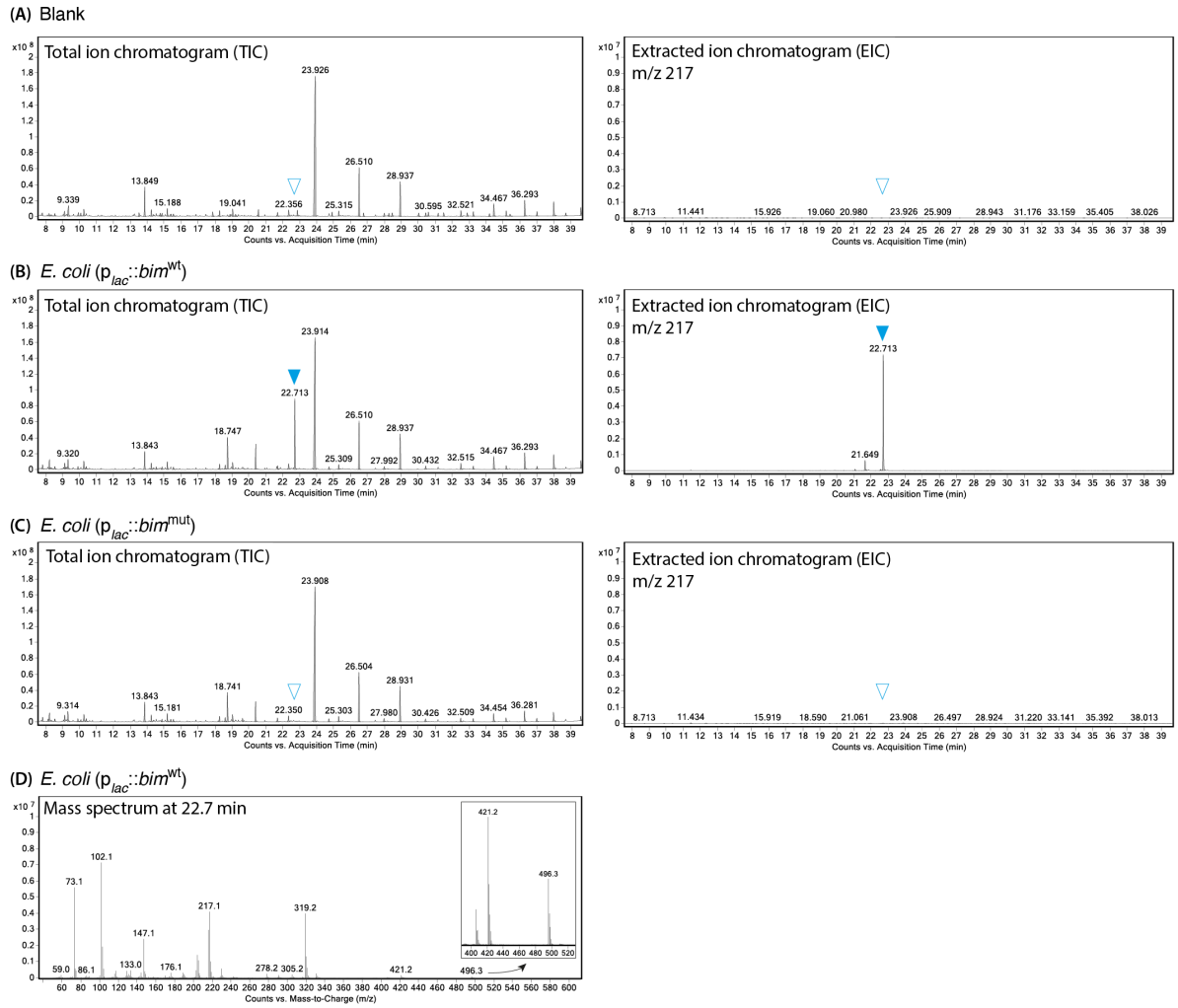


Figure 4.19. TMS derivatisation and GC-MS analysis identified a putative peak and mass spectra of the inhibitor isolated from *E. coli* carrying *bim* cluster. GC-MS analysis results showing total ion chromatogram and extracted ion chromatogram of m/z 217. Input samples were: **(A)** Blank sample of only TMS derivatisation reagents, **(B)** inhibitor eluate from cation exchange resin purification of *E. coli* (p*lac*::*bim*^{wt}) culture, TMS derivatised. Putative signals from the inhibitor are labelled with blue arrowheads, **(C)** no inhibitor eluate from cation exchange resin purification of *E. coli* (p*lac*::*bim*^{mut}) culture, TMS derivatised, **(D)** Mass spectrum of the putative peak of the inhibitor from (B)**.** Solid arrowheads indicate candidate inhibitor peak, while empty arrowheads indicate equivalent positions in no-inhibitor samples.

4.3.18 Analysis of the putative GC-MS spectrum of the inhibitor

I next analysed the putative GC-MS mass spectra of the inhibitor (Figure 4.18D, 19D) to gain more insights into the possible inhibitor molecular identity. GC-MS analysis of TMS-derivatised compounds have been extensively characterised giving a wealth of information to interpret mass spectra and fragmentations patterns (Lai and Fiehn, 2018; Harvey and Vouros, 2020). The largest detected ion with mass of 511 Da was likely the parental molecular ion $[M]^+$ of the TMS-derivatised inhibitor. This was supported by the 496 Da and 421 Da fragment ions, which correspond to commonly detected fragmentation of the parental molecular ion: $[M - 15]^+$ for the loss of CH_3 group from the TMS group and $[M - 90]^+$ for the loss of the TMS-OH group from the TMS-derivatised hydroxyl group, respectively. An additional loss of the TMS-OH group from the 496 Da fragment could yield the 406 Da fragment while an additional loss of the TMS-OH group from the 421 Da fragment could yield the 331 Da fragment.

Smaller fragment ions could be accounted for by characteristic fragmentations of different classes of functional groups (Lai and Fiehn, 2018; Harvey and Vouros, 2020). The 102 Da fragment is characteristic of an amine group, likely secondary amine. The 147 Da fragment is characteristic of a hydroxyl group. The 204, 217, 319 Da fragments are characteristic of sugars, with multiple hydroxyl groups. The higher intensity of 217 Da fragment compared to the 204 Da fragment is suggestive for a furanose form (five-membered ring). The 217 Da fragment corresponds to a 3-carbon fragment with 2 TMS groups, which is usually present in a compound with 3 adjacent hydroxyl groups on a carbon backbone. The 319 Da fragment corresponds to a 4-carbon fragment with 3 TMS groups. The 73 Da fragment corresponds to the free TMS group.

The parental molecular ion suggested that the TMS-derivatised inhibitor has a mass of 511 Da. This likely contains 5 TMS groups, derivatised at 1 amine group and 4 hydroxyl groups. By subtracting the TMS groups (5 x 73 Da) and adding back the H from amine or hydroxyl groups loss from bond formation with TMS (5 x 1 Da), the inhibitor molecule likely has a mass of 151 Da, which is consistent with the mass deduced by LC-MS analysis (Fig 13, 14). Taken together, these results support the putative molecular mass of 151 Da of the inhibitor and suggest that the inhibitor is an iminosugar with 4 hydroxyl groups, possibly in a 5-membered ring form.

4.3.19 The partially purified inhibitor showed inhibitory activity against multiple glycosidases

Another important property of the inhibitor to characterise is its specificity of glycosidase inhibition. I therefore tested the partially purified inhibitor from both *P. syringae* HQ and *E. coli* expressing the *bim* cluster, together with their corresponding mutants as negative controls. Activity of glycosidases were assayed with fluorogenic conjugates of monosaccharide substrates of each particular glycosidase.

I first tested the inhibitor samples against commercially available glycosidases. Inhibitor fractions from both *E. coli* and *P. syringae* inhibited α-galactosidase, β-galactosidase, α-glucosidase and β-glucosidase (Figure 4.20A). Interestingly, the partially purified inhibitor samples isolated from *E. coli* also showed inhibition of β-xylosidase activity, which was not found with the inhibitor samples derived from *P. syringae* (Figure 4.20A). Next, to test biologically relevant targets of the inhibitor, I also tested the inhibition of glycosidase activities found in *N. benthamiana* leaf apoplastic fluid extract. Again, the activity of α-galactosidase, β-galactosidase and β-glucosidase were inhibited (Figure 4.20B).

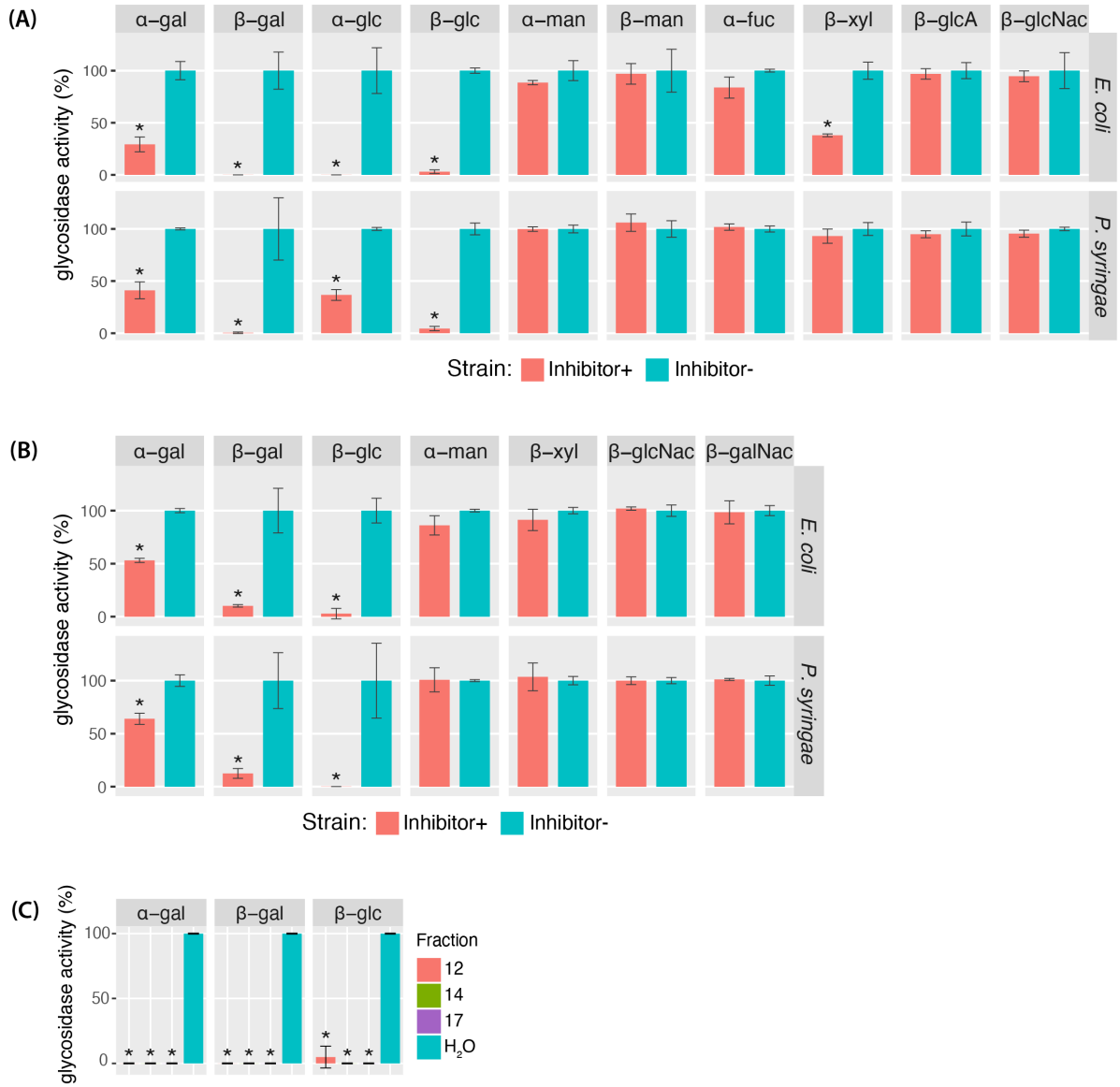


Figure 4.20. The inhibitor isolated from cation exchange purification showed inhibitory activity against multiple glycosidases. Inhibitor eluates (Inhibitor+) from cation exchange resin purification of *E. coli* (p*lac*::*bim*^{wt}) and *P. syringae* HQ cultures, or no inhibitor eluates (Inhibitor-) from cation exchange resin purification of *E. coli* (p*lac*::*bim*^{mut}) and *P. syringae* HQ Δ*bim* cultures were tested for inhibitory activity against a range of glyosidases with 4-MU-conjugated monosaccharide fluorogenic substrates (Table 2). Enzyme activity was expressed as percentage of the activity compared to no inhibitor control. Asterisks denote p-value < 0.05 from t-test comparing inhibitor and no inhibitor in each condition. Mean and standard deviation are shown (n=3). The sources of glycosidase activities were from **(A)** commercially available glycosidases (Table 3), **(B)** *N. benthamiana* leaf apoplastic fluid extract. **(C)** Fractions from HILIC HPLC separation of inhibitor eluate from cation exchange resin purification of *E. coli* (p*lac*::*bim*^{wt}) were tested for inhibition on commercially available glycosidases.

The inhibition of multiple glycosidases could either be because the inhibitor has multiple specificity or there were multiple inhibitors. We, therefore, tested the HILIC isolated inhibitor fractions to determine whether the different glycosidase inhibition could be separated. However, all tested fractions were able to inhibit α-galactosidase, β-galactosidase and β-glucosidase, indicating that these inhibitory activities cannot be separated by HILIC HPLC (Figure 4.20C). However, this still could not rule out the possibility that there were multiple inhibitors because if the inhibitor existed in multiple forms, they are likely to be highly similar and may only differ by stereochemistry in a few positions, making them difficult to separate.

Overall, the isolated inhibitor showed detectable inhibitory activity against α-galactosidase, β-galactosidase, α-glucosidase and β-glucosidase.

4.4 Discussion

In this chapter, I investigated the production, purification, characterisation and identification of the inhibitor. I sought to optimise the production of the inhibitor both in terms of culture conditions and genetic manipulation. Then, I further optimised the purification method to obtain the inhibitor and analytical chemistry methods to detect and characterise the inhibitor.

4.4.1 Optimising inhibitor production

The production of the inhibitor by *P. syringae* and *E. coli* strain can be optimised by testing different culture conditions and strains, taking into account knowledge of the biosynthesis of the inhibitor and its regulation.

First, I investigated the effect of culture media. Since the inhibitor production by *P. syringae* is regulated by type III secretion system (T3SS) regulators, the media that induced the expression of T3SS, such as M9 or MG media, increased inhibitor production. This tight regulation, however, can limit the conditions available for use in inhibitor production. Inhibitor production could be further enhanced through overexpression of HrpL, the sigma factor that activate *bim* cluster expression, or expression in strains that do not produce the T3SS and effectors, which lack competing HrpL-binding sites or complications from overexpression of the T3SS. On the other hand, heterologous expression systems, such as *E. coli* with the constitutively expressed *bim* biosynthetic gene cluster, can be used as an alternative production vessel independent of the native regulation. The media can also affect inhibitor production, for example, if the precursor of the inhibitor synthesis is supplied. However, more detailed knowledge of the inhibitor biosynthetic pathway is required to identify the correct precursor. On the other hand, the screen for substrates that could

increase inhibitor production may provide clues to the biosynthetic pathway or precursor for the inhibitor. Nevertheless, tests using different substrates can still be challenging to interpret since there could be multiple factors involved. For example, utilisation of different substrates may have regulatory effects on gene expression and metabolism (Görke and Stülke, 2008; Rojo, 2010). The interconnected metabolic network can also complicate interpretation of an addition of a substrate, since it can be derived into different metabolic intermediates. I have attempted a screen with a range of substrates as supplements but this yielded inconclusive results (data not shown). In addition, stable isotope-labelled compounds could be used to track the conversion of a precursor in the biosynthetic pathway. For example, the biosynthesis of nectrisine was investigated using stable isotope-labelled ribose and xylose, which enabled confirmation of their roles as biosynthetic precursors based on the presence and distribution pattern of stable isotope labelling (Miyauchi *et al.*, 2015).

In future studies, metabolic engineering approaches could also be used to improve the inhibitor production by the bacteria. The strains can be genetically modified using the knowledge of inhibitor biosynthesis. For example, genes in the biosynthetic pathway of the inhibitor or the pathway leading to the production of precursor of the inhibitor can be overexpressed to increase the flux towards the inhibitor production (Yadav *et al.*, 2012). Furthermore, genes in the competing pathways could be deleted to block the precursor diversion (Yadav *et al.*, 2012). Additionally, genes encoding transporters for inhibitor export can also be overexpressed to pull the flux to promote the production and secretion of the inhibitor (Yadav *et al.*, 2012). In addition, genes involved in the supply of cofactors in the biosynthetic pathway, such as NAD(P)(H), can also be considered target for optimisation (Liu *et al.*, 2018). Another challenge is also to identify the step of the pathway that is a limiting

factor for the inhibitor production since the modification of non-limiting factor would not lead to increased inhibitor production (Woolston, Edgar and Stephanopoulos, 2013). Further improvements may be achieved by enzyme engineering or natural variant screening for better versions of the enzyme with enhanced activity or kinetics (Choi *et al.*, 2019). In this study, overexpression of the *bim* biosynthetic gene cluster could indeed increase inhibitor production in *P. syringae* and confer high level of inhibitor production in *E. coli* heterologous system. However, more knowledge of the biosynthesis of the inhibitor is required for further engineering of other parts of the pathway to improve inhibitor production.

1-deoxynojirimycin (DNJ) is an iminosugar that has been extensively studied, with various strategies for improving microbial production of DNJ reported. For optimisation of culture conditions, studies screening supplementation of a range of carbon sources found different effects on DNJ production in different bacterial strains. For example, sorbitol led to the highest level of DNJ production by *Bacillus amyloquifaciens* (Onose *et al.*, 2013), while sorbose and glucose led to the highest level of DNJ production by *Steptomyces lavendulae* (Wu *et al.*, 2019). In *B. methylotrophicus,* sucrose, glucose and sorbitol all led to high level of DNJ production (Lee *et al.*, 2018). In contrast, lactose and galactose were found to be optimal for DNJ production in other strains of *B. amyloquifaciens* and *B. subtilis* (W. Zhang *et al.*, 2019). For optimisation of biosynthetic enzymes, homologs of GutB1, an alcohol dehydrogenase that participates in DNJ biosynthesis, from different strains were compared to identify the enzyme with highest catalytic efficiency (Wu, Arciola and Horenstein, 2013). Another study screened a random mutagenesis library of GutB1 to identify a clone that exhibit higher activity (Jiang *et al.*, 2015). A range of *Bacillus* strains were also screened, identifying *B. amyloquifaciens* and *B. subtilis* strains that are capable of producing a high level of DNJ (Onose *et al.*, 2013). In

addition, inhibitors of competing metabolic pathways were also used to enhance flux towards DNJ biosynthesis in *S. lavendulae* (Wu *et al.*, 2019). Furthermore, heterologous expression of the biosynthetic gene cluster in *E. coli* was used for DNJ production (Jiang *et al.*, 2015; Rayamajhi *et al.*, 2018). The expression level of each gene in the biosynthetic gene cluster was also fine-tuned by screening different combinations of ribosome-binding sites for the optimal strength that increased DNJ production (Jiang *et al.*, 2015).

4.4.2 Optimising inhibitor purification

Cation exchange resin binding and subsequent elution with NH_4OH is commonly used in iminosugar isolation protocols reported in the literature (Molyneux *et al.*, 2002; Amézqueta and Torres, 2016). Previous study on this particular BGAL inhibitor also established that the inhibitor could be isolated using cation exchange resin and NH_4OH elution (Buscaill *et al.*, 2019). However, I found that the inhibitor is sensitive to basic pH treatment because its exposure to alkaline solutions resulted in a reduced inhibitory activity. The alkaline pH could cause the inhibitor to degrade or undergo other chemical modifications that render it inactive. Either of these cases are not ideal for purification purpose. To overcome this limitation, I adapted the protocol to use HCl for elution instead. This also led to the finding that the inhibitor is stable at acidic pH. The molecular mechanism underlying the acid-stable and base-sensitive property of the inhibitor is still unknown but may become apparent once the identity of the inhibitor is known. This may also provide a clue to the chemical identity of the inhibitor. Interestingly, some pyrrolidine-type iminosugars (5-membered ring) were reported to decompose at high pH (Molyneux, 1993), although no mechanism was described. Furthermore, the stability of the inhibitor at low pH is consistent with its native environment,

the plant apoplast, which usually has acidic pH (~5-6) (Felle, 2001; B. M. O'Leary *et al.*, 2016; Geilfus, 2017).

Another challenge in the isolation by cation exchange resin is salts since cations of salts (such as Na^+, K^+) can also binds to the resin. The presence of high concentration of salts in the eluate can interfere with downstream analysis and chemical reaction or purification. In addition, the high salt also limits the resuspension volume of the isolated inhibitor fraction and hence the concentration of the final fraction. Separating salts from the inhibitor is challenging since both are small and highly hydrophilic. One possible way to alleviate this problem is by desalting with concentrated HCl (Ruud, 2010). However, I demonstrated that although this method was possible, some loss of inhibitor was observed. In addition, the removal of salt was not directly quantified, only visible salt precipitation out of solution was observed. I also tested the reduction of salt in the culture medium, which can allow for more binding of inhibitor to the resin by reducing the competition for resin binding sites between positively charged ions of salts and the inhibitor. The use of low salt medium such as MG instead of M9 medium allowed more input to be loaded onto the cation exchange resin without saturating with cations from salts.

The use of cation exchange resin for inhibitor isolation also co-isolated other positively charged compounds (such as amine containing compounds) making complete isolation of the inhibitor difficult to achieve. Therefore, purification using cation exchange resins needs to be complemented by high-performance liquid chromatography (HPLC), which is a robust and commonly used method for compound purification. My initial trial of HPLC separation of the inhibitor with the commonly used reversed-phase C18 column was unsuccessful. This is consistent with the highly hydrophilic property of the implicated iminosugar nature of the

inhibitor, carrying multiple polar hydroxyl groups and a positive charge on the amine group. A HILIC (hydrophilic interaction liquid chromatography) mode for HPLC is more suitable for separation of polar compounds like carbohydrates since the column stationary phase is more polar and can interact and retain hydrophilic compound better than the hydrophobic C18 column (Nagy, Peng and Pohl, 2017). HILIC has also been used for iminosugar separation (Rodríguez-Sánchez, Quintanilla-López, *et al.*, 2014a; Miyauchi *et al.*, 2015; Nuñez and Horenstein, 2019; Wu *et al.*, 2019). Indeed, HILIC HPLC was found to be more suitable for separation of my inhibitor. Therefore, the purification of the inhibitor was optimised with the coupling of both cation exchange resin and HILIC HPLC. The cation exchange resin is a useful first step for partial purification and enrichment of the inhibitor. It was capable of handling relatively large volume (a few litres) of input and filtering out major unrelated components from relatively complex input sample while capturing the inhibitor. The inhibitor then could be purified from the eluate further with HILIC HPLC.

4.4.3 Characterising inhibitor specificity

One important property of the inhibitor is the specificity and range of glycosidase inhibition. The partially purified inhibitor from cation exchange resin showed inhibition of α-galactosidase, β-galactosidase, α-glucosidase and β-glucosidase activity. This inhibition of multiple glycosidases could either be because multiple inhibitors are present or that the inhibitor is able to interact with multiple enzymes. If the inhibitor existed in multiple forms, they probably only differ by stereochemistry at a few positions. Their structural similarities could make them difficult to separate by HPLC as their elution peaks could be close together. However, with optimal chromatographic conditions, it is possible to separate stereoisomers of iminosugars, such as DNJ and DMJ, isomers with different stereochemistry at one hydroxyl

group (Rodríguez-Sánchez, Quintanilla-López, *et al.*, 2014a). Enzymatic activity inhibition assay coupled with competition of enzymes for inhibitor molecules could be used to test if one or multiple inhibitors exist. Nevertheless, many iminosugars exhibit multiple inhibitory activities against multiple glycosidases (Watson *et al.*, 2001). They may, however, differ in the affinity to different enzymes.

In addition, the substrate range of glycosidases inhibited by the inhibitor could be useful for speculation of the stereochemistry of the structure of the inhibitor molecule (Prichard *et al.*, 2018). For instance, the inhibitor can inhibit both galactosidase and glucosidase activities. Since galactose and glucose share the same stereochemistry at carbon 2 and 3, the inhibitor might have an equivalent configuration at the equivalent part of the molecule that allows it to bind and inhibit the glycosidases acting on these substrates.

Furthermore, the fact that the inhibitor can inhibit multiple enzymes may mean that it could play a role during *P. syringae* infection in the plants by inhibiting multiple glycosidases that are involved in immunity. For example, β-glucosidase can function in various processes including pathogenesis-related protein such as β-glucanase (van Loon, Rep and Pieterse, 2006), or callose degradation (De Storme and Geelen, 2014; Wu *et al.*, 2018), or activation of defence-related secondary metabolites such as glucosinolates or cyanogenic glycosides (Morant *et al.*, 2008). Although not investigated in plant immunity context, iminosugars were found to inhibit the glucosidases responsible for cleavage and activation of glucosinolates (Scofield *et al.*, 1990). Nevertheless, I still need to conduct enzymatic activity assays on these candidate glycosidases and the inhibitor to determine which specific enzymes can be inhibited.

Another interesting observation was the inhibition of β-xylosidase activity by the *E. coli* derived inhibitor but not *P. syringae*-derived inhibitor. This might be due to different metabolic enzymes present in *E. coli* that can modify the inhibitor precursor and give rise to a by-product that inhibits β-xylosidase activity. However, inhibition was not observed for plant β-xylosidase in the apoplast, suggesting potentially specific enzyme case.

4.4.4 Elucidating inhibitor identity

The characterisation of the inhibitor has yielded information on the identity of the inhibitor by integrating both genetic clues from the knowledge of the *bim* biosynthetic gene cluster and chemical clues from the purification and detection of the inhibitor molecule. The knowledge of the *bim* cluster allowed the knock-out strain to be created from the native inhibitor strain of *P. syringae*. In addition, this also allowed the transfer of the cluster to a heterologous system in *E. coli*. Together, they enabled comparative approach for the detection of the inhibitor by using both inhibitor-producing and non-producing strains. Furthermore, the two independent strains provided independent support for the signals detected to be strong putative candidate for the inhibitor.

The purification and detection of the inhibitor molecule also provided additional information on the inhibitor. The isolation by cation exchange resin suggested that the inhibitor contains an amine group, which could carry a positive charge when protonated at acidic pH. This was also confirmed by the ability of the inhibitor to be derivatised by Fmoc-Cl, which reacts with an amine group, on one site on the inhibitor molecule as deduced from the detected mass. The GC-MS mass spectra also identified an indicative fragment of a secondary amine group, consistent with the position of the amine group in the ring. The sugar-like property of the

inhibitor was also suggested by the ability of the inhibitor to be retained in the HILIC column but not the C18 column, suggesting a highly polar nature. The GC-MS mass spectra also identified multiple fragments indicative of multiple hydroxyl groups (potentially 4, with 3 adjacent ones) and potentially of a five-membered ring form.

From both mass spectrometry approaches used, LC-MS and GC-MS, the putatively molecular mass of the inhibitor was consistently identified to be 151 Da. Together with other chemical clues, a putative formula or structure of the inhibitor may be speculated. However, I was still unable to match any structure to the given mass number. For example, 4-deoxy-1,4-imino-D-arabinitol, an iminosugar intermediate in the biosynthesis of nectrisine and DAB-1, has a five-membered ring structure with 4 hydroxyl groups but has a molecular mass of 149 Da, not matching the putative mass of the inhibitor.

Both LC-MS and GC-MS methods used for analysis of the inhibitor required the inhibitor to be chemically derivatised (Fmoc or TMS). These derivatisations were used to overcome the challenge of separation and detection of the inhibitor in these platforms. In addition, they have been used in various other studies of iminosugars (Kim *et al.*, 2003; Rodríguez-Sánchez, Ruiz-Matute, *et al.*, 2014; Amézqueta and Torres, 2016). However, they also posed another limitation since the native inhibitor molecule was not detected. In addition, both derivatisation methods required high pH condition, which I found to be detrimental to the inhibitor. For example, there was a risk that the basic pH caused inhibitor degradation or change in structure and what I detected could be the by-product. Nevertheless, the signal detected from all methods tended to be a single distinguishable peak that were consistent suggesting that it was likely originated from the inhibitor.

The properties of the inhibitor I have identified in this study also allowed comparison with well-characterised natural iminosugars, suggesting that my inhibitor is unlikely any of them. Nojirimycin (NJ, Mw 179 Da) was reported to be unstable at acidic pH (Inouye *et al.*, 1968), while BGAL inhibitor is acid resistant. This instability of NJ was attributed to the presence of the hemiaminal group (amine and hydroxyl group attached to the same carbon) on the carbon 1 position which is prone to rearrangement (Viuff *et al.*, 2015), suggesting that my inhibitor may lack the hydroxyl group on the position 1 carbon. 1-deoxynojirimycin and its galacto- or manno- isomers (DNJ, DGJ and DMJ, Mw 163 Da) are also unlikely since the detected mass from Fmoc derivatisation and LC-MS analysis was different from my inhibitor. Additionally, GC-MS mass spectra and fragmentation patterns of TMS derivatised inhibitor did not match any reported iminosugars (eg. DNJ, DMDP (Mw 163 Da) and fagomine (Mw 147 Da)) (Rodríguez-Sánchez, Ruiz-Matute, *et al.*, 2014). Inhibition range of my inhibitor also rules out some described inhibitors such as nectrisine (Mw 131 Da), which can inhibit mannosidase (Tsujii *et al.*, 1996) and DAB-1 (Mw 133 Da), which can inhibit mannosidase but poorly inhibit β-glucosidase (Nuñez and Horenstein, 2019). Nevertheless, my inhibitor is likely to be relatively simple and the putative molecular mass identified is in the same range of molecular size of these low-molecular weight iminosugars.

Additionally, the biosynthetic enzymes of my inhibitor can also provide clues to the inhibitor property, considering the requirement of the two key biosynthetic enzymes from *bim* cluster (PSPTO0834-5). The alcohol dehydrogenase enzymes homologous to PSPTO0834 are found in multiple bacteria synthesising iminosugars, supporting the idea that my inhibitor could be of similar nature (Clark, Johnson and Horenstein, 2011; Onose *et al.*, 2013; Wu, Arciola and Horenstein, 2013; Nuñez and Horenstein, 2019). The involvement of PSPTO0835 in the

inhibitor biosynthesis also suggests that the inhibitor may not exactly be a previously characterised iminosugar since the homolog was not identified in their proposed biosynthetic pathways. However, I speculate a similarity may exist. The proposed biosynthetic pathways of DNJ and DAB-1 both have the final step that require a reductase to reduce the double bond between the nitrogen and carbon position 1 in the ring (Onose *et al.*, 2013; Nuñez and Horenstein, 2019) and PSPTO0835 is identified as homologous to dihydrofolate reductase which can reduce the C-N double bond in the ring of dihydrofolate (Cao *et al.*, 2018). This suggests that my inhibitor may share the same lack of hydroxyl group at carbon position 1 as DNJ and DAB-1.

Future works for elucidating the structure of the inhibitor can utilise both structural biology and analytical chemistry approaches. Structural biology approach could be used to solve the structure of an enzyme in complex with the inhibitor so that the information of inhibitor structure could be extracted. For example, cryo-EM (electron microscopy) was used to solve the structure of LacZ enzyme in complex with an inhibitor PETG (Bartesaghi *et al.*, 2015). Alternatively, X-ray crystallography approach could be used to solve protein-ligand interaction by crystal soaking or co-crystallisation with the inhibitor (Aguda *et al.*, 2016). Additional analytical chemistry approaches could also be used to characterise the inhibitor identity. High-resolution mass spectrometry (HRMS) can give an accurate mass of the detected inhibitor which can be used to deduce the molecular formula of the inhibitor. This could be coupled with HILIC column chromatography as HILIC-HRMS to allow separation of the native inhibitor molecule prior to mass spectrometry analysis (Wolfender *et al.*, 2019). In addition, NMR (nuclear magnetic resonance) spectroscopy could be used for structure

elucidation of the inhibitor (Molyneux, 1993; Breton and Reynolds, 2013) Finally, the identified structure of the inhibitor should be validated by a chemically synthesised inhibitor.

4.4.5 Conclusion

In this chapter, I have investigated inhibitor production, purification, characterisation and identification. I tested different media for optimising the production of the inhibitor from bacterial culture and tested overexpression of *bim* biosynthetic gene cluster to promote inhibitor production in both *P. syringae* and *E. coli*. The isolation of the inhibitor by cation exchange resin and elution with HCl was optimised to achieve partial purification and enrichment of the inhibitor, which can then be refined further with HILIC HPLC separation. In addition, I also identified two additional important properties of the inhibitor. First, the inhibitor is stable at acidic but not basic pH. Second, the partially purified inhibitor fraction can inhibit multiple glycosidases, including α-galactosidase, β-galactosidase, α-glucosidase and β-glucosidase. Furthermore, I have used LC-MS and GC-MS approach coupled with Fmoc and TMS derivatisation, respectively, to detect the inhibitor, its putative mass of 151 Da and mass spectra. The data support the implicated iminosugar nature of the inhibitor and suggested that the inhibitor could be a five-membered ring with four hydroxyl groups and an amine group. However, the exact structure of the inhibitor remains to be elucidated.

Chapter 5: General discussion

The aim of this work was to investigate the production of BGAL inhibitor by *P. syringae* with respect to its biosynthesis, evolution, regulation, biological roles and molecular identity. I have made progress towards answering these questions and discovered some new leads for further investigation. This section will summarise major findings from this work (Figure 5.1) and discuss future directions.

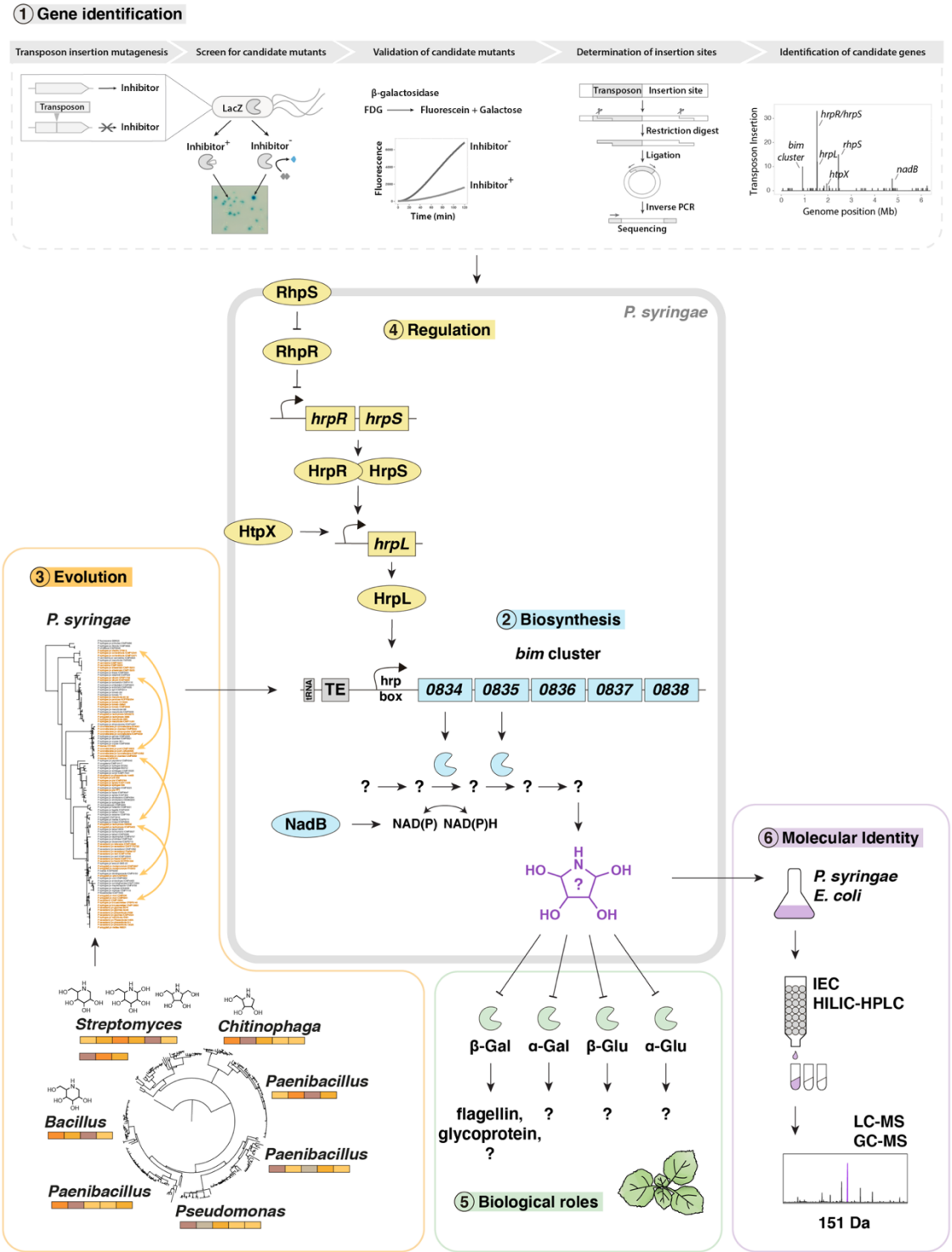


Figure 5.1. Schematic diagram summarising main findings of this thesis.
(1) Identification of genes required for inhibitor production. I developed a high-throughput workflow for screening of BGAL inhibitor mutants (*bim*). First, random transposon insertion mutagenesis was used to generate a library of *P. syringae* mutants, some of which have genes required for inhibitor production disrupted by the transposon. Next, the ability of the mutants to produce the inhibitor was distinguished using the expression of LacZ β-galactosidase together with its chromogenic substrate X-gal as a reporter system for enzymatic activity and inhibition *in situ*. The darker blue color developed from higher enzyme activity in colonies with no inhibitor production allowed isolation of the candidate *bim* mutants from agar plates. These candidates were then validated *in vitro* for their inability to produce the inhibitor by testing their culture supernatant with a 96-well plate based enzymatic activity assay using standard LacZ β-galactosidase and its fluorogenic substrate FDG, which allow for sensitive quantification of enzymatic activity and detection of inhibitor. After confirming the *bim* mutants, transposon insertion sites were identified by inverse PCR and sequencing and then mapped on to the genome to identify candidate genes.
(2) Biosynthesis of the inhibitor. I identified the *bim* gene cluster as the biosynthetic gene cluster of BGAL inhibitor. This cluster consists of 5 genes (PSPTO0834-8) but the first two genes together, PSPTO0834 and PSPTO0835, were already sufficient to confer BGAL inhibitor production in both *P. syringae* and *E. coli*. There are likely other enzymes in the biosynthetic pathway of the inhibitor that I have not yet identified. The precursors and intermediates of the pathway also remain to be identified. In addition, *nadB* was also identified to be involved in inhibitor production likely by impacting the level of NAD(P) required as cofactors for at least two key oxidoreductases enzymes, PSPTO0834 and PSPTO0835.
(3) Evolution of inhibitor production. From the search for homologs of PSPTO0834, I identified a range of potential iminosugar biosynthetic gene clusters with different patterns of recurring genes in various bacterial taxa, both uncharacterised and previously described for production of iminosugars. In *P. syringae*, the *bim* cluster is widespread among strains and is likely to have been acquired through multiple instances of horizontal gene transfer. The *bim* cluster is found in association with transposable elements (TE) adjacent to the tRNA-lysine site in *P. syringae* genomes.
(4) **Regulation of inhibitor production**. The promoter of the *bim* cluster contains a *hrp* box consensus sequence, the binding site for the HrpL sigma factor, explaning the dependence of its expression on HrpL. As part of the HrpL regulon, the expression of the *bim* cluster is induced under the regulatory cascade of the T3SS regulators RhpR/S and HrpR/S. In addition, HtpX, a membrane bound protease, was also identified as a potential positive regulator of the T3SS and inhibitor production required for normal transcription levels of *hrpL* and its regulon.
(5) Biological roles of the inhibitor. I found that the inhibitor produced by the *bim* cluster can suppress not only β-galactosidase but also α-galactosidase, β-glucosidase and α-glucosidase activities. The identification of specific glycosidases that are sensitive to the inhibitor and their substrates may shed light on the biological roles of inhibition of these enzymes during infection. For BGAL1, multiple substrates may be targeted, including glycosylated flagellin and apoplastic glycoproteins.
(6) Molecular identity of the inhibitor. I could produce the inhibitor from *in vitro* cultures of both *P. syringae* and *E. coli* carrying the *bim* cluster. Inhibitor purification can be done using two approaches: cation exchange resin (IEC) and HILIC-HPLC (hydrophilic interaction liquid chromatography-high performance liquid chromatography). For inhibitor detection, I used two approaches combining chemical derivatisation and mass spectrometry, namely Fmoc derivatisation with LC-MS and TMS derivatisation with GC-MS. I was able to identify the putative signals of the inhibitor and deduce a putative molecular mass of the inhibitor as 151 Da. The analyses also suggested that the inhibitor contain an amine group and possibly four hydroxyl groups. However, I could not identify a match with any putative structures or previously characterised iminosugars.

5.1 Identification of genes required for inhibitor production (Figure 5.1.1)

In chapter 2, I developed a high-throughput workflow for screening of BGAL inhibitor mutants (*bim*) to identify genes that are involved in inhibitor production. First, random transposon insertion mutagenesis was used to generate a library of *P. syringae* mutants, some of which have genes required for inhibitor production disrupted by the transposon, rendering them unable to produce the inhibitor. Next, the ability of the mutants to produce the inhibitor was distinguished using the expression of LacZ β-galactosidase together with its chromogenic substrate X-gal as a reporter system for enzymatic activity and inhibition *in situ*. The darker blue color developed from higher enzyme activity in colonies with no inhibitor production allowed isolation of the candidate *bim* mutants from agar plates. These candidates were then validated *in vitro* for their inability to produce the inhibitor by testing their culture supernatant with a 96-well plate based enzymatic activity assay using standard LacZ β-galactosidase and its fluorogenic substrate FDG, which allow for sensitive quantification of enzymatic activity and detection of inhibitor. After confirming the *bim* mutants, transposon insertion sites were identified by inverse PCR and sequencing and then mapped on to the genome to identify candidate genes. In addition, I also incorporated comparative genomics and transcriptional profiling data to further support the candidate genes. From this screen, I successfully identified genes involved in both inhibitor biosynthesis and regulation. However, this screen might still miss other genes that also function in inhibitor production, such as essential genes whose disruption is lethal or functionally redundant genes whose single disruption is not enough to stop inhibitor production. The screen likely also missed genes involved in inhibitor secretion since the intracellular LacZ expression was used for detection of enzyme inhibition. Nevertheless, this approach is useful for screens involving enzymatic

activity or inhibition and may be applied in other systems. For example, a similar approach using β-glycosidase expression in *E. coli* and X-gal substrate on plate was demonstrated for screening of a biosynthetic enzyme GutB1 random mutagenesis library for enhanced activity and ribosome binding site library for optimised expression levels, both to increase production of DNJ inhibitor (Jiang *et al.*, 2015).

5.2 Biosynthesis of the inhibitor (Figure 5.1.2)

In chapter 3, I identified the *bim* gene cluster as the biosynthetic gene cluster of the BGAL inhibitor. This cluster consists of five genes (PSPTO0834-8) but the first two genes together, PSPTO0834 and PSPTO0835, were sufficient to confer BGAL inhibitor production in both *P. syringae* and *E. coli*. The first gene, PSPTO0834, encodes an alcohol dehydrogenase in the L-threonine dehydrogenase family, which is homologous to GutB1, an enzyme implicated in the biosynthesis of an iminosugar inhibitor deoxynojirimycin (DNJ) (Clark, Johnson and Horenstein, 2011). GutB1 catalyses the oxidation of a hydroxyl group to ketone group that then spontaneously reacts with the amine group to form the iminosugar ring structure with an N atom (Figure 1.5A). This is consistent with my hypothesis that the BGAL inhibitor is an iminosugar and can explain the requirement of PSPTO0834 for BGAL inhibitor biosynthesis. The second gene, PSPTO0835, encodes an uncharacterised reductase that belongs to dihydrofolate reductase family. Dihydrofolate reductase enzyme is known to function in the reduction of a double bond between C and N atom in the ring of dihydrofolate (Cao *et al.*, 2018). This reaction is reminiscent of the last step of the proposed DNJ biosynthesis where the double bond between C and N in the iminosugar ring is reduced by an unidentified reductase (Figure 1.5A), suggesting that PSPTO0835 might function in a similar reaction.

In addition to PSPTO0834-5, there are likely other enzymes in the biosynthetic pathway of the inhibitor that I have not yet identified. For example, according to the common framework of previously characterised iminosugar biosynthesis (DNJ and DAB-1) (Clark, Johnson and Horenstein, 2011; Nuñez and Horenstein, 2019), an aminotransferase and a phosphatase are also likely to be involved in the BGAL inhibitor biosynthesis. Furthermore, I also found that fractions isolated from *E. coli* expressing the *bim* cluser, but not *P. syringae,* could inhibit β-xylosidase, suggesting that they may possess slightly different genes that permit biosynthesis of another similar molecule. To identify the remaining genes, targeted manipulation of candidate genes (eg. homologs that could be identified in the *P. syringae* genome) could be used, for example, by knocking out or silencing genes to determine if inhibitor production is impacted. In case of lethality of gene knockout, silencing or inducible silencing could be used. The choice of culture media could also address potential auxotrophy of knockout mutants. For gene redundancy, simultaneous silencing or multiple gene knockout could be used. On the other hand, overexpression may also be used to test for genes in the biosynthetic pathway, which could increase inhibitor production, provided that other parts of the pathway are not limiting factors. In addition, when conducting these tests, it could be useful to not only test whether the inhibitor is produced but also quantify the amount of inhibitor produced since the effect of these genes might be quantitative.

To further characterise the inhibitor biosynthetic enzymes, an assay of purified enzymes can be carried out to test the activity, substrate specificity and product generation. If the full pathway could be identified, an *in vitro* reconstruction of the pathway for inhibitor production could be tested for confirmation. For example, the biosynthesis of phevamineA was reconstructed *in vitro* (O'Neill *et al.*, 2018). However, these assays could be challenging since

the substrate of the pathway has to also be identified. Alternatively, a metabolomics approach could be taken. Metabolite profiling may be used to identify metabolites that could be substrates and products in metabolite extracts of the biosynthetic enzyme mutant, wild type and overexpressor strains; or a mixture of cellular metabolite extract with purified enzymes. For instance, substrates of an enzyme may be predicted from metabolites that are depleted in the metabolome of the strain overexpressing the enzyme or upon incubation of cellular metabolite extract with the purified enzyme (Prosser, Larrouy-Maumus and Carvalho, 2014; Sévin *et al.*, 2017). In addition, stable isotope-labelled compounds could be used to track the conversion of a substrate in the biosynthetic pathway. For example, the biosynthesis of nectrisine was investigated using stable isotope-labelled ribose and xylose, which enabled confirmation of their roles as biosynthetic precursors based on the presence and distribution pattern of stable isotope labelling (Miyauchi *et al.*, 2015).

Apart from the *bim* cluster, *nadB* was also identified to be involved in inhibitor production, although not as an absolute requirement. *nadB* encodes an L-aspartate oxidase functioning in *de novo* biosynthesis of NAD (Kurnasov *et al.*, 2003), which is likely involved in inhibitor production by impacting the level of NAD(P) required as cofactors for at least two key oxidoreductase enzymes, PSPTO0834 and PSPTO0835. This hypothesis could be further confirmed by measurement of cellular NAD(P) levels. It also remains to be determined which form of the cofactor is required for inhibitor biosynthetic enzymes. An enzymatic assay with the purified biosynthetic enzyme may be carried out to test the cofactor they accept.

The discovery of BGAL inhibitor biosynthetic genes and pathways will not only add to the knowledge of the functions of the genes involved but may also inform applications such as metabolic engineering of inhibitor production in bacteria (Kim *et al.*, 2016; Choi *et al.*, 2019;

Ko *et al.*, 2020; Yang *et al.*, 2020). This may include transfer to heterologous systems or enhancement of production in the native host. For example, the biosynthetic enzymes can be overexpressed to boost inhibitor production. The components of the pathway may also be further fine-tuned to improve the production. In addition, the impact of NAD(P) levels on inhibitor production suggests a possibility of overexpressing genes for NAD(P)(H) generation and inhibitor biosynthetic genes to increase the availability of the cofactor for the enzymes to drive inhibitor biosynthesis.

5.3 Evolution of inhibitor production (Figure 5.1.3)

In chapter 3, I explored the evolution of the *bim* cluster and inhibitor production both in *P. syringae* and a broader range of bacterial species. First, since PSPTO0834 and its homologs were found to be involved in biosynthesis of both BGAL inhibitor and previously characterised iminosugars (DNJ (Clark, Johnson and Horenstein, 2011) and DAB-1 (Nuñez and Horenstein, 2019)), I reasoned that this alcohol dehydrogenase group is a key distinguishing factor of iminosugar biosynthesis. Indeed, from the search for homologs of PSPTO0834, I identified a range of potential iminosugar biosynthetic gene clusters in various bacterial taxa, both uncharacterised and previously described for production of iminosugars. In addition, the identified gene clusters contain different patterns of recurring genes including alcohol dehydrogenase, aminotransferase, phosphatase (HAD or FIG family), reductase (RibD family), transporters (MFS, MATE, SulP family) and glycosidases, suggesting possible common evolutionary origin and diversification. This discovery of gene clusters and strains that carry them may lead to identification of new iminosugar-producing bacteria or new iminosugars and their biosynthetic pathways. In addition, with further characterisation, the different

homologs of genes found in these strains could provide resources for metabolic engineering of iminosugar production.

Next, I examined the evolution of the *bim* cluster. From the phylogenetic tree of putative iminosugar biosynthetic gene clusters, the *bim* clusters from *P. syringae* are grouped together in one clade, which is nested within a clade of all *Pseudomonas* species. In addition, *bim* clusters from *P. syringae* are always associated with transposable elements and uniquely contain a conserved *hrp* box sequence in the promoter region for regulation by the type III secretion system regulator HrpL, which is induced during infection in plants (Lam *et al.*, 2014). These findings suggest that the *bim* cluster was transferred once into an ancestor of *Pseudomonas* before being inherited by different *Pseudomonas* species, including once by an ancestor of *P. syringae*. Subsequently, this ancestral *bim* cluster gained the *hrp* box sequence in its promoter as well as the association with transposable elements, which then mediated its horizontal gene transfer among multiple strains of *P. syringae*. Many of these strains maintained the *bim* cluster and passed it on to their progeny strains. This inheritance pattern can be observed in the distribution of the *bim* cluster across the phylogenetic tree of *P. syringae* and the incongruence between the *P. syringae* species tree and *bim* cluster gene PSPTO0834 tree. In addition, the transfer of the *bim* cluster among *P. syringae* strains was likely mediated by a tRNA-lysine site-specific integrase since *bim* clusters are always found next to one of the two conserved tRNA-lysine sites in the genome. Integrases are also commonly known to target specific tRNA genes (Williams, 2002). Since the *bim* cluster was repeatedly acquired and maintained by multiple strains of *P. syringae*, BGAL inhibitor production could confer a selective advantage, possibly through its virulence functions during infection. Similarly, many pathogenicity islands with virulence-related genes, such as genes

for effector proteins, toxin biosynthesis and type III secretion system, also evolved through horizontal gene transfer (Arnold, Pitman and Jackson, 2003).

5.4 Regulation of inhibitor production (Figure 5.1.4)

In chapter 3, I found that BGAL inhibitor production is under the control of RhpS, HrpR, HrpS and HrpL, which are conventional regulators of the type III secretion system (T3SS), a major virulence mechanism of *P. syringae*. Mutants of these T3SS regulators impaired the expression of the *bim* biosynthetic gene cluster. Furthermore, the promoter of the *bim* cluster contains a *hrp* box consensus sequence, the binding site for HrpL sigma factor, consistent with a direct transcriptional regulation by HrpL (Vencato et al., 2006). As part of the HrpL regulon, the expression of the *bim* cluster is induced under the regulatory cascade of T3SS regulators (Xie, Shao and Deng, 2019). Consistently, I observed the induction of *bim* cluster expression, and inhibitor production, both *in planta* during the early hours of infection and *in vitro* in T3SS-inducing minimal medium such as MG medium (Bronstein *et al.*, 2008). Furthermore, the co-regulation of the *bim* cluster with the T3SS suggests that the inhibitor could have a role during infection. This study on the *bim* cluster regulation provides another example of HrpL regulated genes that are not the T3SS or effector proteins.

In addition to conventional T3SS regulators, I also identified HtpX as a potential regulator of T3SS and inhibitor production. HtpX is a membrane-bound protease implicated in proteolytic quality control of membrane proteins (Sakoh, Ito and Akiyama, 2005). I found that HtpX was required for normal transcription of *hrpL* and its regulon, including the *bim* cluster and effector genes. Interestingly, a mutant of *htpX* was also identified in a transposon mutagenesis screen as one of the mutants defective in the induction of T3SS *in plant*a,

consistent with my findings (Xiao, 2005). Thus, HtpX may function as a positive regulator of *hrpL* through its proteolytic activity that either degrades another negative regulator or activates another positive regulator of *hrpL* expression. The identification of proteolytic substrates targeted by HtpX will help reveal the mechanism of HtpX action. Several membrane proteins, especially sensor histidine kinases such as RhpS, CorS and GacS, are involved in regulatory pathways that impact *hrpL* expression so they could be candidate proteins targeted by HtpX (Xie *et al.*, 2020). Furthermore, comparative transcription profiling of wild type bacteria and the *htpX* mutant may also reveal the genes or pathways modulated by HtpX. In addition, it would also be interesting to determine whether HtpX function is related to temperature since its homologs in other bacteria can be either inducible and non-inducible by heat (Sakoh, Ito and Akiyama, 2005). In *P. syringae*, production of the T3SS *hrp* pilli and some toxins, such as coronatine, is repressed at elevated temperature higher than 25°C (Elina Roine *et al.*, 1997; Smirnova *et al.*, 2002). In contrast, I still detected BGAL inhibitor production at 28°C. Overall, this study on HtpX may lead to a discovery of a novel regulator and mechanism that controls T3SS in *P. syringae*.

5.5 Biological roles of the inhibitor (Figure 5.1.5)

In chapter 3, I investigated the biological roles of BGAL inhibitor production by *P. syringae* during infection. Several lines of evidence hint at a possible role of the inhibitor in virulence. First, from my findings on the evolution and regulation discussed above, the *bim* cluster is regulated by the regulatory cascade that also controls the expression of the type III secretion system, a major virulence mechanism of *P. syringae* induced during infection. Current evidence suggests that the *bim* cluster was also horizontally transferred multiple times and became widespread among *P. syringae* strains. Second, the plant BGAL1 enzyme targeted by

the inhibitor has a function in plant defence and BGAL1 knockout in *N. benthamiana* plant led to more susceptibility to *P. syringae* (Buscaill *et al.*, 2019). Third, a previous study found the PSPTO0834 mutant *P. syringae* to be less virulent in Arabidopsis plant infection (Vencato *et al.*, 2006). However, my attempts to test the inhibitor mutant strain did not show a consistent and robust difference from the wild type strain with both infiltration and spray infections on both *N. benthamiana* and *A. thaliana* leaves. There may be possible explanations for this observation. The effect of the inhibitor might be small so more replicates might be needed to produce a statistically significant result. The small effect could also be because the level of inhibitor production by the bacteria was insufficient for complete inhibition of BGAL1 to show a phenotype as strong as the *bgal1* knockout plant. The effect of the inhibitor might be more pronounced at a late stage of infection with high bacterial population and inhibitor accumulation. In addition, the redundancy of multiple virulence mechanisms, including effector proteins and toxins, might mask the effect of the inhibitor. This was also shown in a study on phevamine A toxin production in *P. syringae*, which showed a phenotype only in the genetic background of a strain that lacks effector proteins and coronatine toxin (O'Neill *et al.*, 2018). In my study, I found no significant effects of inhibitor production in infection assays using the *hrpA* knockout *P. syringae* strain, which is unable to secrete effector proteins but still produces coronatine. In future work, I could also test the effect of inhibitor production in a polymutant strain lacking effector proteins and coronatine and the BGAL inhibitor.

I have also discovered possible indications of broader roles of inhibitor production than previously focused on, as multiple glycosidases and substrates can be targeted by the inhibitor production. First, the inhibitor likely functions to suppress BGAL1 activity against more substrates than the modified-viosamine-containing glycosylated flagellin. From my

phylogenetic analysis, the *bim* cluster is not restricted to the *P. syringae* strains with the modified viosamine-glycosylated flagellin that is sensitive to plant BGAL1, suggesting that those strains might use the inhibitor to target BGAL activities on other substrates. In addition, BGAL inhibitor production resulted in an accumulation of galactosylated glycoprotein signals in the leaf apoplastic fluid samples, suggesting that an alteration of apoplastic protein glycosylation profile might be involved in the inhibitor function. Determination of the identity and function of these glycoproteins may shed light on the biological roles of this effect during infection. For example, lectin could be used to isolate glycoproteins with specific glycosylation which could then be identified by mass spectrometry.

From the protein blot probed with RCAI, a galactose binding lectin, I identified the accumulation of high-molecular weight signals reminiscent of the highly glycosylated arabinogalactan proteins of the hydroxyproline-rich glycoprotein (HRGP) family (Malinovsky, Fangel and Willats, 2014), thus this protein family is one of the candidate targets of BGAL. HRGPs are cell wall associated proteins implicated in various processes such as development, growth and signaling (Seifert and Roberts, 2007; Ellis *et al.*, 2010; Malinovsky, Fangel and Willats, 2014). Interestingly, HRGP can be involved in crosslinking and strengthening of cell wall and found in cell wall papillae formed in response to pathogens (Deepak *et al.*, 2010; Underwood, 2012). Expression of genes encoding HRGP was also found to be suppressed by the *P. syringae* effector AvrPto in *Arabidopsis* suggesting a possible role in extracellular defense (Hauck, Thilmony and He, 2003). Another study reported that an HRGP protein in Arabidopsis was required for efficient *Agrobacterium* transformation suggesting a possible role for this protein in facilitating either bacterial attachment or type IV secretion or plant defence signaling (Gaspar *et al.*, 2004). The inhibition of BGAL activity could prevent cleavage

of galactose residues from the glycan of HRGP and alter its functions in plant defence or bacterial colonisation.

Second, the inhibitor produced by the *bim* cluster can suppress not only β-galactosidase but also β-glucosidase, α-glucosidase and α-galactosidase activities. The identification of specific glycosidases that are sensitive to the inhibitor and their substrates may shed light on the biological roles of the inhibition of these enzymes during infection. For example, β-glucosidase activity can function in various plant defence-related roles, such as the pathogenesis-related protein β-glucanase that cleaves β-glucan in the cell wall of filamentous pathogens (van Loon, Rep and Pieterse, 2006), or the degradation of callose deposition (De Storme and Geelen, 2014; Wu *et al.*, 2018), or the activation of defence-related secondary metabolites such as glucosinolates and cyanogenic glycosides (Morant *et al.*, 2008). In addition, β-glucosidase is also involved in the trimming of xyloglucan, a component of the plant cell wall (Sampedro *et al.*, 2017). Xyloglucan heteropolymer processing also requires actions of multiple glycosidases including β-galactosidase (Iglesias *et al.*, 2006; Sampedro *et al.*, 2012). By inhibiting these glycosidases, the inhibitor might have an impact on cell wall dynamics during infection. For α-glucosidase, I did not detect its activity in *N. benthamiana* apoplastic fluid. However, α-glucosidases in the endoplasmic reticulum play an important role in glycoprotein folding and its disruption by iminosugars has been shown to impact N-glycoprotein production (Mega, 2004). The inhibitor secreted by *P. syringae* in the apoplast needs to be able to enter the plant cells and accumulate at a sufficient level to have the inhibitory effect on glycoprotein processing. For α-galactosidase, its roles are not well-characterised but have been suggested to be involved in cell wall modification during cell expansion (Chrost *et al.*, 2007). Overall, it would be interesting to test which of these

additional enzymes are also suppressed by the inhibitor. Furthermore, the inhibitor might have different affinities towards different enzymes so some enzymes could be preferential targets of the inhibitor.

The inhibitor could also have functions in contexts that were not tested in my leaf infection assay, such as in other tissues or stages of development of the plant or in interactions with other microbes or insects. Multiple co-infecting pathogens can compete and exclude one another or cooperate to promote their colonisation in plant tissues (Abdullah *et al.*, 2017). For example, a fungal plant pathogen was shown to secrete an effector protein with antimicrobial activities to modulate host microbiomes by suppressing antagonistic bacteria (Snelders *et al.*, 2020). Microbial production of secondary metabolites have also been studied in complex interactions within a microbiome that affects disease outcomes (Carrión *et al.*, 2019). The inhibitor might act as an antibiotic or inhibit glycosidases used by another microbe to gain nutrients. The inhibitor might also provide protection against plant defence to other microbes, such as plant β-glucanase against fungal cell wall, so they can synergistically promote plant colonisation. Alternatively, interactions with insects could be involved. Iminosugar inhibitors are produced by many plants to function as a defence against herbivores by inhibiting their digestive enzymes (Stiitz, 1999). The inhibitor may be produced in infected tissue to deter consumption by herbivores. On the other hand, since some insects can act as a vector for *P. syringae* (Stavrinides, McCloskey and Ochman, 2009), the inhibitor might also suppress plant defence against insects, such as glucosinolate activation, or enhance survival in the insect gut.

Overall, the study on the roles of glycosidase inhibitor in plant-pathogen interaction highlights the potentially important and diverse contributions of glycobiology that remain to be

uncovered. This might lead to discovery of new mechanisms of virulence or new functions of glycans and glycosidases in plant defence.

5.6 Molecular identity of the inhibitor (Figure 5.1.6)

In chapter 4, I investigated inhibitor production, purification and detection in order to characterise the molecular identity and properties of the inhibitor. The knowledge of the *bim* biosynthetic gene cluster enabled genetic manipulations to generate a knock-out strain of the native *P. syringae* host and an overexpression strain of the heterologous *E. coli* host. Together, inhibitor-producing and non-producing strains of both bacterial hosts allowed me to conduct a comparative analysis for inhibitor detection. For inhibitor production, I could produce the inhibitor from *in vitro* cultures of both *P. syringae* and *E. coli* carrying the *bim* cluster.

For inhibitor purification, I tested two approaches that can be used to isolate the inhibitor. First, cation exchange resin can be used to bind the inhibitor, then the resin can be washed with water and ethanol to remove other impurities and finally the inhibitor can be eluted from the resin with HCl solution. This cation exchange resin method is commonly used for isolation of iminosugars since the amine group can be protonated and carry a positive charge for binding with the resin (Molyneux *et al.*, 2002; Amézqueta and Torres, 2016). Although elution was normally done with a basic NH_4OH solution, I avoided NH_4OH because I found that basic pH inactivates my inhibitor. One limitation of this cation exchange resin approach is the co-isolation of positively charged compounds, such as amine containing compounds, so only partial purification was possible. Second, a HILIC-HPLC (hydrophilic interaction liquid chromatography - high performance liquid chromatography) can also be used for separation of the inhibitor. The HILIC-HPLC is suitable for use with highly polar hydrophilic compounds

(Nagy, Peng and Pohl, 2017) and has been used for many iminosugars studies (Rodríguez-Sánchez, Quintanilla-López, *et al.*, 2014a; Miyauchi *et al.*, 2015; Nuñez and Horenstein, 2019; Wu *et al.*, 2019). Separation using HILIC-HPLC of the eluate from the cation exchange resin led to further inhibitor enrichment. Therefore, the purification of the inhibitor was optimised with the coupling of both cation exchange resin and HILIC-HPLC.

For inhibitor detection, I used two approaches combining chemical derivatisation and mass spectrometry, namely Fmoc derivatisation with LC-MS and TMS derivatisation with GC-MS. These techniques have also been used in various other studies of iminosugars (Kim *et al.*, 2003; Rodríguez-Sánchez, Ruiz-Matute, *et al.*, 2014; Amézqueta and Torres, 2016). Using these analyses to compare samples from the inhibitor-producing and non-producing strains, I was able to identify the putative signals of the inhibitor and deduce a putative molecular mass of the inhibitor as 151 Da. The analyses also yielded additional information on the chemical composition of the inhibitor, such as the presence of an amine group and possibly four hydroxyl groups. However, I could not identify a match with any putative structures or previously characterised iminosugars. In addition, one main limitation of my detection approaches is the requirement of chemical derivatisation, with Fmoc or TMS, at high pH conditions, which could be detrimental to the inhibitor. There is a risk that the basic pH caused inhibitor degradation or changes in structure and what I detected could be a by-product. I also did not detect the native inhibitor molecule with the approaches used. Nevertheless, based on my characterisation, the inhibitor is likely to be a relatively simple low-molecular weight iminosugar.

Future work for elucidating the structure of the inhibitor may involve both structural biology and analytical chemistry approaches. Structural biology approaches could be used to solve

the structure of the inhibitor in complex with its enzyme target. For example, cryo-EM (electron microscopy) was used to solve the structure of LacZ enzyme in complex with the inhibitor phenylethyl-β-D-galactopyranoside (PETG) (Bartesaghi *et al.*, 2015). Alternatively, X-ray crystallography approach could be used to solve protein-ligand interaction by crystal soaking or co-crystallisation with the inhibitor (Aguda *et al.*, 2016). Additionally, analytical chemistry approaches could also be used to characterise the inhibitor identity. High-resolution mass spectrometry (HRMS) can give an accurate mass of the detected inhibitor which can be used to deduce the molecular formula of the inhibitor. This could be coupled with HILIC column chromatography as HILIC-HRMS to allow separation of the native inhibitor molecule prior to mass spectrometry analysis (Wolfender *et al.*, 2019). In addition, NMR (nuclear magnetic resonance) spectroscopy could be used for structure elucidation of the inhibitor (Molyneux, 1993; Breton and Reynolds, 2013). Finally, the identified structure of the inhibitor could be validated by a chemically synthesised inhibitor. The successful completion of this work will elucidate a new iminosugar molecule produced by *P. syringae* and the approach may be useful for future iminosugar discoveries.

5.7 Conclusion

I have investigated the production of the BGAL inhibitor by *P. syringae* with respect to its biosynthesis, evolution, regulation, biological roles and molecular identity. I first developed a screen for *P. syringae* BGAL inhibitor mutants (*bim*) that led to the identification of a number of genes involved in inhibitor production. For the biosynthesis of the inhibitor, I have identified the *bim* cluster as the biosynthetic gene cluster of the BGAL inhibitor. I found that the *bim* cluster genes PSPTO0834 and PSPTO0835 together were sufficient to initiate inhibitor production in both *P. syringae* and *E. coli*. Although the full biosynthetic pathway of the

inhibitor remains to be characterised, overall, it is likely to be similar to the framework of previously described iminosugars. Next, I investigated the evolution of the *bim* cluster and found putative iminosugar biosynthetic gene clusters homologs in a range of bacterial species. The *bim* cluster was likely horizontally transferred to *P. syringae* and subsequently transferred between different strains to become widespread. Furthermore, I also characterised the regulation of inhibitor production in *P. syringae* by the conventional type III secretion system (T3SS) regulators (RhpS, HrpR, HrpS, HrpL) and identified HtpX as a new potential regulator of inhibitor production and T3SS. Additionally, I explored the role of the BGAL inhibitor in bacterial virulence during plant infection. Although no robust virulence phenotype was observed, I have discussed other possibilities that can be explored in future work and have identified additional potential targets of the inhibitor such as the modification of apoplastic glycoproteins. Moreover, I found that the inhibitor fraction can inhibit multiple glycosidases, including α-galactosidase, β-galactosidase, α-glucosidase and β-glucosidase, suggesting that multiple pathways in host plants could be targeted. Identification of additional enzymes targeted by the inhibitor could reveal new players in plant-pathogen interactions. Finally, I have also made progress in the characterisation of the inhibitor. The isolation of the inhibitor by cation exchange resin and elution with HCl was optimised to achieve partial purification and enrichment of the inhibitor, which can then be refined further with HILIC HPLC separation. In addition, I also found that the inhibitor is stable at acidic but not basic pH. Furthermore, I have used LC-MS and GC-MS approach coupled with Fmoc and TMS derivatisation, respectively, to detect the inhibitor, its putative mass of 151 Da and mass spectra. Current data support the implicated iminosugar nature of the inhibitor and suggest that the inhibitor could be a five-membered ring with four hydroxyl groups and an amine group. The molecular structure of the inhibitor remains to be elucidated.

References

Abdullah, A. S. *et al.* (2017) 'Host–multi-pathogen warfare: pathogen interactions in co-infected plants', *Frontiers in Plant Science*, 8, pp. 1–12. doi: 10.3389/fpls.2017.01806.

Abramovitch, R. B., Anderson, J. C. and Martin, G. B. (2006) 'Bacterial elicitation and evasion of plant innate immunity', *Nature Reviews Molecular Cell Biology*, 7(8), pp. 601–611. doi: 10.1038/nrm1984.

Agarwala, R. *et al.* (2016) 'Database resources of the National Center for Biotechnology Information', *Nucleic Acids Research*. doi: 10.1093/nar/gkv1290.

Aguda, A. H. *et al.* (2016) 'Affinity crystallography: a new approach to extracting high-affinity enzyme inhibitors from natural extracts', *Journal of Natural Products*, 79(8), pp. 1962–1970. doi: 10.1021/acs.jnatprod.6b00215.

Aguilera, S. *et al.* (2017) 'Temperature-mediated biosynthesis of the phytotoxin phaseolotoxin by *Pseudomonas syringae* pv. p*haseolicola* depends on the autoregulated expression of the *phtABC* genes', *PLoS ONE*, 12(6), pp. 1–15. doi: 10.1371/journal.pone.0178441.

Ahn, Y. O. *et al.* (2007) 'Functional genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 35', *Phytochemistry*, 68(11), pp. 1510–1520. doi: 10.1016/j.phytochem.2007.03.021.

Ahuja, I., Kissen, R. and Bones, A. M. (2012) 'Phytoalexins in defense against pathogens', *Trends in Plant Science*, 17(2), pp. 73–90. doi: 10.1016/j.tplants.2011.11.002.

Alfano, J. R. *et al.* (2000) 'The *Pseudomonas syringae* Hrp pathogenicity island has a tripartite mosaic structure composed of a cluster of type III secretion genes bounded by exchangeable effector and conserved effector loci that contribute to parasitic fitness and pathogenicity in plants', *Proceedings of the National Academy of Sciences of the United States of America*, 97(9), pp. 4856–4861. doi: 10.1073/pnas.97.9.4856.

Allard-Massicotte, R. *et al.* (2016) '*Bacillus subtilis* early colonization of *Arabidopsis thaliana* roots involves multiple chemotaxis receptors', *mBio*, 7(6), pp. 1–10. doi: 10.1128/mBio.01664-16.

Alonzi, D. S. *et al.* (2017) 'Iminosugar antivirals: The therapeutic sweet spot', *Biochemical Society Transactions*, 45(2), pp. 571–582. doi: 10.1042/BST20160182.

Altschul, S. F. *et al.* (1997) 'Gapped BLAST and PSI-BLAST: A new generation of protein database search programs', *Nucleic Acids Research*. doi: 10.1093/nar/25.17.3389.

Amézqueta, S. and Torres, J. L. (2016) 'Advances in the analysis of iminocyclitols: Methods, sources and bioavailability', *Talanta*, 151, pp. 157–171. doi: 10.1016/j.talanta.2016.01.034.

Anderson, C. T. and Kieber, J. J. (2020) 'Dynamic construction, perception, and remodeling of plant cell walls', *Annual Review of Plant Biology*, 71, pp. 39–69. doi: 10.1146/annurev-arplant-081519-035846.

Andriotis, V. M. E., Rejzek, M., Barclay, E., *et al.* (2016) 'Cell wall degradation is required for normal starch mobilisation in barley endosperm', *Scientific Reports*, 6(August), pp. 1–15. doi: 10.1038/srep33215.

Andriotis, V. M. E., Rejzek, M., Rugen, M. D., *et al.* (2016) 'Iminosugar inhibitors of carbohydrate-active enzymes that underpin cereal grain germination and endosperm metabolism', *Biochemical Society Transactions*, 44, pp. 159–165. doi: 10.1042/BST20150222.

Añorga, M. *et al.* (2020) 'Genes *ptz* and *idi*, coding for cytokinin biosynthesis enzymes, are essential for tumorigenesis and in planta growth by *P. syringae* pv. *savastanoi* NCPPB 3335', *Frontiers in Plant Science*. doi: 10.3389/fpls.2020.01294.

Arnold, D. L., Pitman, A. and Jackson, R. W. (2003) 'Pathogenicity and other genomic islands in plant pathogenic bacteria', *Molecular Plant Pathology*, 4(5), pp. 407–420. doi: 10.1046/j.1364-3703.2003.00187.x.

Arnold, D. L. and Preston, G. M. (2019) '*Pseudomonas syringae*: Enterprising epiphyte and stealthy parasite', *Microbiology (United Kingdom)*, 165(3), pp. 251–253. doi: 10.1099/mic.0.000715.

Arrebola, E. *et al.* (2011) 'Chemical and metabolic aspects of antimetabolite toxins produced by *Pseudomonas syringae* pathovars', *Toxins*, 3(9), pp. 1089–1110. doi: 10.3390/toxins3091089.

Arrebola, E. *et al.* (2015) 'Cellulose production in *Pseudomonas syringae* pv. *syringae*: A compromise between epiphytic and pathogenic lifestyles', *FEMS Microbiology Ecology*, 91(7), pp. 1–12. doi: 10.1093/femsec/fiv071.

Asano, N. *et al.* (2000) 'Sugar-mimic glycosidase inhibitors: Natural occurrence, biological activity and prospects for therapeutic application', *Tetrahedron Asymmetry*, 11(8), pp. 1645–1680. doi: 10.1016/S0957-4166(00)00113-0.

Aslam, S. N. *et al.* (2008) 'Bacterial polysaccharides suppress induced innate immunity by calcium chelation', *Current Biology*, 18(14), pp. 1078–1083. doi: 10.1016/j.cub.2008.06.061.

Bacete, L. *et al.* (2018) 'Plant cell wall-mediated immunity: cell wall changes trigger disease resistance responses', *Plant Journal*, 93(4), pp. 614–636. doi: 10.1111/tpj.13807.

Baltrus, D. A. *et al.* (2011) 'Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 *Pseudomonas syringae* isolates', *PLoS Pathogens*, 7(7). doi: 10.1371/journal.ppat.1002132.

Baltrus, D. A., McCann, H. C. and Guttman, D. S. (2017a) 'Evolution, genomics and epidemiology of *Pseudomonas syringae*: Challenges in Bacterial Molecular Plant Pathology', *Molecular Plant Pathology*, 18(1), pp. 152–168. doi: 10.1111/mpp.12506.

Baltrus, D. A., McCann, H. C. and Guttman, D. S. (2017b) 'Evolution, genomics and epidemiology of *Pseudomonas syringae*', *Molecular Plant Pathology*, 18(1), pp. 152–168. doi: 10.1111/mpp.12506.

Bartesaghi, A. *et al.* (2015) '2.2 å resolution cryo-EM structure of β-galactosidase in complex with a cell-permeant inhibitor', *Science*. doi: 10.1126/science.aab1576.

Bateman, A. *et al.* (2021) 'UniProt: the universal protein knowledgebase in 2021', *Nucleic Acids Research*, 49(D1), pp. D480–D489. doi: 10.1093/nar/gkaa1100.

Bellincampi, D., Cervone, F. and Lionetti, V. (2014) 'Plant cell wall dynamics and wall-related susceptibility in plant-pathogen interactions', *Frontiers in Plant Science*, 5(MAY), pp. 1–8. doi: 10.3389/fpls.2014.00228.

Bender, C L, Alarcón-Chaidez, F. and Gross, D. C. (1999) '*Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases.', *Microbiology and molecular biology reviews : MMBR*, 63(2), pp. 266–92. doi: 1092-2172/99.

Bender, Carol L., Alarcón-Chaidez, F. and Gross, D. C. (1999) '*Pseudomonas syringae* phytotoxins: mode of action, regulation, and biosynthesis by peptide and polyketide synthetases', *Microbiology and Molecular Biology Reviews*, 63(2), pp. 266–292. doi: 10.1128/mmbr.63.2.266-292.1999.

Berge, O. *et al.* (2002) '*Paenibacillus graminis* sp. nov. and *Paenibacillus odorifer* sp. nov., isolated from plant roots, soil and food', *International Journal of Systematic and Evolutionary Microbiology*, 52(2), pp. 607–616. doi: 10.1099/00207713-52-2-607.

Berti, A. D. *et al.* (2007) 'Identification of a biosynthetic gene cluster and the six associated lipopeptides involved in swarming motility of *Pseudomonas syringae* pv. *tomato* DC3000', *Journal of Bacteriology*, 189(17), pp. 6312–6323. doi: 10.1128/JB.00725-07.

Bohlin, J. *et al.* (2017) 'The nucleotide composition of microbial genomes indicates differential patterns of selection on core and accessory genomes', *BMC Genomics*. doi: 10.1186/s12864-017-3543-7.

Bozkurt, T. O. *et al.* (2011) '*Phytophthora infestans* effector AVRblb2 prevents secretion of a plant immune protease at the haustorial interface', *Proceedings of the National Academy of Sciences*, 108(51), pp. 20832–20837. doi: 10.1073/pnas.1112708109.

Breton, R. C. and Reynolds, W. F. (2013) 'Using NMR to identify and characterize natural products', *Natural Product Reports*, 30(4), pp. 501–524. doi: 10.1039/c2np20104f.

Bretz, J. *et al.* (2002) 'Lon protease functions as a negative regulator of type III protein secretion in *Pseudomonas syringae*', *Molecular Microbiology*, 45(2), pp. 397–409. doi: 10.1046/j.1365-2958.2002.03008.x.

Bronstein, P. A. *et al.* (2008) 'Global transcriptional responses of *Pseudomonas syringae* DC3000 to changes in iron bioavailability in vitro', *BMC Microbiology*, 8, pp. 1–15. doi: 10.1186/1471-2180-8-209.

Brummell, D. A. and Harpster, M. H. (2001) 'Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants', *Plant Molecular Biology*, 47(1–2), pp. 311–340. doi: 10.1023/A:1010656104304.

Buell, C. R. *et al.* (2003) 'The complete genome sequence of the Arabidopsis and tomato pathogen

Pseudomonas syringae pv. *tomato* DC3000', *Proc Natl Acad Sci U S A*, 100(18), pp. 10181–10186. doi: 10.1073/pnas.1731982100.

Buscaill, P. *et al.* (2019) 'Glycosidase and glycan polymorphism control hydrolytic release of immunogenic flagellin peptides', *Science*. doi: 10.1126/science.aav0748.

Camacho, C. *et al.* (2009) 'BLAST+: Architecture and applications', *BMC Bioinformatics*. doi: 10.1186/1471-2105-10-421.

Campos, M. L. *et al.* (2018) 'The role of antimicrobial peptides in plant immunity', *Journal of Experimental Botany*, 69(21), pp. 4997–5011. doi: 10.1093/jxb/ery294.

Cao, H. *et al.* (2018) 'The crystal structure of a tetrahydrofolate-bound dihydrofolate reductase reveals the origin of slow product release', *Communications Biology*, 1(1), pp. 1–10. doi: 10.1038/s42003-018-0236-y.

Capra, E. J. and Laub, M. T. (2012) 'Evolution of two-component signal transduction systems', *Annual Review of Microbiology*, 66, pp. 325–347. doi: 10.1146/annurev-micro-092611-150039.

Carrión, V. J. *et al.* (2012) 'The mbo operon is specific and essential for biosynthesis of mangotoxin in *Pseudomonas syringae*', *PLoS ONE*, 7(5). doi: 10.1371/journal.pone.0036709.

Carrión, V. J. *et al.* (2014) 'Mangotoxin production of *Pseudomonas syringae* pv. *syringae* is regulated by MgoA', *BMC Microbiology*, 14(1), pp. 1–13. doi: 10.1186/1471-2180-14-46.

Carrión, V. J. *et al.* (2019) 'Pathogen-induced activation of disease-suppressive functions in the endophytic root microbiome', *Science*, 366(6465), pp. 606–612. doi: 10.1126/science.aaw9285.

Chandrasekar, B. *et al.* (2014) 'Broad-range glycosidase activity profiling', *Molecular & Cellular Proteomics*, 13(10), pp. 2787–2800. doi: 10.1074/mcp.O114.041616.

Chandrasekar, B. (2018) *From glycosidase activity profiling to inhibitor discovery in the plant pathogen Pseudomonas syringae*. University of Bonn.

Chandrasekar, B. and van der Hoorn, R. A. L. (2016) 'Beta galactosidases in Arabidopsis and tomato-a mini review', *Biochemical Society Transactions*, 44(1), pp. 150–158. doi: 10.1042/BST20150217.

Chao, M. C. *et al.* (2016) 'The design and analysis of transposon insertion sequencing experiments', *Nature Reviews Microbiology*. doi: 10.1038/nrmicro.2015.7.

Chatterjee, A. *et al.* (2003) 'GacA, the response regulator of a two-component system, acts as a master regulator in *Pseudomonas syringae* pv. *tomato* DC3000 by controlling regulatory RNA, transcriptional activators, and alternate sigma factors', *Molecular Plant-Microbe Interactions*, 16(12), pp. 1106–1117. doi: 10.1094/MPMI.2003.16.12.1106.

Chen, X. H. *et al.* (2007) 'Comparative analysis of the complete genome sequence of the plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42', *Nature Biotechnology*. doi: 10.1038/nbt1325.

Chen, X. L. *et al.* (2014) 'N-Glycosylation of effector proteins by an α-1,3-Mannosyltransferase is required for the rice blast fungus to evade host innate immunity', *Plant Cell*, 26(3), pp. 1360–1376. doi: 10.1105/tpc.114.123588.

Choi, I. D. *et al.* (2008) 'Novel Ree1 regulates the expression of ENO1 via the Snf1 complex pathway in Saccharomyces cerevisiae', *Biochemical and Biophysical Research Communications*. doi: 10.1016/j.bbrc.2008.09.146.

Choi, K. R. *et al.* (2019) 'Systems metabolic engineering strategies: integrating systems and synthetic biology with metabolic engineering', *Trends in Biotechnology*, 37(8), pp. 817–837. doi: 10.1016/j.tibtech.2019.01.003.

Chrost, B. *et al.* (2007) 'An α-galactosidase with an essential function during leaf development', *Planta*, 225(2), pp. 311–320. doi: 10.1007/s00425-006-0350-9.

Clark, L. F., Johnson, J. V. and Horenstein, N. A. (2011) 'Identification of a gene cluster that initiates Azasugar biosynthesis in *Bacillus amyloliquefaciens*', *ChemBioChem*, 12(14), pp. 2147–2150. doi: 10.1002/cbic.201100347.

Cui, H., Tsuda, K. and Parker, J. E. (2015) 'Effector-triggered immunity: From pathogen perception to robust defense', *Annual Review of Plant Biology*, 66, pp. 487–511. doi: 10.1146/annurev-arplant-050213-040012.

Davies, G. and Henrissat, B. (1995) 'Structures and mechanisms of glycosyl hydrolases', *Structure*, 3(9), pp. 853–859. doi: 10.1016/S0969-2126(01)00220-9.

Dean, G. H. *et al.* (2007) 'The Arabidopsis MUM2 gene encodes a β-galactosidase required for the production of seed coat mucilage with correct hydration properties', *Plant Cell*, 19(12), pp. 4007–4021. doi: 10.1105/tpc.107.050609.

Deepak, S. *et al.* (2010) 'Hydroxyproline-rich glycoproteins and plant defence', *Journal of Phytopathology*, 158(9), pp. 585–593. doi: 10.1111/j.1439-0434.2010.01669.x.

Deng, X. *et al.* (2009) '*Pseudomonas syringae* pv. *phaseolicola* mutants compromised for type III secretion system gene induction', *Mol.Plant Microbe Interact.*, 22, pp. 964–976.

Deng, X. *et al.* (2014) 'Molecular mechanisms of two-component system RhpRS regulating type III secretion system in *Pseudomonas syringae*', *Nucleic Acids Research*, 42(18), pp. 11472–11486. doi: 10.1093/nar/gku865.

Deslandes, L. and Rivas, S. (2012) 'Catch me if you can: Bacterial effectors and plant targets', *Trends in Plant Science*, 17(11), pp. 644–655. doi: 10.1016/j.tplants.2012.06.011.

Devries, C. A. *et al.* (1995) 'Genetic linkage in *Pseudomonas aeruginosa* of *algT* and *nadB:* Mutation in *nadB* does not affect NAD biosynthesis or alginate production', *Gene*, 156(1), pp. 63–67. doi: 10.1016/0378-1119(95)00028-5.

Dietz, K. J. *et al.* (2000) 'Extracellular β-glucosidase activity in barley involved in the hydrolysis of ABA glucose conjugate in leaves', *Journal of Experimental Botany*, 51(346), pp. 937–944. doi: 10.1093/jxb/51.346.937.

Dillon, M. M. *et al.* (2017) 'Recombination of ecologically and evolutionarily significant loci maintains genetic cohesion in the *Pseudomonas syringae* species complex', *bioRxiv*, pp. 1–28. doi: 10.1101/227413.

Dillon, M. M. *et al.* (2019) 'Molecular evolution of *Pseudomonas syringae* type III secreted effector proteins', *Frontiers in Plant Science*, 10, pp. 1–18. doi: 10.3389/fpls.2019.00418.

Ding, P. and Ding, Y. (2020) 'Stories of salicylic acid: a plant defense hormone', *Trends in Plant Science*, 25(6), pp. 549–565. doi: 10.1016/j.tplants.2020.01.004.

Djami-Tchatchou, A. T. *et al.* (2020) 'Dual role of auxin in regulating plant defense and bacterial virulence gene expression during *Pseudomonas syringae Pto*DC3000 pathogenesis', *Molecular Plant-Microbe Interactions*, 33(8), pp. 1059–1071. doi: 10.1094/MPMI-02-20-0047-R.

Dodds, P. N. and Rathjen, J. P. (2010) 'Plant immunity: Towards an integrated view of plant pathogen interactions', *Nature Reviews Genetics*, 11(8), pp. 539–548. doi: 10.1038/nrg2812.

Doehlemann, G. and Hemetsberger, C. (2013) 'Apoplastic immunity and its suppression by filamentous plant pathogens', *New Phytologist*, 198(4), pp. 1001–1016. doi: 10.1111/nph.12277.

Donati, I. *et al.* (2018) 'Pathways of flower infection and pollen-mediated dispersion of *Pseudomonas syringae* pv. *actinidiae*, the causal agent of kiwifruit bacterial canker', *Horticulture Research*. doi: 10.1038/s41438-018-0058-6.

Du, D. *et al.* (2015) 'Structure, mechanism and cooperation of bacterial multidrug transporters', *Current Opinion in Structural Biology*, 33, pp. 76–91. doi: 10.1016/j.sbi.2015.07.015.

Dwevedi, A. and Kayastha, A. M. (2010) 'Plant β -galactosidases : physiological significance and recent advances in technological applications', *Journal of Plant Biochemistry & Biotechnology*, 19, pp. 9–20.

Dworkin, J. (2014) 'The medium is the message: Interspecies and interkingdom signaling by peptidoglycan and related bacterial glycans', *Annual Review of Microbiology*, 68, pp. 137–154. doi: 10.1146/annurev-micro-091213-112844.

El-Gebali, S. *et al.* (2019) 'The Pfam protein families database in 2019', *Nucleic Acids Research*. doi: 10.1093/nar/gky995.

Ellis, M. *et al.* (2010) 'Arabinogalactan-proteins: Key regulators at the cell surface?', *Plant Physiology*, 153(2), pp. 403–419. doi: 10.1104/pp.110.156000.

Emms, D. M. and Kelly, S. (2015) 'OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy', *Genome Biology*, 16(1), p. 157. doi: 10.1186/s13059-

015-0721-2.

Emms, D. M. and Kelly, S. (2019) 'OrthoFinder: Phylogenetic orthology inference for comparative genomics', *Genome Biology*, 20(1), pp. 1–14. doi: 10.1186/s13059-019-1832-y.

Erbs, G. and Newman, M. A. (2012) 'The role of lipopolysaccharide and peptidoglycan, two glycosylated bacterial microbe-associated molecular patterns (MAMPs), in plant innate immunity', *Molecular Plant Pathology*, 13(1), pp. 95–104. doi: 10.1111/j.1364-3703.2011.00730.x.

Felle, H. H. (2001) 'pH: Signal and messenger in plant cells', *Plant Biology*. doi: 10.1055/s-2001-19372.

Fernando, G.-B. *et al.* (2014) 'Biosynthetic pathways to glycosidase inhibitors', *Current Chemical Biology*, 8(1), pp. 10–16.

Ferreira, A. O. *et al.* (2006) 'Whole-genome expression profiling defines the HrpL regulon of *Pseudomonas syringae* pv. *tomato* DC3000, allows de novo reconstruction of the Hrp cis clement, and identifies novel coregulated genes.', *Molecular plant-microbe interactions : MPMI*, 19(11), pp. 1167–1179. doi: 10.1094/MPMI-19-1167.

Ficicioglu, C. (2008) 'Review of miglustat for clinical management in Gaucher disease type I', *Therapeutics and Clinical Risk Management*, 4(2), pp. 425–431. doi: 10.2147/tcrm.s6865.

Figurski, D. H. and Helinski, D. R. (1979) 'Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans, *Proc. Nati. Acad. Sc*, 76(4), pp. 1648–1652.

Fishman, M. R. *et al.* (2017) 'Ca^{2+}-induced two-component system CvsSR regulates the type IIIsecretion system and the extracytoplasmic function sigma factor Algu in *Pseudomonas syringae* pv. *tomato* DC3000', *Journal of Bacteriology*, 200(5), pp. 1–21. doi: 10.1128/JB.00538-17.

Fones, H. *et al.* (2010) 'Metal hyperaccumulation armors plants against disease', *PLOS Pathogens*, 6(9), p. e1001093. doi.org/10.1371/journal.ppat.1001093.

Fouts, D. E. *et al.* (2002) 'Genomewide identification of *Pseudomonas syringae* pv. *tomato* DC3000 promoters controlled by the HrpL alternative sigma factor', *Proceedings of the National Academy of Sciences*, 99(4), pp. 2275–2280. doi: 10.1073/pnas.032514099.

Gaspar, Y. M. *et al.* (2004) 'Characterization of the Arabidopsis lysine-rich arabinogalactan-protein AtAGP17 mutant (*rat1*) that results in a decreased efficiency of Agrobacterium transformation', *Plant Physiology*, 135(4), pp. 2162–2171. doi: 10.1104/pp.104.045542.

Geilfus, C. M. (2017) 'The pH of the apoplast: dynamic factor with functional impact under stress', *Molecular Plant*, 10(11), pp. 1371–1386. doi: 10.1016/j.molp.2017.09.018.

Geng, X. *et al.* (2014) 'The phytotoxin coronatine is a multifunctional component of the virulence armament of *Pseudomonas syringae*', *Planta*, 240(6), pp. 1149–1165. doi: 10.1007/s00425-014-2151-x.

Geoghegan, I., Steinberg, G. and Gurr, S. (2017) 'The role of the fungal cell wall in the infection of plants', *Trends in Microbiology*, 25(12), pp. 957–967. doi: 10.1016/j.tim.2017.05.015.

Gibson, D. G. *et al.* (2009) 'Enzymatic assembly of DNA molecules up to several hundred kilobases', *Nature Methods*, 6(5), pp. 343–345. doi: 10.1038/nmeth.1318.

Giovannoni, J. (2004) 'Genetic regulation of fruit development and ripening', *The Plant Cell,* 16, pp. 170–181. doi: 10.1105/tpc.019158.Fruit.

Gloster, T. M. and Davies, G. J. (2010) 'Glycosidase inhibition: Assessing mimicry of the transition state', *Organic and Biomolecular Chemistry*, 8(2), pp. 305–320. doi: 10.1039/b915870g.

Goettig, P. (2016) 'Effects of glycosylation on the enzymatic activity and mechanisms of proteases', *International Journal of Molecular Sciences*, 17(12), pp. 1–24. doi: 10.3390/ijms17121969.

Gomila, M. *et al.* (2017) 'Clarification of taxonomic status within the *Pseudomonas syringae* species group based on a phylogenomic analysis', *Frontiers in Microbiology*, 8, pp. 1–13. doi: 10.3389/fmicb.2017.02422.

Görke, B. and Stülke, J. (2008) 'Carbon catabolite repression in bacteria: Many ways to make the most out of nutrients', *Nature Reviews Microbiology*, 6(8), pp. 613–624. doi: 10.1038/nrmicro1932.

Grady, E. N. *et al.* (2016) 'Current knowledge and perspectives of *Paenibacillus*: A review', *Microbial Cell Factories*, 15(1), pp. 1–18. doi: 10.1186/s12934-016-0603-7.

Green, B. *et al.* (2012) 'Insertion site preference of Mu, Tn5, and Tn7 transposons', *Mobile DNA*, 3(1), pp. 1–6. doi: 10.1186/1759-8753-3-3.

Gross, H. and Loper, J. E. (2009) 'Genomics of secondary metabolite production by *Pseudomonas spp*.', *Natural Product Reports*, 26(11), p. 1408. doi: 10.1039/b817075b.

Hardick, D. J. *et al.* (1992) 'Glucose is a Precursor of 1-deoxynojirimycin and 1-deoxymannonojirimycin in *Streptomyces subrutilus*', *Tetrahedron*, 48(30), pp. 6285–6296. doi: 10.1016/S0040-4020(01)88220-X.

Hardick, D. J. and Hutchinson, D. W. (1993) 'The biosynthesis of 1-deoxynojirimycin in *Bacillus subtilis* var *niger*', *Tetrahedron*, 49(30), pp. 6707–6716. doi: 10.1016/S0040-4020(01)81840-8.

Harit, V. K. and Ramesh, N. G. (2016) 'Amino-functionalized iminocyclitols: synthetic glycomimetics of medicinal interest', *RSC Adv.*, 6(111), pp. 109528–109607. doi: 10.1039/C6RA23513A.

Harvey, D. J. and Vouros, P. (2020) 'Mass spectrometric fragmentation of trimethylsilyl and related alkylsilyl derivatives', *Mass Spectrometry Reviews*, 39(1–2), pp. 105–211. doi: 10.1002/mas.21590.

Hauck, P., Thilmony, R. and He, S. Y. (2003) 'A *Pseudomonas syringae* type III effector suppresses cell wall-based extracellular defense in susceptible Arabidopsis plants', *Proceedings of the National Academy of Sciences of the United States of America*, 100(14), pp. 8577–8582. doi: 10.1073/pnas.1431173100.

Häweker, H. *et al.* (2010) 'Pattern recognition receptors require N-glycosylation to mediate plant immunity', *Journal of Biological Chemistry*, 285(7), pp. 4629–4636. doi: 10.1074/jbc.M109.063073.

Heeb, S. and Haas, D. (2001) 'Regulatory roles of the GacS / GacA two-component system in plant-associated and other Gram-negative bacteria', *Molecular Plant-Microbe Interactions*, 14(12), pp. 1351–1363. doi: 10.1094/MPMI.2001.14.12.1351.

Hématy, K., Cherk, C. and Somerville, S. (2009) 'Host-pathogen warfare at the plant cell wall', *Current Opinion in Plant Biology*, 12(4), pp. 406–413. doi: 10.1016/j.pbi.2009.06.007.

Henrissat, B. and Grenoble, F. (1991) 'A classification of glycosyl hydrolases based on sequence similarities amino acid', *Journal of Biochemistry*, 280, pp. 309–316.

Heredia-Ponce, Z. *et al.* (2020) 'Biological role of EPS from *Pseudomonas syringae* pv. *syringae* UMAF0158 extracellular matrix, focusing on a Psl-like polysaccharide', *npj Biofilms and Microbiomes*, 6(1). doi: 10.1038/s41522-020-00148-6.

Hernandez, M. N. and Lindow, S. E. (2019) '*Pseudomonas syringae* increases water availability in leaf microenvironments via production of hygroscopic syringafactin', *Applied and Environmental Microbiology*, 85(18), pp. 1–14. doi: 10.1101/626630.

Hmelo, L. R. *et al.* (2015) 'Precision-engineering the *Pseudomonas aeruginosa* genome with two-step allelic exchange', *Nature Protocols*, 10(11), pp. 1820–1841. doi: 10.1038/nprot.2015.115.

Van Der Hoorn, R. A. L. and Jones, J. D. G. (2004) 'The plant proteolytic machinery and its role in defence', *Current Opinion in Plant Biology*, 7(4), pp. 400–407. doi: 10.1016/j.pbi.2004.04.003.

Horne, G. *et al.* (2011) 'Iminosugars past, present and future: Medicines for tomorrow', *Drug Discovery Today*, 16(3–4), pp. 107–118. doi: 10.1016/j.drudis.2010.08.017.

Hossain, F. and Andreana, P. R. (2019) 'Developments in carbohydrate-based cancer therapeutics', *Pharmaceuticals*, 12(2). doi: 10.3390/ph12020084.

Hou, S. *et al.* (2019) 'Damage-associated molecular pattern-triggered immunity in plants', *Frontiers in Plant Science*, 10(May). doi: 10.3389/fpls.2019.00646.

Hutcheson, S. W. *et al.* (2001) 'Enhancer-binding proteins HrpR and HrpS interact to regulate *hrp*-encoded type III protein secretion in *Pseudomonas syringae* strains', *Journal of Bacteriology*, 183(19), pp. 5589–5598. doi: 10.1128/JB.183.19.5589-5598.2001.

Hwang, M. *et al.* (2005) 'Phylogenetic characterization of virulence and resistance phenotypes of *Pseudomonas syringae*', *Applied and Environmental Microbiology*, 71(9), pp. 5182–5191. doi: 10.1128/AEM.71.9.5182.

Ichinose, Y., Taguchi, F. and Mukaihara, T. (2013) 'Pathogenicity and virulence factors of *Pseudomonas syringae*', *Journal of General Plant Pathology*, 79(5), pp. 285–296. doi: 10.1007/s10327-013-0452-8.

Iglesias, N. *et al.* (2006) 'Apoplastic glycosidases active against xyloglucan oligosaccharides of *Arabidopsis thaliana*', *Plant and Cell Physiology*, 47(1), pp. 55–63. doi: 10.1093/pcp/pci223.

Inouye, S. *et al.* (1968) 'Structure and synthesis of nojirimycin', *Tetrahedron*, 24(5), pp. 2125–2144. doi: 10.1016/0040-4020(68)88115-3.

Itakura, Y. *et al.* (2007) 'Systematic comparison of oligosaccharide specificity of ricinus communis agglutinin i and erythrina lectins: a search by frontal affinity chromatography', *Journal of Biochemistry*, 142(4), pp. 459–469. doi: 10.1093/jb/mvm153.

Jashni, M. K. *et al.* (2015) 'The battle in the apoplast: further insights into the roles of proteases and their inhibitors in plant–pathogen interactions', *Frontiers in Plant Science*, 6, pp. 1–7. doi: 10.3389/fpls.2015.00584.

Jiang, P. *et al.* (2015) 'Design and application of a novel high-throughput screening technique for 1-deoxynojirimycin', *Scientific Reports*, 5, pp. 1–6. doi: 10.1038/srep08563.

Jones, J. D. G. and Dangl, J. L. (2006) 'The plant immune system', *Nature*, 444(7117), pp. 323–329. doi: 10.1038/nature05286.

Jovanovic, M. *et al.* (2011) 'Regulation of the co-evolved HrpR and HrpS AAA+ proteins required for *Pseudomonas syringae* pathogenicity', *Nature Communications*, 2(1). doi: 10.1038/ncomms1177.

Juers, D. H., Matthews, B. W. and Huber, R. E. (2012) 'LacZ β-galactosidase: Structure and function of an enzyme of historical and molecular biological importance', *Protein Science*, 21(12), pp. 1792–1807. doi: 10.1002/pro.2165.

Juge, N. (2006) 'Plant protein inhibitors of cell wall degrading enzymes', *Trends in Plant Science*, 11(7), pp. 359–367. doi: 10.1016/j.tplants.2006.05.006.

Kang, K. D. *et al.* (2011) 'Identification of the genes involved in 1-deoxynojirimycin synthesis in *Bacillus subtilis* MORI 3K-85', *Journal of Microbiology*, 49(3), pp. 431–440. doi: 10.1007/s12275-011-1238-3.

Kayakiri, H. *et al.* (1989) 'Structure of Kifunensine, a New Immunomodulator Isolated from an Actinomycete', *Journal of Organic Chemistry*, 54(17), pp. 4015–4016. doi: 10.1021/jo00278a003.

Kim, B. J. *et al.* (2009) 'Effect of iron concentration on the growth rate of *Pseudomonas syringae* and the expression of virulence factors in *hrp*-inducing minimal medium', *Applied and Environmental Microbiology*, 75(9), pp. 2720–2726. doi: 10.1128/AEM.02738-08.

Kim, H. U. *et al.* (2016) 'Metabolic engineering with systems biology tools to optimize production of prokaryotic secondary metabolites', *Natural Product Reports*, 33(8), pp. 933–941. doi: 10.1039/c6np00019c.

Kim, J. W. *et al.* (2003) 'Determination of 1-deoxynojirimycin in *Morus alba* L. leaves by derivatization with 9-fluorenylmethyl chloroformate followed by reversed-phase high-performance liquid chromatography', *Journal of Chromatography A*, 1002(1–2), pp. 93–99. doi: 10.1016/S0021-9673(03)00728-3.

Ko, Y. S. *et al.* (2020) 'Tools and strategies of systems metabolic engineering for the development of microbial cell factories for chemical production', *Chemical Society Reviews*, 49(14), pp. 4615–4636. doi: 10.1039/d0cs00155d.

Kong, Y. *et al.* (2015) 'Galactose-depleted xyloglucan is dysfunctional and leads to dwarfism in Arabidopsis', *Plant Physiology*, 167(4), pp. 1296–1306. doi: 10.1104/pp.114.255943.

Koonin, E. V, Makarova, K. S. and Aravind, L. (2001) 'Horizontal gene transfer in prokaryotes: quantification and classification', *Annual Reviews of Microbiology*, 55, pp.709-742.

Kotake, T., Dina, S. and Konishi, T. (2005) 'Molecular cloning of a β-galactosidase from radish that specifically hydrolyzes β-(1-> 3)-and β-(1->6)-galactosyl residues of arabinogalactan protein', *Plant Physiology*, 138(3), pp. 1563–1576. doi: 10.1104/pp.105.062562.1.

Kourelis, J. and van der Hoorn, R. A. L. (2018) 'Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for r protein function', *The Plant Cell*, 30. doi: 10.1105/tpc.17.00579.

Kriechbaum, R. *et al.* (2020) 'BGAL1 depletion boosts the level of β-galactosylation of N- and O-glycans in

N. benthamiana', *Plant Biotechnology Journal*, 18(7), pp. 1537–1549. doi: 10.1111/pbi.13316.

Kubicek, C. P., Starr, T. L. and Glass, N. L. (2014) 'Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi', *Annual Review of Phytopathology*, 52, pp. 427–451. doi: 10.1146/annurev-phyto-102313-045831.

Kurnasov, O. *et al.* (2003) 'NAD biosynthesis: identification of the tryptophan to quinolinate pathway in bacteria', *Chemistry and Biology*. doi: 10.1016/j.chembiol.2003.11.011.

Kvitko, B. H. *et al.* (2009) 'Deletions in the repertoire of *Pseudomonas syringae* pv. *tomato* DC3000 type III secretion effector genes reveal functional overlap among effectors', *PLoS Pathogens*, 5(4). doi: 10.1371/journal.ppat.1000388.

Kvitko, B. H. and Collmer, A. (2011) 'Construction of *Pseudomonas syringae* pv. *tomato* DC3000 mutant and polymutant strains.', *Methods in molecular biology (Clifton, N.J.)*. doi: 10.1007/978-1-61737-998-7_10.

Lai, Z. and Fiehn, O. (2018) 'Mass spectral fragmentation of trimethylsilylated small molecules', *Mass Spectrometry Reviews*. doi: 10.1002/mas.21518.

Lam, H. N. *et al.* (2014) 'Global analysis of the HrpL regulon in the plant pathogen *Pseudomonas syringae* pv. *tomato* DC3000 reveals new regulon members with diverse functions', *PLoS ONE*. doi: 10.1371/journal.pone.0106115.

Lamb, C. and Dixon, R. A. (1997) 'The oxidative burst in plant disease resistance', *Annual Review of Plant Biology*, 48, pp. 251–275. doi: 10.1146/annurev.arplant.48.1.251.

Laue, H. *et al.* (2006) 'Contribution of alginate and levan production to biofilm formation by *Pseudomonas syringae*', *Microbiology*, 152(10), pp. 2909–2918. doi: 10.1099/mic.0.28875-0.

Lee, H. *et al.* (2018) 'Culture Optimization Strategy for 1-Deoxynojirimycin-producing *Bacillus methylotrophicus* K26 Isolated from Korean Fermented Soybean Paste, Doenjang', *Biotechnology and Bioprocess Engineering*, 23(4), pp. 424–431. doi: 10.1007/s12257-018-0159-y.

Lee, K. H. *et al.* (2006) 'Activation of glucosidase via stress-induced polymerization rapidly increases active pools of abscisic acid', *Cell*, 126(6), pp. 1109–1120. doi: 10.1016/j.cell.2006.07.034.

Letunic, I. and Bork, P. (2019) 'Interactive Tree of Life (iTOL) v4: Recent updates and new developments', *Nucleic Acids Research*. doi: 10.1093/nar/gkz239.

Li, C. Y. *et al.* (2016) 'Reclassification of *Enterobacter oryziphilus* and *Enterobacter oryzendophyticus* as *Kosakonia oryziphila* comb. nov. and *Kosakonia oryzendophytica* comb. nov', *International Journal of Systematic and Evolutionary Microbiology*. doi: 10.1099/ijsem.0.001054.

Li, K. *et al.* (2019) 'Two components of the *rhpPC* operon coordinately regulate the type III secretion system and bacterial fitness in *Pseudomonas savastanoi* pv. *phaseolicola*', *PLoS Pathogens*, 15(4), pp. 1–27. doi: 10.1371/journal.ppat.1007673.

Lillelund, V. H. *et al.* (2002) 'Recent developments of transition-state analogue glycosidase inhibitors of non-natural product origin', *Chemical Reviews*, 102(2), pp. 515–553. doi: 10.1021/cr000433k.

Lin, B. *et al.* (2020) 'Role of protein glycosylation in host-pathogen interaction', *Cells*, 9(4), pp. 1–24. doi: 10.3390/cells9041022.

Lin, T. H., Huang, S. C. and Shaw, G. C. (2012) 'Reexamining transcriptional regulation of the *Bacillus subtilis htpx* gene and the *Ykrk* gene, encoding a novel type of transcriptional regulator, and redefining the *ykrk*', *Journal of Bacteriology*, 194(24), pp. 6758–6765. doi: 10.1128/JB.01258-12.

Lindeberg, M. *et al.* (2006) 'Closing the circle on the discovery of genes encoding hrp regulon members and type iii secretion system effectors in the genomes of three model *Pseudomonas syringae* strains', *Molecular Plant-Microbe Interactions*, 19(11), pp. 1151–1158. doi: 10.1094/MPMI-19-1151.

Lindeberg, M., Cunnac, S. and Collmer, A. (2012) '*Pseudomonas syringae* type III effector repertoires: Last words in endless arguments', *Trends in Microbiology*, 20(4), pp. 199–208. doi: 10.1016/j.tim.2012.01.003.

Liu, J. *et al.* (2018) 'Redox cofactor engineering in industrial microorganisms: strategies, recent applications and future directions', *Journal of Industrial Microbiology and Biotechnology*, 45(5), pp. 313–327. doi:

10.1007/s10295-018-2031-7.

Liu, X. *et al.* (2014) 'Host-induced bacterial cell wall decomposition mediates pattern-triggered immunity in Arabidopsis', *eLife*, 3, pp. 1–24. doi: 10.7554/elife.01990.

Liu, Y. G. and Whittier, R. F. (1995) 'Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking', *Genomics*. doi: 10.1016/0888-7543(95)80010-J.

van Loon, L. C., Rep, M. and Pieterse, C. M. J. (2006) 'Significance of inducible defense-related proteins in infected plants', *Annual Review of Phytopathology*, 44(1), pp. 135–162. doi: 10.1146/annurev.phyto.44.070505.143425.

De Lorenzo, V. *et al.* (1990) 'Mini-Tn5 transposoon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria', *Journal of Bacteriology*, 172(11), pp. 6568–6572. doi: 10.1128/jb.172.11.6568-6572.1990.

Lovelace, Amelia H., Smith, A. and Kvitko, B. H. (2018) 'Pattern-triggered immunity alters the transcriptional regulation of virulence-associated genes and induces the sulfur starvation response in *Pseudomonas syringae* pv. *tomato* DC3000', *Molecular Plant-Microbe Interactions*, 31(7), pp. 750–765. doi: 10.1094/MPMI-01-18-0008-R.

Lu, S. *et al.* (2020) 'CDD/SPARCLE: The conserved domain database in 2020', *Nucleic Acids Research*. doi: 10.1093/nar/gkz991.

Luo, F. *et al.* (2020) 'Unveiling of swainsonine biosynthesis via a multibranched pathway in fungi', *ACS Chemical Biology*, 15(9), pp. 2476–2484. doi: 10.1021/acschembio.0c00466.

Luo, Y. *et al.* (2018) 'Complete genome sequence of industrial biocontrol strain *Paenibacillus polymyxa* HY96-2 and further analysis of its biocontrol mechanism', *Frontiers in Microbiology*. doi: 10.3389/fmicb.2018.01520.

Ma, Z. *et al.* (2015) 'A *Phytophthora sojae* glycoside hydrolase 12 protein is a major virulence factor during soybean infection and is recognized as a PAMP', *Plant Cell*, 27(7), pp. 2057–2072. doi: 10.1105/tpc.15.00390.

Ma, Z. *et al.* (2017) 'A paralogous decoy protects *Phytophthora sojae* apoplastic effector PsXEG1 from a host inhibitor', *Science*, 355(6326), pp. 710–714. doi: 10.1126/science.aai7919.

Madeira, F. *et al.* (2019) 'The EMBL-EBI search and sequence analysis tools APIs in 2019', *Nucleic Acids Research*. doi: 10.1093/nar/gkz268.

Malinovsky, F. G., Fangel, J. U. and Willats, W. G. T. (2014) 'The role of the cell wall in plant immunity', *Frontiers in Plant Science*, 5, pp. 1–12. doi: 10.3389/fpls.2014.00178.

Mansfield, J. *et al.* (2012) 'Top 10 plant pathogenic bacteria in molecular plant pathology', *Molecular Plant Pathology*, 13(6), pp. 614–629. doi: 10.1111/j.1364-3703.2012.00804.x.

Markel, E. *et al.* (2016) 'AlgU controls expression of virulence genes in *Pseudomonas syringae* pv. *tomato* DC3000', *Journal of Bacteriology*, 198(17), pp. 2330–2344. doi: 10.1128/JB.00276-16.

Maruri-López, I. *et al.* (2019) 'Intra and extracellular journey of the phytohormone salicylic acid', *Frontiers in Plant Science*, 10, pp. 1–11. doi: 10.3389/fpls.2019.00423.

McCafferty, E. H. and Scott, L. J. (2019) 'Migalastat: a review in Fabry disease', *Drugs*, 79(5), pp. 543–554. doi: 10.1007/s40265-019-01090-4.

McClerklin, S. A. *et al.* (2017) 'Indole-3-acetaldehyde dehydrogenase-dependent auxin synthesis contributes to virulence of *Pseudomonas syringae* strain DC3000', *bioRxiv*, pp. 1–24. doi: 10.1101/173302.

McCraw, S. L. *et al.* (2016) 'GABA (γ-Aminobutyric Acid) uptake via the GABA permease GabP represses virulence gene expression in *Pseudomonas syringae* pv. t*omato* DC3000', *Molecular Plant-Microbe Interactions*. doi: 10.1094/MPMI-08-16-0172-R.

Mega, T. (2004) 'Conversion of the carbohydrate structures of glycoproteins in roots of *Raphanus sativu*s using several glycosidase inhibitors', *Journal of Biochemistry*, 136(4), pp. 525–531. doi: 10.1093/jb/mvh153.

Minic, Z. (2008) 'Physiological roles of plant glycoside hydrolases', *Planta*, 227(4), pp. 723–740. doi: 10.1007/s00425-007-0668-y.

Minic, Z. and Jouanin, L. (2006) 'Plant glycoside hydrolases involved in cell wall polysaccharide degradation', *Plant Physiology and Biochemistry*, 44(7–9), pp. 435–449. doi: 10.1016/j.plaphy.2006.08.001.

Misas-Villamil, J. C. *et al.* (2013) '*Pseudomonas syringae* pv. *syringae* uses proteasome inhibitor syringolin A to colonize from wound infection sites', *PLoS Pathogens*, 9(3), pp. 10–16. doi: 10.1371/journal.ppat.1003281.

Misas-Villamil, J. C., van der Hoorn, R. A. L. and Doehlemann, G. (2016) 'Papain-like cysteine proteases as hubs in plant immunity', *New Phytologist*, 212(4), pp. 902–907. doi: 10.1111/nph.14117.

Misas-Villamil, J. C. and van der Hoorn, R. AL (2008) 'Enzyme-inhibitor interactions at the plant-pathogen interface', *Current Opinion in Plant Biology*, 11(4), pp. 380–388. doi: 10.1016/j.pbi.2008.04.007.

Miyake, Y. and Ebata, M. (1988) 'Isolation and properties of a new B-galactosidase inhbitor galactostatin from *Streptomyces lydicus*', *Agric. Biol. Chem.*, 52(1), pp. 153–158.

Miyauchi, R. *et al.* (2015) 'Biosynthesis of nectrisine in *Thelonectria discophora* SANK 18292', *Phytochemistry*, 116(1), pp. 87–93. doi: 10.1016/j.phytochem.2015.03.011.

Miyauchi, R. *et al.* (2016) 'Nectrisine biosynthesis genes in *Thelonectria discophora* SANK 18292: Identification and functional analysis', *Applied and Environmental Microbiology*, 82(21), pp. 6414–6422. doi: 10.1128/AEM.01709-16.

Mole, B. M. *et al.* (2007) 'Global virulence regulation networks in phytopathogenic bacteria', *Trends in Microbiology*, 15(8), pp. 363–371. doi: 10.1016/j.tim.2007.06.005.

Molyneux, R. J. (1993) 'Isolation, characterization and analysis of polyhydroxy alkaloids', *Phytochemical Analysis*, 4(5), pp. 193–204. doi: 10.1002/pca.2800040502.

Molyneux, R. J. *et al.* (2002) 'Polyhydroxy alkaloids: Chromatographic analysis', *Journal of Chromatography A*, 967(1), pp. 57–74. doi: 10.1016/S0021-9673(01)01558-8.

Morant, A. V. *et al.* (2008) 'β-Glucosidases as detonators of plant chemical defense', *Phytochemistry*, 69(9), pp. 1795–1813. doi: 10.1016/j.phytochem.2008.03.006.

Morimoto, K. and Van Der Hoorn, R. A. L. (2016) 'The increasing impact of activity-based protein profiling in plant science', *Plant and Cell Physiology*, 57(3), pp. 446–461. doi: 10.1093/pcp/pcw003.

Mucyn, T. S. *et al.* (2014) 'Variable suites of non-effector genes are co-regulated in the type III secretion virulence regulon across the *Pseudomonas syringae* Phylogeny', *PLoS Pathogens*, 10(1). doi: 10.1371/journal.ppat.1003807.

Mueller, A. N. *et al.* (2013) 'Compatibility in the *Ustilago maydis*-maize interaction requires inhibition of host cysteine proteases by the fungal effector Pit2', *PLoS Pathogens*, 9(2). doi: 10.1371/journal.ppat.1003177.

Nagahashi, G. *et al.* (1990) 'Inhibition of cell wall-associated enzymes in vitro and in vivo with sugar analogs', *Plant Physiology*, 92(2), pp. 413–418. doi: 10.1104/pp.92.2.413.

Nagy, G., Peng, T. and Pohl, N. L. B. (2017) 'Recent liquid chromatographic approaches and developments for the separation and purification of carbohydrates', *Analytical Methods*, 9(24), pp. 3579–3593. doi: 10.1039/c7ay01094j.

Nash, R. J. *et al.* (2011) 'Iminosugars as therapeutic agents: Recent advances and promising trends', *Future Medicinal Chemistry*, 3(12), pp. 1513–1521. doi: 10.4155/fmc.11.117.

Naumann, T. A., Wicklow, D. T. and Price, N. P. J. (2011) 'Identification of a chitinase-modifying protein from *Fusarium verticillioides*: Truncation of a host resistance protein by a fungalysin metalloprotease', *Journal of Biological Chemistry*, 286(41), pp. 35358–35366. doi: 10.1074/jbc.M111.279646.

Newman, M.-A. *et al.* (2013) 'MAMP (microbe-associated molecular pattern) triggered immunity in plants', *Frontiers in Plant Science*, 4, pp. 1–14. doi: 10.3389/fpls.2013.00139.

Nuñez, C. and Horenstein, N. A. (2019) 'Functional analysis of a gene cluster from *Chitinophaga pinensis* involved in biosynthesis of the pyrrolidine azasugar DAB-1', *Journal of Natural Products*, 82(12), pp. 3401–3409. doi: 10.1021/acs.jnatprod.9b00758.

O'Brien, H. E., Thakur, S. and Guttman, D. S. (2011) 'Evolution of plant pathogenesis in *Pseudomonas syringae* : a genomics perspective', *Annual Review of Phytopathology*, 49(1), pp. 269–289. doi: 10.1146/annurev-phyto-

072910-095242.

O'Leary, B. M. *et al.* (2016) 'Early changes in apoplast composition associated with defence and disease in interactions between Phaseolus vulgaris and the halo blight pathogen *Pseudomonas syringae* pv. *phaseolicola*', *Plant Cell and Environment*, 39(10), pp. 2172–2184. doi: 10.1111/pce.12770.

O'Leary, N. A. *et al.* (2016) 'Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation', *Nucleic Acids Research*. doi: 10.1093/nar/gkv1189.

O'Neill, E. M. *et al.* (2018) 'Phevamine A, a small molecule that suppresses plant immune responses', *Proceedings of the National Academy of Sciences of the United States of America*, 115(41), pp. E9514–E9522. doi: 10.1073/pnas.1803779115.

Obranić, S., Babić, F. and Maravić-Vlahoviček, G. (2013) 'Improvement of pBBR1MCS plasmids, a very useful series of broad-host-range cloning vectors', *Plasmid*. doi: 10.1016/j.plasmid.2013.04.001.

Ochman, Howard, Lawrence, Jeffrey G. and Groisman, Eduardo A. (2000) 'Lateral gene transfer and the nature of bacterial innovation', *Nature*, 405(6784), pp. 299–304.

Oerke, E. C. (2006) 'Crop losses to pests', *Journal of Agricultural Science*, 144(1), pp. 31–43. doi: 10.1017/S0021859605005708.

Omote, H. *et al.* (2006) 'The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations', *Trends in Pharmacological Sciences*, 27(11), pp. 587–593. doi: 10.1016/j.tips.2006.09.001.

Onose, S. *et al.* (2013) 'Production of the α-glycosidase inhibitor 1-deoxynojirimycin from *Bacillus* species', *Food Chemistry*, 138(1), pp. 516–523. doi: 10.1016/j.foodchem.2012.11.012.

Van Opijnen, T. and Camilli, A. (2013) 'Transposon insertion sequencing: A new tool for systems-level analysis of microorganisms', *Nature Reviews Microbiology*. doi: 10.1038/nrmicro3033.

Ortiz-Martín, I. *et al.* (2010) 'Positive regulation of the Hrp type III secretion system in *Pseudomonas syringae* pv. *phaseolicola*', *Molecular Plant-Microbe Interactions*, 23(5), pp. 665–681. doi: 10.1094/MPMI-23-5-0665.

Page, D. R. and Grossniklaus, U. (2002) 'The art and design of genetic screens: *Arabidopsis thaliana*', *Nature Reviews Genetics*. doi: 10.1038/nrg730.

Pao, S. S., Paulsen, I. T. and Saier, M. H. (1998) 'Major facilitator superfamily', *Microbiology and molecular biology reviews : MMBR*, 62(1), pp. 1–34.

Parales, R. E. and Harwood, C. S. (1993) 'Construction and use of a new broad-host-range *lacZ* transcriptional fusion vector, pHRP309, for Gram - bacteria', *Gene*, 133(1), pp. 23–30. doi: 10.1016/0378-1119(93)90220-W.

Parenti, G. (2009) 'Treating lysosomal storage diseases with pharmacological chaperones: From concept to clinics', *EMBO Molecular Medicine*, 1(5), pp. 268–279. doi: 10.1002/emmm.200900036.

Park, D. H. *et al.* (2010) 'Mutations in γ-aminobutyric acid (GABA) transaminase genes in plants or *Pseudomonas syringae* reduce bacterial virulence', *Plant Journal*. doi: 10.1111/j.1365-313X.2010.04327.x.

Paulus, J. K. *et al.* (2020) 'Extracellular proteolytic cascade in tomato activates immune protease Rcr3', *Proceedings of the National Academy of Sciences of the United States of America*, 117(29), pp. 17409–17417. doi: 10.1073/pnas.1921101117.

Peña, M. J. *et al.* (2004) 'The galactose residues of xyloglucan are essential to maintain mechanical strength of the primary cell walls in arabidopsis during growth', *Plant Physiology*, 134(1), pp. 443–451. doi: 10.1104/pp.103.027508.

Perinbam, K. *et al.* (2020) 'A shift in central metabolism accompanies virulence activation in *Pseudomonas aeruginosa*', *mBio*. doi: 10.1128/mBio.02730-18.

Pfeilmeier, S., Caly, D. L. and Malone, J. G. (2016) 'Bacterial pathogenesis of plants: future challenges from a microbial perspective', *Molecular Plant Pathology*, 17(8), pp. 1298–1313. doi: 10.1111/mpp.12427.

Piasecka, A., Jedrzejczak-Rey, N. and Bednarek, P. (2015) 'Secondary metabolites in plant innate immunity: Conserved function of divergent chemicals', *New Phytologist*, 206(3), pp. 948–964. doi: 10.1111/nph.13325.

Pitman, A. R. *et al.* (2005) 'Exposure to host resistance mechanisms drives evolution of bacterial virulence in

plants', *Current Biology*, 15(24), pp. 2230–2235. doi: 10.1016/j.cub.2005.10.074.

Pollock, D. D. and Larkin, J. C. (2004) 'Estimating the degree of saturation in mutant screens', *Genetics*. doi: 10.1534/genetics.103.024430.

Preston, G. *et al.* (1998) 'Negative regulation of hrp genes in *Pseudomonas syringae* by HrpV', *Journal of Bacteriology*, 180(17), pp. 4532–4537. doi: 10.1128/jb.180.17.4532-4537.1998.

Preston, G. M. (2000) '*Pseudomonas syringae* pv. *tomato*: the right pathogen, of the right plant, at the right time', *Molecular Plant Pathology*, 1(5), pp. 263–275. doi: 10.1046/j.1364-3703.2000.00036.x.

Prichard, K. *et al.* (2018) 'Biological activities of 3,4,5-trihydroxypiperidines and their N- and O-derivatives', *Chemical Biology and Drug Design*, 92(1), pp. 1171–1197. doi: 10.1111/cbdd.13182.

Prosser, G. A., Larrouy-Maumus, G. and Carvalho, L. P. S. (2014) 'Metabolomic strategies for the identification of new enzyme functions and metabolic pathways', *EMBO reports*, 15(6), pp. 657–669. doi: 10.15252/embr.201338283.

Raaymakers, T. M. and Van Den Ackerveken, G. (2016) 'Extracellular recognition of oomycetes during biotrophic infection of plants', *Frontiers in Plant Science*, 7, pp. 1–12. doi: 10.3389/fpls.2016.00906.

Radhakrishnan, R., Hashem, A. and Abd Allah, E. F. (2017) '*Bacillus*: A biological tool for crop improvement through bio-molecular changes in adverse environments', *Frontiers in Physiology*. doi: 10.3389/fphys.2017.00667.

Ramírez, V. *et al.* (2013) 'An extracellular subtilase switch for immune priming in Arabidopsis', *PLoS Pathogens*, 9(6). doi: 10.1371/journal.ppat.1003445.

Rapicavoli, J. N. *et al.* (2018) 'Lipopolysaccharide O-antigen delays plant innate immune recognition of *Xylella fastidiosa*', *Nature Communications*, 9(1). doi: 10.1038/s41467-018-02861-5.

Ravenhall, M. *et al.* (2015) 'Inferring horizontal gene transfer', *PLoS Computational Biology*, 11(5), pp. 1–16. doi: 10.1371/journal.pcbi.1004095.

Rayamajhi, V. *et al.* (2018) 'Improved production of 1-deoxynojirymicin in *Escherichia coli* through metabolic engineering', *World Journal of Microbiology and Biotechnology*, 34(6), pp. 1–11. doi: 10.1007/s11274-018-2462-3.

Ritz, C. *et al.* (2015) 'Dose-response analysis using R', *PLoS ONE*. doi: 10.1371/journal.pone.0146021.

Rodríguez-Sánchez, S. *et al.* (2011) 'A derivatization procedure for the simultaneous analysis of iminosugars and other low molecular weight carbohydrates by GC-MS in mulberry (*Morus sp.*)', *Food Chemistry*, 126(1), pp. 353–359. doi: 10.1016/j.foodchem.2010.10.097.

Rodríguez-Sánchez, S., Ruiz-Matute, A. I., *et al.* (2014) 'Characterization of trimethylsilyl ethers of iminosugars by gas chromatography-mass spectrometry', *Journal of Chromatography A*, 1372, pp. 221–227. doi: 10.1016/j.chroma.2014.10.084.

Rodríguez-Sánchez, S., Quintanilla-López, J. E., *et al.* (2014) 'Evaluation of different hydrophilic stationary phases for the simultaneous determination of iminosugars and other low molecular weight carbohydrates in vegetable extracts by liquid chromatography tandem mass spectrometry', *Journal of Chromatography A*, 1372, pp. 81–90. doi: 10.1016/j.chroma.2014.10.079.

Roine, E. *et al.* (1997) 'Hrp pilus: An hrp-dependent bacterial surface appendage produced by *Pseudomonas syringae* pv. *tomato* DC3000', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.94.7.3459.

Rojo, F. (2010) 'Carbon catabolite repression in *Pseudomonas*: Optimizing metabolic versatility and interactions with the environment', *FEMS Microbiology Reviews*, 34(5), pp. 658–684. doi: 10.1111/j.1574-6976.2010.00218.x.

Rooney, H. C. E. *et al.* (2005) '*Cladosporium* Avr2 inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance', *Science*, 308(5729), pp. 1783–1786. doi: 10.1126/science.1111404.

Rovenich, H., Zuccaro, A. and Thomma, B. P. H. J. (2016) 'Convergent evolution of filamentous microbes towards evasion of glycan-triggered immunity', *New Phytologist*, 212(4), pp. 896–901. doi:

10.1111/nph.14064.

Rugen, M. D. *et al.* (2018) 'A chemical genetic screen reveals that iminosugar inhibitors of plant glucosylceramide synthase inhibit root growth in Arabidopsis and cereals', *Scientific Reports*, 8(1), pp. 1–12. doi: 10.1038/s41598-018-34749-1.

Ruud, J. (2010) 'United States Patent : 3871965', 2(12), pp. 4–6.

Saijo, Y., Loo, E. P. and Yasuda, S. (2017) 'Pattern recognition receptors and signaling in plant-microbe interactions', *The Plant Journal*, pp. 592–613. doi: 10.1111/tpj.13808.

Sakoh, M., Ito, K. and Akiyama, Y. (2005) 'Proteolytic activity of HtpX, a membrane-bound and stress-controlled protease from *Escherichia coli*', *Journal of Biological Chemistry*, 280(39), pp. 33305–33310. doi: 10.1074/jbc.M506180200.

Sampedro, J. *et al.* (2012) 'AtBGAL10 is the main xyloglucan β-galactosidase in arabidopsis, and its absence results in unusual xyloglucan subunits and growth defects', *Plant Physiology*, 158(3), pp. 1146–1157. doi: 10.1104/pp.111.192195.

Sampedro, J. *et al.* (2017) 'Soluble and membrane-bound β-glucosidases are involved in trimming the Xyloglucan backbone', *Plant Physiology*, 173(2), pp. 1017–1030. doi: 10.1104/pp.16.01713.

Sandhu, A. (2008) *Phyllosphere-mediated bacterial degradation of airborne phenol*. Iowa State University.

Savary, S. *et al.* (2019) 'The global burden of pathogens and pests on major food crops', *Nature Ecology and Evolution*, 3(3), pp. 430–439. doi: 10.1038/s41559-018-0793-y.

Schäfer, A. *et al.* (1994) 'Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*', *Gene*. doi: 10.1016/0378-1119(94)90324-7.

Schaller, A. *et al.* (2018) 'From structure to function – a family portrait of plant subtilases', *New Phytologist*, 218(3), pp. 901–915. doi: 10.1111/nph.14582.

Schellenberg, B., Ramel, C. and Dudler, R. (2010) '*Pseudomonas syringae* virulence factor syringolin a counteracts stomatal immunity by proteasome inhibition', *Molecular Plant-Microbe Interactions*, 23(10), pp. 1287–1293. doi: 10.1094/MPMI-04-10-0094.

Schindelin, J. *et al.* (2012) 'Fiji: An open-source platform for biological-image analysis', *Nature Methods*. doi: 10.1038/nmeth.2019.

Scofield, A. M. *et al.* (1990) 'Inhibition of thioglucosidase-catalysed glucosinolate hydrolysis by castanospermine and related alkaloids', *Phytochemistry*, 29(1), pp. 107–109. doi: 10.1016/0031-9422(90)89021-Z.

Scott, L. J. and Spencer, C. M. (2000) 'Miglitol: A review of its therapeutic potential in type 2 diabetes mellitus', *Drugs*, 59(3), pp. 521–549. doi: 10.2165/00003495-200059030-00012.

Seifert, G. J. and Roberts, K. (2007) 'The biology of arabinogalactan proteins', *Annual Review of Plant Biology*, 58, pp. 137–161. doi: 10.1146/annurev.arplant.58.032806.103801.

Sellstedt, A. and Richau, K. H. (2013) 'Aspects of nitrogen-fixing actinobacteria, in particular free-living and symbiotic *Frankia*', *FEMS Microbiology Letters*, 342(2), pp. 179–186. doi: 10.1111/1574-6968.12116.

Seo, M.-J. *et al.* (2013) 'Isolation of the putative biosynthetic gene cluster of 1-deoxynojirimycin by *Bacillus amyloliquefaciens* 140N, its production and application to the fermentation of soybean paste', *Bioscience, Biotechnology, and Biochemistry*, 77(2), pp. 398–401. doi: 10.1271/bbb.120753.

Sévin, D. C. *et al.* (2017) 'Nontargeted in vitro metabolomics for high-throughput identification of novel enzymes in Escherichia coli', *Nature Methods*, 14(2), pp. 187–194. doi: 10.1038/nmeth.4103.

Shabab, M. *et al.* (2008) 'Fungal effector protein AVR2 targets diversifying defense-related cys proteases of tomato', *Plant Cell*, 20(4), pp. 1169–1183. doi: 10.1105/tpc.107.056325.

Shibano, M. *et al.* (2004) 'Biosynthesis of 1-deoxynojirimycin in *Commelina communis*: A difference between the microorganisms and plants', *Phytochemistry*, 65(19), pp. 2661–2665. doi: 10.1016/j.phytochem.2004.08.013.

Shindo, T. *et al.* (2016) 'Screen of non-annotated small secreted proteins of *pseudomonas syringae* reveals a virulence factor that inhibits tomato immune proteases', *PLoS Pathogens*, 12(9), pp. 1–24. doi: 10.1371/journal.ppat.1005874.

Silby, M. W. *et al.* (2011) '*Pseudomonas* genomes: Diverse and adaptable', *FEMS Microbiology Reviews*, 35(4), pp. 652–680. doi: 10.1111/j.1574-6976.2011.00269.x.

Singh, B. *et al.* (2014) 'An alpha-glucosidase inhibitor from an endophytic *Cladosporium sp*. with potential as a biocontrol agent', *Applied Biochemistry and Biotechnology*, 175(4), pp. 2020–2034. doi: 10.1007/s12010-014-1325-0.

Slavokhotova, A. A. *et al.* (2014) 'Novel mode of action of plant defense peptides - hevein-like antimicrobial peptides from wheat inhibit fungal metalloproteases', *FEBS Journal*, 281(20), pp. 4754–4764. doi: 10.1111/febs.13015.

Smirnova, A. V. *et al.* (2002) 'Control of temperature-responsive synthesis of the phytotoxin coronatine in *Pseudomonas syringae* by the unconventional two-component system CorRPS', *Journal of Molecular Microbiology and Biotechnology*, 4(3), pp. 191–196.

Smith, D. L. and Gross, K. C. (2000) 'A family of at least seven β-galactosidase genes is expressed during tomato fruit development', *Plant Physiology*, 123(3), pp. 1173–1183. doi: DOI 10.1104/pp.123.3.1173.

Snelders, N. C. *et al.* (2020) 'Microbiome manipulation by a soil-borne fungal plant pathogen using effector proteins', *Nature Plants*, 6. doi: 10.1038/s41477-020-00799-5.

Sreedharan, A. *et al.* (2006) 'CorR regulates multiple components of virulence in *Pseudomonas syringae* pv. *tomato* DC3000', *Mol.Plant Microbe Interact.*, 19, pp. 768–779.

Stanley, D. *et al.* (2011) 'The role of α-glucosidase in germinating barley grains', *Plant Physiology*, 155(2), pp. 932–943. doi: 10.1104/pp.110.168328.

Stavrinides, J., McCloskey, J. K. and Ochman, H. (2009) 'Pea aphid as both host and vector for the phytopathogenic bacterium *Pseudomonas syringae*', *Applied and Environmental Microbiology*, 75(7), pp. 2230–2235. doi: 10.1128/AEM.02860-08.

Stecher, G., Tamura, K. and Kumar, S. (2020) 'Molecular evolutionary genetics analysis (MEGA) for macOS', *Molecular Biology and Evolution*. doi: 10.1093/molbev/msz312.

Steele, H. *et al.* (2010) 'Analysis of the natural infection of European horse chestnut (*Aesculus hippocastanum*) by *Pseudomonas syringae* pv. *aesculi*', *Plant Pathology*. doi: 10.1111/j.1365-3059.2010.02354.x.

Stiitz, A. E. (1999) *Iminosugars as glycosidase inhibitors:nojirimycin and beyond*, *Wiley-VCH : New York*.

De Storme, N. and Geelen, D. (2014) 'Callose homeostasis at plasmodesmata: Molecular regulators and developmental relevance', *Frontiers in Plant Science*, 5(APR), pp. 1–23. doi: 10.3389/fpls.2014.00138.

Strange, R. N. and Scott, P. R. (2005) 'Plant disease: a threat to global food security', *Annual Review of Phytopathology*, 43(1), pp. 83–116. doi: 10.1146/annurev.phyto.43.113004.133839.

Strasser, R. *et al.* (2007) 'A unique β1,3-galactosyltransferase is indispensable for the biosynthesis of N-glycans containing Lewis a structures in *Arabidopsis thaliana*', *Plant Cell*, 19(7), pp. 2278–2292. doi: 10.1105/tpc.107.052985.

Strasser, R. (2016) 'Plant protein glycosylation', *Glycobiology*, 26(9), pp. 926–939. doi: 10.1093/glycob/cww023.

Strus, M. *et al.* (2016) 'Effects of the selected iminosugar derivatives on *Pseudomonas aeruginosa* biofilm formation', *Microbial Drug Resistance*, 22(8), pp. 638–645. doi: 10.1089/mdr.2015.0231.

Stütz, A. E. (ed.) (1998) *Iminosugars as Glycosidase Inhibitors*. Wiley. doi: 10.1002/3527601740.

Taguchi, F. *et al.* (2009) 'Glycosylation of flagellin from *Pseudomonas syringae* pv. *tabaci* 6605 contributes to evasion of host tobacco plant surveillance system', *Physiological and Molecular Plant Pathology*, 74(1), pp. 11–17. doi: 10.1016/j.pmpp.2009.08.001.

Taguchi, F. *et al.* (2010) 'Defects in flagellin glycosylation affect the virulence of *Pseudomonas syringae* pv. *tabaci* 6605', *Microbiology*, 156(1), pp. 72–80. doi: 10.1099/mic.0.030700-0.

Tang, D., Wang, G. and Zhou, J.-M. (2017) 'Receptor kinases in plant-pathogen interactions: more than pattern recognition', *The Plant Cell*, 29(4), pp. 618–637. doi: 10.1105/tpc.16.00891.

Tang, X., Xiao, Y. and Zhou, J. M. (2006) 'Regulation of the type III secretion system in phytopathogenic bacteria', *Mol Plant Microbe Interact*, 19(11), pp. 1159–1166. doi: 10.1094/MPMI-19-1159.

Tanthanuch, W. *et al.* (2008) 'Genomic and expression analysis of glycosyl hydrolase family 35 genes from rice (*Oryza sativa* L.)', *BMC Plant Biology*, 8. doi: 10.1186/1471-2229-8-84.

Thakur, S., Weir, B. S. and Guttman, D. (2016) 'Phytopathogen genome announcement: draft genome sequences of 62 *Pseudomonas syringae* type and pathotype strains', *Molecular Plant-Microbe Interactions*, 29(4), p. MPMI-01-16-0013-TA. doi: 10.1094/MPMI-01-16-0013-TA.

Thomas, E. L. and van der Hoorn, R. A. L. (2018) 'Ten prominent host proteases in plant-pathogen interactions', *International Journal of Molecular Sciences*, 19(2), pp. 1–12. doi: 10.3390/ijms19020639.

Tian, M. *et al.* (2004) 'A kazal-like extracellular serine protease inhibitor from *Phytophthora infestans* targets the tomato pathogenesis-related protease P69B', *Journal of Biological Chemistry*, 279(25), pp. 26370–26377. doi: 10.1074/jbc.M400941200.

Tian, M. *et al.* (2007) 'A *Phytophthora infestans* cystatin-like protein targets a novel tomato papain-like apoplastic protease', *Plant Physiology*, 143(1), pp. 364–377. doi: 10.1104/pp.106.090050.

Tian, M., Benedetti, B. and Kamoun, S. (2005) 'A second kazal-like protease inhibitor from *Phytophthora infestans* inhibits and interacts with the apoplastic pathogenesis-related protease P69B of tomato', *Plant Physiology*, 138(3), pp. 1785–1793. doi: 10.1104/pp.105.061226.

Tierens, K. F. M. J. *et al.* (2001) 'Study of the role of antimicrobial glucosinolate-derived isothiocyanates in resistance of Arabidopsis to microbial pathogens', *Plant Physiology*, 125(4), pp. 1688–1699. doi: 10.1104/pp.125.4.1688.

Timmusk, S., Grantcharova, N. and Wagner, E. G. H. (2005) '*Paenibacillus polymyxa* invades plant roots and forms biofilms', *Applied and Environmental Microbiology*, 71(11), pp. 7292–7300. doi: 10.1128/AEM.71.11.7292-7300.2005.

Toruño, T. Y., Stergiopoulos, I. and Coaker, G. (2016) 'Plant-pathogen effectors: cellular probes interfering with plant defenses in spatial and temporal manners', *Annual Review of Phytopathology*, 54, pp. 419–441. doi: 10.1146/annurev-phyto-080615-100204.

Trda, L. *et al.* (2015) 'Perception of pathogenic or beneficial bacteria and their evasion of host immunity: pattern recognition receptors in the frontline', *Frontiers in Plant Science*, 6, pp. 1–11. doi: 10.3389/fpls.2015.00219.

Tsujii, E. *et al.* (1996) 'Nectrisine is a potent inhibitor of α-glucosidases, demonstrating activities similarly at enzyme and cellular levels', *Biochemical and Biophysical Research Communications*, 220(2), pp. 459–466. doi: 10.1006/bbrc.1996.0427.

Ullrich, M. *et al.* (1995) 'A modified two-component regulatory system is involved in temperature-dependent biosynthesis of the *Pseudomonas syringae* phytotoxin coronatine', 177(21), pp. 6160–6169.

Underwood, W. (2012) 'The plant cell wall: A dynamic barrier against pathogen invasion', *Frontiers in Plant Science*, 3, pp. 1–6. doi: 10.3389/fpls.2012.00085.

Varki, A. (2017) 'Biological roles of glycans', *Glycobiology*, 27(1), pp. 3–49. doi: 10.1093/glycob/cww086.

Vencato, M. *et al.* (2006) 'Bioinformatics-enabled identification of the Hrpl regulon and type III secretion system effector proteins of *Pseudomonas syringae* pv. *phaseolicola* 1448A', *Molecular Plant-Microbe Interactions*, 19(11), pp. 1193–1206. doi: 10.1094/MPMI-19-1193.

Viuff, A. H. *et al.* (2015) 'Stable analogues of nojirimycin-synthesis and biological evaluation of nojiristegine and manno-nojiristegine', *Organic and Biomolecular Chemistry*, 13(37), pp. 9637–9658. doi: 10.1039/c5ob01281c.

Voigt, C. A. (2014) 'Callose-mediated resistance to pathogenic intruders in plant defense-related papillae', *Frontiers in Plant Science*, 5, pp. 1–6. doi: 10.3389/fpls.2014.00168.

Vorwerk, S., Somerville, S. and Somerville, C. (2004) 'The role of plant cell wall polysaccharide composition in

disease resistance', *Trends in Plant Science*, 9(4), pp. 203–209. doi: 10.1016/j.tplants.2004.02.005.

Waite, C. *et al.* (2017) 'Negative autogenous control of the master type III secretion system regulator HrpL in *Pseudomonas syringae*', 8(1), pp. 1–16.

Wakabayashi, T. *et al.* (2015) 'Planteose as a storage carbohydrate required for early stage of germination of *Orobanche minor* and its metabolism as a possible target for selective control', *Journal of Experimental Botany*, 66(11), pp. 3085–3097. doi: 10.1093/jxb/erv116.

Wang, C. *et al.* (2019) 'Agrobacteria reprogram virulence gene expression by controlled release of host-conjugated signals', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1903695116.

Wang, S. *et al.* (2021) 'Cleavage of a pathogen apoplastic protein by plant subtilases activates host immunity', *New Phytologist*, doi: 10.1111/nph.17120.

Ward, N. and Moreno-Hagelsieb, G. (2014) 'Quickly finding orthologs as reciprocal best hits with BLAT, LAST, and UBLAST: How much do we miss?', *PLoS ONE*. doi: 10.1371/journal.pone.0101850.

Watanabe, S. *et al.* (1995) 'Isolation of 2R,5R-Dihydroxymethyl-3R,4R-dihydroxypyrrolidine (DMDP) from the Fermentation Broth of *Streptomyces sp*. KSC-5791', *Bioscience, Biotechnology, and Biochemistry*, 59(5), pp. 936–937. doi: 10.1271/bbb.59.936.

Watson, A. A. *et al.* (2001) 'Polyhydroxylated alkaloids - Natural occurrence and therapeutic applications', *Phytochemistry*, 56(3), pp. 265–295. doi: 10.1016/S0031-9422(00)00451-9.

Wei, C. F. *et al.* (2007) 'A *Pseudomonas syringae* pv. *tomato* DC3000 mutant lacking the type III effector HopQ1-1 is able to cause disease in the model plant *Nicotiana benthamiana*', *Plant Journal*. doi: 10.1111/j.1365-313X.2007.03126.x.

Weingart, H. *et al.* (2001) 'The role of ethylene production in virulence of *Pseudomonas syringae* pvs. glycinea and *phaseolicola*', *Phytopathology*. doi: 10.1094/PHYTO.2001.91.5.511.

Williams, K. P. (2002) 'Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: Sublocation preference of integrase subfamilies', *Nucleic Acids Research*, 30(4), pp. 866–875. doi: 10.1093/nar/30.4.866.

Winchester, B. G. (2009) 'Iminosugars: from botanical curiosities to licensed drugs', *Tetrahedron Asymmetry*, 20(6–8), pp. 645–651. doi: 10.1016/j.tetasy.2009.02.048.

Winsor, G. L. *et al.* (2016) 'Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database', *Nucleic Acids Research*. doi: 10.1093/nar/gkv1227.

Wolfender, J. L. *et al.* (2019) 'Accelerating metabolite identification in natural product research: toward an ideal combination of liquid chromatography-high-resolution tandem mass spectrometry and nmr profiling, in silico databases, and chemometrics', *Analytical Chemistry*, 91(1), pp. 704–742. doi: 10.1021/acs.analchem.8b05112.

Woolston, B. M., Edgar, S. and Stephanopoulos, G. (2013) 'Metabolic engineering: Past and future', *Annual Review of Chemical and Biomolecular Engineering*, 4, pp. 259–288. doi: 10.1146/annurev-chembioeng-061312-103312.

Wrodnigg, T. M., Steiner, A. J. and Ueberbacher, B. J. (2008) 'Natural and synthetic iminosugars as carbohydrate processing enzyme inhibitors for cancer therapy', *Anti-cancer agents in medicinal chemistry*, 8(1), pp. 77–85. doi: 10.2174/187152008783330851.

Wu, H. *et al.* (2019) 'A novel strategy to regulate 1-deoxynojirimycin production based on its biosynthetic pathway in S*treptomyces lavendulae*', *Frontiers in Microbiology*, 10(AUG), pp. 1–12. doi: 10.3389/fmicb.2019.01968.

Wu, S. W. *et al.* (2018) 'Callose balancing at plasmodesmata', *Journal of Experimental Botany*, 69(22), pp. 5325–5339. doi: 10.1093/jxb/ery317.

Wu, Y., Arciola, J. and Horenstein, N. (2013) 'Medium-chain dehydrogenases with new specificity: amino mannitol dehydrogenases on the azasugar biosynthetic pathway', *Protein & Peptide Letters*, 21(1), pp. 10–14. doi: 10.2174/092986652101131219093413.

Xia, Y. *et al.* (2020) 'N-glycosylation shields *Phytophthora sojae* apoplastic effector PsXEG1 from a specific host aspartic protease', *Proceedings of the National Academy of Sciences of the United States of America*, 117(44), pp. 27685–27693. doi: 10.1073/pnas.2012149117.

Xiang, T. *et al.* (2008) '*Pseudomonas syringae* effector AvrPto blocks innate immunity by targeting receptor kinases', *Current Biology*. doi: 10.1016/j.cub.2007.12.020.

Xiao, Y. (2005) *Regulation of type III secretion system in Pseudomonas syringae*. Kansas State University.

Xie, Y. *et al.* (2020) 'Signal transduction schemes in *Pseudomonas syringae*', *Computational and Structural Biotechnology Journal*, 18, pp. 3415–3424. doi: 10.1016/j.csbj.2020.10.039.

Xie, Y., Shao, X. and Deng, X. (2019) 'Regulation of type III secretion system in *Pseudomonas syringae*', *Environmental Microbiology*, 21(12), pp. 4465–4477. doi: 10.1111/1462-2920.14779.

Xin, X.-F. and He, S. Y. (2013) '*Pseudomonas syringae* pv. *tomato* DC3000: a model pathogen for probing disease susceptibility and hormone signaling in plants', *Annual Review of Phytopathology*, 51(1), pp. 473–498. doi: 10.1146/annurev-phyto-082712-102321.

Xin, X. F., Kvitko, B. and He, S. Y. (2018) '*Pseudomonas syringae*: What it takes to be a pathogen', *Nature Reviews Microbiology*, 16(5), pp. 316–328. doi: 10.1038/nrmicro.2018.17.

Ya'kobovitz, M. K., Butters, T. D. and Cohen, E. (2016) 'Inhibition of α-glucosidase activity by N-deoxynojirimycin analogs in several insect phloem sap feeders', *Insect Science*, 23(1), pp. 59–67. doi: 10.1111/1744-7917.12229.

Yadav, V. G. *et al.* (2012) 'The future of metabolic engineering and synthetic biology: Towards a systematic practice', *Metabolic Engineering*, 14(3), pp. 233–241. doi: 10.1016/j.ymben.2012.02.001.

Yamamoto, M. *et al.* (2011) 'Identification of genes involved in the glycosylation of modified viosamine of flagellins in P*seudomonas syringae* by mass spectrometry', *Genes*. doi: 10.3390/genes2040788.

Yan, Q. *et al.* (2020) 'Ancient co-option of an amino acid ABC transporter locus in *Pseudomonas syringae* for host signal-dependent virulence gene regulation', *PLoS Pathogens*, 16(7), pp. 1–26. doi: 10.1371/journal.ppat.1008680.

Yang, D. *et al.* (2020) 'Metabolic Engineering of *Escherichia coli* for Natural Product Biosynthesis', *Trends in Biotechnology*, 38(7), pp. 745–765. doi: 10.1016/j.tibtech.2019.11.007.

Yoo, Y. *et al.* (2019) 'Inhibitory effect of *Bacillus velezensis* on biofilm formation by *Streptococcus mutans*', *Journal of Biotechnology*, 298, pp. 57–63. doi: 10.1016/j.jbiotec.2019.04.009.

Yu, J. *et al.* (1999) 'Involvement of the exopolysaccharide alginate in the virulence and epiphytic fitness of *Pseudomonas syringae* pv. *syringae*', *Molecular Microbiology*, 33(4), pp. 712–720. doi: 10.1046/j.1365-2958.1999.01516.x.

Zhang, J. H., Dawe, G. and Stemmer, W. P. C. (1997) 'Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.94.9.4504.

Zhang, K. *et al.* (2019) 'Overexpressing the myrosinase gene *tgg1* enhances stomatal defense against *pseudomonas syringae* and delays flowering in Arabidopsis', *Frontiers in Plant Science*, 10, pp. 1–14. doi: 10.3389/fpls.2019.01230.

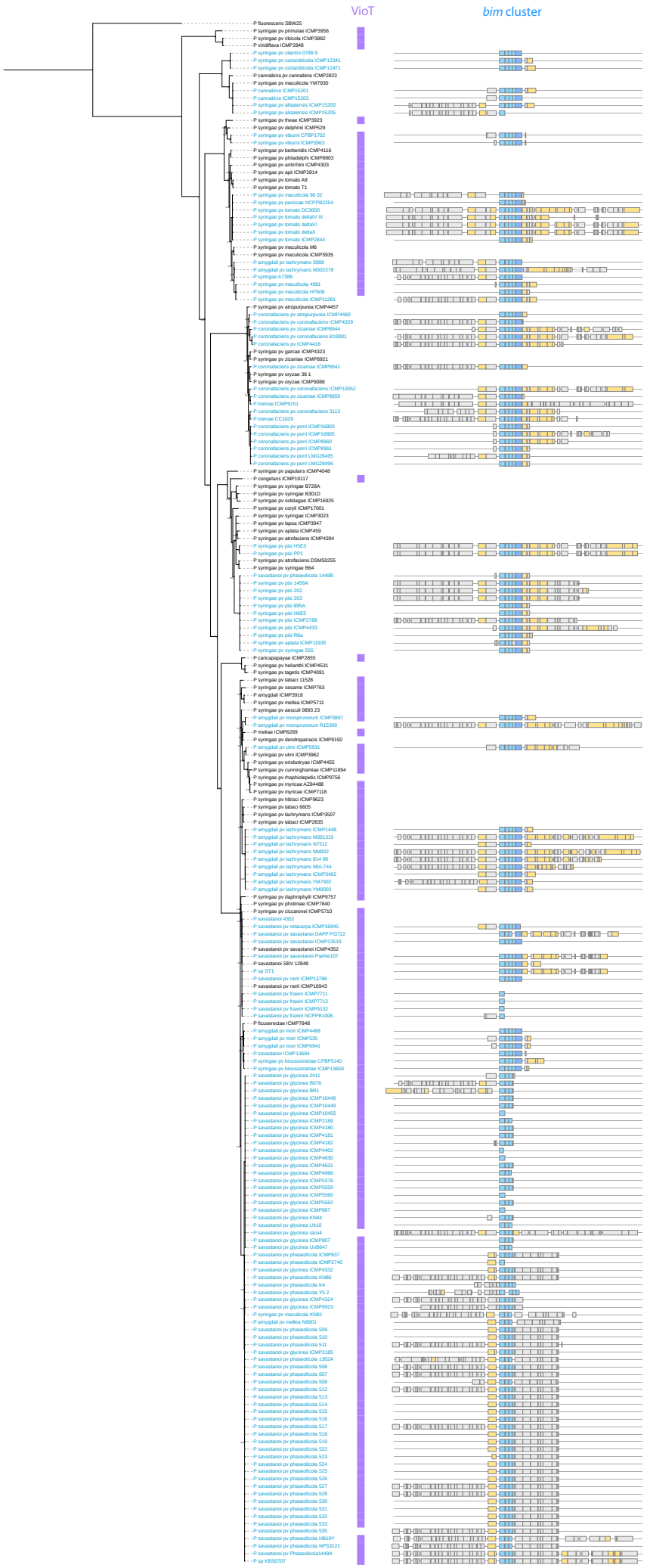
Zhang, W. *et al.* (2019) 'An overview of the biological production of 1-deoxynojirimycin: current status and future perspective', *Applied Microbiology and Biotechnology*, 103(23–24), pp. 9335–9344. doi: 10.1007/s00253-019-10191-9.

Appendix 1

Homologs of the alcohol dehydrogenase PSPTO0834 are present in different putative biosynthetic gene clusters for iminosugars in a wide range of bacterial groups.

A phylogenetic tree was reconstructed using amino acid sequence of homologs of PSPTO0834 identified by PSI-BLAST. The amino acid sequences were aligned with MUSCLE and partial deletion with 95% cutoff was used to remove gaps in the alignment. Maximum Likelihood (ML) phylogenetic tree was reconstructed. Branch support values were calculated from Bootstrap method with 100 replications. Branches with less than 0.5 branch support value were collapsed. Putative gene clusters containing each homolog of the alcohol dehydrogenase in each genome were assigned by identifying a stretch of consecutive genes with the same orientation on the genome (same strand) as the alcohol dehydrogenase gene both up and downstream until there is a gene with an opposite orientation. Gene annotation is shown as rectangles. Notable genes of interest are highlighted with different colours. Association with transposable elements (indicated with solid black circle) was defined as the presence of annotated transposase, integrase, recombinase or conjugative elements within five genes distance from the first and last gene in the cluster of interest. It is also worth noting that many available genome sequences were not complete assemblies so fragmented genomic regions and incompletely annotated genes are common. Genes of interest include homologs of the alcohol dehydrogenase PSPTO0834 (ADH), aminotransferase (AMT), phosphatase (HAD or FIG family), transporter (MFS or MATE or SulP family), reductase (RibD family), glycosidase (GH) and protein of unknown function (DUF1349).





Appendix 2

The bim cluster is distributed among most major phylogroups of *P. syringae*. The phylogenetic tree was reconstructed using concatenated coding DNA sequences of 4 conserved housekeeping genes (rpoD, gyrB, gltA and gapA). *P. fluorescens* SBW25 was included as an outgroup. The DNA sequences were aligned with MUSCLE and partial deletion with 95% cutoff was used to remove gaps in the alignment. Maximum Likelihood (ML) phylogenetic tree was reconstructed. Branch support values were calculated from Bootstrap method with 100 replications. Branches with less than 0.5 branch support value were collapsed. Strains containing the bim cluster are labelled in blue. Phylogroup classification is displayed on each branch. VioT solid squares indicate that the strain has a transferase for modified viosamine in flagellin glycosylation. Genomic regions (20 kbp) around the bim cluster are displayed. Annotated genes are shown as rectangles. The bim cluster genes PSPTO0834-7 are labelled in light blue, PSPTO0838 is labelled in dark blue. Annotated transposase, integrases, recombinases or conjugative elements are labelled in yellow. Incomplete assembly of genome is common as can be seen when the genome region has no gene annotations (rectangles) displayed.