Computational modeling of virulence genes expression : A study of gene regulation during plant pathogenesis

Shiny Martis Badiadka

N°d'ordre NNT : 2020LYSEI100

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
Institut National des sciences appliquées

Ecole Doctorale ED 341
Ecosystèmes Evolution Modélisation Microbiologie

Spécialité/ discipline de doctorat :

Biomath-Bioinfo-Génomique evolutive

Soutenue publiquement le 06/11/2020, par :
Shiny MARTIS BADIADKA

Computational modeling of virulence genes expression: a study of gene regulation during plant pathogenesis

---

Devant le jury composé de :

GENIN,	Stephane	Directeur de Recherche	CNRS	Rapporteur
de-JONG,	Hidde	Directeur de Recherche	INRIA	Rapporteur
REVERCHON,	Sylvie	Professeur	INSA, Lyon	Directrice de thèse
MEYER,	Sam	Maitre de conferences	INSA, Lyon	Co-directeur de thèse

Département FEDORA – INSA Lyon - Ecoles Doctorales – Quinquennal 2016-2020

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	**CHIMIE DE LYON** **http://www.edchimie-lyon.fr** Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage **secretariat@edchimie-lyon.fr** INSA : R. GOURDON	**M. Stéphane DANIELE** Institut de recherches sur la catalyse et l'environnement de Lyon IRCELYON-UMR 5256 Équipe CDFA 2 Avenue Albert EINSTEIN 69 626 Villeurbanne CEDEX **directeur@edchimie-lyon.fr**
E.E.A.	**ÉLECTRONIQUE, ÉLECTROTECHNIQUE, AUTOMATIQUE** **http://edeea.ec-lyon.fr** Sec. : M.C. HAVGOUDOUKIAN **ecole-doctorale.eea@ec-lyon.fr**	**M. Gérard SCORLETTI** École Centrale de Lyon 36 Avenue Guy DE COLLONGUE 69 134 Écully Tél : 04.72.18.60.97 Fax 04.78.43.37.17 **gerard.scorletti@ec-lyon.fr**
E2M2	**ÉVOLUTION, ÉCOSYSTÈME, MICROBIOLOGIE, MODÉLISATION** **http://e2m2.universite-lyon.fr** Sec. : Sylvie ROBERJOT Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 INSA : H. CHARLES **secretariat.e2m2@univ-lyon1.fr**	**M. Philippe NORMAND** UMR 5557 Lab. d'Ecologie Microbienne Université Claude Bernard Lyon 1 Bâtiment Mendel 43, boulevard du 11 Novembre 1918 69 622 Villeurbanne CEDEX **philippe.normand@univ-lyon1.fr**
EDISS	**INTERDISCIPLINAIRE SCIENCES-SANTÉ** **http://www.ediss-lyon.fr** Sec. : Sylvie ROBERJOT Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 INSA : M. LAGARDE **secretariat.ediss@univ-lyon1.fr**	**Mme Sylvie RICARD-BLUM** Institut de Chimie et Biochimie Moléculaires et Supramoléculaires (ICBMS) - UMR 5246 CNRS - Université Lyon 1 Bâtiment Curien - 3ème étage Nord 43 Boulevard du 11 novembre 1918 69622 Villeurbanne Cedex Tel : +33(0)4 72 44 82 32 **sylvie.ricard-blum@univ-lyon1.fr**
INFOMATHS	**INFORMATIQUE ET MATHÉMATIQUES** **http://edinfomaths.universite-lyon.fr** Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage Tél : 04.72.43.80.46 **infomaths@univ-lyon1.fr**	**M. Hamamache KHEDDOUCI** Bât. Nautibus 43, Boulevard du 11 novembre 1918 69 622 Villeurbanne Cedex France Tel : 04.72.44.83.69 **hamamache.kheddouci@univ-lyon1.fr**
Matériaux	**MATÉRIAUX DE LYON** **http://ed34.universite-lyon.fr** Sec. : Stéphanie CAUVIN Tél : 04.72.43.71.70 Bât. Direction **ed.materiaux@insa-lyon.fr**	**M. Jean-Yves BUFFIÈRE** INSA de Lyon MATEIS - Bât. Saint-Exupéry 7 Avenue Jean CAPELLE 69 621 Villeurbanne CEDEX Tél : 04.72.43.71.70 Fax : 04.72.43.85.28 **jean-yves.buffiere@insa-lyon.fr**
MEGA	**MÉCANIQUE, ÉNERGÉTIQUE, GÉNIE CIVIL, ACOUSTIQUE** **http://edmega.universite-lyon.fr** Sec. : Stéphanie CAUVIN Tél : 04.72.43.71.70 Bât. Direction **mega@insa-lyon.fr**	**M. Jocelyn BONJOUR** INSA de Lyon Laboratoire CETHIL Bâtiment Sadi-Carnot 9, rue de la Physique 69 621 Villeurbanne CEDEX **jocelyn.bonjour@insa-lyon.fr**
ScSo	**ScSo*** **http://ed483.univ-lyon2.fr** Sec. : Véronique GUICHARD INSA : J.Y. TOUSSAINT Tél : 04.78.69.72.76 **veronique.cervantes@univ-lyon2.fr**	**M. Christian MONTES** Université Lyon 2 86 Rue Pasteur 69 365 Lyon CEDEX 07 **christian.montes@univ-lyon2.fr**

*ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

Computational modeling of virulence genes expression: a study of gene regulation during plant pathogenesis

Abstract

Plant pathogens are a major threat to food security all around the world. Virulence genes enable the pathogens to disease the host and pose threats. We study and computationally model the expression of virulence in pathogens. The *pel* genes, a major virulence factor in pectinolytic bacteria which infect plants and cause soft rot disease. The Pel enzymes released by the pathogen result in the maceration of tissue in crops and facilitate pathogenesis. We model the *pel* genes in our model organism *Dickeya dadantii*. We study *pelD/E* genes which encode proteins with similar enzymatic activities, but also share the same set of transcriptional regulators, thus raising the question of the benefit of this duplication and the specific role of these two genes. We model their expression with respect to the KdgR repressor and CRP activator. We develop a quantitative dynamic kinetic model of this process that reproduces the observed behavior of the two genes and explains their specific role in the infection. The regulatory part of the model is based on experimental values, and fitted kinetic parameters are subjected to systematic evaluation. The model is robust and has multiple applications for studying the pathogenesis of pectinolytic bacteria. In addition, we explore the instance of carbon catabolite repression observed during the regulation of virulence genes *Dickeya dadantii*.

The novel antibiotic seconeolitsin is characterized in the gram negative model organism *Dickeya dadantii*. Growth kinetics and MIC studies are carried out. The variations in the level of supercoiling in the presence of seconeolitsin and novobiocin are quantified and analyzed. The effect of supercoiling on the expression of genes, especially the *pelE* gene is studied using the novel antibiotic seconeolitsin. The study explores the role of supercoiling as a gene regulator.

Acknowledgements

I thank Stephane Genin and Hidde de-Jong for accepting to be the rapporteurs. I am greatly thankful to my thesis committee Hans Geiselman, Elisabeth Kay and Ivan Junier for their critical analysis and inputs. I am grateful to Michel droux for the collaboration and teaching me to work with HPLC. Most of all I am thankful to Sam and Sylvie, my directors for accepting me as a student, guiding me and making my thesis happen.

I extend my gratitude to the laboratory MAP and all the colleagues who helped me with familiarization of the new culture and work. I thank the lab director William Nasser for answering every query of mine and the guidance.

I am sincerely thankful to Lucie, Laurent and Florence from the Organic and Bioorganic chemistry team at ICBMS laboratory for the synthesis of seceoneolitsin.

I thank the medical caregivers at Eduard Herriot hospital Lyon for saving me a lot of pain and helping me get healthy. I am grateful to my family, friends and everyone who supported me, helped and led me towards success.

It has been a difficult journey made easier with kindness from many people. I am ever grateful to ALL!

Table of contents:

List of figures:

Figure 3.3.18: The effect of seconeolitsin on *pelE* expression at stationary state. A represents the luminescence recorded along the growth period represented in B. The bacteria is grown in M63 medium. C represents the luminescence recorded after the seconeolitsin shock treatment

Figure 3.3.19: The standard curve for the quantification of *pAW, pelE* and *gyrB* external normalizer

Figure 3.3.20: Quantification of the level of genes expression with qRT-PCR method. The expression of pelE, gyrB and topA genes are shown in the figure. Calculations carried out based on standard curve estimation method. The obtained result is proportional and hence multiplied by an arbitrary constant to obtain a meaningful number. The results can be compared based on their levels and are not absolute

List of tables:

List of abbreviations:

Etracytoplasmic function (ECF)

ROS (Reactive Oxygen Species)

catabolite control protein A (CcpA)

catabolite responsive elements (*cre*)

Tricarboxylic acid cycle (TCA)

Catabolite repressor protein (CRP)

Carbon catabolite repression (CCR)

Cyclic adenosine monophosphate (cAMP)

NAPs (Nucleoid Associated Proteins)

Transcription factor binding site (TFBS)

Ordinary differential equation (ODE).

5-keto-4-deoxyuronate (DKI)

Factor for inversion stimulation (Fis)

Endo-pectate lyases (Pels)

2-keto-3-deoxygluconate (KDG)

Polygalacturonate (PGA)

2 -Chloroacetaldehyde (CAA)

Dimethyl sulfoxide (DMSO)

Dichloromethane (DCM)

α-chloroethyl chloroformate (ACE-Cl)

Optical Density (OD)

Green fluorescent protein (GFP)

Minimum Inhibitory Concentrations (MIC)

LB (Liquid Broth)

List of parameters:

KA - The dissociation constant for the binding of activator molecules at the DNA binding site

f - The fold change in the expression of genes when compared to the absence of activator to the complete activation

KR - The concentration of repressor molecules required to bind half of the capacity of repressor binding sites available

A - maximum growth achieved by the bacteria under study

μ_m - growth rate of the bacteria under study

Λ - lag time observed for the onset of exponential bacterial growth

1. Pathogenesis in plants is a global threat

Plants, the green abundance on earth are most important for the vitality of living beings and balance in ecosystems. They are the lungs of earth purifying the air we breathe. Plants are the basis of all the food sources and a major share of the burden of food security of the planet is on plant resources. Ensuring plant health ensures the sustenance of organisms dependent on them. Plants are attacked by pests and pathogens threatening their lives and posing a major food sustainability challenge all over the world. Plant pathogens are responsible for a significant amount of crop losses worldwide every year. Agricultural researchers are working on full swing to meet the needs of the growing human population. Production of enough food for the world population gets challenging due to pathogenic threat, given that there is a need of about 50% increase production to meet the needs in coming 30 years [1]. It is essential to get over pathogenic threats and help farmers feed world hunger and ensure food security. A study on pathogen threat on major food crops enumerates the loss caused every year to be over 20% and emphasizes on the prioritization of crop health management to improve the sustainability of agroecosystems in delivering services to societies [2]. A vast variety of plant pathogens exist and their effects range from mild symptoms to catastrophes affecting large areas of planted food crops. Raising of the resistant strains of phytopathogens are also causing difficulties in disease management. Especially bacterial plant pathogens are particularly difficult to control due to the scarcity of chemical control agents for bacteria with the exception of antibiotics. However, the use of antibiotics is restricted in many countries due to the high potential for evolution of antibiotic resistance in plant pathogens and its eventual transmission to human pathogens. In addition, bacteria show an incredible ability to disseminate and adapt to changing environments like colonization in a new host, geographic invasion of new areas, etc. Climate change is also likely to accelerate the evolution of pathogens and increase their diversity making it difficult to tackle them [3]. To migrate or evolve are two options for plant pathogens under conditions of climate change. Therefore, the ways for combating bacterial plant diseases are limited and it is a laborious task to prevent the evolving threat.

Fighting plant pathogens is necessary to overcome the loss caused by them. It is necessary to study pathogenesis in plants, define the biological problem and develop measures to fight it. In addition to identification of pathogens, identification of the disease causing virulence mechanism is essential to understand the mode of infection and to estimate the level of threat posed [4]. Virulence refers to the uphand of the pathogen being able to break down the host barrier and establish itself in the host and flourish by destroying the host. Pathogens have various virulence mechanisms in order to fight the difficulties in the environments posing threat to their lives. Disease

eradication mechanisms can be developed based on our understanding of the virulence system of pathogens and thus ensuring their eradication.

In this context, there is a critical need for comprehensively understanding the process of bacterial pathogenesis in plants in order to develop novel knowledge-based strategies to control plant bacterial diseases. *Dickeya spp.* are among the top 10 most important bacterial pathogens in agriculture limiting the crop yield and quality [5]. *Dickeya* are now considered as emergent pathogens in European countries [6]. The apoplast is a nutrient-limited environment that is guarded by plant defenses, so *Dickeya* have evolved with an intricate strategy to successfully colonize this niche. In the course of this thesis, we propose a method to study the virulence mechanism by modeling the virulence gene expression in the model organism of choice, *Dickeya dadantii*. We choose the *pel* genes as the model virulence genes to study their mechanism and develop a transcriptional regulation model to elucidate the dynamics. The computational model developed will be a unique tool that can be used in research to understand the pathogenic virulence system and its action mechanism.

2. Bacterial virulence genes

Pathogenesis in plants evolves as an interactive two-way communication. Pathogens sense the plant environment signals and respond to establish themselves. Plants also sense their presence and activate the defense mechanisms [7]. Timely and appropriate response of pathogens to plant defense mechanisms end up in successful pathogenesis. Pathogens infect hosts in progressive manner. They sense the host environments and employ their virulence system to overcome the host defense. They undergo metabolic shifts to obtain their energy. Its combined action of sensing the environmental stimulus and adapting to the extremities leads to a successful infection

Virulence is defined as an ability of a pathogenic organism to infect and propagate disease in the host organism. Molecules that aid the pathogen in disease and colonize in the host are called virulence factors. The virulence factors are cytosolic, secretory or membrane associated [8]. The adaptive, metabolic and physiological changes occurring in pathogens as a response to the host environment are facilitated by membrane associated cytosolic factors. The adhesion and evasion along the host epithelial cell are accomplished by membrane associated virulence factors. Once the pathogen confines itself in the host, the innate and adaptive immune response of the host is tackled by the bacterial arsenal consisting of secreted factors. In extracellular pathogens, the secreted virulence factors act synergistically to kill the host cells [8]. Genes responsible for the expression of virulence factors and facilitation of pathogenicity are broadly classified as virulence genes. The entire pool of genes functioning with respect to virulence action consists of a virulome.

The definition of virulence genes tend to vary from study to study, but it is evident that any genes associated with the pathogenic lifestyle of the organism can be classified as virulence genes. More specifically, based on their connection to virulence they can be categorized as i) True virulence genes ii) Virulence associated genes and iii) Virulence lifestyle genes [9]. Figure 2.1 shows this classification.

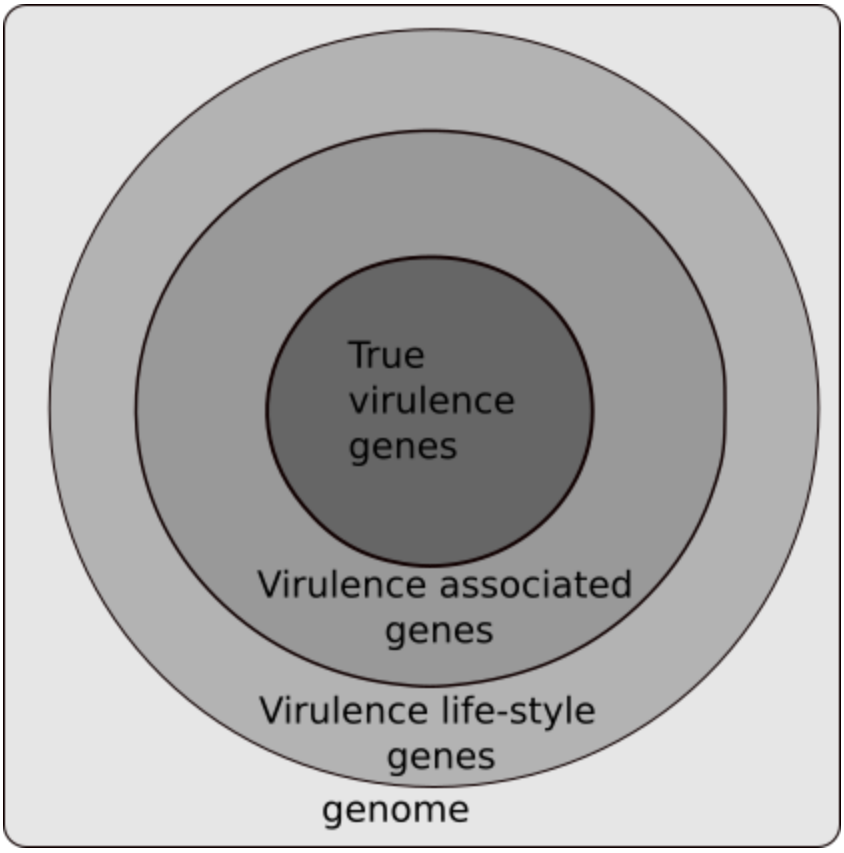


Figure 1.2.1: The categorization of Virulence genes based on their functionalities [9]

Based on the strict definition of virulence, fewer genes qualify to be called virulence genes. In Figure 2.1 concentric circles represent the classification of genes based on their function in virulence action. The square box represents the entire genome of a pathogen. The innermost circle refers to the genes coding to factors directly involved in pathogenesis causing pathological damage and are absent in non-pathogenic strains. These kinds of genes are called true virulence genes. The second circle from inside represents the genes indirectly involved in disease propagation are called virulence-associated genes. They are further classified as follows

a. Processing virulence genes which function in post-translational modification of virulence factors, folding of the virulence gene proteins, secretion of virulence factors, *etc.*
b. Auxiliary virulence gene responsible for activation of virulence factors
c. Regulatory virulence genes helping in regulation of the true virulence genes

In the outermost circle an additional set of genes are referred as virulence life-style genes. They consist of the following type of genes

a. Adaptive genes expressing factors enabling the colonization of the host nd the survival to hostile conditions encountered in the host
b. Evasion genes expressing factors involved in evasion of the host immune system

Altogether the involvement of genes with the virulence function speaks about the importance of the gene for the pathogenesis process. Understanding this basic classification increases chances at targeted disease control models. We will be focusing on the virulence genes in phytopathogen *Dickeya dadantii* which will be discussed in the coming section.

3. The Dickeya Model

Dickeya dadantii is a Gram-negative bacillus belonging to the family *Pectobacteriaceae* which causes soft rot disease in plants. We chose this potential phytopathogen to be the model to study the regulation of virulence genes. In order to effectively study the model organism, understanding of its pathogenesis process is more important. Figure 3.1 depicts the process of plant infection by *Dickeya dadantii* with respect to location and process involved. These bacteria are extracellular necrotrophic pathogens propagating in plant apoplast and destroying plant cell-wall resulting in soft rot disease. There are three main steps in *Dickeya* infection process

1. Adhesion to the plant surface and entry in to the tissue at an opening like wound
2. Colonization in the plant cell apoplastic space
3. Maceration of the Plant tissue and disease establishment

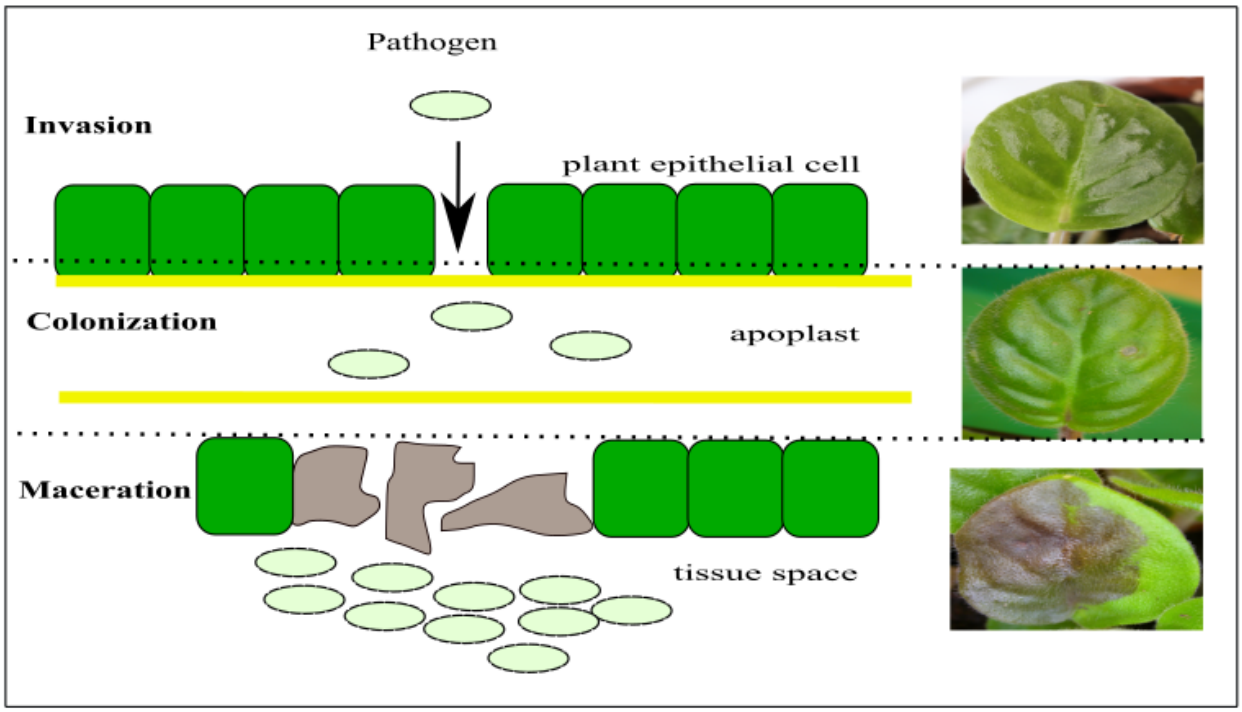


Figure 1.3.1: The *Dickeya dadantii* infection process in Saintpaulia ionantha leaf

In order to achieve each of these stages, a delicate sensing of the environment and accurate virulence action to overcome the difficulties is observed. The employment of virulence strategies must be tightly coordinated with the bacterial growth to rationalize the energy efficiency, disease development and evade the host defense and bacterial

survival [10]. **Adhesion** can be occurring in a leaf, stem or roots or any part of the plant open to the pathogen and easy for the pathogen to live on. The exopolysaccharides encoding genes *bcsABCD*, *wza-wzb-wzc* are identified to be virulence genes playing a major role in the adhesion process [11], [12]. After adhesion, the bacteria reciprocate to the plant signals and invade the plant apoplast looking for energy sources and destroying the primary defense barrier of the plants. The *sapABCDF* transporter is responsible for specific entry of the antimicrobial peptides from the plant allowing their degradation by bacterial cytoplasmic proteases, The *arnB-T* genes involved in modification of the liposaccharide reduce the sensibility to the plant antimicrobial peptides [13], [14]. In this step, most of the incompetent bacteria fail as the plant cell wall renforced by callose deposition limiting bacterial invasion. The competent pathogen successfully invading into the plant tissue will then suit itself to proliferate on the available nutrient sources by colonizing and overcoming the host defense and adapt itself to the metabolic and environmental changes [15]. During the colonization process the bacterial population density increases, bacteria overcome the plant defense and cause plant tissue maceration resulting in soft rot disease. Thus the cycle of pathogenesis is complete with the bacteria establishing itself in the host and diseasing the host. Bacterial virulence factors enable the bacteria to propagate pathogenesis. Expression of virulence and adaptive genes in accurate time breaks the host defense and sets in the bacterial attack and helps the bacteria in growing in the host environment.

In order to achieve each of these stages, a delicate sensing of the environment and accurate virulence action to overcome the difficulties is observed. The virulence genes action at correct timing is vital for establishing pathogenesis. There are various stages of onset of virulence action depending on the stress or threat experienced by the bacteria. After the adhesion on plant surface, continuous monitoring of change in environment is required to identify an opening into the plant cell apoplast [12]. Motility related genes eg *motAB, cheAWDRBYZ* and flagellar genes are expressed at this instance of time and entry in plant apoplast is mediated by jasmonic acid, which is produced by wounded tissue and may enable the bacterial cells to move toward plant wounds and facilitate systemic invasion [16]. In the plant apoplast, multiple efflux pumps are activated to enable the survival of bacteria and to over the plant defense system. Bacteria also adapt to the change of environment inside the apoplast for example the type VI secretion system and type 6 effectors like *impBCGHJKAL, rhsABC* are triggered by the sensed **acidic stress** signals [17]. Reactive oxygen species, the first line of plant defense reaction are overcome by *sodAC, katGE, ahpCF, indABC, msrA* virulence genes [18]–[21]. Finally, the bacteria establishes themself in plant apoplast utilizing the energy from the macromolecules like pectins degraded by their metabolic changes acquired in the host environment. The degradation of macromolecules tends to

breakdown of the plant cell wall and thus resulting in an **osmotic stress** which is sensed by the bacteria that respond by producing the osmoprotectant molecules glutamine, glutamate and α-glucosylglycerate [22]. The table below lists out the various virulence and adaptive genes involved in various stages of infection

Table 1.1. Selected virulence genes involved at various stages of interaction [10]

Gene ID	Gene name	Stage of interaction	Description	References
ABF001 7612	bcsABCD	Epiphytic	Cellulose microfibrils synthesis, adhesion to plant surface	(Jahn et al., 2011)
ABF002 0470	wza-wzb-wz c…	Epiphytic	Exopolysaccharide synthesis, adhesion to plant surface	(Condemine et al., 1999)
ABF002 0788	hecBA2/ cdiBA2	Epiphytic	Type V autotransporter contact dependent inhibition system contributing to bacterial adhesion to plant surface	(Aoki et al., 2010)
ABF001 7267	rhsA	Epiphytic	Intercellular competition system	(Koskiniemi et al., 2013)
ABF001 8746	rhsB	Epiphytic	Intercellular competition system	
ABF004 6548	rhsC	Epiphytic	Intercellular competition system	
ABF001 5869	impBCGHJK AL	Epiphytic	Type VI secretion system involved in Rhs secretion	
ABF001 9839	rhlA	Epiphytic	Biosurfactant synthesis, colonization of plant surface	(Hommais et al., 2008)
ABF001 8761	motAB-cheA WDRBYZ	Epiphytic and other	Flagella motor and chemotaxis, entry into plant apoplast	(Antunez-Lamas et al., 2009a)
ABF002 0713	cfa	Early	Cyclopropane fatty acid synthesis, resistance to acidic pH	Reverchon et al., unpublished
ABF001 4750	asr	Early	Acid shock periplasmic protein, resistance to acidic pH	
ABF001 6248	arnB-T	Early	Enzymes involved in LPS modification, resistance to cationic antimicrobial peptides	(Costechareyre et al., 2013)

ABF001 9265	sapABCDF	Intermediate	Antimicrobial peptide import system, resistance to antimicrobial peptides	(Lopez-Solanilla et al., 1998)
ABF001 5241	ohr	Intermediate	Thiol peroxidase, resistance to lipid hydroperoxide	(Reverchon et al., 2010)
ABF001 6112	sodA	Intermediate	Superoxide dismutase, resistance to oxidative stress	(Santos et al., 2001)
ABF001 7092	sodC	Intermediate	Superoxide dismutase, resistance to oxidative stress	
ABF001 5800	katG	Intermediate	Catalase-peroxidase, resistance to oxidative stress	(Miguel et al., 2000)
ABF001 9342	katE	Intermediate	Catalase-peroxidase, resistance to oxidative stress	
ABF001 8570	ahpCF	Intermediate	Alkyl hydroperoxide reductase, resistance to oxidative stress	
ABF001 6084	indABC	Intermediate	Indigoidine biosynthesis, resistance to oxidative stress	(Reverchon et al., 2002)
ABF001 8996	ibpS	Intermediate	Secreted metal binding protein, resistance to oxidative stress	(Liu et al., 2019)
ABF001 4816	msrA	Intermediate	methionine sulfoxide reductase, resistance to oxidative stress	(El Hassouni et al., 1999)
ABF001 8866	acsF-A	Intermediate	Achromobactin synthesis, iron scavenging	(Franza and Expert, 2013)
ABF001 8856	cbrABCD	Intermediate	Achromobactin transport system, iron scavenging	
ABF001 9156	fct-cbsCEBAP	Intermediate	Chrysobactin synthesis and receptor, iron scavenging	
ABF001 4651	ftnA	Intermediate	Maxi-ferritin, iron storage, iron homeostasis	(Boughammoura et al., 2008)
ABF004 6787	bfr	Intermediate	Maxi-ferritin, iron storage, iron homeostasis	
ABF001 8914	dps	Intermediate	Mini-ferritin, iron storage, iron homeostasis	(Boughammoura et al., 2012)

ABF0020090	sufABCDSE	Intermediate	Iron-sulfur cluster assembly system	(Expert et al., 2008)
ABF0018171	iscRSUA	Intermediate	Iron-sulfur cluster assembly system	(Rincon-Enriquez et al., 2008)
ABF0019593	hrpA-hrpE	Intermediate	Type III secretion system	(Yang et al., 2002)
ABF0020784	hrpN	Intermediate	Harpin secreted by the type III system	
ABF0019009	hrpW	Intermediate	Harpin secreted by the type III system	
ABF0019012	dspEF	Intermediate	Type III effectors	
ABF0018072	budAB	Intermediate	Butanediol fermentative pathway	(Effantin et al., 2011)
ABF0016562	iaaMH	Intermediate Late	Auxin production	(Yang et al., 2007)
ABF0019643	pelA-pelE-pelD	Late	Pectate lyases involved in pectin degradation	(Robert-Baudouy et al., 2000)
ABF0020840	pelB-pelC	Late	Pectate lyases involved in pectin degradation	
ABF0018773	pelL	Late	Pectate lyases involved in pectin degradation	
ABF0018192	ganEFGABC	Late	Degradation of galactan, a major component of the ramified regions of pectin	(Delangle et al., 2007)
ABF0018772	celZ	Late	Cellulase involved in cellulose degradation	(Py et al., 1991)
ABF0018229	outC-outO	Late	Type II secretion system required for pectate lyases and cellulase secretion	(Condemine et al., 1992)
ABF0017123	prtG-inh-prtDEFBCA	Late	Proteases G, A, B, C and their type I secretion system PrtDEF	(Letoffe et al., 1990)

Pectin metabolism a major virulence mechanism in dickeya

Pectin metabolism helps the bacteria to derive energy and establish itself in the apoplast. We study the pectin degrading enzyme Pels as virulence factors since they are very essential for obtaining energy. The apoplast is scarce in easy forms of energy like glucose and Dickeya can use pectin present in plant cell-wall as a carbon source. Degradation of the pectin of the plant tissue results in the symptoms of soft rot.

Bacterial Pectin metabolism

Pectin is a complex polygalacturonic acid polymer found in plant cell walls and is considered as the plant cell wall ciment. In the plant cell wall, pectin is mostly found in the primary cell wall intercalated between the cellulosic fibers which can be observed in Figure 3.2. Given that cellulose is organized in macro-fibrils and is very complex to break down, pectin is an easier molecule to be targeted and broken down by pathogens [23].Degradation of pectin destroys the intactness of the plant cell wall and also serves as an energy rich carbohydrate source [24], [25]. Pectin is sequentially de-esterified and depolymerized into oligosaccharides using pectinases including pectate lyases enzymes as depicted in Figure 3.3. Oligosaccharides are uptaken into the cell and converted to 2 keto-3 deoxygluconate (KDG) for energy synthesis..

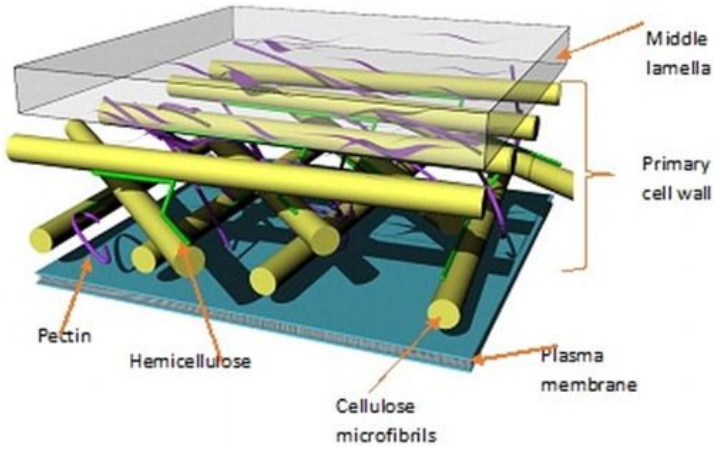


Figure 1.3.2: The cross sectional depiction of the plant cell wall and its components [24]

The metabolite KDG is found common to the energy derivation from pectin as an energy source [25]. In herbivorous animals, anaerobic bacteria sustaining the fermentation of pectin are found in the gastrointestinal tract [26]. Various microbes are known to derive their energy from the degraded pectin products in plants and herbivorous animals. Bacteroides thetaiotaomicron breaks down pectin in the herbivorous gut [27]. Ramen pectinolytic bacteria *Lachnospira multiparus, Prevotella ruminicola* and *Butyrivibrio fibrisolvens* assists in degradation and fermentation of pectin products [28].

Pectin degradation by *Dickeya dadantii*

Dickeya is a well studied species of phytopathogen with pectin degradation capacity. Different classes of pectinases are produced by the *Dickeya* species assisting pectin degradation and causing soft rot disease in plants. We study the pectin degradation in our model organism, pectin degrading phytopathogen *Dickeya dadantii*. Pectin is mainly a polymer of galacturonic acid with ramifications. The breakdown of this complex polymer is a multistep process achieved by different enzymes with specific functions. A first set of enzymes comprising the PaeX, PaeY esterases (pectin acetylesterases) and PemA, PemB (pectin methylesterases) aims to remove these ramifications to allow effective action of other enzymes [29]. The linear polymer accommodating the enzymes efficiently is then degraded by several pectate lyases as well as minor enzymes such as pectin lyases and polygalacturonases. There are about 9 extracellular pectate lyases in *Dickeya dadantii*. These extracellular enzymes are secreted by two type two secretion systems, Stt (which secretes only pectin lyase PnlH) and Out (which secretes all the other enzymes [30]. Pectate lyases are more or less recent acquisitions by horizontal transfer and duplication events. For example, *pelB* and *pelC* are present in *Pectobacterium* and *Dickeya* while *pelA, pelE* and *pelD* exist only in *Dickeya* [31]. The *pelBC* are located adjacent to each other and *pelAED* are located together in the *Dickeya* chromosome. The diversity of pectate lyases is believed to be responsible for the board host spectrum for *D. dadantii*, allowing it to cope with different pectin structures and breaking them down easily [32].

The degradation then continues intracellularly as shown in Figure 3.3.

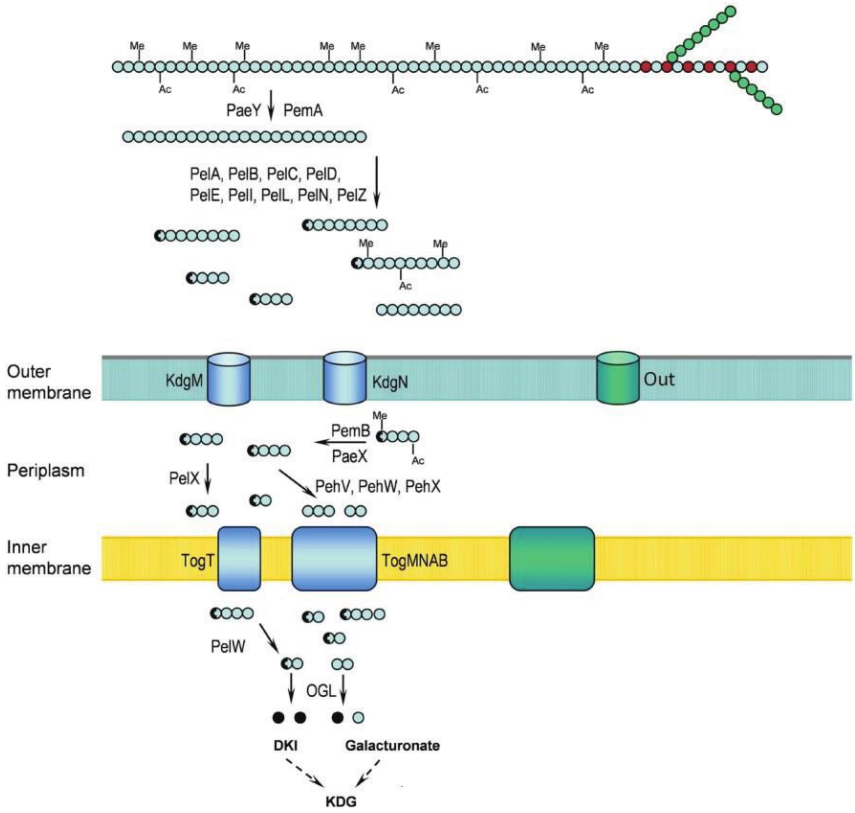


Figure 1.3.3: The enzymes in the pectin metabolism pathway [33]

4. Regulation of pectinolysis in *Dickeya dadantii*

In this study, the *pel* genes in *Dickeya dadantii* are our main virulence genes to be studied for their regulations and then modeled for their expressions by the regulators. Since pectin degradation sets the disease in plants when infected by *Dickeya dadantii* they are one of the main virulence genes for the infection. The production of extracellular Pel enzymes is finely regulated to ensure the success of the infection. The infection strategy for *Dickeya dadantii* is indeed to start with an asymptomatic phase, multiplying in intercellular spaces, until upper hand in infection conditions are met with respect to the plant defense mechanism [34]. Then *Dickeya dadantii* suddenly induces massive production of pectinases to degrade plant cells before an effective response from the host plant has been established [34]. Thus, this timing when the plant immune system is compromised is very important for the maceration by *pels* and successfully establishing the disease. Although pectinases are the most numerous and most important in virulence of *Dickeya dadantii*, 4 proteases and one cellulase enzymes are also identified in *Dickeya* species acting in the cell wall breaking down process [10]. This regulation is carried out by many transcription factors and nucleoid associated proteins (NAPs) in a timely manner to facilitate pathogenesis. There are numerous regulators for the virulence factors in *Dickeya dadantii.* They vary in their functions and mode of actions. The major factors of *pel* regulation are KdgR, PecS, PecT, CRP, H-NS, FIS, IHF, Fur, GacA, SlyA,MfbR, the Vfm quorum sensing system [30]. HFQ, ProQ and IHF are also essential for the virulence of *Dickeya dadantii*. The different levels of virulence gene regulation in Dickeya dadantii is discussed below.

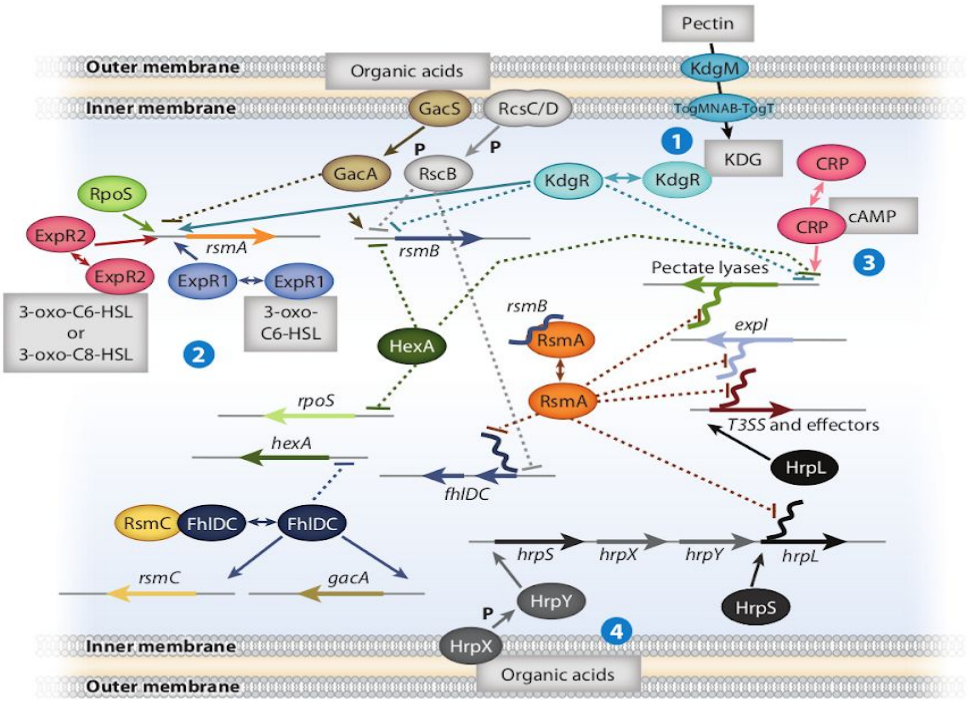


Figure 1.4.1 The various gene regulation instances in *Dickeya dadantii* [30]

Dickeya virulence gene regulation is achieved mainly at transcriptional and post transcriptional levels. The various common regulations are showed in the Figure 3.4 which are explained in the following paragraphs.
The different instances of transcriptional level regulations are

- NAP's
- Global transcriptional regulator: CRP
- Specific transcription factors: KdgR, PecS, PecT
- Two component system: GacA/ GacS
- Sigma factors
- Supercoiling

NAP's

Nucleoid architectures determined by NAPs have a significant impact on gene transcription and DNA replication [35], [36]. Recent findings have highlighted one of the NAPs, Factor for inversion stimulation (Fis), as a key regulator of virulence in several pathogenic bacteria such as pathogenic *Escherichia coli* strains, *Shigella flexneri, Salmonella, Vibrio cholerae* and *Dickeya dadantii* [37]–[39]. Fis acts as global regulator of growth by enhancing synthesis and transfer of ribosomal RNAs [40]. It is also known to vary the topological state of DNA by repressing DNA gyrase, stimulating topoisomerase I and binding on to the DNA [41]. Thus, Fis is known to have a role in maintaining local supercoiling state of DNA [42]. Moreover, the synthesis of *fis* is very growth dependent, increasing its weight as a regulator acting dependent to the growth stage of bacteria [43]. Hence, Fis is responsible for a sequence dependent or growth phase-dependent regulation of its target genes. Fis appears to be a versatile regulator that is very efficient in controlling basic cellular processes and specific genetic programs such as virulence programs in *Dickeya dadantii*. H-NS is a global modulator of gene expression in *Dickeya dadantii,* responsive to environmental conditions including temperature, pH osmolarity and oxidative stresses [44]. H-NS is often referred to as a 'universal repressor' or a 'modulator of environmentally regulated gene expression' [36]. *pel* gene expression is repressed by H-NS upon binding to the regulatory regions [45]. H-NS modulates the expression of *pel* genes in response to acidic and oxidative stress [10].

Global transcriptional regulator: CRP

CRP (Cyclic Amp Receptor Protein) is a key transcription factor in bacteria and a major global regulator. CRP is known to be one of the very important global regulators in bacteria. In *Escherichia coli* transcriptomic analysis many genes linked to metabolism are known to be regulated by CRP. CRP mainly activates sugar catabolism [46]. CRP is also known for the heavy role in virulence gene regulation [47]. In *Dickeya dadantii* it is known to be one of the main activators of virulence genes. CRP is extensively studied in *Escherichia coli* in terms of Carbon catabolite repression (CCR). It is studied for its reagulational activity in *Dickeya* species and identified as a primary activator of pectinolysis genes [48]. CRP binds to one or two cAMP molecules to form a functionally active complex [49]. Upon binding to cAMP the conformation of CRP changes and it acquires the ability to bind to the DNA regulates the gene transcription [50]. The regulation of the activity of CRP is controlled at the post-translational level via the control of the level of cAMP molecules available to activate CRP. The level of cAMP is very dependent on the glycolytic flux and growth of the bacteria making it a very good regulation mechanism of CRP. The cAMP molecule also relays environmental signals to regulatory outcomes [51]. It gets activated up on receiving environmental signals which then activates the genes corresponding to the signals. The regulation of cAMP

concentration is carried out at its level of synthesis and degradation [52]. cAMP is degraded by single phosphodiesterase CpdA which converts cAMP to AMP which can later be converted to ATP [53]. This elucidates the role of energy linked to the CRP-cAMP regulation and metabolism.

CRP was initially studied for its activating function which consists of two possible mechanisms. The first mechanism requires contact with the C-terminal domain of the α subunit of RNA polymerase. Presence of CRP binding site on the DNA double helix on the same side as RNA polymerase binding site. RNAP binds between the binding sites of CRP and the transcription initiation site. The second one needs a contact between both N and C terminal domains of the RNAP α unit and is only possible with a distance of about 41.5 bp between the CRP site and the transcription initiation site [54]. CRP is also capable of bending DNA with a 90° to 130° angle, which can be used in activation mechanisms. On the *pel* genes studied in *Dickeya dadantii*, *pelD* and *pelE*, both mechanisms of CRP activation are present. Thus, in *Dickeya dadantii*, CRP is also a central regulator of virulence. The increase in cAMP levels indicates the depletion of sugar resources and thus the need to initiate degradation of pectin. CRP therefore regulates the majority of pectin degradation pathway genes encoding both extracellular enzymes (*pel, pemA*) and intracellular enzymes (*ogl, kdu,kdg*) [48]. Though CRP mostly acts as an activator, it is a self repressor of *crp* gene and known to suppress a few enzymes [46]. Although catabolic repression is an almost universal mechanism in bacteria, it is not always CRP that is responsible for this. The regulator CcpA is also known to play the role of CRP in a few bacteria especially Gram-positive bacteria *Bacillus subtilis*. In *Pseudomonas,* a ortholog of CRP is called Vfr (Virulence factor regulator) which acts regardless of cAMP [55]. In conclusion, the CRP protein may have been a regulator of virulence before to also become a regulator of metabolism [56]. Like in many bacteria, CRP displays virulence-metabolism regulation in *Dickeya dadantii*.

Two component system: GacA/ GacS

Two component system is the most known gene regulation mechanism helping the bacteria to adapt to a frequently changing environment. It consists of a sensor with kinase activity, and a responding regulator. The kinases are generally integral membrane proteins, which senses specific environment signals and modulates the activity of cognate regulators by reversible phosphorylation. The regulator upon phosphorylation has its DNA affinity modulated and can act as a trans-activator protein for the target genes. The name two compartment symbolizes the extracellular detection of environmental stimuli and kinase mediated intracellular signal transduction of gene regulation. The advantage of this system is that it is possible to simultaneously regulate unlinked genes and their expression to target the stimulus. Two component signal transduction systems regulate the stress responding virulence genes in a global manner. The GacS/GacA two component system in *Dickeya dadantii* is an example of

two component systems responding to organic acids [57]. GacA is the response regulator of the putative sensor kinase GacS. GacA is activated by autophosphorylated GacS. Activated GacAworks as a transcriptional activator and is responsible for the activation of pectate lyases production etc. [58], [59]. GacA is also known to regulate the biofilm-pellicle formation. Studies also show that the GacA/GacS system is involved in regulation of type III secretion systems (T3SS). T3SS is controlled by the Gac-Rsm regulatory network and is crucial for establishment of virulence mechanism of bacterium within the plant host [75].

Transcription factors

- **KdgR**

KdgR is a transcriptional factor of the IclR family acting as a major regulator of virulence in *D. dadantii* [60]. It's a repressor which directly controls about 50 genes accounting for 1% of the total genes in this bacterium. Target gene products of KdgR are almost exclusively linked to the breakdown of pectin including *pels* [61]. In the absence of pectin, KdgR binds to DNA at the level of the promoter sequences and prevents transcription in general by occlusion of the RNA polymerase binding site [62]. The intermediate product of pectin degradation 2-keto 3-deoxygluconate (KDG) binds on KdgR and prevents its interaction with DNA. KdgR is therefore the central player in response to the presence of pectin. KdgR regulator is also present in other bacteria like *Salmonella, Escherichia, Xanthomonas,* etc. Contrary to *Dickeya* and *Pectobacterium* in *Xanthomonas oryzae* KdgR controls the type III hrp secretory system helping the bacteria in infection process [63]. KdgR is also an important regulator in Zoo Pathogenic *Enterobacteria* [61]. KdgR is also identified to be one of the regulators involved in the virulence function in *S. enterica* surviving in soft rot induced by *Dickeya* and *Pectobacterium* [64]–[66]. This makes the KdgR regulators to be an important factor regulating the phytopathogenic virulence.

- **PecS**

PecS is a transcription factor of the MarR family which prevents the early expression of numerous virulence factors in *Dickeya dadantii* [67]. In the process of pathogenesis, it is vital to express virulence factors at right to fight the host defense. Thus, the PecS factor plays an important role in the timing of expression of virulence factors. PecS targets are varied. PecS is known to suppress the production and export extracellular enzymes [68]. In *Dickeya dadantii* about 140 genes responding to PecS have been identified Including a variety of virulence genes like *pelE* and *pelD*. PecS is also thought to be a sensor of stress like oxidative stress and in response regulating corresponding virulence genes [69].

- **PecT**

PecT is a transcription factor of the LysR family involved in the virulent thermoregulation of *Dickeya dadantii* [70]. Classical genetics approaches have identified a relatively large number of important genes regulated by PecT, including the genes of extracellular enzymes like *pelE* and *pelD*. It is also known to have an effect on synthesis of exopolysaccharides [11]. The action of pecT on *pel* genes is dependent on the level of DNA supercoiling. PecT has a higher affinity towards relaxed DNA. However, at high temperature, DNA becomes loose by changes in supercoiling and PecT's affinity for *pel* regulatory regions increases, which helps suppress the expression of *pel* at high temperature [71]. In addition, the expression of pecT is repressed by H-NS, allowing it to serve as an intermediary in the regulation of *pel* genes by the global regulator H-NS [72].

Sigma factor mediated regulation

Sigma factor mediated regulation is an important mechanism of transcriptional regulation. The regulation of T3SS is an example of alternative sigma factor mediated regulation of virulence factors. HrpL is a master regulator and alternative sigma factor controlling the expression of structural and effector genes of T3SS [73]. RpoN (σ54) in conjugation with enhancer protein HrpS activates the hrpL expression [73], [74] . The transcription of hrpS is upregulated by the two component signal transduction system (TCSS), HrpX/HrpY [73]. The hrpL transcription is down regulated by PNPase by reducing the stability of hrpL mRNA [75].

Supercoiling

We have discussed the transcriptional regulation based on the interaction between transcriptional factors, metabolites, transcriptional elements and the DNA in the above sections. Another peculiar kind of regulation is observed in recent years based on the structure of the DNA. The transcriptional regulation gene is influenced by the physical state of the chromosome, i.e. supercoiling level on the chromosome. The supercoiling level of the bacterial chromosome seems to be modified by severe environmental changes in pH, temperature, osmolarity and oxidative stress, as well as nutritional deprivation. Under these stressful conditions, the nucleoid in bacteria is organized differently by certain NAPs (Nucleoid Associated Proteins) which also protect the bacterial genome and regulate transcription in order to survive through these growth-limiting and potentially lethal conditions [76]. NAPs organize and mediate the Packaging of genomic DNA, thus the chromosomal architecture. They act by binding directly on DNA and are abundantly present in the bacterial nucleoid [77]. The DNA supercoiling levels determined by the chromosomal organization of NAPs is also known to have a role in regulating specific genes. Many examples of virulence genes regulated by supercoiling have come to light in recent times [78]. In *Dickeya dadantii* supercoiling

is known to play a role in regulating virulence genes like *pelE, etc*. In the motive to understand the regulation of virulence genes, understanding the role of supercoiling is also very important. Since virulence genes are spread out in a chromosome and have different regulators acting on them, searching for a central regulator which could globally modify genes is very justified. It is because of the time constraints on bacteria to reciprocate to a stressful situation in a host and react by synthesizing essential adaptive factors to overcome the stress. Supercoiling being a fundamental characteristic of any bacteria it can be a suitable factor fulfilling the role of a regulator. There has also been proof that various stress conditions have changed the supercoiling levels of bacteria. A detailed study on the role of supercoiling has been discussed in chapter three of this thesis.

Post transcriptional level regulation

Virulence genes in *Dickeya dadantii* are often observed to be regulated at post transcriptional level. It is one of the main gene regulatory mechanisms. Rsm system is an example of post transcriptional gene regulatory mechanisms consisting of RsmA, rsmB and RsmC. It regulated the expression of genes for the extracellular enzymes as well as the type 3 secretory system. RsmA is a small RNA-binding protein, up on binding lowers the half life of mRNA species where as rsmB is an untranslated regulatory RNA that binds to RsmA and neutralizes its negative regulatory effect by forming an inactive ribonucleoprotein complex [79]. RsmC controls the production of RsmA and rsmB RNA by positively regulating rsmA and negatively controlling rsmB [80]. The expression of hrpL the sigma factor regulator discussed earlier is also controlled at the post-translational level by the RsmA/rsmB system [57].

As we have seen so far, *Dickeya dadantii* is equipped with a large number of efficient regulators modulating and managing the expression of virulence factors. *Dickeya dadantii* has also acquired a few virulence genes by horizontal transfer to sustain itself in the host environment. The regulators that control these virulence factors are also of various origins. Some are global virulence regulators who have mostly continued to regulate the same genes (KdgR, PecT). Others are ancestral regulators including the regulon which has undergone major modifications to adapt to the constraints of the plant host (FIS, CRP). Finally, *Dickeya dadantii* has also integrated new regulators acquired by horizontal transfer (PecS).

We focus on the regulation of the two *pel* genes pelD and pelE known to be the result of gene duplication [81]. The corresponding enzymes PelD and PelE have indeed highly similar sequences (78% amino acid identity, 89% amino acid similarity) and are biochemically nearly identical [82]. It was found that the expression of *pelE* and *pelD* is regulated by the same set of TFs, the main ones being the activator CRP and repressor

KdgR [62], [83]. *pelD* is most strongly inducible by pectin but has a very low pectin-independent expression, while *pelE*, although inducible by pectin, has a high expression level even in the absence of pectin, *pelE* acts as an initiator of pectin degradation generating the initial inducers of *pel* gene expression (DKI, DKII, KDG) required to prevent the binding of the repressor KdgR to its operators [84]. In this study, the necessity of the presence of two same functional genes in the model organism *Dickeya dadantii* and the varying expression pattern in spite of having the same transcription factors is investigated using the transcriptional model developed with respect to repressor KdgR and activator CRP. This model can also be used to study the instance of carbon catabolite repression where CRP acts as a virulence gene regulator. The dynamics of switching in carbon source from glucose to pectin is also illustrated in the model. This model is then used as a tool to extrapolate to different other pathogens to study the transcriptional regulation of their virulence genes.

5. Carbon catabolite repression

In gram-negative bacteria a well known instance in the glucose+lactose medium in *Escherichia coli* preferring glucose first is a classic example of CCR, called *lac* operon model [85], [86]. The presence of glucose inhibits the *lac* operon accessorizing genes required for lactose uptake by the bacteria. Thus, this phenomenon where the presence of one kind of carbohydrate represses expression of certain genes or operons, making it difficult to uptake the latter carbon source constitutes the process of CCR. It generally refers to the repression of certain genes by the presence of particular carbon sources in the medium [52]. CCR is studied extensively in *Escherichia coli* but the mechanism is not completely understood. In *Escherichia coli*, CCR is explained by the Inducer exclusion model [87] and the cAMP model [88]. Figure 3.3 depicts the *lac* operon in *Escherichia coli*.

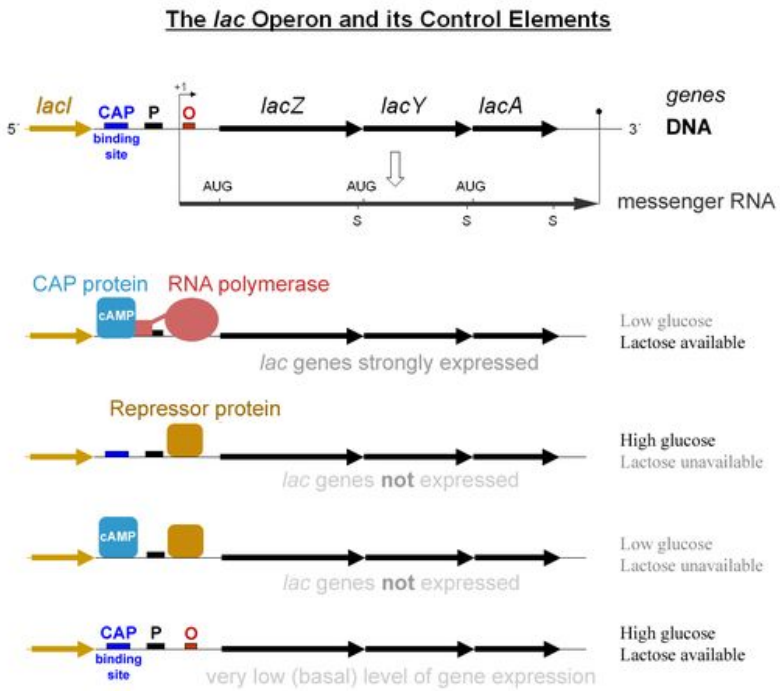


Figure 1.5.1: The *lac* operon in *Escherichia coli* and the different instances of the regulation of *lac* operon [89]

The cAMP model

The cAMP model accounts for the absence of cAMP inside the cell to activate CRP, the central regulator of the *lac* operon. In the abundance of glucose, the phosphorylated form of enzyme EIIAglc binds and thereby inhibits the sugar transport system preventing the entry of inducing sugar. Depletion of glucose leads to the emergence of

phosphorylated form of the enzyme EIIAglc activates adenylate cyclase leading to the formation of metabolite cAMP [90]. Upon binding, cAMP activates the regulator CRP providing the activating boost for the LAC operon expression. By this theory, it is expected to have an activation of the gene when exogenous cAMP is added to the medium at any point during the presence of glucose. But the experiments have shown otherwise [91]. Hence the quest for further explanation continued.

Inducer exclusion model

The Inducer exclusion model explains the exclusion of inducer molecule, lactose, from entering the cell due to the lack of permease enzyme LacY [91]. In the abundance of glucose, the phosphorylated form of EIIAglc enzyme binds to carbohydrate permease enzymes and inhibits it [91]. Thus, the presence of EIIAglc enzyme inhibits the sugar transport system preventing the entry of inducing sugar. The presence of inducible sugar inside the cell leads to gene expression. In the case of *lac* operon, the preferable source of carbon is glucose as expected in most CCR instances and an inducible secondary source of carbon is Lactose. The inhibited lactose permease is incapable of importing lactose from the extracellular region. Thus, only glucose is consumed and when glucose is depleted the phosphorylated form of EIIAglc enzyme activates adenylate cyclase leading to the formation of metabolite cAMP which activates the gene responsible for the consumption of secondary sources of carbon lactose.

Induction of the *lac* genes is caused by the absence of glucose. Glucose is also responsible to have an effect on the import of exogenous cAMP for the CRP activator to be active and induce the gene. Thus, glycolytic flux, conversion of glucose to glucose-6-phosphate is also known to have an effect on the CCR mechanism [92]. Interestingly, the bacterial phosphoenolpyruvate (PEP): carbohydrate phosphotransferase system (PTS), which catalyzes the uptake and concomitant phosphorylation of numerous carbohydrates, plays a major role in bacterial CCR [90]. The repression of *lacZ* expression by glucose is not due to the reduction of the cAMP-CRP level but due to an inducer exclusion mechanism which is mediated by the phosphoenolpyruvate-dependent sugar phosphotransferase system [93]. Thus, it is significant that the metabolic state of the cell and global physiology is somehow linked to the optimized carbon utilizing mechanism. This concludes that CCR is controlled by the presence of the primary source of carbon, activation of CRP and metabolic flux.

Carbohydrate metabolism is very crucial for bacterial survival. It determines the energy availability for the ongoing metabolic activities of the bacteria. The switch in carbon sources is interconnected to a complex regulatory network which eventually results in gene regulation. Thus, this crucial metabolic regulation can be exploited to regulate genes involved in optimizing the bacterial growth in difficult environments. For example,

Cyclic AMP receptor protein (CRP) regulates the expression of cold shock responsive genes *cspA, cspB, cspG and cspI*, members of the cspA family, in *Escherichia coli* [94]. Some specific bacteria have evolved to utilize the regulatory action of CRP in the absence of cAMP to make it an evolutionary choice and effective regulator. Evidence suggests the utilization of global regulator CRP as a general regulator for acquired horizontal genes with no prior regulatory mechanism in the acquired organism [95]. This makes CRP an important regulator for the acquired virulence genes due to the evolutionary traits observed in bacteria and thus for bacteria survival in the new environmental niche. In the instance of CCR by CRP regulations cAMP acts as a cofactor activating CRP up on binding, making cAMP an important metabolite in CRP regulation.

6. Transcriptional regulation models

The gene regulation models effective tools to understand the gene expression and control. Computational tools are used to best represent the system and utilize them in understanding gene expression and regulation systems. A successful model should incorporate all the input parameters and fit to the experimental known system data and give new biological insight to the system. The prevalence of the model depends on the new insights obtained based on the existing data and not merely recapitulate what is already known.

Various approaches are used to model gene expressions based on the system being subjected. Broadly, they can be categorized as **discrete or continuous** expression regulatory models. Discrete models represent time, space and state as discrete sets of values, for example, Boolean models. Discrete models have a predefined number of states for the regulation. Continuous models represent the behavior of a system over continuous values of time, for example, differential equation models utilize continuous values and provide smoother representation of dynamical changes occurring in the gene regulatory system [96].

Based on the data available to describe the gene expression system, **statistical or analytical techniques** are employed to develop the gene regulatory models. Statistical models are more suited for data sets comprising the transcriptome data, representing the expression levels of several thousand genes. Probabilistic models such as Neural, Boolean and Bayesian network models are used to represent graph based regulatory links between the gene regulation interactome [97], [98]. Possible regulatory relations in different conditions are correlated by statistical analysis over the transcriptome and highlighted over the set of similar function genes. Shared motifs involved in transcriptional control and co expressing genes are identified as a part of this analysis. Statistical analysis ultimately reveals the regulatory networks underlying the expression mechanism of the given data. It covers a large fraction of genes documented in the transcriptome data and provides a bigger picture of gene regulation in the transcriptome. Given the recent advances in gene annotations, statistical models are fairly developed and studied [99]–[101]. The limitations of transcriptome statistical models is that they are not capable of explaining the complex relation between transcription factors, RNA polymerase and other regulator proteins involved in gene regulation. Statistical models are well appreciated in DNA sequence based modeling of gene regulation.

Analytical approach of gene regulation represented by a variety of mathematical models generally focuses on a smaller number of genes. These kinds of models include data

regarding binding of RNA polymerase and transcription factors to the DNA site on the gene along with inhibitory, enhancing and cooperative interactions between transcription factors and other regulatory factors like mRNA, protein degradation and mRNA dynamics like translation and degradation parameters. This approach needs extensive knowledge of the system components and their interactions and good understanding of the system structure to formulate well-educated hypotheses.

These models summarise the experimental data, represent relation between model components and infer new scientific testable hypotheses and enable the scientists to find new properties of the system that can add scientific values to the whole understanding and lead to the development of new experimental techniques to obtain further insight into the system and incorporating novel elements to the existing model [102]–[104]. The difference between different mathematical models describing regulation of genes have specific features characterizing them. **Deterministic and stochastic** nature of the model is one of the often observed characteristics. The change to an independent variable in the deterministic model has predictable yet reproducible impact on the dependent variables. Stochastic models are structured to capture the erratic behavior of biological systems influenced by intrinsic or extrinsic influencer noise signals. There is less uniformity in terms of model constituents in analytical models since they describe different individual gene expression systems [96].
There are three major classes of analytical approach models of gene regulation namely

a) Thermodynamic model
b) Differential equation model
c) Boolean model

The three approaches are represented in Figure 4.1 and explained below.

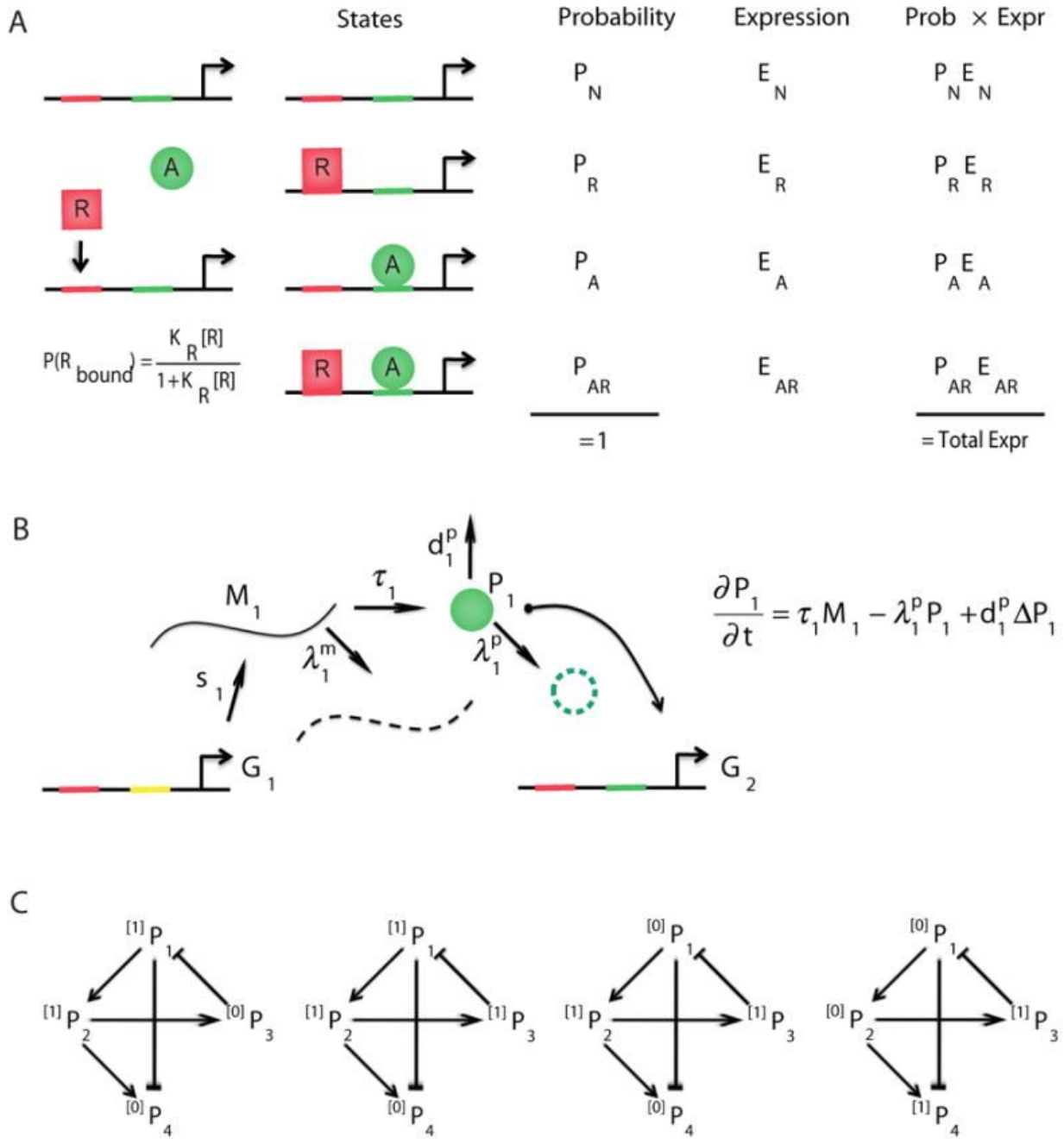


Figure1.6.1: Different mathematical models of gene regulation [96]

b) Thermodynamic or fractional occupancy model of gene expression

The image A represents a simplified enhancer region with two binding sites. One for repressor (R) and another for activators (A). Binding efficiency for the repressor site is represented by the mathematical equation $P_{R(bound)}$. All possible four states of this enhancer region are shown in the second column. The probability of these states occurring as a function of protein concentration and affinities to the binding site is represented in the third column. The fourth column represents the efficiency of gene expression with respect to four individual states. The last column represents the total expression occurring due to individual state.

c) Differential equation model of gene expression

This particular system of gene expression is modeled with respect to gene 1 (G1) involves expression of mRNA (M1) and translation of protein (P1), which regulates gene 2 (G2). There are parameters with respect to mRNA turnover and protein diffusion.
mRNA and protein synthesis, degradation and diffusion events are shown at left. The equations representing the model are in the right. Each equation represents the dynamics of each molecular constituent. s_1 is the rate of transcription, τ is the rate of translation and λ is the rate of degradation of the protein.

d) Boolean model of gene expression

The Boolean model of gene expression generally describes the gene regulatory network involved in regulation. P_1, P_2, P_3 and P_4 represent the four different proteins involved in the network. The direct arrows indicate activation and blunt arrows depict repression. There are two states of gene expression and components namely active and inactive leading to the regulatory action of the circuit.

Traditionally most of the gene regulatory models are ODE representing the gene expression as a function of protein dynamics and effect of factors involved. This needs a vast accumulation of knowledge of the system under consideration . When there is not enough information about the system to develop an ODE, an approach of randomization is used to develop regulatory models. Randomized parameter values are used to obtain a model representing a little known system at its best. For example, a tool named RACIPE employs a self-consistent scheme to randomize all kinetic parameters for a mathematical model [105]. Randomized parameter models employ parameter sensitivity and parameter space estimation analysis to obtain an idea about the range and relation between different parameters [106]–[108]. Randomized parameter models help us to interrogate the dynamical property of a gene circuit by randomization and extracting further insights of the system from the developed model.

Modeling gene expression with differential equations

The modeling of gene expression with differential equations is employed to develop the transcriptional regulation model of virulence genes in this study. The differential equation model is explained in Figure 4.2 below. It consists of three processes. The gene transcription to synthesize mRNA, translation of mRNA to synthesize proteins and the process of degradation of mRNA and proteins. There are rate constants at which mRNA and proteins are synthesized C and L, respectively. The degradation rate constants for mRNA and protein are V and U, respectively. Depending on the kind of gene being modeled there is always an involvement of feedback loop considered for autoregulation by the proteins. It could be negative if the protein synthesized is an auto repressor and positive if the synthesized protein causes activation of genes. In the differential equation model the process is differentiated with respect to time to obtain a rate factor for example rate of expression of genes. We will look into the detailed modeling approach in the coming chapters.

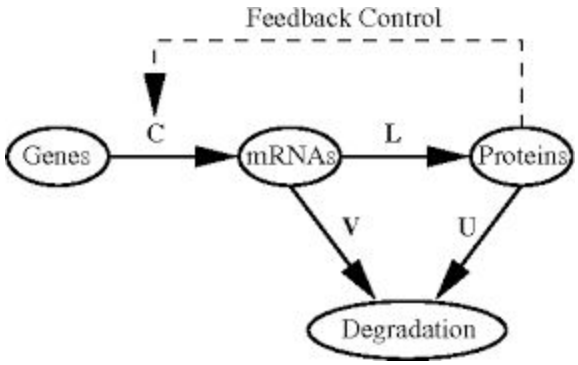


Figure 1.6.2: The differential equation gene expression model [109]

Transcriptional regulation models

Bacteria have devoted 10% of their genes to regulatory proteins to control the regulation of the genes [110]. Due to the extreme instability of mRNA most of the gene regulation occurs at a transcription level [111]. Such form of regulation has to be highly energy efficient and effective. Transcriptional regulation is most often carried out by proteins and metabolites binding to DNA and RNA. The success of regulation depends on how these regulators perceive the signal delivered to them by the recognition system. The three main mechanism of signal perception by transcriptional regulators

i) Small molecule messengers interaction with regulation proteins

The lac operon is an example of signal perception by the CRP regulators by the small molecule cAMP.

ii) Regulation by protein phosphorylation
There have been vast examples for the control exerted by the phosphotransferase enzymes on the DNA or RNA binding transcription factors [112], [113].

iii) Regulation by protein interactions with sensory proteins
The examples of control of sigma factors by their dedicated antagonists and anti-antagonists derived from the bacterial sensory systems

Understanding the interaction of transcription factors and knowing their functions is the basis to any transcriptional models. It can be predicted based on the sequence similarity approach. Various databases have documented these information like RegulonDB version 9.0: High-level integration of gene regulation, co-expression, motif clustering and beyond, coCyc: Fusing model organism databases with systems biology, PRODORIC: Prokaryotic database of gene regulation, RegTransBase: a database of regulatory sequences and interactions based on literature and resource for investigating transcriptional regulation in prokaryotes, CollecTF: a database of experimentally validated transcription factor-binding sites in bacteria.

The key to successful modelling is to obtain all the useful information from the data sources to transform them to understanding and working knowledge. Figure 4.3 shows an image of all the related sources of data with respect to transcription regulation. It hints on the usefulness of various data sources and their utilization in developing various aspects of a gene transcriptional model.

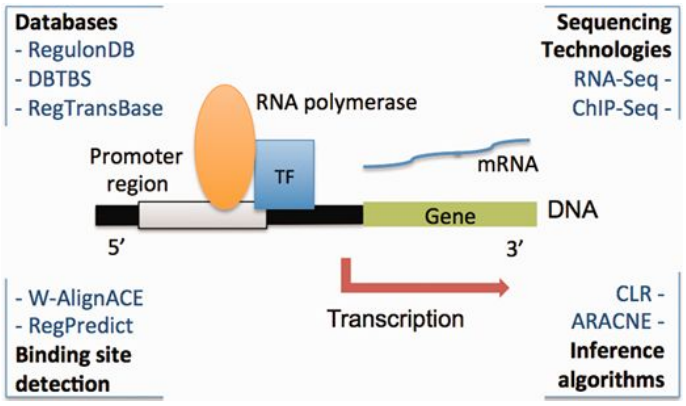


Figure 1.6.3: Technologies, tools and resources for transcriptional regulatory network modeling and reconstruction [114]

The information provided by these vast stored data helps in interpreting and predicting the relationship between various components involved in any gene transcription. Developing transcription regulatory networks to indicate all the connections and relations has been a main tool in developing transcription regulation models.

So far, the transcriptional regulatory models developed can be categorized as two types: Sequenced based models and Transcription network based models. Sequence based models depend on the very content of gene sequence and probable interaction with respect to the sequence and properties of DNA sequence. The binding of transcription factors to the sequence is based on the defined site on the sequence and its affinity [115]. The revolution in the DNA sequencing has led to significant development in the sequence based transcription models and their applications [116]–[118]. Predictions of transcription factor binding to specific sequences as well as resulting DNA conformations and their effect on the transcription is also studied and employed in therapeutic use [119], [120].

The transcription regulatory network based models are very specific to a gene expression system and incorporate all the components of the gene expression system compared to the sequence based system. Genetic regulatory networks consist of a set of genes, proteins, small molecules, and their mutual regulatory interactions. There are three approaches to developing a transcriptional gene regulatory model based on the genetic regulatory network [114].

Template based modeling
Template based modeling works with the sequenced DNA template and its properties. Based on the known minimal information a template is drawn which includes the number of genes and possible transcription factors involved. Any known interaction between the genes is also documented. Then search algorithm tools like blast and other data repositories are used to derive more information based on the sequence similarity. The available information is used to draw a more realistic regulatory network which is then validated with experiments.

Transcription factor binding site (TFBS) based modeling
Transcription factor binding site based modeling of transcriptional regulation depends on the prediction of sequence binding site for possible transcription factors and their involvement in the regulation process. Orthologous search in the known similar species predicts the possible sites for transcription factors. Then the whole genome is searched for known transcription factors and a regulatory link between the genes is drawn.

Denovo reverse engineering

Denovo reverse engineering is based on the gene expression data and identifying the relation between the gene expression patterns. The link between up and down regulation of genes and transcription factor is reverse engineering to obtain a gene regulatory network to understand the transcriptional regulation of expression.

We attempt to develop a simplified transcriptional regulation model partially based on the gene regulatory network and we introduce the growth dynamics into the system. The thermodynamic relation between the transcription factors and gene expression is modeled along with the known understanding of metabolic networks involved in the gene regulation system. In reality the metabolic networks are complex, we consider the known simplified network relevant for the model construction.

The functional properties of a transcriptomic circuit is defined by the interactions between its components. The level of cooperativity among components of the transcription unit is described in terms of binding efficiency and priority given to functional interactions. When it comes to transcriptional regulation, the aspect of cooperativity is defined by the affinity of transcription factors interacting with DNA. Most of the binding sites for the transcription factors are located in the cis (the region in the proximity of promoter of the gene) regulatory region. Thus, cis-regulatory region formulates the regulation with respect to the interaction and affinity of transcription factors to the DNA. The binding affinity varies with respect to the DNA architecture and molecule. This variation describes the mechanism of transcriptional regulation. The transcription factors can either activate or repress the gene. The fold change in the gene expression as a result of activation or repression by transcription factors can be thermodynamically modeled and represented in the form of an ordinary differential equation (ODE).

Thermodynamic models have proved efficient in quantitative analysis of bacterial transcription regulation [121]. We use a thermodynamic modeling approach to obtain the transcriptional regulation model with the integration of transcription factors in the form of ODE. Intending to keep the model as simple and efficient we have incorporated a simple repression and activation as a part of regulation.

Simple activator model

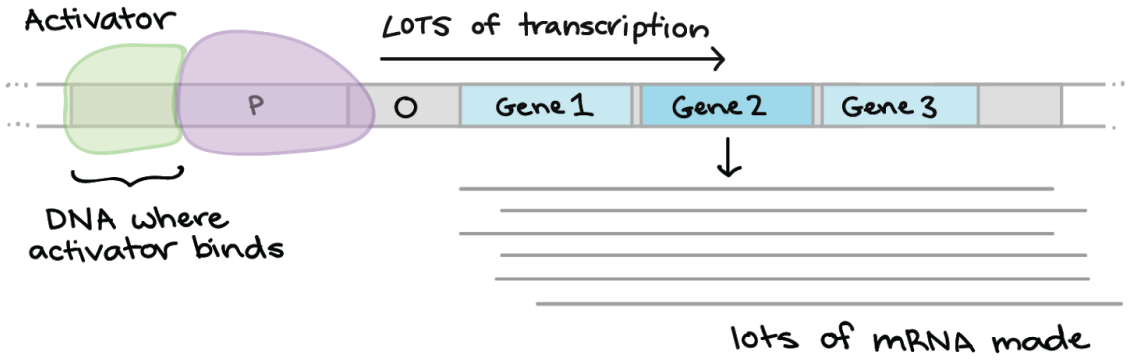


Figure 1.6.4: Depiction of a simple activator in transcription process [122]

Simple activator molecule represents the activation of gene expression when the activator molecule binds on the designated binding site in the promoter region as shown in Figure 4.4. The activator molecule can enhance the recruitment of RNA polymerase and enable boosting to the gene transcription. Simple activator model defines the activation obtained to be proportional to the activator molecules bound on to the site. The regulation factor is given by the term

$$\left[\frac{1+\frac{[A]*f}{KA}}{1+\frac{[A]}{KA}} \right]$$

Where [A] represents the concentration of the activator molecules and the *KA* represents the dissociation constant for the binding of activator molecules at the DNA binding site. It is defined as the concentration of molecules required to bind half of the available capacity of activator binding sites. The activator factor *f* represents fold change in the expression of genes when compared to the absence of activator to the complete activation.

Simple repressor model

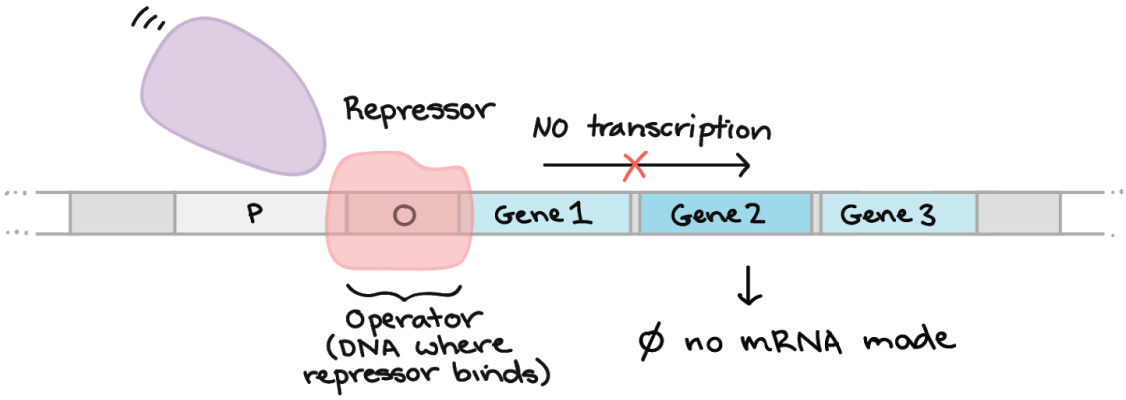


Figure 1.6.5: Depiction of a simple repressor in transcription process [122]

Simple repressor model consists of a single repressor binding in the vicinity of the promoter region and repressing the expression of genes as shown in Figure 4.5. The extent binding of the repressor molecules define the extent of repression on the gene. The binding factor is represented by the binding affinity of the repressor molecule to the binding site in the DNA molecule. The regulation factor is given by the term

$$\left[1 + \frac{[R]}{KR}\right]^{-1}$$

Where [R] is the concentration of repressor molecules. The dissociation constant *KR* is defined as the concentration of repressor molecules required to bind half of the capacity of repressor binding sites available. The repressor factor is not applicable to the simple repressor model since the ratio of unbound to the completely bound gene expression is mathematically infinite, but in actual biological situations there is some residual expression possible but cannot be quantified.

The model also studies the relation between the gene expression and the growth of the bacteria. The dependability of the transcription factors and the growth dynamics is incorporated into the model.

7. References

[1] S. Chakraborty and A. C. Newton, “Climate change, plant diseases and food security: an overview: Climate change and food security,” *Plant Pathol.*, vol. 60, no. 1, pp. 2–14, Feb. 2011, doi: 10.1111/j.1365-3059.2010.02411.x.

[2] S. Savary, L. Willocquet, S. J. Pethybridge, P. Esker, N. McRoberts, and A. Nelson, “The global burden of pathogens and pests on major food crops,” *Nat. Ecol. Evol.*, vol. 3, no. 3, pp. 430–439, 2019, doi: 10.1038/s41559-018-0793-y.

[3] S. Chakraborty, “Migrate or evolve: options for plant pathogens under climate change,” *Glob. Change Biol.*, vol. 19, no. 7, pp. 1985–2000, Jul. 2013, doi: 10.1111/gcb.12205.

[4] R. N. Strange and P. R. Scott, “Plant disease: a threat to global food security,” *Annu. Rev. Phytopathol.*, vol. 43, pp. 83–116, 2005, doi: 10.1146/annurev.phyto.43.113004.133839.

[5] J. Mansfield *et al.*, “Top 10 plant pathogenic bacteria in molecular plant pathology: Top 10 plant pathogenic bacteria,” *Mol. Plant Pathol.*, vol. 13, no. 6, pp. 614–629, Aug. 2012, doi: 10.1111/j.1364-3703.2012.00804.x.

[6] I. K. Toth *et al.*, “Dickeya species: an emerging problem for potato production in Europe,” *Plant Pathol.*, vol. 60, no. 3, pp. 385–399, 2011, doi: 10.1111/j.1365-3059.2011.02427.x.

[7] Q. M. Imran and B.-W. Yun, “Pathogen-induced Defense Strategies in Plants,” *J. Crop Sci. Biotechnol.*, vol. 23, no. 2, pp. 97–105, Mar. 2020, doi: 10.1007/s12892-019-0352-0.

[8] A. K. Sharma *et al.*, “Bacterial Virulence Factors: Secreted for Survival,” *Indian J. Microbiol.*, vol. 57, no. 1, pp. 1–10, Mar. 2017, doi: 10.1007/s12088-016-0625-1.

[9] T. M. Wassenaar and W. Gaastra, “Bacterial virulence: can we draw the line?,” *FEMS Microbiol. Lett.*, vol. 201, no. 1, pp. 1–7, Jul. 2001, doi: 10.1111/j.1574-6968.2001.tb10724.x.

[10] S. Reverchon and W. Nasser, “Dickeya ecology, environment sensing and regulation of virulence programme,” *Environ. Microbiol. Rep.*, vol. 5, no. 5, pp. 622–636, Oct. 2013, doi: 10.1111/1758-2229.12073.

[11] G. Condemine, A. Castillo, F. Passeri, and C. Enard, “The PecT repressor coregulates synthesis of exopolysaccharides and virulence factors in Erwinia chrysanthemi,” *Mol. Plant-Microbe Interact. MPMI*, vol. 12, no. 1, pp. 45–52, Jan. 1999, doi: 10.1094/MPMI.1999.12.1.45.

[12] C. E. Jahn, D. A. Selimi, J. D. Barak, and A. O. Charkowski, “The Dickeya dadantii biofilm matrix consists of cellulose nanofibres, and is an emergent property dependent upon the type III secretion system and the cellulose synthesis operon,” *Microbiology*, vol. 157, no. 10, pp. 2733–2744, Oct. 2011, doi: 10.1099/mic.0.051003-0.

[13] E. López-Solanilla, F. García-Olmedo, and P. Rodríguez-Palenzuela, “Inactivation of the sapA to sapF locus of Erwinia chrysanthemi reveals common features in plant and animal bacterial pathogenesis.,” *Plant Cell*, vol. 10, no. 6, pp. 917–924, Jun. 1998.

[14] D. Costechareyre, J.-F. Chich, J.-M. Strub, Y. Rahbé, and G. Condemine, “Transcriptome of Dickeya dadantii Infecting Acyrthosiphon pisum Reveals a Strong Defense against Antimicrobial Peptides,” *PLOS ONE*, vol. 8, no. 1, p. e54118, Jan. 2013, doi: 10.1371/journal.pone.0054118.

[15] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, “Cell Biology of Infection,” *Mol. Biol. Cell 4th Ed.*, 2002, Accessed: Aug. 07, 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK26833/.

[16] M. Antúnez-Lamas *et al.*, “Role of motility and chemotaxis in the pathogenesis of Dickeya dadantii 3937 (ex Erwinia chrysanthemi 3937),” *Microbiol. Read. Engl.*, vol. 155, no. Pt 2, pp. 434–442, Feb. 2009, doi: 10.1099/mic.0.022244-0.

[17] S. Koskiniemi *et al.*, “Rhs proteins from diverse bacteria mediate intercellular competition,” *Proc. Natl. Acad. Sci.*, vol. 110, no. 17, pp. 7032–7037, Apr. 2013, doi: 10.1073/pnas.1300627110.

[18] S. R, F. T, L. Ml, S. C, T. D, and E. D, “Essential role of superoxide dismutase on the pathogenicity of Erwinia chrysanthemi strain 3937.,” *Mol. Plant-Microbe Interact. MPMI*, vol. 14, no. 6, pp. 758–767, Jun. 2001, doi: 10.1094/mpmi.2001.14.6.758.
[19] E. Miguel *et al.*, “Evidence against a direct antimicrobial role of H2O2 in the infection of plants by Erwinia chrysanthemi,” *Mol. Plant-Microbe Interact. MPMI*, vol. 13, no. 4, pp. 421–429, Apr. 2000, doi: 10.1094/MPMI.2000.13.4.421.
[20] “Characterization of Indigoidine Biosynthetic Genes in Erwinia chrysanthemi and Role of This Blue Pigment in Pathogenicity | Journal of Bacteriology.” https://jb.asm.org/content/184/3/654 (accessed Aug. 07, 2020).
[21] M. E. Hassouni, J. P. Chambost, D. Expert, F. V. Gijsegem, and F. Barras, “The minimal gene set member msrA, encoding peptide methionine sulfoxide reductase, is a virulence determinant of the plant pathogen Erwinia chrysanthemi,” *Proc. Natl. Acad. Sci.*, vol. 96, no. 3, pp. 887–892, Feb. 1999, doi: 10.1073/pnas.96.3.887.
[22] B. C. Freeman, “The role of water stress in plant disease resistance and the impact of water stress on the global transcriptome and survival mechanisms of the phytopathogen *Pseudomonas syringae*,” Doctor of Philosophy, Iowa State University, Digital Repository, Ames, 2009.
[23] A. G. J. Voragen, G.-J. Coenen, R. P. Verhoef, and H. A. Schols, “Pectin, a versatile polysaccharide present in plant cell walls,” *Struct. Chem.*, vol. 20, no. 2, p. 263, Mar. 2009, doi: 10.1007/s11224-009-9442-z.
[24] Y. Xing, P. Jones, M. Bosch, I. Donnison, M. Spear, and G. Ormondroyd, “Exploring design principles of biological and living building envelopes: what can we learn from plant cell walls?,” *Intell. Build. Int.*, vol. 10, no. 2, pp. 78–102, Apr. 2018, doi: 10.1080/17508975.2017.1394808.
[25] J. K. Hobbs *et al.*, “KdgF, the missing link in the microbial metabolism of uronate sugars from pectin and alginate,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 22, pp. 6188–6193, May 2016, doi: 10.1073/pnas.1524214113.
[26] I. Hanning and S. Diaz-Sanchez, “The functionality of the gastrointestinal microbiome in non-human animals,” *Microbiome*, vol. 3, Nov. 2015, doi: 10.1186/s40168-015-0113-6.
[27] D. Ndeh *et al.*, “Complex pectin metabolism by gut bacteria reveals novel catalytic functions,” *Nature*, vol. 544, no. 7648, pp. 65–70, 06 2017, doi: 10.1038/nature21725.
[28] D. Dusková and M. Marounek, “Fermentation of pectin and glucose, and activity of pectin-degrading enzymes in the rumen bacterium Lachnospira multiparus,” *Lett. Appl. Microbiol.*, vol. 33, no. 2, pp. 159–163, Aug. 2001, doi: 10.1046/j.1472-765x.2001.00970.x.
[29] N. Hugouvieux-Cotte-Pattat, G. Condemine, W. Nasser, and S. Reverchon, “Regulation of pectinolysis in Erwinia chrysanthemi,” *Annu. Rev. Microbiol.*, vol. 50, pp. 213–257, 1996, doi: 10.1146/annurev.micro.50.1.213.
[30] A. Charkowski *et al.*, “The role of secretion systems and small molecules in soft-rot Enterobacteriaceae pathogenicity,” *Annu. Rev. Phytopathol.*, vol. 50, pp. 425–449, 2012, doi: 10.1146/annurev-phyto-081211-173013.
[31] N. Hugouvieux-Cotte-Pattat and J. Robert-Baudouy, “Analysis of the regulation of the pelBC genes in Erwinia chrysanthemi 3937,” *Mol. Microbiol.*, vol. 6, no. 16, pp. 2363–2376, 1992, doi: 10.1111/j.1365-2958.1992.tb01411.x.
[32] N. Hugouvieux-Cotte-Pattat, G. Condemine, and V. E. Shevchik, “Bacterial pectate lyases, structural and functional diversity,” *Environ. Microbiol. Rep.*, vol. 6, no. 5, pp. 427–440, 2014, doi: 10.1111/1758-2229.12166.
[33] A. Duprey, “Régulation de la transcription des gènes de virulence bactériens: dynamique des complexes nucléoprotéïques,” p. 157.
[34] I. K. Toth, K. S. Bell, M. C. Holeva, and P. R. J. Birch, “Soft rot erwiniae: from genes to genomes,” *Mol. Plant Pathol.*, vol. 4, no. 1, pp. 17–30, Jan. 2003, doi: 10.1046/j.1364-3703.2003.00149.x.

[35] D. F. Browning, D. C. Grainger, and S. J. Busby, “Effects of nucleoid-associated proteins on bacterial chromosome structure and gene expression,” *Curr. Opin. Microbiol.*, vol. 13, no. 6, pp. 773–780, Dec. 2010, doi: 10.1016/j.mib.2010.09.013.
[36] S. C. Dillon and C. J. Dorman, “Bacterial nucleoid-associated proteins, nucleoid structure and gene expression,” *Nat. Rev. Microbiol.*, vol. 8, no. 3, pp. 185–195, Mar. 2010, doi: 10.1038/nrmicro2261.
[37] M. Falconi, G. Prosseda, M. Giangrossi, E. Beghetto, and B. Colonna, “Involvement of FIS in the H-NS-mediated regulation of virF gene of Shigella and enteroinvasive Escherichia coli,” *Mol. Microbiol.*, vol. 42, no. 2, pp. 439–452, Oct. 2001, doi: 10.1046/j.1365-2958.2001.02646.x.
[38] A. Kelly, M. D. Goldberg, R. K. Carroll, V. Danino, J. C. D. Hinton, and C. J. Dorman, “A global role for Fis in the transcriptional control of metabolism and type III secretion in Salmonella enterica serovar Typhimurium,” *Microbiology,* vol. 150, no. 7, pp. 2037–2053, 2004, doi: 10.1099/mic.0.27209-0.
[39] D. H. Lenz and B. L. Bassler, “The small nucleoid protein Fis is involved in Vibrio cholerae quorum sensing,” *Mol. Microbiol.*, vol. 63, no. 3, pp. 859–871, Feb. 2007, doi: 10.1111/j.1365-2958.2006.05545.x.
[40] S. E. Finkel and R. C. Johnson, “The Fis protein: it’s not just for DNA inversion anymore,” *Mol. Microbiol.*, vol. 6, no. 22, pp. 3257–3265, Nov. 1992, doi: 10.1111/j.1365-2958.1992.tb02193.x.
[41] R. Schneider, A. Travers, T. Kutateladze, and G. Muskhelishvili, “A DNA architectural protein couples cellular physiology and DNA topology in Escherichia coli,” *Mol. Microbiol.*, vol. 34, no. 5, pp. 953–964, Dec. 1999, doi: 10.1046/j.1365-2958.1999.01656.x.
[42] A. Travers and G. Muskhelishvili, “DNA supercoiling - a global transcriptional regulator for enterobacterial growth?,” *Nat. Rev. Microbiol.*, vol. 3, no. 2, pp. 157–169, Feb. 2005, doi: 10.1038/nrmicro1088.
[43] M. D. Bradley, M. B. Beach, A. P. J. de Koning, T. S. Pratt, and R. Osuna, “Effects of Fis on Escherichia coli gene expression during different growth stages,” *Microbiol. Read. Engl.*, vol. 153, no. Pt 9, pp. 2922–2940, Sep. 2007, doi: 10.1099/mic.0.2007/008565-0.
[44] T. Atlung and H. Ingmer, “H-NS: a modulator of environmentally regulated gene expression,” *Mol. Microbiol.*, vol. 24, no. 1, pp. 7–17, 1997, doi: 10.1046/j.1365-2958.1997.3151679.x.
[45] W. Nasser and S. Reverchon, “H-NS-dependent activation of pectate lyases synthesis in the phytopathogenic bacterium Erwinia chrysanthemi is mediated by the PecT repressor,” *Mol. Microbiol.*, vol. 43, no. 3, pp. 733–748, 2002, doi: 10.1046/j.1365-2958.2002.02782.x.
[46] G. Gosset, Z. Zhang, S. Nayyar, W. A. Cuevas, and M. H. Saier, “Transcriptome Analysis of Crp-Dependent Catabolite Control of Gene Expression in Escherichia coli,” *J. Bacteriol.*, vol. 186, no. 11, pp. 3516–3524, Jun. 2004, doi: 10.1128/JB.186.11.3516-3524.2004.
[47] D. C. Grainger, D. Hurd, M. Harrison, J. Holdstock, and S. J. W. Busby, “Studies of the distribution of Escherichia coli cAMP-receptor protein and RNA polymerase along the E. coli chromosome,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 49, pp. 17693–17698, Dec. 2005, doi: 10.1073/pnas.0506687102.
[48] S. Reverchon, D. Expert, J. Robert-Baudouy, and W. Nasser, “The cyclic AMP receptor protein is the main activator of pectinolysis genes in Erwinia chrysanthemi,” *J. Bacteriol.*, vol. 179, no. 11, pp. 3500–3508, Jun. 1997, doi: 10.1128/jb.179.11.3500-3508.1997.
[49] Y. Tutar, “Syn,anti, and finally both conformations of cyclic AMP are involved in the CRP-dependent transcription initiation mechanism inE. coli lac operon,” *Cell Biochem. Funct.*, vol. 26, no. 4, pp. 399–405, Jun. 2008, doi: 10.1002/cbf.1462.
[50] J. L. Botsford and J. G. Harman, “Cyclic AMP in prokaryotes,” *Microbiol. Rev.*, vol. 56, no. 1, pp. 100–122, Mar. 1992.
[51] K. A. McDonough and A. Rodriguez, “The myriad roles of cyclic AMP in microbial

pathogens, from signal to sword," *Nat. Rev. Microbiol.*, vol. 10, no. 1, pp. 27–38, Nov. 2011, doi: 10.1038/nrmicro2688.
[52] B. Görke and J. Stülke, "Carbon catabolite repression in bacteria: many ways to make the most out of nutrients," *Nat. Rev. Microbiol.*, vol. 6, no. 8, pp. 613–624, Aug. 2008, doi: 10.1038/nrmicro1932.
[53] D. Richani *et al.*, "Participation of the adenosine salvage pathway and cyclic AMP modulation in oocyte energy metabolism," *Sci. Rep.*, vol. 9, no. 1, Art. no. 1, Dec. 2019, doi: 10.1038/s41598-019-54693-y.
[54] W. Niu, Y. Kim, G. Tau, T. Heyduk, and R. H. Ebright, "Transcription Activation at Class II CAP-Dependent Promoters: Two Interactions between CAP and RNA Polymerase," *Cell*, vol. 87, no. 6, pp. 1123–1134, Dec. 1996.
[55] E. L. Fuchs, E. D. Brutinel, E. R. Klem, A. R. Fehr, T. L. Yahr, and M. C. Wolfgang, "In vitro and in vivo characterization of the Pseudomonas aeruginosa cyclic AMP (cAMP) phosphodiesterase CpdA, required for cAMP homeostasis and virulence factor regulation," *J. Bacteriol.*, vol. 192, no. 11, pp. 2779–2790, Jun. 2010, doi: 10.1128/JB.00168-10.
[56] J. A. Lemos, M. M. Nascimento, V. K. Lin, J. Abranches, and R. A. Burne, "Global regulation by (p)ppGpp and CodY in Streptococcus mutans," *J. Bacteriol.*, vol. 190, no. 15, pp. 5291–5299, Aug. 2008, doi: 10.1128/JB.00288-08.
[57] S. Yang *et al.*, "Dynamic regulation of GacA in type III secretion, pectinase gene expression, pellicle formation, and pathogenicity of Dickeya dadantii (Erwinia chrysanthemi 3937)," *Mol. Plant-Microbe Interact. MPMI*, vol. 21, no. 1, pp. 133–142, Jan. 2008, doi: 10.1094/MPMI-21-1-0133.
[58] A. Chatterjee, Y. Cui, and A. K. Chatterjee, "Regulation of Erwinia carotovora hrpL(Ecc) (sigma-L(Ecc)), which encodes an extracytoplasmic function subfamily of sigma factor required for expression of the HRP regulon," *Mol. Plant-Microbe Interact. MPMI*, vol. 15, no. 9, pp. 971–980, Sep. 2002, doi: 10.1094/MPMI.2002.15.9.971.
[59] A. Chatterjee, Y. Cui, H. Yang, A. Collmer, J. R. Alfano, and A. K. Chatterjee, "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," *Mol. Plant-Microbe Interactions*®, vol. 16, no. 12, pp. 1106–1117, Dec. 2003, doi: 10.1094/MPMI.2003.16.12.1106.
[60] S. Reverchon, W. Nasser, and J. Robert-Baudouy, "Characterization of kdgR, a gene of Erwinia chrysanthemi that regulates pectin degradation," *Mol. Microbiol.*, vol. 5, no. 9, pp. 2203–2216, Sep. 1991, doi: 10.1111/j.1365-2958.1991.tb02150.x.
[61] D. A. Rodionov, M. S. Gelfand, and N. Hugouvieux-Cotte-Pattat, "Comparative genomics of the KdgR regulon in Erwinia chrysanthemi 3937 and other gamma-proteobacteria," *Microbiol. Read. Engl.*, vol. 150, no. Pt 11, pp. 3571–3590, Nov. 2004, doi: 10.1099/mic.0.27041-0.
[62] W. Nasser, S. Reverchon, G. Condemine, and J. Robert-Baudouy, "Specific interactions of Erwinia chrysanthemi KdgR repressor with different operators of genes involved in pectinolysis," *J. Mol. Biol.*, vol. 236, no. 2, pp. 427–440, Feb. 1994, doi: 10.1006/jmbi.1994.1155.
[63] Y. Lu, I. M. Rashidul, H. Hirata, and S. Tsuyumu, "KdgR, an ICIR Family Transcriptional Regulator, Inhibits Virulence Mainly by Repression of hrp Genes in Xanthomonas oryzae pv. oryzae▽," *J. Bacteriol.*, vol. 193, no. 23, pp. 6674–6682, Dec. 2011, doi: 10.1128/JB.05714-11.
[64] D. M. Goudeau, C. T. Parker, Y. Zhou, S. Sela, Y. Kroupitski, and M. T. Brandl, "The Salmonella Transcriptome in Lettuce and Cilantro Soft Rot Reveals a Niche Overlap with the Animal Host Intestine," *Appl. Environ. Microbiol.*, vol. 79, no. 1, pp. 250–262, Jan. 2013, doi: 10.1128/AEM.02290-12.
[65] G. Kwan, A. O. Charkowski, and J. D. Barak, "Salmonella enterica suppresses

Pectobacterium carotovorum subsp. carotovorum population and soft rot progression by acidifying the microaerophilic environment," *mBio*, vol. 4, no. 1, pp. e00557-00512, Feb. 2013, doi: 10.1128/mBio.00557-12.
[66] A. S. George, I. Salas González, G. L. Lorca, and M. Teplitski, "Contribution of the Salmonella enterica KdgR Regulon to Persistence of the Pathogen in Vegetable Soft Rots," *Appl. Environ. Microbiol.*, vol. 82, no. 4, pp. 1353–1360, 15 2016, doi: 10.1128/AEM.03355-15.
[67] N. Mhedbi-Hajri, P. Malfatti, J. Pédron, S. Gaubert, S. Reverchon, and F. Van Gijsegem, "PecS is an important player in the regulatory network governing the coordinated expression of virulence genes during the interaction between Dickeya dadantii 3937 and plants," *Environ. Microbiol.*, vol. 13, no. 11, pp. 2901–2914, Nov. 2011, doi: 10.1111/j.1462-2920.2011.02566.x.
[68] F. Hommais *et al.*, "PecS Is a Global Regulator of the Symptomatic Phase in the Phytopathogenic Bacterium Erwinia chrysanthemi 3937," *J. Bacteriol.*, vol. 190, no. 22, pp. 7508–7522, Nov. 2008, doi: 10.1128/JB.00553-08.
[69] I. C. Perera and A. Grove, "Urate is a ligand for the transcriptional regulator PecS," *J. Mol. Biol.*, vol. 402, no. 3, pp. 539–551, Sep. 2010, doi: 10.1016/j.jmb.2010.07.053.
[70] E. Hérault, S. Reverchon, and W. Nasser, "Role of the LysR-type transcriptional regulator PecT and DNA supercoiling in the thermoregulation of *pel* genes, the major virulence factors in *Dickeya dadantii*: *Dickeya dadantii* PecT protein and virulence thermoregulation," *Environ. Microbiol.*, vol. 16, no. 3, pp. 734–745, Mar. 2014, doi: 10.1111/1462-2920.12198.
[71] E. Hérault, S. Reverchon, and W. Nasser, "Role of the LysR-type transcriptional regulator PecT and DNA supercoiling in the thermoregulation of pel genes, the major virulence factors in Dickeya dadantii," *Environ. Microbiol.*, vol. 16, no. 3, pp. 734–745, Mar. 2014, doi: 10.1111/1462-2920.12198.
[72] W. Nasser and S. Reverchon, "H-NS-dependent activation of pectate lyases synthesis in the phytopathogenic bacterium Erwinia chrysanthemi is mediated by the PecT repressor," *Mol. Microbiol.*, vol. 43, no. 3, pp. 733–748, Feb. 2002, doi: 10.1046/j.1365-2958.2002.02782.x.
[73] X. Tang, Y. Xiao, and J.-M. Zhou, "Regulation of the Type III Secretion System in Phytopathogenic Bacteria," *Mol. Plant-Microbe Interactions®*, vol. 19, no. 11, pp. 1159–1166, Nov. 2006, doi: 10.1094/MPMI-19-1159.
[74] M.-N. Yap, C.-H. Yang, J. D. Barak, C. E. Jahn, and A. O. Charkowski, "The Erwinia chrysanthemi Type III Secretion System Is Required for Multicellular Behavior," *J. Bacteriol.*, vol. 187, no. 2, pp. 639–648, Jan. 2005, doi: 10.1128/JB.187.2.639-648.2005.
[75] Y. Cui, A. Chatterjee, and A. K. Chatterjee, "Effects of the two-component system comprising GacA and GacS of Erwinia carotovora subsp. carotovora on the production of global regulatory rsmB RNA, extracellular enzymes, and harpinEcc," *Mol. Plant-Microbe Interact. MPMI*, vol. 14, no. 4, pp. 516–526, Apr. 2001, doi: 10.1094/MPMI.2001.14.4.516.
[76] R. Amit, A. B. Oppenheim, and J. Stavans, "Increased bending rigidity of single DNA molecules by H-NS, a temperature and osmolarity sensor," *Biophys. J.*, vol. 84, no. 4, pp. 2467–2473, Apr. 2003, doi: 10.1016/S0006-3495(03)75051-6.
[77] T. A. Azam and A. Ishihama, "Twelve species of the nucleoid-associated protein from Escherichia coli. Sequence recognition specificity and DNA binding affinity," *J. Biol. Chem.*, vol. 274, no. 46, pp. 33105–33113, Nov. 1999, doi: 10.1074/jbc.274.46.33105.
[78] S. Martis B, R. Forquet, S. Reverchon, W. Nasser, and S. Meyer, "DNA Supercoiling: an Ancestral Regulator of Gene Expression in Pathogenic Bacteria?," *Comput. Struct. Biotechnol. J.*, vol. 17, pp. 1047–1055, 2019, doi: 10.1016/j.csbj.2019.07.013.
[79] Y. Cui, A. Chatterjee, Y. Liu, C. K. Dumenyo, and A. K. Chatterjee, "Identification of a global repressor gene, rsmA, of Erwinia carotovora subsp. carotovora that controls extracellular enzymes, N-(3-oxohexanoyl)-L-homoserine lactone, and pathogenicity in

soft-rotting Erwinia spp.," *J. Bacteriol.*, vol. 177, no. 17, pp. 5108–5115, Sep. 1995, doi: 10.1128/jb.177.17.5108-5115.1995.
[80] Y. Cui, A. Mukherjee, C. K. Dumenyo, Y. Liu, and A. K. Chatterjee, "rsmC of the Soft-Rotting Bacterium Erwinia carotovora subsp. carotovora Negatively Controls Extracellular Enzyme and HarpinEcc Production and Virulence by Modulating Levels of Regulatory RNA (rsmB) and RNA-Binding Protein (RsmA)," *J. Bacteriol.*, vol. 181, no. 19, pp. 6042–6052, Oct. 1999, doi: 10.1128/JB.181.19.6042-6052.1999.
[81] L. Nachin and F. Barras, "External pH: an environmental signal that helps to rationalize pel gene duplication in Erwinia chrysanthemi," *Mol. Plant-Microbe Interact. MPMI*, vol. 13, no. 8, pp. 882–886, Aug. 2000, doi: 10.1094/MPMI.2000.13.8.882.
[82] F. Tardy, W. Nasser, J. Robert-Baudouy, and N. Hugouvieux-Cotte-Pattat, "Comparative analysis of the five major Erwinia chrysanthemi pectate lyases: enzyme characteristics and potential inhibitors.," *J. Bacteriol.*, vol. 179, no. 8, pp. 2503–2511, 1997, doi: 10.1128/JB.179.8.2503-2511.1997.
[83] W. Nasser, J. Robert-Baudouy, and S. Reverchon, "Antagonistic effect of CRP and KdgR in the transcription control of the Erwinia chrysanthemi pectinolysis genes," *Mol. Microbiol.*, vol. 26, no. 5, pp. 1071–1082, Dec. 1997, doi: 10.1046/j.1365-2958.1997.6472020.x.
[84] A. Duprey, W. Nasser, S. Léonard, C. Brochier-Armanet, and S. Reverchon, "Transcriptional start site turnover in the evolution of bacterial paralogous genes - the pelE-pelD virulence genes in Dickeya," *FEBS J.*, vol. 283, no. 22, pp. 4192–4207, 2016, doi: 10.1111/febs.13921.
[85] J. Monod, "The Growth of Bacterial Cultures," p. 25.
[86] B. Magasanik, "Catabolite Repression," *Cold Spring Harb. Symp. Quant. Biol.*, vol. 26, pp. 249–256, Jan. 1961, doi: 10.1101/SQB.1961.026.01.031.
[87] M. H. Saier and S. Roseman, "Inducer exclusion and repression of enzyme synthesis in mutants of Salmonella typhimurium defective in enzyme I of the phosphoenolpyruvate: sugar phosphotransferase system," *J. Biol. Chem.*, vol. 247, no. 3, pp. 972–975, Feb. 1972.
[88] R. R. Chaudhuri *et al.*, "Comprehensive assignment of roles for Salmonella typhimurium genes in intestinal colonization of food-producing animals," *PLoS Genet.*, vol. 9, no. 4, p. e1003456, Apr. 2013, doi: 10.1371/journal.pgen.1003456.
[89] "*lac* operon," *Wikipedia*. Aug. 04, 2020, Accessed: Aug. 09, 2020. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Lac_operon&oldid=971168681.
[90] J. Deutscher, C. Francke, and P. W. Postma, "How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria," *Microbiol. Mol. Biol. Rev. MMBR*, vol. 70, no. 4, pp. 939–1031, Dec. 2006, doi: 10.1128/MMBR.00024-06.
[91] B. M. Hogema *et al.*, "Inducer exclusion by glucose 6-phosphate in Escherichia coli," *Mol. Microbiol.*, vol. 28, no. 4, pp. 755–765, May 1998, doi: 10.1046/j.1365-2958.1998.00833.x.
[92] K. E. R. Hollinshead and D. A. Tennant, "Mitochondrial metabolic remodeling in response to genetic and environmental perturbations," *WIREs Syst. Biol. Med.*, vol. 8, no. 4, pp. 272–285, 2016, doi: 10.1002/wsbm.1334.
[93] T. Inada, K. Kimata, and H. Aiba, "Mechanism responsible for glucose–lactose diauxie in Escherichia coli: challenge to the cAMP model," *Genes Cells*, vol. 1, no. 3, pp. 293–301, 1996, doi: 10.1046/j.1365-2443.1996.24025.x.
[94] S. Uppal and N. Jawali, "Cyclic AMP receptor protein (CRP) regulates the expression of cspA, cspB, cspG and cspI, members of cspA family, in Escherichia coli," *Arch. Microbiol.*, vol. 197, no. 3, pp. 497–501, Apr. 2015, doi: 10.1007/s00203-015-1085-4.
[95] G. Soberón-Chávez, L. D. Alcaraz, E. Morales, G. Y. Ponce-Soto, and L. Servín-González, "The Transcriptional Regulators of the CRP Family Regulate Different Essential Bacterial Functions and Can Be Inherited Vertically and Horizontally," *Front. Microbiol.*, vol. 8, May 2017, doi: 10.3389/fmicb.2017.00959.

[96] A. Ay and D. N. Arnosti, “Mathematical modeling of gene expression: a guide for the perplexed biologist,” *Crit. Rev. Biochem. Mol. Biol.*, vol. 46, no. 2, pp. 137–151, Apr. 2011, doi: 10.3109/10409238.2011.556597.
[97] H. de Jong, “Modeling and simulation of genetic regulatory systems: a literature review,” *J. Comput. Biol. J. Comput. Mol. Cell Biol.*, vol. 9, no. 1, pp. 67–103, 2002, doi: 10.1089/10665270252833208.
[98] N. Friedman, M. Linial, I. Nachman, and D. Pe’er, “Using Bayesian networks to analyze expression data,” *J. Comput. Biol. J. Comput. Mol. Cell Biol.*, vol. 7, no. 3–4, pp. 601–620, 2000, doi: 10.1089/106652700750050961.
[99] N. Joly *et al.*, “Managing membrane stress: the phage shock protein (Psp) response, from molecular mechanisms to physiology,” *FEMS Microbiol. Rev.*, vol. 34, no. 5, pp. 797–827, Sep. 2010, doi: 10.1111/j.1574-6976.2010.00240.x.
[100] P. Model, G. Jovanovic, and J. Dworkin, “The Escherichia coli phage-shock-protein (psp) operon,” *Mol. Microbiol.*, vol. 24, no. 2, pp. 255–261, Apr. 1997, doi: 10.1046/j.1365-2958.1997.3481712.x.
[101] A. J. Darwin, “Stress relief during host infection: The phage shock protein response supports bacterial virulence in various ways,” *PLoS Pathog.*, vol. 9, no. 7, p. e1003388, 2013, doi: 10.1371/journal.ppat.1003388.
[102] W. D. Fakhouri, A. Ay, R. Sayal, J. Dresch, E. Dayringer, and D. N. Arnosti, “Deciphering a transcriptional regulatory code: modeling short-range repression in the Drosophila embryo,” *Mol. Syst. Biol.*, vol. 6, p. 341, 2010, doi: 10.1038/msb.2009.97.
[103] C. H. Yuh, H. Bolouri, and E. H. Davidson, “Cis-regulatory logic in the endo16 gene: switching from a specification to a differentiation mode of control,” *Dev. Camb. Engl.*, vol. 128, no. 5, pp. 617–629, Mar. 2001.
[104] J. Jaeger *et al.*, “Dynamic control of positional information in the early Drosophila embryo,” *Nature*, vol. 430, no. 6997, pp. 368–371, Jul. 2004, doi: 10.1038/nature02678.
[105] B. Huang, D. Jia, J. Feng, H. Levine, J. N. Onuchic, and M. Lu, “RACIPE: a computational tool for modeling gene regulatory circuits using randomization,” *BMC Syst. Biol.*, vol. 12, no. 1, p. 74, 19 2018, doi: 10.1186/s12918-018-0594-6.
[106] X.-J. Feng, S. Hooshangi, D. Chen, G. Li, R. Weiss, and H. Rabitz, “Optimizing genetic circuits by global sensitivity analysis,” *Biophys. J.*, vol. 87, no. 4, pp. 2195–2202, Oct. 2004, doi: 10.1529/biophysj.104.044131.
[107] E. Meir, G. von Dassow, E. Munro, and G. M. Odell, “Robustness, flexibility, and the role of lateral inhibition in the neurogenic network,” *Curr. Biol. CB*, vol. 12, no. 10, pp. 778–786, May 2002, doi: 10.1016/s0960-9822(02)00839-4.
[108] M. Leon, M. L. Woods, A. J. H. Fedorec, and C. P. Barnes, “A computational method for the investigation of multistable systems and its application to genetic switches,” *BMC Syst. Biol.*, vol. 10, no. 1, p. 130, 07 2016, doi: 10.1186/s12918-016-0375-z.
[109] T. Chen, H. L. He, and G. M. Church, “Modeling gene expression with differential equations,” in *Biocomputing ’99*, 0 vols., WORLD SCIENTIFIC, 1998, pp. 29–40.
[110] E. P. Greenberg, “Bacterial genomics. Pump up the versatility,” *Nature*, vol. 406, no. 6799, pp. 947–948, Aug. 2000, doi: 10.1038/35023203.
[111] C. M. Arraiano, F. Mauxion, S. C. Viegas, R. G. Matos, and B. Séraphin, “Intracellular ribonucleases involved in transcript processing and decay: precision tools for RNA,” *Biochim. Biophys. Acta*, vol. 1829, no. 6–7, pp. 491–513, Jul. 2013, doi: 10.1016/j.bbagrm.2013.03.009.
[112] R. Gao and A. M. Stock, “Molecular strategies for phosphorylation-mediated regulation of response regulator activity,” *Curr. Opin. Microbiol.*, vol. 13, no. 2, pp. 160–167, Apr. 2010, doi: 10.1016/j.mib.2009.12.009.
[113] P. Joyet *et al.*, “Transcription regulators controlled by interaction with enzyme IIB components of the phosphoenolpyruvate: sugar phosphotransferase system,” *Biochim.*

Biophys. Acta, vol. 1834, no. 7, pp. 1415–1424, Jul. 2013, doi: 10.1016/j.bbapap.2013.01.004.
[114] J. P. Faria, R. Overbeek, F. Xia, M. Rocha, I. Rocha, and C. S. Henry, “Genome-scale bacterial transcriptional regulatory networks: reconstruction and integrated analysis with metabolic models,” *Brief. Bioinform.*, vol. 15, no. 4, pp. 592–611, Jul. 2014, doi: 10.1093/bib/bbs071.
[115] E. Segal, T. Raveh-Sadka, M. Schroeder, U. Unnerstall, and U. Gaul, “Predicting expression patterns from regulatory sequence in Drosophila segmentation,” *Nature*, vol. 451, no. 7178, pp. 535–540, Jan. 2008, doi: 10.1038/nature06496.
[116] A. Y. Tulchinsky, N. A. Johnson, W. B. Watt, and A. H. Porter, “Hybrid incompatibility arises in a sequence-based bioenergetic model of transcription factor binding,” *Genetics*, vol. 198, no. 3, pp. 1155–1166, Nov. 2014, doi: 10.1534/genetics.114.168112.
[117] S. L. Barnes, N. M. Belliveau, W. T. Ireland, J. B. Kinney, and R. Phillips, “Mapping DNA sequence to transcription factor binding energy in vivo,” *PLOS Comput. Biol.*, vol. 15, no. 2, p. e1006226, Feb. 2019, doi: 10.1371/journal.pcbi.1006226.
[118] M. Slattery, T. Zhou, L. Yang, A. C. Dantas Machado, R. Gordân, and R. Rohs, “Absence of a simple code: how transcription factors read the genome,” *Trends Biochem. Sci.*, vol. 39, no. 9, pp. 381–399, Sep. 2014, doi: 10.1016/j.tibs.2014.07.002.
[119] K. D. MacIsaac, K. A. Lo, W. Gordon, S. Motola, T. Mazor, and E. Fraenkel, “A Quantitative Model of Transcriptional Regulation Reveals the Influence of Binding Location on Expression,” *PLOS Comput. Biol.*, vol. 6, no. 4, p. e1000773, Apr. 2010, doi: 10.1371/journal.pcbi.1000773.
[120] T. Zhou *et al.*, “Quantitative modeling of transcription factor binding specificities using DNA shape,” *Proc. Natl. Acad. Sci.*, vol. 112, no. 15, pp. 4654–4659, Apr. 2015, doi: 10.1073/pnas.1422023112.
[121] L. Bintu *et al.*, “Transcriptional regulation by the numbers: models,” *Curr. Opin. Genet. Dev.*, vol. 15, no. 2, pp. 116–124, Apr. 2005, doi: 10.1016/j.gde.2005.02.007.
[122] “Overview: Gene regulation in bacteria (article),” *Khan Academy*. https://www.khanacademy.org/science/biology/gene-regulation/gene-regulation-in-bacteria/a/overview-gene-regulation-in-bacteria (accessed Aug. 09, 2020).

Chapter II: Computational modeling of regulation of *pelE* and *pelD* genes with respect to local transcription factor KdgR and global transcription factor CRP

1.Introduction

Modeling and simulation of microbial cell processes is an important aspect in biological studies. Combination of experimental works with mathematical modeling makes it possible to provide meaningful and quantitative interpretation of the experimental results and also reveal new aspects of microbial physiology. As we have discussed in the previous chapter, we intend to model the expression of virulence *pel* genes in *Dickeya dadantii*. The ability of the model organism to utilize pectin present in the plant cell wall makes it a competent phytopathogen. Initiation of *pel* gene expression is a key step in triggering the virulence of *Dickeya dadantii* providing a source of energy as well as helps the bacterium break down the primary defence layer of the plant cells, the cell wall [1]. The de-polymerization of pectin is accomplished by a variety of pectinases produced by pathogens. Among these pectinases, endo-pectate lyases (Pels) are known to play the major role in the soft rot disease caused by this bacterial species. PelABCDEILNZ are the extracellular acting *pels*, they act on the long chain galacturonate polymers. PelX acts on the oligogalacturonides in the periplasm of the bacteria and PelW acts on smaller galacturonates in the inner cell membrane [2].

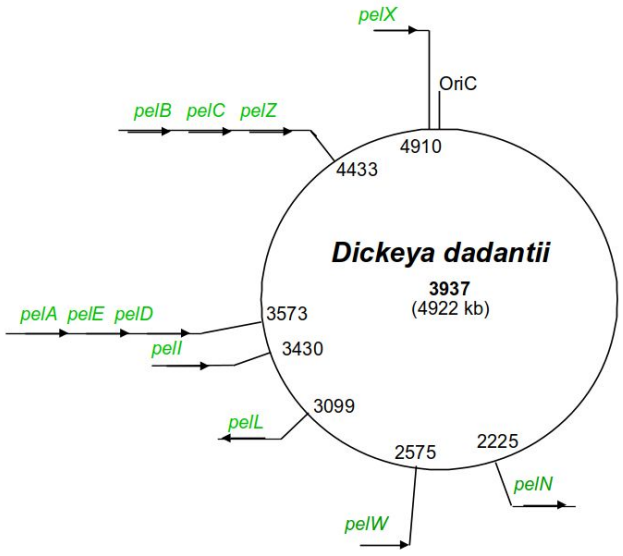


Figure 2.1.1: The locations of *pel* genes along the *Dickeya dadantii* chromosome

Locations of the *pels* are spread across the Dickeya chromosome as shown in Figure 2.1.. *pelAED* and *pelBC* are located together in the chromosome [3]. Multiple PAIs are located in the *Dickeya dadantii* chromosome which is unusual to PAIs generally found in plasmids [4]. It can be inferred that these virulence genes are essential for this pathogen as they are integrated within the chromosome unlike certain opportunistic pathogens having them in their plasmids. Expression of all these the *pel* genes involved in pectin degradation is controlled by a complex regulation system. Previous studies have led to the identification of several regulators (KdgR, PecS, PecT, CRP, H-NS and ExpR) controlling the synthesis of Pels in response to various signals, such as the presence of pectin [5]–[13]. The *pel* genes are strongly repressed by KdgR, a protein regulating most of the genes involved in pectinolysis and activated by a cAMP receptor protein (CRP) activator [14].

An earlier study in our lab modeled the *pel* expression with respect to KdgR repressor [15]. The model defined the dynamics of *pel* expression and pectin degradation. The *pel* enzymes randomly cleave pectic polymers such as polygalacturonate (PGA), de-esterified pectin, or low esterified pectin to unsaturated oligo-galacturonates (UGA). into the cell and is converted to 2-keto-3-deoxygluconate (KDG) [15]. KDG binds to KgdR, a main repressor of *pel* genes and forms a complex [KdgR.KDG] and relieves the *pel* gene repression. The model presented a few drawbacks with respect to certain assumptions made. One of the main drawbacks is that the previous model assumed the import of UGA to be a very slow process, thus, predicting the accumulation of KDG to be slower and reaching a peak simultaneous to the *pel* expression peaks. This assumption was inferred from the hypothesis that KdgR is the only regulator of *pel* expression, which interacts with the metabolite KDG. In reality, the slow intake of primary carbon sources in a pathogenic environment becomes suspicious, the pathogen is always in search of carbon sources to meet the energy required during pathogenesis progression. In order to clarify this doubt we developed an inhouse method to quantify KDG over time and understand the mystery in dynamics of UGA uptake. The paper next explains the quantification technique of KDG using a HPLC platform.

2. Quantification of KDG using HPLC

Separation and quantification of 2-Keto-3-deoxy-gluconate (KDG) a major metabolite in pectin and alginate degradation pathways, previous derivatization with *O*-phenylenediamine (OPD) using HPLC with photometric and fluorimetric serial detection

Shiny MARTIS B[a], Michel DROUX[a,*], Florelle DEBOUDARD[a], William NASSER[a], Sam MEYER[a], Sylvie REVERCHON[a]

[a]Univ Lyon, Université Claude Bernard Lyon 1, INSA-Lyon, CNRS, UMR5240 MAP, F-69621, France

[*] Corresponding author: michel.droux@univ-lyon1.fr
Biomolecule platform

Highlights

A novel HPLC method is reported for quantification of KDG, a metabolite originating from alginate and pectin polysaccharide degradation pathways.

Abstract

A rapid and sensitive High Performance Liquid Chromatography (HPLC) method with photometric and fluorescence detection is developed for routine analysis of 2-Keto-3-deoxy-gluconate (KDG), a catabolite product of pectin and alginate. These polysaccharides are both used for biofuel production and to generate high-value-added products. HPLC is performed, after derivatization of the carbonyl groups of the metabolite with *O*-phenylenediamine (OPD), using a linear gradient of trifluoroacetic acid and acetonitrile. Quantification is accomplished with an internal standard method. The gradient is optimized to distinguish KDG from its close structural analogues such as 5-keto-4-deoxyuronate (DKI) and 2,5-diketo-3-deoxygluconate (DKII). The proposed method is simple, highly sensitive and accurate for time course analysis of pectin or alginate degradation.

Keywords: KDG, HPLC, *O*-phenylenediamine, fluorescence, pectin or alginate degradation

Introduction

The major obstacle to use plant biomass for the production of fuels, chemicals, and bioproducts, is our current lack of knowledge of how to effectively deconstruct cell-wall polymers for their subsequent use as feedstocks. Pectin plays an important role in biomass recalcitrance [1]. Pectins are complex branched polysaccharides containing high amount of partly methyl-esterified galacturonic acid and low amount of rhamnose and carry arabinose and galactose as major neutral sugars. Due to their structural complexity, they are modifiable by many different enzymes, including polygalacturonases, pectate lyases, and esterases. Pectin-degrading enzymes and -modifying enzymes may be use in a wide variety of applications to modulate pectin properties or produce pectin derivatives for food industry [2], as well as for pharmaceutical industry. Pectin and pectin-derived oligosaccharides acts as anti-metastatic [3], cholesterol inhibiting [4], prebiotic and probiotic encapsulant [5]. In addition, pectin is being widely used as a component of bio-nano-packaging [6,7].

In this context, numerous plant pathogenic microorganisms such as soft-rot bacteria (*Dickeya* species and *Pectobacterium*species) or necrotic fungi (*Aspergillus species, Botritys species*) are effective in pectin degradation and represent a large resource of pectinases that generated KDG metabolite [8-10]. In addition, novel metabolic pathways for hexuronate utilisation in various proteobacteria have been recently described and all converge to the KDG metabolite formation [11,12].

Similarly, algae are a large group of marine vegetation that also immerge as feedstock for biofuel production [13,14]. Alginate is a unique structural polysaccharide in algae, abundant in the cell wall ensuring resistance to mechanical stress. It is a linear copolymer of α-L-guluronate and its C5 epimer β-D-mannuronate [15]. Alginate degrading enzymes include different types of alginate lyases generating various alginate oligosaccharides which are further converted to KDG. Alginate lyases are present in a wide range of marine and terrestrial bacteria [16]. In addition to biofuel production, alginate derivatives can be used as therapeutic agents such as anticoagulants [17] and tumor inhibition [18]. Furthermore, alginate lyase also shows great potential application in treatment of cystic fibrosis by degrading the polysaccharide biofilm of bacterium [19]. KdgF was identifed as the missing link in the microbial metabolism of uronate sugars from pectin and alginate|20] (Figure 1)

Therefore, quantification of KDG is of wide interest to analyse the metabolic flux of hexuronate, pectin and alginate from numerous microorganisms.

In this paper, we took advantage of our expertise on the plant pathogenic bacteria *Dickeya dadantii* and the availability of both chemical and genetic resources in order to develop methods for monitoring and optimizing the breakdown of pectic polymers into oligomers or monomers of interest. This bacterium attacks a wide range of plant species, including many crops of economic importance. Soft rot, the visible symptom, is mainly due to the degradation of pectins by pectate lyase (Pel) activity generating unsaturated oligogalacturonate [21,22]. *Dickeya* also produces low amount of polygalacturonases generating saturated oligo-galacturonates (pehN, pehVWX, pehK) [23], |24]. *D. dadantii* can utilize pectin as its sole carbon and energy source (Figure 1) [21]. The catabolic pathway of pectins (pectinolysis) has been characterized at the genetic and biochemical level. These studies have revealed that KDG, a catabolite product of pectin, plays a key role in pectinolysis by inactivating KdgR, the main

repressor of this catabolic pathway [25-27]. The positive feedback loop between the extracellular pectin degradation by Pels and Pehs and the intracellular inactivation of KdgR by the metabolite KDG has been modeled [7,28]. Refining this type of model requires a precise determination of the concentration of KDG present in the bacteria's growth environment. The conventional technique to determine the concentration of KDG so far is thiobarbituric assay [29]. Detection of KDG can be carried on by cellulose thin-layer chromatography by spraying with either periodate-thiobarbituric acid or *o*-phenylenediamine [30]. These methods are not very accurate for biological application since they are not very specific to KDG alone, but rather quantify DKI, DKII and KDG indistinctly. Here, we present a new method to quantify this metabolite using HPLC. The novel technique presented here allows quantifying KDG with increased precision and specificity with respect to functionally related compounds, as required for further regulatory models.

Materials and Methods

Chemicals

KDG (95%), OPD (O-phenylenediamine), HPLC grade Trifluoroacetic acid (TFA) were from Sigma Aldrich (France). Partially purified DKI and DKII (90%) was a gift by Guy Condemine. Optima LC-MS acetonitrile (ACN) was from Fisher UK. Polygalacturonate dipecta (PGA) was from Agdia laboratory (France). All HPLC elutants, standards and assays were prepared using MilliQ pure water with a specific resistance of at least 18 m Ω .

Bacterial strains and culture conditions

Table 1: *Dickeya dadantii* strains used in the study

Strain	Genotype	Reference
3937	Wild Type	[31]
A2395	*kduI::kan*	[32]
A797	*kduD::kan*	[33]
A3999	*kdgK::kan*	[34]

Dickeya dadantii (strain 3937) and its derived *kduI*, *kduD* and *kdgK* mutants were used for all the experiments [21] (Table 1). Bacteria were grown at 30°C in M63 minimal medium supplemented with a carbon source: glucose (0.2% w/v) alone or in addition with PGA (0.4% w/v). Cells were harvested in the late exponential growth phase when the bacterial population has reached its maximum.

Sample collection

Bacteria were separated from the medium by centrifugation at 5000 *g* for 8 minutes. Cells and cultures medium samples were freeze dried with liquid nitrogen and stored separately at -80°C. Bacterial pellet was resuspended with a 10 ml cold solution of 2.5 % (v/v) TFA, 50% (v/v) acetonitrile. The dissolved lysed pellet was maintained at -80°C for 15 minutes and thawed at room temperature. The solution was then centrifuged at 20000 *g* for 5 minutes. The clear supernatant containing the metabolites (10 ml) was separated from the pellet and then lyophilized overnight. The dry powder was dissolved in 1 ml of 0.1 M HCl and centrifuged again to remove particles. Culture medium (10 ml) was lyophilized and re-suspended in 1 ml 0.1 M HCl.

Sample derivatization

200 µL of sample was incubated for 60 minutes at 60°C with 20 µL solution of 100 mM OPD prepared in 0.1 M HCl with agitation in an Eppendorf thermomixer as described previously [35], (Figure S1). Samples turn yellowish. Appropriate amounts of the metabolites-adduct were then injected into HPLC for analysis.

HPLC system and gradient elution

The HPLC system consisted of an Alliance HT Waters 2795 separation module and detection was performed with a Waters 2996 photodiode array detector interconnected with a fluorimeter SFM25 (Kontron). The OPD-adducts to be determined were passed through a 4.6x250-mm Uptisphere 5 µM HDOC18 column (Interchim) at 35°C and a flow rate of 1 mL/min. The elution protocol employed linear gradient as follow: initial 0 to 5 min, 100%A; 5 to 12 min, 75%A, 25%B; 12 to 20 min, 75%A, 25%B; 20 min to 30 min 75%A, 25% B to 100 %B; 30 min to 35 min, 100%B, then, 35 min to 37 min 100%B to 100%A and re-equilibrated for 3 min in 100%A before a new run. Solvent A consisted of 0.005% trifluoroacetic acid (TFA) and solvent B contained 60% acetonitrile. The OPD adduct(s) was detected using a photodiode array detector (Waters 2996) set up at 338 nm, and with a SFM25 fluorimeter (Kontron) equipped with a 8 µL flow cell. Emitted fluorescence of the OPD adduct(s) was measured with an excitation wavelength of 338 nm and an emission wavelength of 414 nm (slit sizes, 15 and 10 nm respectively, and lamp power of 400 W). Quantification of metabolites was carried out by measuring peak areas using the chromatography data system Mass-Link software and assigned intensity to the peak (Waters).

Results and discusssion

The derivatization and subsequent HPLC procedure described here represents a successful improvement over the thiobarbituric method developed previously [24]. We used OPD as a probe for the carbonyl groups of metabolites such as KDG, DKI and DKII (Figure S1). Figure 2 shows a representative chromatogram of KDG standard solution. The standard KDG has a retention time of around 16 minutes. From the estimated area of the respective peaks, the fluorescence intensity is approximately 2500-fold more sensitive than the UV absorbance intensity. Thus, fluorescence is more accurate for the resolution and quantitation of the

metabolite of interest. The ODP adduct of the carbonyl group of KDG is characterized by a typical spectrum with a maximum absorption in the limit of the visible range at 338 nm. Upon excitation at this wavelength, a maximum emission is measured at 414 nm (Figure 3A and B).

The unsaturated digalacturonate pathway has DKI and DKII as intermediate metabolites before the formation of KDG (Figure 1). The properties of derivatization based on the presence of a carbonyl group led us to hypothesize that these metabolites would be derivatized also in our experimental conditions. We performed the same procedure of analysis starting from a partially pure DKI/DKII mixture (obtained by purification from *kduI/kduD* mutants, G. Condemine, personal communication). The closeness of the structures of DKI and DKII makes it difficult to differentiate them. Consistently, the corresponding chromatogram reveals a major single peak with a retention time of about 21 minutes for these compounds (Figure 4). We hypothesized that the signature spectrum of the KDG molecule would be comparable to that of DKI or DKII since the molecular structures are closely similar (see Figure 1). Only the predominant peak present in the absorbance chromatogram (Figure 4A) shows the typical spectrum signature of KDG as described in Figure 3A. We therefore optimized the gradient to discriminate the KDG from other similar structured metabolites produced in the media culture like DKI and DKII, allowing quantification from the intensity of a single specific peak (Figure 2).

Linearity

Linearity of the method was confirmed by preparing KDG standard curves for the analytical range of 0-5 nanomoles (Figure 5). Statistical analysis using Least Square Regression (LSR) indicated excellent linearity in the mentioned range, with a correlation of $R^2 \geq 0.993$ for all standard curves with fluorescence detection and $R^2 > 0.988$ for the UV detection. The slopes of the calibration curves were 0.0026 and 6.64 for absorbance and fluorescence intensity detection, respectively (Figure 5). From the respective slopes, the fluorescence intensity detection is more sensitive than the absorbance intensity detection, with a ratio of around 2500-fold.

Limit of Detection and Limit of Quantification

The range of application of the method was verified by analyzing replicate samples containing the metabolite at various concentrations. The Limit of Detection (LOD) and Limit of Quantification (LOQ) were calculated by the linear regression method: the LOD is given by $3S_a/b$ and LOQ is given by $10S_a/b$, where S_a is the standard error of the intercept and b is the slope of the calibration curve [36]. Based on these equations for absorbance and fluorescence detection (Fig. 5), the calculated LOD values were found to be 0.093 nanomoles and 0.077 nanomoles respectively, and LOQ values were found to be 0.28 nanomoles and 0.23 nanomoles respectively, in our experimental condition settings.

Biological Sample Analysis

In order to test the method for KDG quantification, we used the *Dickeya dadantii* wild type strain 3937 and its *kduI*, *kduD* and *kdgK* mutant derivatives. Bacteria were grown on a glucose minimal medium supplemented or not with polygalacturonate (PGA), and samples were collected after 10 hours of growth, corresponding to the transition to stationary phase. The characteristic peak of KDG appeared only in the presence of addition of PGA in the medium, and was absent in cultures performed in glucose, as expected. Figure 6 depicts the chromatogram for the wild type bacterial samples with and without PGA in the isocratic phase of the gradient to amplify the image resolution of KDG-OPD adduct molecule. The full chromatogram for the wild type strain is depicted (Figure S2). KDG is eluted at around 16 min in a symmetrical peak while DKI and DKII accumulated in the *kduI*, and *kduD* mutants, respectively at 21 min (Figure S3 and S4). As expected, in the *kdgK* mutant, a unique peak at 16 min is highlighted in presence of polygalacturonate in the medium (Figure S5).

Quantitation of metabolites

The KDG concentration was quantified in the wild type *D. dadantii* strain and the three mutant derivatives (*kduI*, *kduD* and *kdgK)* to study the formation of intracellular concentration of KDG (Figure 7). The intracellular quantification was carried out based on the assumption of a cell volume to be 3.6 $\mu L.mL^{-1}.OD^{-1}$ where optical density (OD) is measured at 600 nm [37]. In the medium supplemented with a PGA source the maximum quantification of KDG is observed in the *kdgK* mutant, as expected since it accumulates the KDG metabolite (up to 233 mM). The *kduI* and *kduD* mutants accumulated KDG at higher level than wild type cells (10- to 30-times respectively), due to the increased catabolism of saturated galacturonate in the hexuronate pathway. The Pel enzymatic activities assayed in the mutants also are consistent with the elevated levels of KDG (Table S1).

DKI and DKII could not be further differentiated on the chromatogram due to their structural similarity and unavailability of pure DKI and DKII compounds. We quantified DKI and DKII based on the KDG calibration curve considering that the absorbance and fluorescence slopes would be similar. In *kduI* mutants DKI is not converted into DKII, and its concentration was quantified around 15mM. In the *kduD* mutant the combined concentration of DKI+DKII was 31mM. The *kdgK* mutant had no detectable DKI or DKII molecules.

.

Conclusion

To demonstrate that our technique is specific for KDG alone, we showed that the retention time for the derivatives DKI/DKII is 21 minutes whereas the retention time for KDG is around 16 minutes. Nonetheless, this method is solid enough to eliminate the suspicion of cross derivatization between KDG and DKI/DKII. We do not consider galacturonate pathway since the similarity in the structure of compounds are not significant for the possibility of cross derivatization.

The fluorescent technique was 2500-fold more sensitive over the absorbance detection. The results also indicate the higher reliability of the fluorescence technique over the absorbance detection method since it has a lower coefficient of variation. Nonetheless, the combination of the two techniques provide further validation of the results and stronger quantification.

Thus, we describe a robust method for the quantification of the metabolite KDG from biological samples. This method will help in the analysis and quantitative modelling of gene regulatory networks where KDG is known to have any effect [7,28]. Moreover, quantifying metabolite KDG flux is critical for biofuel production during pectin and alginate degradation or for other high value products derived from these two polymers [20]. There is also potential scope for the industrial application of this technique in metabolic engineering of compounds like Vitamin C from KDG [38].

Acknowledgements
Authors are grateful to G Condemine and N Hugouvieux-Cotte-Pattat for providing the DKI/DKII samples, and for discussion and critical readings.

Funding
BQR INSA 2016 (to S. M.)

References

[1] M.E. Himmel, Biomass recalcitrance: deconstructing the plant cell wall for bioenergy. Oxford, 2008 568 pages Blackwell.

[2] T. Vanitha, M. Khan, Role of Pectin in Food Processing and Food Packaging. February 22nd 2019. doi: 10.5772/intechopen.83677.

[3] Y. Kimura, Prevention of cancer chemotherapy drug-induced adverse reaction, antitumor and antimetastatic activities by natural products, In Bioactive Natural Products (Part I), Studies in Natural Products Chemistry 28 Part I 559-586 (2003). doi: 10.1016/S1572-5995(03)80149-7.

.

[4] J.J. Cerda, S.J. Normann, M.P. Sullivan, C.W. Burgin, F.L. Robbins, S. Vathada, P. Leelachaikul, Inhibition of atherosclerosis by dietary pectin in microswine with sustained hypercholesterolemia, Circulation 89(3) (1994) 1247-1253. doi:10.1161/01.cir.89.3.1247.

[5] A. Yasmin, M.S. Butt, M. Afzaal, M. Baak, M.T. Nadeem, M.Z. Shahid, Prebiotics, gut microbiota and metabolic risks: Unveiling the relationship, Journal of Functional Foods 17 (2015) 189-201. doi.org/10.1016/j.jff.2015.05.004.

[6] F. Naqash, F.A. Masoodi, S.A. Rather, S.M. Wani, A. Gani, Emerging concepts in the nutraceutical and functional properties of pectin-A Review, Carbohydrate Polymers 168 (2017) 227-239. doi.org/10.1016/j.carbpol.2017.03.058.

[7] J.A. Sepulchre, S. Reverchon, J.L. Gouzé, W. Nasser, Modeling the bioconversion of polysaccharides in a continuous reactor: A case study of the production of oligogalacturonates by *Dickeya dadantii*, J Biol Chem. 294(5) (2019) 1753-1762. doi:10.1074/jbc.RA118.004615.

[8] G. Bethke, J. Glazebrook, Measuring pectin properties to track cell wall alterations during plant-pathogen interactions, Methods Mol Biol. 1991 (2019) 55-60. doi:10.1007/978-1-4939-9458-8_6.

[9] H. Culleton, V. McKie, R.P. de Vries, Physiological and molecular aspects of degradation of plant polysaccharides by fungi: What have we learned from Aspergillus?, Biotechnol J. 8(8) (2013) 884–894. doi: 10.1002/biot.201200382.

[10] L. Zhang, R.J. Lubbers, A. Simon, J.H. Stassen, P.R. Vargas Ribera, M. Viaud, J.A. van Kan, A novel Zn_2Cys_6 transcription factor BcGaaR regulates D-galacturonic acid utilization in *Botrytis cinerea*, Mol Microbiol. 100 (2016) 247-262. doi:10.1111mmi.13314.

[11] J.T. Bouvier, N.V. Sernova, S. Ghasempur, I.A. Rodionova, M.W. Vetting, N.F. Al-Obaidi, S.C. Almo, J.A. Gerlt, D.A. Rodionov, Novel metabolic pathways and regulons

for hexuronate utilization in proteobacteria, J Bacteriol. 201(2) (2018) e00431-18. doi:10.1128/JB.00431-18.

[12] J. Kuivanen, A. Biz, P. Richard, Microbial hexuronate catabolism in biotechnology, AMB Express *9*(1) (2019) 16. doi.org/10.1186/s13568-019-0737-1.

|13] N. Wei, J. Quarterman, Y. Jin, Marine macroalgae: an untapped resource for producing fuels and chemicals, Trends Biotechnol. 31 (2013) 70–77. doi: 10.1016/j.tibtech.2012.10.009.

[14] A.J. Wargacki, E. Leonard, M.N. Win, D.D. Regitsky, C.N.Santos, P.B. Kim, S.R. Cooper, R.M. Raisner, A. Herman, A.B. Sivitz, A. Lakshmanaswamy, Y. Kashiyama, D. Baker, Y. Yoshikuni, An engineered microbial platform for direct biofuel production from brown macroalgae, Science 335 (2012) 308–313. doi: 10.1126/science.1214547.

[15] B. Zhu, H. Yin, Alginate lyase: Review of major sources and classification, properties, structure-function analysis and applications, Bioengineered 6(3) (2015) 125-131. doi:10.1080/21655979.2015.1030543.

|16] S.Q Ji, B. Wang, M. Lu, F.L. Li, *Defluviitalea phaphyphila sp. nov*., a novel thermophilic bacterium that degrades brown algae, Appl Environ Microbiol. 82(3) (2015) 868-877. doi:10.1128/AEM.03297-15.

[17] H. Huang, H. Chen, X. Wang, F. Qiu, H. Liu, J. Lu, L. Tong, Y. Yang, X. Wang, H. Wu, Degradable and bioadhesive alginate-based composites: an effective hemostatic agent, ACS Biomater. Sci. Eng. 5(10) (**2019**) 5498-5505. doi.org/10.1021/acsbiomaterials.9b01120.

[18] T. Han, L. Zhang, X. Yu, S. Wang, C. Xu, H. Yin, S. Wang, Alginate oligosaccharide attenuates α 2,6-sialylation modification to inhibit prostate cancer cell growth via the Hippo/YAP pathway, Cell death & disease 10(5) (2019) 374. doi.org/10.1038/s41419-019-1560-y.

[19] N. Blanco-Cabra, B. Paetzold, T. Ferrar, R. Mazzolini, E. Torrents, L. Serrano, M. LLuch-Senar, Characterization of different alginate lyases for dissolving *Pseudomonas aeruginosa* biofilms, Sci Rep. 10(1) (2020) 9390. doi.org/10.1038/s41598-020-66293-2.

[20] J.K. Hobbs, S.M. Lee, M. Robb, F. Hof, C. Barr, K.T. Abe, J-H. Hehemann, R. McLean, D.W. Abbott, A.B. Boraston, KdgF, the missing link in the microbial metabolism of uronate sugars from pectin and alginate. Proc Natl Acad Sci U S A. 113(22) (2016) 6188-6193. doi:10.1073/pnas.1524214113.

|21] N. Hugouvieux-Cotte-Pattat, G. Condemine, W. Nasser, S. Reverchon, Regulation of pectinolysis in *Erwinia chrysanthemi,* Annu Rev Microbiol. 50 (1996) 213–257. doi: 10.1146/annurev.micro.50.1.213.

[22] N. Hugouvieux-Cotte-Pattat, G. Condemine, V.E. Shevchik, Bacterial pectate lyases, structural and functional diversity: Bacterial pectate lyases, Environmental Microbiology Reports. 6(5) (2014) 427–440. doi: 10.1111/1758-2229.12166.

[23] N. Hugouvieux-Cotte-Pattat, V.E. Shevchik, W. Nasser, PehN, a polygalacturonase homologue with a low hydrolase activity, is coregulated with the other *Erwinia chrysanthemi* polygalacturonases, J Bacteriol. 184(10) (2002) 2664–2673. doi: 10.1128/JB.184.10.2664-2673.2002.

|24] W. Nasser, V.E. Shevchik, N. Hugouvieux-Cotte-Pattat, Analysis of three clustered polygalacturonase genes in *Erwinia chrysanthemi* 3937 revealed an anti-repressor function for the PecS regulator, Mol Microbiol. 34(4) (1999) 641-650. doi:10.1046/j.1365-2958.1999.01609.x.

[25] S. Reverchon, W. Nasser, J. Robert‐Baudouy, Characterization of *kdgR* , a gene of *Erwinia chrysanthemi* that regulates pectin degradation, Mol Microbiol. 5 (1991) 2203-2216. doi:10.1111/j.1365-2958.1991.tb02150.x.

[26] W. Nasser, S. Reverchon, J. Robert-Baudouy, Purification and functional characterization of the KdgR protein, a major repressor of pectinolysis genes of *Erwinia chrysanthemi*, Mol Microbiol. 6(2) (1992) 257-265. doi: 10.1111/j.1365-2958.1992.tb02007.x.

[27] W. Nasser, S. Reverchon, G. Condemine, J. Robert-Baudouy. Specific interactions of *Erwinia chrysanthemi* KdgR repressor with different operators of genes involved in pectinolysis, J Mol Biol. 236(2) (1994) 427-440. doi: 10.1006/jmbi.1994.1155.

[28] W.D. Kepseu, J.A. Sepulchre, S. Reverchon, W. Nasser, Toward a quantitative modeling of the synthesis of the pectate lyases, essential virulence factors in *Dickeya dadantii*, J Biol Chem. 285(37) (2010) 28565-28576. doi: 10.1074/jbc.M110.114710.

|29] A. Weissbach, J. Hurwitz, The formation of 2-keto-3-deoxyheptonic acid in extracts of *Escherichia coli* B. I. Identification, J Biol Chem. 234(4) (1959) 705–709. PMID: 13654246.

[30] K. Kersters, J. De Ley, 2-Keto-3-deoxy-D-gluconate, Methods Enzymol. 41 (1975) 99–101. doi.org/10.1016/S0076-6879(75)41024-2.

[31] A. Kotoujansky, M. Lemattre, P. Boistard, Utilization of a thermosensitive episome bearing transposon TN10 to Isolate Hfr Donor Strains of *Erwinia Carotovora* Subsp. *Chrysanthemi*, J Bacteriol. 150(1) (1982) 122-31.

[32] G. Condemine, J. Robert-Baudouy, Analysis of an *Erwinia chrysanthemi* gene cluster involved in pectin degradation, Mol Microbiol. 5(9) (1991) 2191-2202. doi: 1, 0.1111/j.1365-2958.1991.tb02149.

[33] G. Condemine, N. Hugouvieux-Cotte-Pattat, J. Robert-Baudouy, Isolation of *Erwinia chrysanthemi kduD* mutants altered in pectin degradation, J Bacteriol. 165 (3) (1986) 937-941. doi: 10.1128/jb.165.3.937-941.1986.

[34] N. Hugouvieux-Cotte-Pattat, W. Nasser, J. Robert-Baudouy, Molecular characterization of the *Erwinia chrysanthemi kdgK* gene involved in pectin degradation, J Bacteriol. 176 (8) (1994) 2386-2392. doi: 10.1128/jb.176.8.2386-2392.1994.

[35] G.J. Stijntjes, J.M. TeKoppele, N.P.E .Vermeuleun, High-performance liquid-chromatography fluorescence assay of pyruvic-acid to determine cysteine conjugate beta-lyase activity – Application to S-1,2-dichlorovinyl-L-cysteine and S-2-benzothiazolyl-L-cysteine, Anal Biochem. 206(2) (1992) 334-343. doi: 10.1016/0003-2697(92)90375-h.

[36] A. Shrivastava, V. Gupta, Methods for the determination of limit of detection and limit of quantitation of the analytical methods, Chronicles of Young Scientists 2(1) (2011) 21. doi: 10.4103/2229-5186.79345.

[37] B. Volkmer, M. Heinemann, Condition-dependent cell volume and concentration of *Escherichia coli* to facilitate data conversion for systems biology modeling, *PLoS One* 6(7) (2011). doi: 10.1371/journal.pone.0023126.

[38] T. Dodge, M. Rashid, F. Valle, Metabolically engineered bacterial strains having enhanced 2-keto-D-gluconate accumulation, US20050106693A1 (2005).

List of Figures

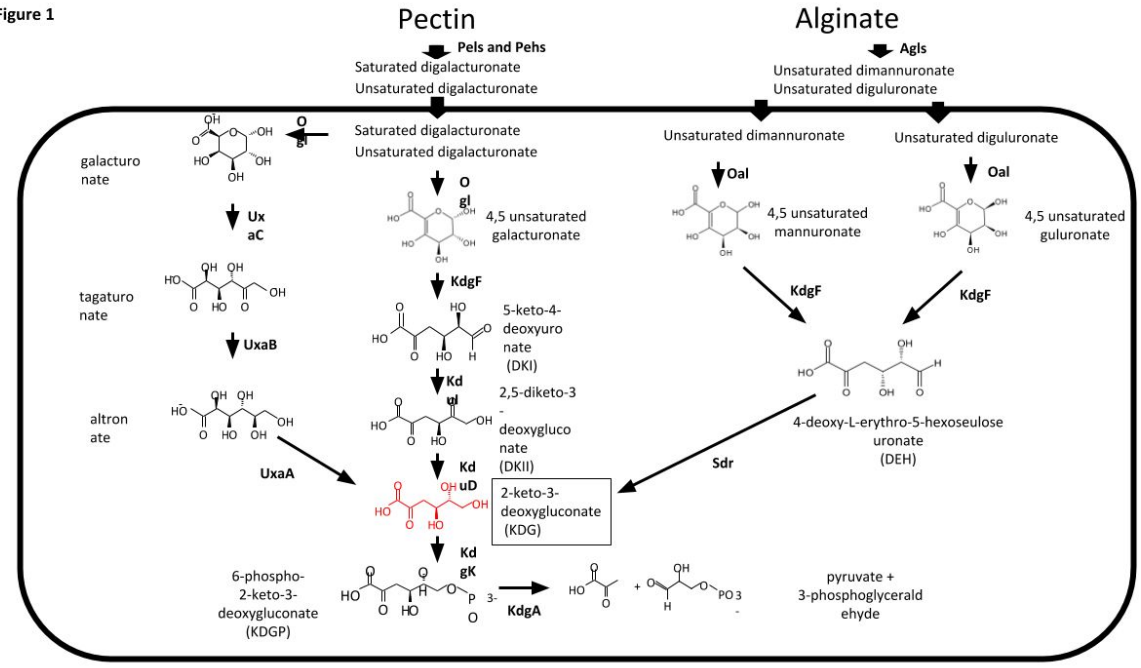


Figure 1: Schematic metabolic pathways leading to KDG (2-Keto-3-deoxy-gluconate) production from pectin and alginate degradation.

Pectin degradation involved pectate lyases and polygalacturonases that cleaved the polygalacturonate, the demethylated pectin, into unsaturated and saturated digalacturonate, respectively. Inside the bacteria, these coumpounds are converted by Ogl enzyme into galacturonate and unsaturated galacturonate (for saturated digalacturonate) and into two molecules of unsaturated galacturonate (for unsaturated digalacturonate). KdgF converts unsaturated galacturonate into DKI. KduI enzyme then converts DKI to DKII and KduD subsequently converts DKII to KDG, 2-Keto-3-deoxy-gluconate. This compound is then phosphorylated by a specific kinase KdgK and finally cleaves by KdgA to yield pyruvate and phospho-glycerate. The galacturonate is degraded to KDG by the enzymes UxaC, UxaB, UxaA in the hexuronate pathway.

Alginate degradation involved alginate lyases that cleaved the corresponding polymer into unsaturated dimannuronate and unsaturated digulunonate. Inside the bacteria, these compounds are converted by oligouronide lyase Oal into unsaturated mannuronate and unsaturated guluronate monomers. These 4,5-unsaturated monouronates are converted to DEH (4-deoxy-L-erythro-5-hexoseulose uronate) by the enzyme KdgF. DEH is then reduced into KDG by the Sdr enzyme.

Enzymes stand for: Pels, pectate lyases; Pehs, polygalacturonases; Ogl, oligogalacturonate lyase; KdgF, responsible for linearisation of 4,5-unsaturated monouronates to linear ketonized forms; KduI, 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase; KduD, 2-dehydro-3-deoxy-D-gluconate 5-dehydrogenase; KdgK, 2-dehydro-3-deoxygluconokinase; KdgA, 2-dehydro-3-deoxy-phosphogluconate aldolase; UxaC, D-galacturonate isomerase; UxaB,

Altronate oxidoreductase; UxaA, Altronate dehydratase; Agls, alginate lyases; Oal, oligouronide lyase; Sdr, 4-deoxy-L-erythro-5-hexoseulose uronate reductase.

Figure 2

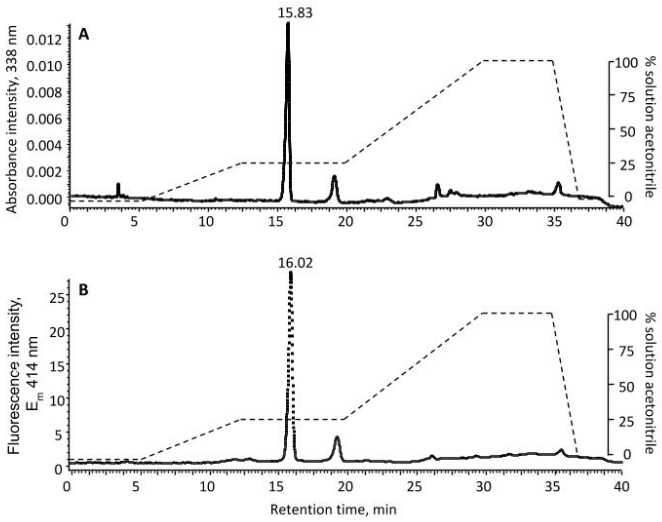


Figure 2 : Chromatogram for the detection of 1 nanomole of standard KDG-OPD molecule adduct. A) Absorbance and B) Fluorescence detection.

Figure 3

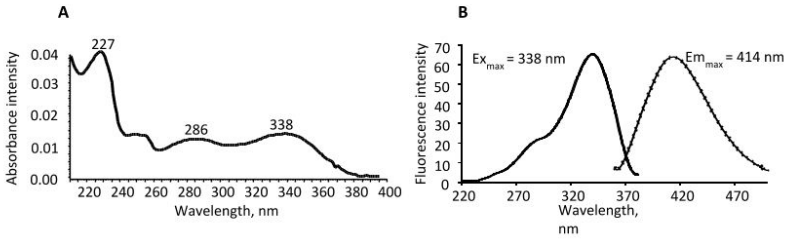


Figure 3 : A) Characteristic of the absorbance spectrum of 1 nanomole of the KDG-OPD molecule adduct derived from the peak previously isolated by HPLC. B) Excitation and emission peaks of the purified KDG-OPD molecule adduct purified above. Spectra were performed in the elution buffer using a SFM25 spectrofluorimeter (Kontron).

Figure 4

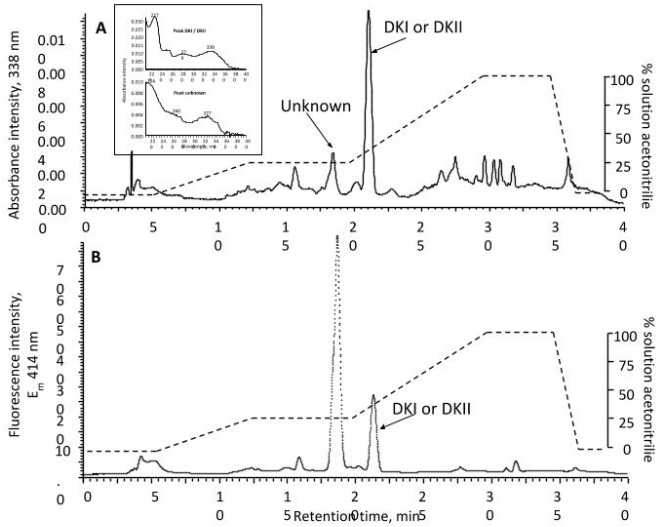


Figure 4 : Chromatogram of the partially pure DKI or DKII samples. A) Absorbance and B) Fluorescence detections. Inserts showed spectra of the DKI or DKII –OPD, and of the minor unknown metabolite. Only, the peak eluted at 21 min, based on the similarity of the structure with KDG showed the typical spectrum as reported in figure 3. It should be noticed that the unknown metabolite adduct is characterized by a greater yield in fluorescence emission upon excitation at 338 nm.

Figure 5

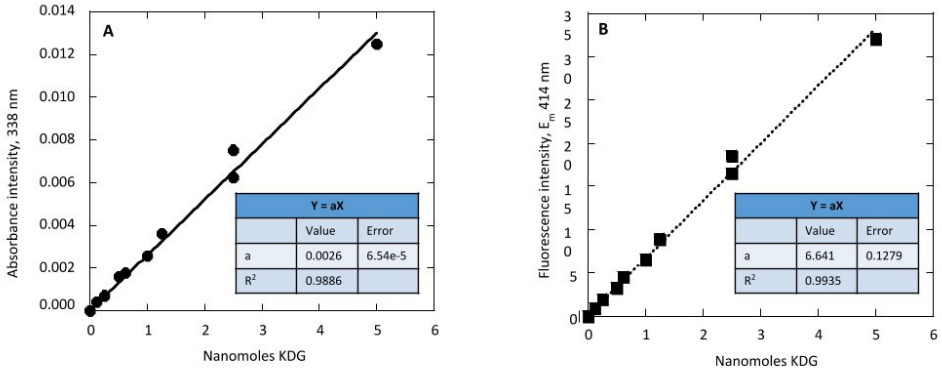


Figure 5: Standard curves for the pure KDG metabolite. A) Absorbance intensity at 338 nm and B) Fluorescence intensity at 414 nm upon excitation at 338 nm. Inserts reported the value of the linear slope (a) with R^2 being the regression coefficient. As noticed, the fluorescence intensity is more accurate than the absorbance intensity detection in our experimental set-up of the fluorimeter (2500-fold more sensitive).

Figure 6

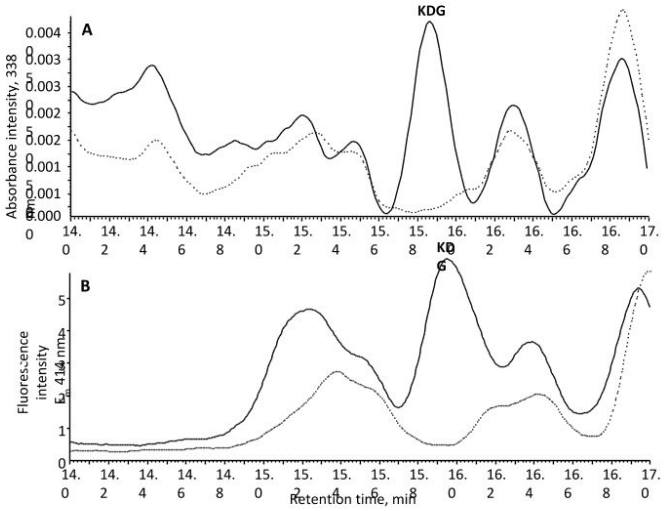


Figure 6: Chromatogram of sample from wild-type bacterial cells grown in the presence (continuous black line) and absence (gray dotted line) of pectin. The chromatogram is in the isocratic phase of the gradient to amplify the image for resolution of KDG-OPD adduct molecule. A) Absorbance intensity at 338 nm and B) Fluorescence intensity at 414 nm.

Figure 7

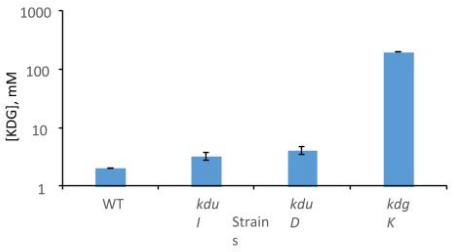


Figure 7: Logarithmic representation of KDG concentrations measured in wild type *Dickeya dadantii* and *kduI*, *kduD*, *kdgK* mutants affected from steps of the pathway of unsaturated digalacturonate (Figure 1). Error bars indicate the standard error observed in experimental quantifications over biological triplicates.

Supplementary Information

Table S1 : Pel activity in WT, *kdgK, kduI, kduD* mutants

Strain	Growth medium	Pel activity
WT	Glucose	0.3
	Glucose +PGA	7.4
kdgK	Glucose	0.5
	Glucose +PGA	37.0
kduD	Glucose	0.2
	Glucose +PGA	15.0
kduI	Glucose	0.2
	Glucose +PGA	12.0

Representative pectate lyase production of *Dickeya dadantii* wild type strain and mutant strain grown in minimal medium supplemented with glucose or glucose + polygalacturonate. Samples were taken in late stationary phase. Pel specific activity is expressed as µmol of unsaturated product liberated per min per mg of bacterial dry weight. These values were reported from literature and measures performed in this study.

Figure S1: OPD derivatization of carbonyl groups of metabolites such as KDG, DKI, DKII

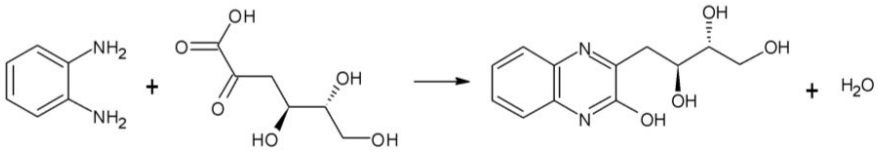


Figure S5: Chromatogram of sample from the *kdgK* mutant

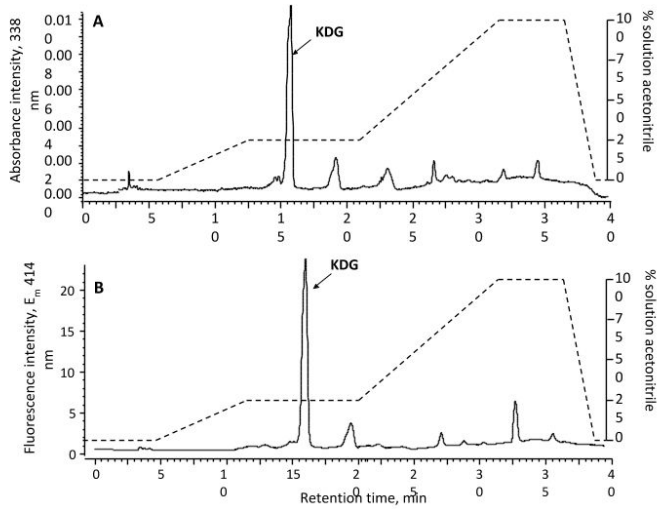


Figure S4: Chromatogram of sample from the *kduD* mutant

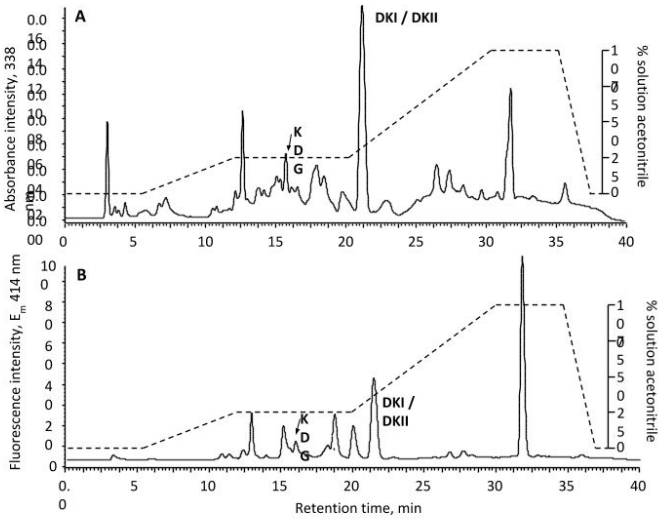


Figure S3: Chromatogram of sample from *kduI* mutant

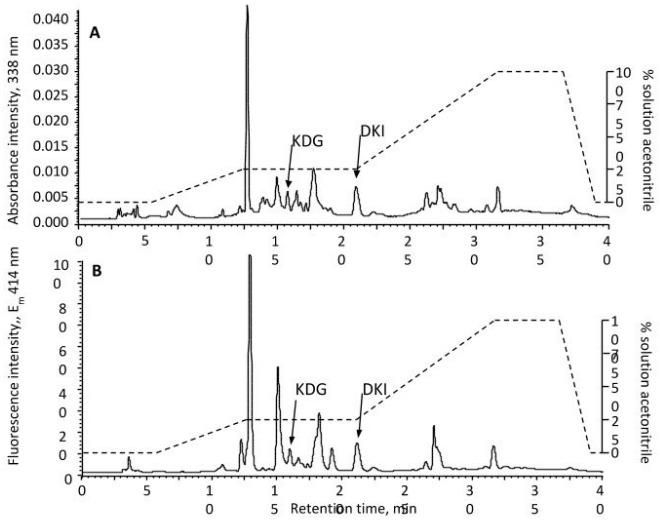
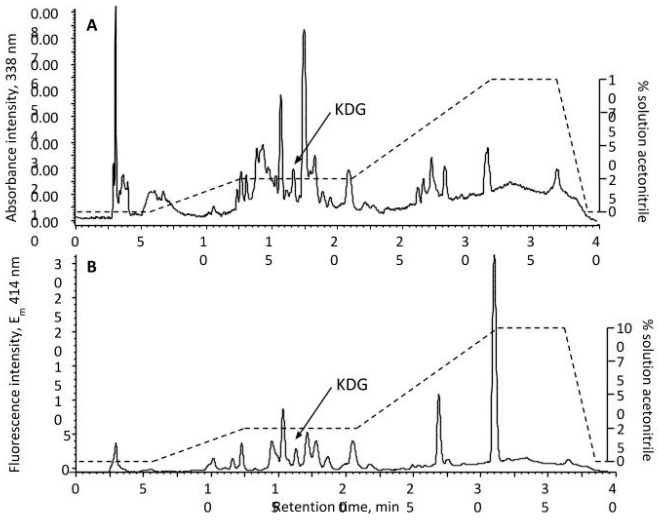


Figure S2: Chromatogram of sample from the wild type strain



Conclusion

KDG metabolite can initiate the expression of the *pel* gene by relieving the KdgR repressor but there has to be another regulator tuning the *pel* expression. Looking closely into the process of Pel synthesis, it combines metabolic and genetic regulations with respect to switch in carbon source and virulence action. In order to obtain the peak expression around the growth transition period, there had to be an activator to provide the necessary boost in the level of expression. Based on the literature, we hypothesized CRP to be the regulator responsible for the timing of *pel* expression. The presence of dual sugar sources and metabolic switch between sugar sources added up more weight to our choice. In order to incorporate the regulation by CRP we had to measure the dynamics of cAMP concentrations. So we developed a HPLC mediated method to quantify cAMP using fluorescence detection. We have developed a direct measurement technique for cAMP metabolites using the HPLC-fluorescent detection platform. The method developed is quick, easy and doesn't involve harmful exposure to toxic substances. The advantage of the method is it doesn't involve any pre-separation or tagging of the samples and can be applied on the crude sample extract. Though, there is another HPLC method reported earlier measuring the traces of cAMP from locust samples employing the 2 chloroacetaldehyde derivatization [16], in our technique of separation of cAMP, the signature chromatogram is very precise. It is defined by its specific elution time. The possibility of cross interference of similar compounds are ruled out by the symmetry and separation of the peak in the chromatogram. The distinguishable separation of the peaks is achieved by using two different buffers methanol and trifluoroacetic acid and the optimization of their gradient.

pelA,B,C,D,E, etc. are the various divergent types of the pectinolytic genes present in pectinolytic bacteria *Dickeya dadantii* among which *pelE*, *pelD* and *pelA* are known to play important roles in the virulence mechanism [2]. Several mechanisms contribute to the divergence in the gene regulation process. *pelE* and *pelD* genes are the product of a recent gene duplication event and are the most important among the group of *pels* on the account of virulence. Though gene duplication plays key roles in organismal evolution, presence of two similar genes with the same set of regulators and function is questionable. The corresponding enzymes PelD and PelE have indeed highly similar sequences (78% amino acid identity, 89% amino acid similarity) and are biochemically nearly identical. It was found that the expression of *pelE* and *pelD* is regulated by the same set of TFs, the main ones being the activator CRP and the repressors KdgR. We chose paralog genes *pelE* and *pelD* genes as our model genes because of their peculiar similarity in function and strong virulence effect. We had experimental insights to the differential expression of *pelE* and *pelD* in spite being regulated by the same set of transcription factors [17]. *pelD* is most strongly inducible by pectin but has a very low

pectin-independent expression, while *pelE*, although inducible by pectin, has a high expression level even in the absence of pectin. *pelE* acts as an initiator of pectin degradation generating the initial inducers of *pel* gene expression (DKI, DKII, KDG) required to prevent the binding of the repressor KdgR to its operators [17]. We look into the necessity of the presence of two same functional genes in the model organism *Dickeya dadantii* and the varying expression pattern in spite of having the same transcription factors is investigated using the transcriptional model with respect to repressor KdgR and activator CRP. The paper describes the model and the model results.

3.Carbon catabolite repression in pectin digestion by phytopathogen *Dickeya dadantii*

Shiny Martis B., Michel Droux, William Nasser, Sylvie Reverchon and Sam Meyer

Université de Lyon, INSA Lyon, Université Lyon 1, CNRS UMR 5240, Laboratoire de Microbiologie, Adaptation et Pathogénie, 11 avenue Jean Capelle, 69621 Villeurbanne, France

Running title: Pectin catabolic repression in *Dickeya*

Keywords: pathogenicity, pectinolysis, transcriptional regulation, quantitative modelling, catabolite repression

Introduction

Dickeya dadantii is a soft rotting Gram-negative bacterium that attacks a wide range of plant species, including many crops of economical importance. In a first asymptomatic phase of infection, *D. dadantii* colonizes the intercellular space (apoplast) where it grows on available simple sugars produced by the plant. The symptomatic phase, maceration, is characterized by the degradation of pectin present in the plant cell walls, resulting in soft rot, the main visible symptom [1]. The success of the infection process crucially depends on *D. dadantii*'s ability to switch from glucose to pectin catabolism in a timely and efficient manner, in order to overcome attacks by the plant's defense systems [2]. Thus the pectin metabolism plays an important role in the virulence action. With degradation of pectin the plant cell wall disintegrates allowing access to the bacterium causing pathogenesis.

The depolymerization of the pectin polysaccharide is achieved by a variety of pectinases. Among these, endo-pectate lyases (Pels) are known to play a prominent role and considered as major virulence factors [3], especially those encoded by the *pelD* and *pelE* paralogous genes [4]. In presence of pectin, the expression of *pel* genes is induced via a positive feedback loop: the binding of the main repressor KdgR at their promoters is relieved in presence of its cofactor KDG, itself resulting from the degradation of pectin by Pel enzymes [5]. This nonlinear mechanism has been formerly proposed to explain the strong boost in *pel* expression required for a fast switch to the maceration phase [5]. However, culture experiments showed that pectin is degraded early during bacterial growth (Kepseu et al., 2010 [5] and Fig. 1), whereas *pel* expression peaks only hours later, close to the transition to stationary phase (Fig. 1D). In this paper, we propose a new model for this regulatory loop, where the expression of pel genes is triggered not by the degradation of pectin alone, but rather by the shift in metabolic uptake from glucose to pectin. We show that this shift is controlled by a carbon catabolite repression (CCR) mechanism that presents many similarities but also interesting differences with the classical glucose/lactose example in *Escherichia coli*, in particular regarding dynamical properties required for a successful infection.

Salmonella enterica [6], *Vibrio cholerae* [7], *Helicobacter pylori* [8], *Pseudomonas aeruginosa* [9], *Escherichia coli* and *Shigella spp* [10] are known to have virulence genes assisting in their pathogenicity propagation as a response to metabolic changes. Studies with these examples have revealed the evidence that virulence and metabolism have an interconnection with respect to pathogenic action [11]. In order to understand the propagation and establishment of infection, it is necessary to unravel the regulatory link between metabolism and virulence. In this context, the *D. dadantii* model under study constitutes an attractive model to study the relationship between pathogenic growth and metabolism [12]. The proposed novel modelling integrates the localized action of transcription factors at virulence gene promoters with global metabolic pathways into a single dynamical and quantitative framework.

Although the regulation of pectin degrading *pel* genes involves the combined action of at least a dozen different transcription factors and nucleoid-associated proteins [12], the strongest of these regulators monitored in the cell are KdgR and cAMP receptor protein (CRP) [13], [14]. Since the latter is also the main catabolite activator protein controlling the glucose metabolic switch in *E. coli*, we hypothesized that the combined action of these two regulators could explain the time course of *pel* expression through variations in the concentration of their respective metabolite co-factors, KDG and cAMP. To validate this hypothesis, we developed a new highly sensitive method to measure the latter concentrations, based on a derivatization of the compounds followed by HPLC quantification, which are presented in the first sections. We then present a quantitative dynamical model of the regulation pathway. The binding affinities of both regulators with their cofactors as well as *pel* operators were formerly measured [5], [15] as well as other key parameters of the system, requiring only 5 adjustable parameters in the model. The latter explains the dynamics of Pel production during bacterial growth when both glucose and pectin are available, and the specific role of the recently diverged paralogous genes *pelE* and *pelD* in pectinolysis, explaining why this divergence could constitute an evolutionary advantage for pathogenesis.

Results

Pectin degradation and KDG concentration peak occur before Pel production boost

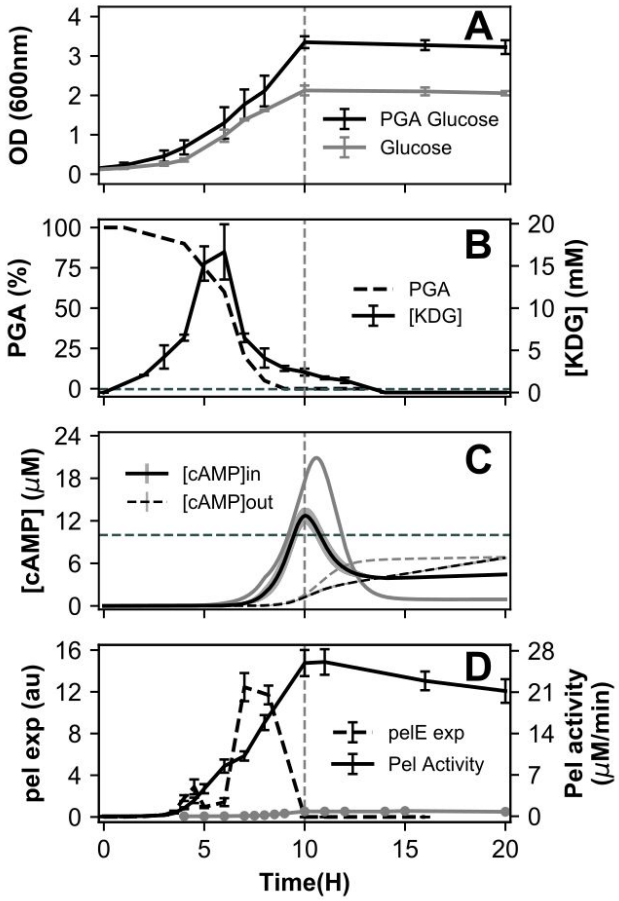


Figure 1. (A) Growth curves of *D. dadantii* in M63 minimal medium supplemented with glucose (gray) or glucose+PGA (black). (B) Time course of PGA degradation (data from Kepseu et al 2010) and intracellular KDG concentration measured by HPLC. The dashed horizontal line (0.4 mM) indicates the KDG-KdgR affinity: above this level, repression is relieved and *pel* expression is strongly induced. (C) Time course of measured extracellular cAMP concentration (dashed line) and inferred intracellular cAMP concentration (solid line, with shaded confidence area), for cells grown in glucose (gray) or glucose+PGA (black). The dashed horizontal line (10 μM) indicates the cAMP-CRP binding affinity: above this level, *pel* expression is activated. Raw data points are shown in Supplementary Figures S2 and S3. (D) *pelE* expression measured by qRT-PCR (dashed line) and total Pel enzymatic activity in glucose (gray) or glucose+PGA (black) media.

The production of pectate lyases was monitored during *D. dadantii* growth in minimal medium supplemented with glucose or glucose+polygalacturonate (PGA), a simple form

of pectin (Fig. 1). The latter acts as an inducer, boosting the production of enzymes by a factor of around 30, which was attributed to the indirect activation of *pel* genes by the metabolite KDG resulting from PGA degradation: when KDG binds the regulator KdgR, it relieves transcriptional repression by the latter, thus triggering a positive feedback loop of pectin degradation (Fig. 2). However, this scenario is challenged by the dynamics of PGA degradation (Fig. 1B, data from Kempsey), which occurs extremely rapidly after around 5h of growth, hours before most Pel enzymes are produced (close to transition). To explain this delay, a very slow import of the extracellular pectin degradation products (oligogalacturonides, UGA) and/or subsequent intracellular conversion into KDG inducer was invoked (Fig. 2, the detailed pathway of pectin degradation is shown in Supplementary Figure S1); yet this scenario seems unrealistic, since depolymerisation rather than import is thought to be the rate-limiting step of this pathway, and because such a slow kinetics would be a major obstacle for the fast activation of *pel* genes by pectin required for efficient plant infection .

Resolving this discrepancy required a direct measurement of intracellular KDG concentration. This was achieved using a new method based on the derivatization of the carbonyl group of the molecule with o-phenylenediamine and subsequent HPLC quantification. Elution with a nonlinear increasing gradient of trifluoroacetic acid and acetonitrile allows separating KDG from its close structural analogues, as described in [16]. Samples collected at regular time intervals clearly indicate that the KDG inducer peak occurs almost immediately after PGA degradation, after around 6h of growth (Fig. 1B), confirming that the import of oligosaccharides and further degradation steps into KDG are extremely fast and cannot be the limiting steps for strong *pel* expression. Most strikingly, the lifetime of KDG is also quite short in the cell, since the concentration has already dropped to half its value one hour after its peak time (the growth rate is by far insufficient to explain this decrease through a dilution effect). This profile should be compared to the expression profile of the major pectate lyase gene *pelE* measured by qRT-PCR (Fig. 1D), which peaks almost entirely after the concentration of KDG inducer has dropped. Since the cellular concentration of the KdgR repressor does not vary considerably with time [17], these observations show that the timing of *pel* expression cannot be explained by the KDG-KdgR regulatory pathway alone. The question then arises to propose an alternate or additional regulatory mechanism.

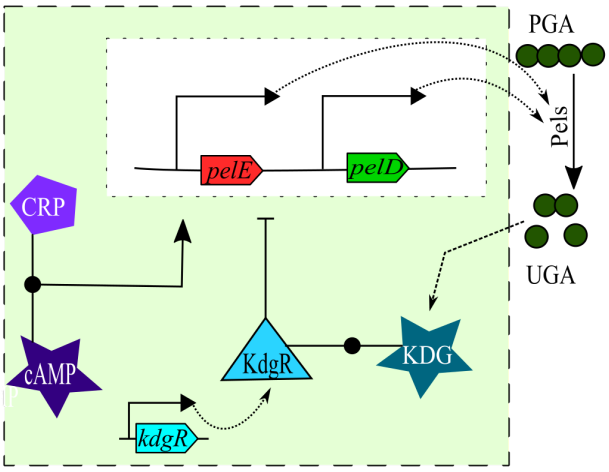


Fig 2: Diagram of the components and reactions considered in the model. The two metabolites, cAMP and KDG, indirectly activate the expression of *pel* genes by binding their associated regulators: cAMP allows CRP to bind and activate the promoters, whereas KDG relieves KdgR repression, respectively. Pel enzymes are instantaneously exported and degrade PGA (pectin) into UGA, which is imported and converted into KDG, triggering a positive feedback loop. KdgR is synthesized at a constant rate.

pel genes exhibit a particularly complex regulation, probably optimized for the specific requirements of virulence where any error in their expression time and strength results in the detection, attack, and ultimately, destruction of the pathogen by the host's defense systems [18]. This regulation involves global transcription factors (CRP) and nucleoid-associated proteins (Fis, H-NS, IHF), which act in combination with DNA supercoiling [19], as well as more classical regulators of the pectin catabolic pathway (KdgR, PecT, PecS, MfbR) [3]. Several of these regulators might contribute to relate *pel* expression to the metabolic state of the cell, such as the Fis repressor that is mostly present in early exponential phase or DNA supercoiling [19]. However, the latters' quantitative effect on *pel* expression is known to be significantly milder than that of CRP and KdgR [20], [21], which are considered as the main regulators and whose binding and regulatory properties on several *pel* promoters were therefore analysed with much detail, both *in vitro* [15], [22] and *in vivo* [13]. We took advantage of this accumulated knowledge to develop a quantitative modelling of the regulatory dynamics of the system. This started by quantifying the second main contribution of the regulatory signal, *i.e.*, the concentration of metabolite cAMP, which triggers a conformational change in CRP that allows it to bind most of its operator sites and play its activating role (Fig. 2 and Tutar et al., 2008 [23]).

cAMP measurements are carried out widely using accumulation assay kits with fluorescence or radio detection, using probes such as antibodies [24]. FRET techniques

are also employed to increase the real time precision [25], as well as various reporter gene assays [24]. However these techniques are indirect; some of them have limited sensitivity as well as specificity due to cross reactivity of the antibodies [26]. Instead, we adapted the strategy used above for KDG, in order to develop a new HPLC methodology for direct cAMP quantification, which can be carried with usual lab equipment, is less time-consuming and avoids the toxicity associated with radioactive probes, while allowing good sensitivity and precision. We used 2-chloroacetaldehyde to derivatize the amino group on the cAMP molecule, as already proposed in an earlier study [27]but with elution carried in a specifically developed buffer of trifluoroacetic acid (TFA) and methanol, where cAMP can be separated from other compounds of similar structures. It can either be detected based on UV emission resulting from derivatization, but also based on its fluorescence emission at 418 nm after excitation at the absorption wavelength 278 nm, ensuring both maximal sensitivity and specificity (see Materials and Methods for all details).

Samples were collected at regular time intervals during growth, and cAMP concentrations were quantified in the extracellular medium. Internal cAMP concentrations were then inferred from these latter by means of a kinetic model of cAMP export and import, as previously proposed ([28], see Supplementary Information). The resulting intracellular cAMP concentration profile (Fig. 1C) entirely matches previous values obtained by both direct (Makman 1965, Kao 2004) [29], [30] and indirect measurements in *E. coli* [28], with a very low level during exponential growth in glucose, a sharp peak at the transition to stationary phase, and lower but nonzero concentrations in the latter.

Combined regulation of *pel* genes through KDG and cAMP metabolites

The cAMP concentration time course was monitored during growth on glucose with or without PGA (Fig. 1C). The profiles are qualitatively similar, except that the peak is somewhat sharper when glucose is the only sugar source present, as we expected. In both conditions, the binding of a significant fraction of CRP proteins present in the cell (the affinity is indicated as a dashed horizontal line), subsequent promoter binding by the cAMP-CRP complex and Pel enzyme production boost are thus expected to occur around the transition to stationary phase, as indeed observed (Fig. 1D). In summary, while KDG acts as an inducer of Pel production (as visible in the relative amounts of enzymes produced in the two media), the timing of this production is thus rather mostly controlled by cAMP, in contrast with previous hypotheses [5]. Note that, for the concentration of PGA used in these experiments, the levels of internal KDG concentration reached in the cell are around 40 times higher than those required to bind KdgR and relieve the repression (0.4 mM, horizontal dashed line in Fig. 1B), explaining why the inducing effect can last hours after KDG concentration has started to decrease.

In other words, only a tiny fraction of Pels, produced very early, is sufficient to degrade the whole amount of pectin present in our culture medium.

We now translate these qualitative observations into a quantitative model, schematised in Fig. 2, which explicitly incorporates two pel genes with major virulence effect, *pelD* and *pelE*. To facilitate the understanding, we first describe the model without distinguishing these two genes which share the same regulators and produce enzymes of similar activity, leaving the analysis of their differential effects to the next section.

The dynamics of all components (bacterial cell density and concentrations of enzymes, regulators and metabolites inside/outside the cells) are simulated over time, according to a set of coupled differential equations (see Supplementary Information). Most parameters were obtained from experimental knowledge (Supplementary Table 2), and the 5 remaining parameters (Pel production rates and kinetic parameters controlling KDG synthesis and degradation, see Supplementary Table 3) were numerically tuned to reproduce the observations of Fig. 1.

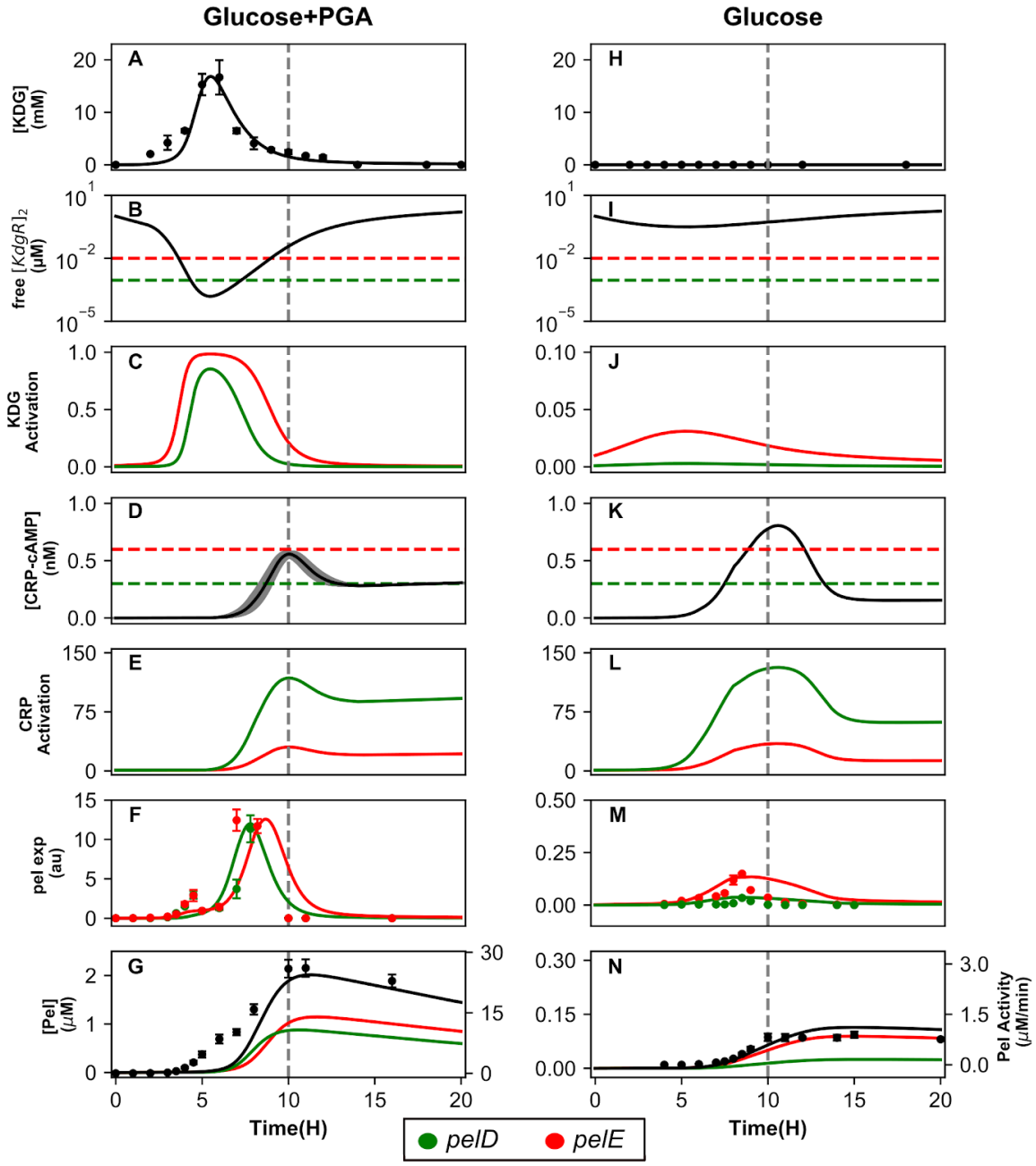


Fig 3: Modelling of *pel* regulatory pathways and expression in bacteria grown in minimal medium supplemented with glucose+PGA (left column, A-G) or glucose (right, H-N). Model results are shown as solid lines, and experimental data as dots. The transition time is shown as a dashed gray vertical line. **(A)** KDG intracellular concentration superposed with HPLC direct measurement data exhibiting a peak after around 6 hours of growth; **(B)** concentration of intracellular free KdgR dimer, with horizontal dashed lines showing the binding affinities of the dimer for the *pelE* (red) and *pelD* (green) promoters, indicating the thresholds of repression; **(C)** *pelE* and *pelD* (depicted in red

and green respectively throughout the figure) activation curves based on KdgR binding, exhibiting a de-repression in the middle of exponential phase due to high intracellular levels of KDG; **(D)** intracellular concentration of the cAMP-CRP complex inferred from HPLC measurement of cAMP in the medium (Fig. 1B); the gray area is the 95% confidence interval due to experimental variations, and dashed horizontal lines indicate the binding affinities for *pelD/E* promoters; **(E)** *pel* activation curves based on cAMP-CRP binding, with a boost occuring at transition; **(F)** *pel* expression time course based on combined regulation by KdgR and CRP, superposed with qRT-PCR data points peaking short before transition time; **(G)** extracellular concentration of Pel enzymes, either from individual genes (red and green) or the combination of both (black), superposed with measurements of enzymatic activity (dots) reflecting total Pel enzymes, including those not considered in our modelling. **(H-N)** Same legends as A-G: in absence of PGA, the dilution of KdgR during growth is sufficient to partly relieve *pelE* repression, while cAMP induces expression of both *pels* close to transition. The timing and expression levels of both genes is accurately reproduced by the model in both conditions, as well as the inducing effect of PGA.

The essential steps and results of the modelling are illustrated on Fig. 3. The KDG profile results from the degradation of PGA by Pel enzymes previously produced and exported (A). While KdgR is synthesized in the cell at a constant rate, it is mostly bound by KDG when the latter is present in sufficient quantity (over 0.4 mM), resulting in the unbinding of *pel* gene promoters (B) and relieve of KdgR repression (C) compared to cells grown with glucose only (J). The expression of *pel* genes is computed based on a thermodynamic model of transcription [31], with experimental values of KdgR/CRP affinities and activation factors (Table 1). To illustrate the effect of each regulatory signal, we compute activation profiles (Fig. 3 C, E, J, L), representing the gene expression fold-change due to repressor (<1) or activator (>1) binding, and proportional to the statistical weight of the unbound or bound state of the corresponding regulatory protein (here KdgR and CRP) respectively. Because the KdgR-KDG regulatory system involves a positive feedback loop of highly nonlinear behaviour, reproducing the measured KDG peak is the most sensitive step in the numerical tuning procedure, since it depends on several kinetic parameters of unknown value (Supplementary Table S3), whereas the regulatory part of the model entirely relies on knowledge-based parameter values.

In order to keep the modelling as simple as possible, we considered the internal cAMP concentration profile as an input information obtained by the method described above (Fig. 1C). cAMP binds CRP with a measured affinity, assuming a constant concentration of the protein in the cell, resulting in a concentration peak of the cAMP-CRP complex at the transition (Fig. 3D). In turn, this complex binds and activates *pel* gene promoters based on measured affinities and activating factors [15] (Fig. 3E).

The peak in *pel* expression thus results from the combination of the two latter regulatory signals, which are separated in time but exhibit a short overlap after around 7-8 hours of growth, in relatively good agreement with our qRT-PCR data both in presence and absence of pectin (Fig. 3 F and M). Similarly, the production level and timing of total Pel enzymes match the experimental profiles, with an approximate 30-fold induction by PGA (Figs. 3 G and N). Note that the measured Pel enzymatic activity exhibits a slow accumulation already at earlier time than in the model, and matching with weak early peaks of *pelE/D* expression after around 4h. Although this activity profile contains contributions from various *pel* genes not considered explicitly in the model and responding to similar regulatory signals [32], this discrepancy may be due to additional regulators modulating *pel* expression and disregarded here (see next section).

Quantitative differences in regulatory sequences underpin qualitatively different roles for paralogous genes

Promoter	***pelD***	***pelE***
KdgR binding affinity (nM)	0.9	10
cAMP-CRP binding affinity (nM)	0.3	0.6
cAMP-CRP activation factor	181	62
PGA induction factor (exp)	330	83
PGA induction factor (model)	324	93

Table 1: Experimental regulatory parameter values, and expression levels measured and obtained by the model after optimisation, for *pelD* and *pelE*.

In our modelling, we explicitly considered two separate *pel* genes, *pelD* and *pelE*, which play a major role in virulence as observed from pathogenicity assays [33], and were therefore studied extensively [34]. These paralogous genes result from a recent duplication event from a common ancestor present in other *Dickeya* species [4], they respond to the same regulators and encode proteins of similar enzymatic activity, yet were shown to play different roles during bacterial growth and plant infection [4]. Since the regulatory parameters are experimentally available and differ between the two promoters (Tab. 1) due to point mutations [4], we addressed the question, whether the modelling could help understanding their respective effect.

The main known difference between these genes is the level of induction by pectin. Whereas they exhibit a comparable expression level in presence of PGA, *pelE* still has a significant level of expression in absence of inducer, whereas that of *pelD* is undetectable [4], [35]. In the modelling, any observed difference in behaviour results

primarily from the respective binding affinities of the KdgR repressor, which differ by a factor of ten (1 nM for *pelD*, 10 nM for *pelE* promoter). With a cellular concentration of KdgR in the micromolar range as measured *in vivo* [5], both promoters are mostly bound by the repressor in absence of PGA in the medium (Fig. 3I); however, because of the affinity difference, unbinding events are then far more frequent at the *pelE* than *pelD* promoter, hence the relatively higher basal expression level of the former (Fig. 3J). Still in the glucose medium, at mid-exponential phase, fast growth results in KdgR dilution by cell divisions, thus weakening its repressive effect on *pelE* and favouring an expression peak before transition time (Fig. 3I). In the presence of PGA, most of KdgR is bound by KDG in the mid-exponential phase, which reduces the concentration of free KdgR below the levels required for the binding of both promoters, and fully activates their expression (Fig. 3C). The very different induction rates of the two genes thus result from their different basal activity in absence of inducer.

Activation by CRP is also different at the two promoters, but the binding affinities differ only by a factor of 2 (Table 1 and Fig. 3D), resulting in profiles with a similar shape characterized by a sharp peak at transition to stationary phase; however, this peak has a much stronger magnitude for *pelD* due to the three-fold higher activation factor of CRP compared to *pelE*.

Altogether, the combined regulation by KdgR and CRP induces both genes almost simultaneously, as observed, shortly before transition, i.e., when induction by KDG is still significant while the exhaustion of glucose has already triggered the induction by cAMP. The reader may note a small discrepancy in *pelE* expression time, which is predicted 1-2 hours too late. Such a deviation was not unexpected though, since *pel* promoters are sensitive to many regulators disregarded in our simplified model [12], as said above. In particular, we note that *pelE* is one of the *D. dadantii* promoters most strongly repressed by DNA supercoiling relaxation [19], [35] which occurs precisely at the transition to stationary phase (revue) and is thus expected to displace the *pelE* expression peak toward earlier time, as observed.

Discussion

Regulation of *pel* gene expression during plant infection

The presented model quantitatively reproduces the expression dynamics of *pel* genes and pectin degradation, in a culture medium composed of a combination of glucose and pectin. One of the most surprising features of our observations is that most Pel enzymes are produced *after* pectin has been entirely degraded, and they remain therefore unused. This behaviour might be rationalized, if we consider the differences between the growth conditions in our stable culture and in the course of plant infection, which is the main context of evolutionary selection.

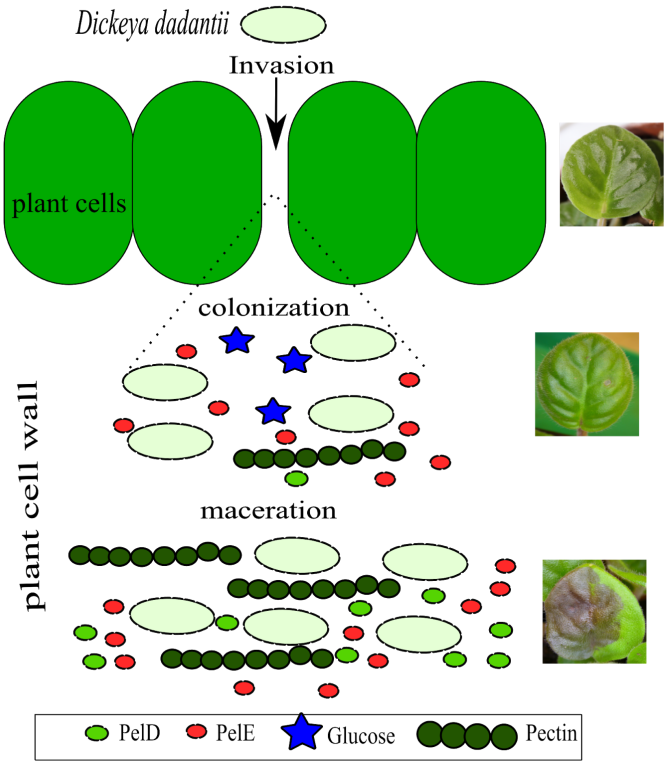


Fig. 5: Schematic depiction of the successive steps of plant infection by *D. dadantii*: invasion, colonisation, maceration (top to bottom). During colonisation (asymptomatic), the bacteria grow on glucose in the plant apoplast. Maceration is characterised by a sudden and massive production of Pels responsible for pectin degradation (visible as soft rot).

The availability of carbon sources in this process is summarised in Fig. 5. After invading the plant, e.g. through a wound, the bacterium experiences a long asymptomatic stage of colonisation of the apoplast (intercellular space) where it mostly grows on simple sugars produced by the plant. The maceration phase starts when the available glucose is exhausted, and pectin from the plant cell walls starts to be degraded, resulting in a sudden and rapid extension of soft rot in the plant tissues associated with the Pel burst. The efficiency of this drastic transition is critical for the success of the infection, since oligogalacturonates (UGA), the product of Pel enzymatic activity, are the main signal inducing the plant defense reactions. Any production of Pels in significant amounts thus triggers a survival race between the plant and the bacteria; if this event occurs before reaching sufficient bacterial density, or if the production boost is not sufficient, bacterial cells will ultimately be destroyed. Does our model give insights into the regulation events associated with this intrinsically dynamical process?

Our culture conditions mimic quite closely those encountered during the colonisation stage, with bacteria mostly growing on simple sugars, with a level of Pel production remaining sufficiently low to avoid a significant reaction from the plant. Crucially, we have shown that the transition to a high Pel production regime (and thus to the symptomatic phase) is triggered, not by an initial event of pectin degradation as suggested earlier [5], but rather by the exhaustion of glucose in the medium (and possibly of other simple sugars in the plant), resulting in a peak of cAMP and *pel* expression boost. The expression of key virulence factors is thus intrinsically related to a metabolic switch of the bacteria.

In the second stage of infection however, the conditions in the plant deviate from those of our stable culture medium. While in the latter, as said above, the low amount of available pectin was already entirely consumed way before the main peak of *pel* expression, in the plant the production of Pels induces a sustained supply of additional pectin as the soft rot propagates in the plant tissues, more similar to the conditions of a continuous bioreactor [36]. Based on our modeling, this difference has a significant regulatory consequence for *pel* genes.

Here, the activating effects of KDG and cAMP were essentially separated in time (6h and 10h, Fig. 3 C and E), with KDG activation already very weak at the time of *pel* expression (8-9h). In contrast, with a sustained supply of pectin, KdgR would remain unbound when cAMP peaks, and the simultaneous activation by the two pathways would then induce a considerably stronger expression boost, as required for an efficient infection. Additionally, the different regulatory properties of *pelE* and *pelD* imply that they will be differently affected by this boost. *pelE* always plays a crucial role as an inducer of the Pel/KDG feedback loop, because of its higher basal expression level in absence of pectin compared to *pelD* [4], whereas in our data, the two genes have a comparable maximal expression level (Fig. 3F). However, in the case of a simultaneous activation by both pathways, we expect *pelD* to be expressed at a much (approximately three-fold) higher level than *pelE*, because it is much more inducible by CRP (Fig. 3E), and thus to play the major role in the massive production of Pels during the maceration phase. In summary, in contrast to our culture conditions where *pelE* plays a predominant role, we expect the two genes to play a qualitatively distinct but equally crucial role in the plant, which is fully consistent with the observation that both *pelD* and *pelE* mutants exhibit phenotypes with strongly reduced virulence [33]. The coexistence of these two genes may thus provide a significant evolutionary advantage, possibly explaining their conservation in several *Dickeya* species (*D. dadantii*, *D. solani*, *D. zeae*, *D. dianthicola*) compared to close relatives (e.g., *D. paradisiaca*).

Carbon catabolite repression in *Dickeya dadantii*

The competitive uptake of either carbon source presents many similarities, but also interesting differences with the classical example of carbon catabolite repression (CCR) involved in the glucose/lactose switch of *E. coli* [37], [38].

The first apparent and previously identified difference is the monotonous aspect of the growth curve in the mixed medium (Fig. 1A), which contrasts with the diauxic growth of *E. coli* in glucose+lactose [37]. Even though pectin is a complex polysaccharide, the consumption of which depends on the production, export and activity of many degrading enzymes, this more continuous pattern previously suggested that pectin consumption starts while glucose is still available [5], whereas *E. coli* uses lactose only after glucose exhaustion [37]. Our KDG quantification data confirm this hypothesis, and even demonstrate that in our culture conditions, pectin is already converted into KDG after 5 hours, only half the entire duration of growth. Can we relate this behaviour to the underlying regulatory properties of the *pel* genes, and extend the comparison to those of the *lac* operon in *E. coli*?

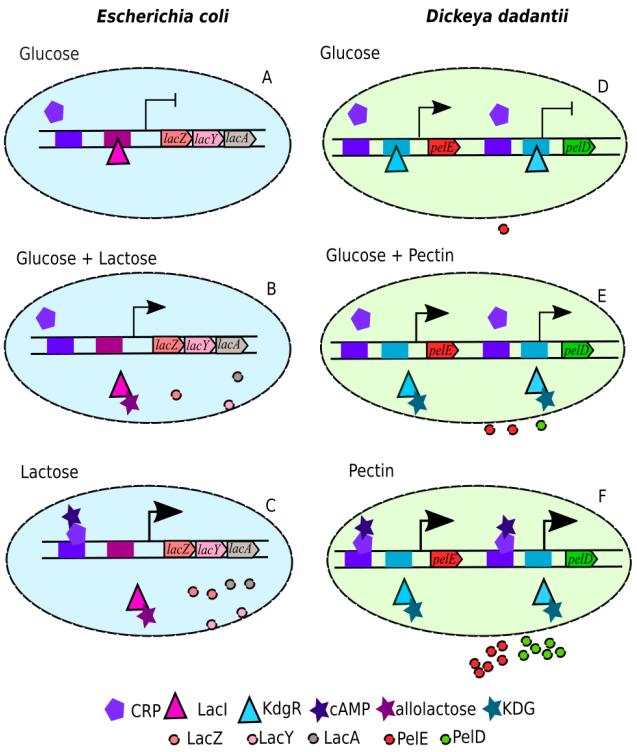


Fig. 4: Schematic depiction of the *lac* operon regulation in *E. coli* (left) and *pelED* regulation in *D. dadantii* (right), in presence of glucose, lactose/pectin, or both sources of carbon, illustrating carbon catabolite repression in both systems. Promoters are shown as arrows, with expression levels indicated by arrowhead sizes (and full

repression by ⊣). Genes are indicated along the DNA double helix, and regulator binding sites are shown as coloured boxes.

Fig. 5 gives a schematic depiction of the regulatory systems of these operons controlling the catabolism of the secondary sugar sources in both species (lactose and pectin, respectively). These systems present a conspicuous similarity, with an activation by cAMP-CRP and repression by a pathway-specific regulator, LacI and KdgR respectively, which can be driven away by the degradation product of the sugar, allolactose or KDG respectively. On the other hand, the modes of action of the enzymes involved in the two catabolic pathways are very different, since pectin is a polysaccharide degraded outside the cells by Pels, whereas lactose is a disaccharide and is imported when LacY is present at the cell membrane [39].

In presence of glucose alone (top), the activator CRP is unbound, but the repressor inhibits the expression of *pel*/*lac* genes. In *D. dadantii* however, the presence of a promoter with low repressor affinity (*pelE*) results in a low basal expression and a less drastic repressive effect. When both sources of carbon are present simultaneously (middle), a moderate level of expression can be sustained by the release of the repressor. When only the secondary source is present (bottom), the expression is boosted by the binding of cAMP-CRP.

As visible in Fig. 4, among the two *pel* genes considered in this study, *pelD* is the one with a regulatory system most comparable to the *lac* promoter, exhibiting a strict dependence on the availability of the secondary carbon source. In this figure showing a stationary behaviour in each medium, the difference between the two *pel* genes appears quite limited however, and their coexistence would probably not constitute a key advantage if the bacterium lived in a stable or slowly changing environment, as non-pathogenic *E. coli* does most of the time. On the other hand, when both sources of carbon are present (middle panel), the depicted partial expression requires an initial expression event, in order to start producing allolactose/KDG and sustain subsequent transcription. In *E. coli*, because of the absence of a significant basal level of *lac* expression, this initiation step may be very long, partly explaining why glucose is entirely consumed first (diauxic growth). In *D. dadantii*, the basal expression of *pelE* drastically reduces this initiation time, facilitating the switch in carbon source and resulting in continuous growth. Interestingly, measurements in a *pelE* mutant strain indeed exhibited a severe delay in the expression of *pel* genes, due to a much longer initiation step of the pectin degradation feedback loop [4]. As discussed in the previous section, the evolutionary advantage of having coexisting *pelD* and *pelE* genes is thus intimately related to the dynamical requirements of *pel* expression in the context of pathogenesis. This expression is part of a race with the host's defense systems where time plays a prominent role compared to the requirements of non-pathogenic bacteria which are rather selected for resource optimisation [12].

Materials and Methods

Bacterial Strain and Culture Conditions

Wild type strain of *D. dadantii* 3937 was used for all of the experiments described. The cultures were grown at 30 °C in M63 minimal salt medium supplemented with 0.2% (w/v) glucose as primary carbon source and 4% (w/v) PGA for pectin degradation studies. Liquid cultures were grown in a shaking incubator (125 r.p.m.).

Gene activity assays

pelE and *pelD* expression were measured by qRT-PCR as described in [5]. Pectate lyase enzymatic activity assays were performed on tokenized cell extracts in 1 mL cell culture harvested at predetermined time points along the growth of the bacteria. The degradation of PGA is monitored by absorption spectrometry of unsaturated oligogalacturonides at 230nm [40]. Pel activity is expressed as µmol of unsaturated products liberated per minute and per ml of enzymatic extract, as described in Kepseu et al [5]. All experiments were carried with two biological replicates.

Bacterial samples preparation for KDG and cAMP quantification

The bacterial culture samples were harvested at predetermined time points after measuring the optical density (OD), centrifuged at 8000 rpm for 6 minutes to separate cells from supernatant medium. Bacterial pellet was freeze dried with liquid nitrogen for storing. Cells and medium samples were stored separately at -80°C. In order to process for quantification the cell pellet was resuspended in 10 ml cold solution of 2.5 % (v/v) TFA, 50% (v/v) acetonitrile. The dissolved lysed pellet was maintained at -80°C for 15 minutes and thawed at room temperature. The solution was then centrifuged at 20.000*g* for 5 minutes. The clear supernatant containing the metabolites (10 ml) was separated from the pellet and then lyophilized overnight. The dry powder was dissolved in 1 ml of 0.1 M HCl and centrifuged again to remove particles. Culture medium is directly made up to .1 M HCl with 1 M HCl. The measurements were carried in two biological replicates in glucose+PGA medium (and one in the glucose control medium).

KDG quantification using HPLC

KDG metabolite is quantified using a novel HPLC technique developed inhouse as described in [16]. The carbonyl group is derivatized with o-phenylenediamine and eluted with buffers consisting of 0.005% trifluoroacetic acid and 60% acetonitrile.

cAMP quantification

Bacterial cell and suspension medium samples extracts are suspended in 0.1 M HCl. Equal amount of sample in 0.1 M HCl and 1 M sodium acetate is mixed and 10% v/v of 2-chloroacetaldehyde solution is added to the mixture and incubated at 80°C for 30 min with continuous shaking (Fig. S4A and sanoki 1989 derivatization of adenine with CAA). The appropriate amounts of incubated sample are injected onto an Uptisphere C18 column (Interchim) for separation of the purine derivative adducts using the following procedure: Buffer A was composed of 0.005 % trifluoroacetic acid (TFA) in water and

Buffer B of 100% methanol. The flow rate was set-up at 35 °C, and run at 1 ml/min using the following gradient: initial 90% A, 10% B; linear gradient from 0 to 20 min up to 70% A, 30% B; 22 min 100% B; 25 min 25% B then re-equilibration to initial condition from 28 to 30 min (see the profile in Fig. 6). The mean retention time for cAMP was between 13.04 to 13.10 min. The associated peak is well-separated from neighbor peaks characteristic of other compounds of similar structure (Fig. 6B), and can be further distinguished by its characteristic absorption/emission spectrum identical to that of the purified molecule (Fig. S4B). The quantification is performed from a standard curve (Supplementary Figure S4C). The model used to infer intracellular cAMP concentrations is described in Supplementary Information.

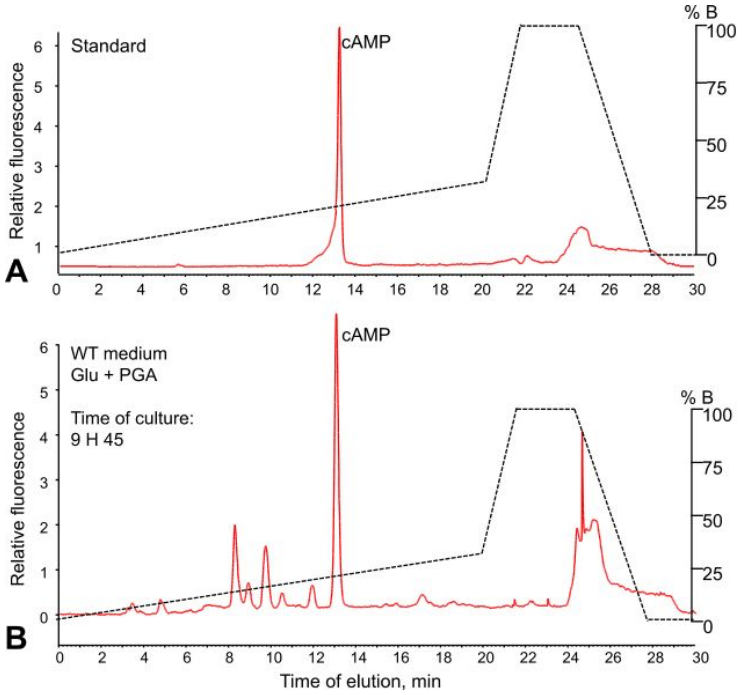


Figure 6: Elution profile for cAMP quantification by fluorescence detection at 418 nm after specific excitation at 278 nm. **(A)** Purified cAMP sample (XX nanomoles). **(B)** Sample from bacteria grown on M63 medium supplemented with glucose and PGA. The black dashed line indicates the proportion of buffer B (right axis).

Model

The system dynamics is simulated with a set of deterministic ordinary differential equations relating the concentrations of all species (see Supplementary Information). Bacterial growth follows a logistic equation where the maximal growth rate and final density depend linearly on the available nutrients (sucrose and/or PGA) [5]. All binding reactions are described as thermodynamic equilibria, in particular those between the metabolites (KDG and cAMP) and the regulatory proteins (KdgR and CRP dimers, respectively), and between the latter and the two *pel* promoters. Promoter activities follow a thermodynamic model of transcription (Bintu) involving the activator cAMP-CRP and repressor KdgR (Table 1). All enzymatic reactions are treated with

Michaelis-Menten kinetics. The list of model parameters is provided in Supplementary Table S2. All thermodynamic and regulatory parameters were obtained experimentally [5], [15], [21], [41] as well as most kinetic parameters. The 5 remaining parameters (Supplementary Table S3) were numerically estimated by fitting a set of observed quantities (*pel* expression peak time and ratio in presence and absence of PGA, KDG concentration peak time and magnitude) using the truncated Newton method. The dynamical system is simulated with the Euler algorithm, using a constant timestep of 0.3 min ensuring numerical stability. All data analyses and computations were carried in Python using NumPy, Scipy and Pandas libraries.

References

[1] N. Hugouvieux-Cotte-Pattat, G. Condemine, E. Gueguen, and V. E. Shevchik, "Dickeya Plant Pathogens," in *eLS*, American Cancer Society, 2020, pp. 1–10.

[2] S. Hassan, V. E. Shevchik, X. Robert, and N. Hugouvieux-Cotte-Pattat, "PelN Is a New Pectate Lyase of Dickeya dadantii with Unusual Characteristics," *J. Bacteriol.*, vol. 195, no. 10, pp. 2197–2206, May 2013, doi: 10.1128/JB.02118-12.

[3] N. Hugouvieux-Cotte-Pattat, G. Condemine, W. Nasser, and S. Reverchon, "Regulation of pectinolysis in erwinia chrysanthemi," *Annu. Rev. Microbiol.*, vol. 50, no. 1, pp. 213–257, Oct. 1996, doi: 10.1146/annurev.micro.50.1.213.

[4] A. Duprey, W. Nasser, S. Léonard, C. Brochier-Armanet, and S. Reverchon, "Transcriptional start site turnover in the evolution of bacterial paralogous genes - the pelE-pelD virulence genes in Dickeya," *FEBS J.*, vol. 283, no. 22, pp. 4192–4207, 2016, doi: 10.1111/febs.13921.

[5] W. D. Kepseu, J.-A. Sepulchre, S. Reverchon, and W. Nasser, "Toward a quantitative modeling of the synthesis of the pectate lyases, essential virulence factors in Dickeya dadantii," *J. Biol. Chem.*, vol. 285, no. 37, pp. 28565–28576, Sep. 2010, doi: 10.1074/jbc.M110.114710.

[6] S. E. Winter *et al.*, "Gut inflammation provides a respiratory electron acceptor for Salmonella," *Nature*, vol. 467, no. 7314, pp. 426–429, Sep. 2010, doi: 10.1038/nature09415.

[7] S. Almagro-Moreno and E. F. Boyd, "Sialic Acid Catabolism Confers a Competitive Advantage to Pathogenic Vibrio cholerae in the Mouse Intestine," *Infect. Immun.*, vol. 77, no. 9, pp. 3807–3816, Sep. 2009, doi: 10.1128/IAI.00279-09.

[8] "H. pylori virulence factors: influence on immune system and pathology. - Abstract - Europe PMC." https://europepmc.org/article/PMC/3918698 (accessed Sep. 10, 2020).

[9] D. A. D'Argenio *et al.*, "Growth phenotypes of Pseudomonas aeruginosa lasR mutants adapted to the airways of cystic fibrosis patients," *Mol. Microbiol.*, vol. 64, no. 2, pp. 512–533, Apr. 2007, doi: 10.1111/j.1365-2958.2007.05678.x.

[10] M. Touchon *et al.*, "Organised Genome Dynamics in the Escherichia coli Species Results in Highly Diverse Adaptive Paths," *PLoS Genet.*, vol. 5, no. 1, Jan. 2009, doi: 10.1371/journal.pgen.1000344.

[11] L. Rohmer, D. Hocquet, and S. I. Miller, "Are pathogenic bacteria just looking for food? Metabolism and microbial pathogenesis," *Trends Microbiol.*, vol. 19, no. 7, pp. 341–348, Jul. 2011, doi: 10.1016/j.tim.2011.04.003.

[12] S. Reverchon and W. Nasser, "Dickeya ecology, environment sensing and regulation of virulence programme," *Environ. Microbiol. Rep.*, vol. 5, no. 5, pp. 622–636, 2013, doi: 10.1111/1758-2229.12073.

[13] S. Reverchon, D. Expert, J. Robert-Baudouy, and W. Nasser, "The cyclic AMP

receptor protein is the main activator of pectinolysis genes in Erwinia chrysanthemi.," *J. Bacteriol.*, vol. 179, no. 11, pp. 3500–3508, Jun. 1997, doi: 10.1128/jb.179.11.3500-3508.1997.
[14] S. Reverchon, W. Nasser, and J. Robert-Baudouy, "Characterization of kdgR, a gene of Erwinia chrysanthemi that regulates pectin degradation," *Mol. Microbiol.*, vol. 5, no. 9, pp. 2203–2216, Sep. 1991, doi: 10.1111/j.1365-2958.1991.tb02150.x.
[15] W. Nasser, J. Robert-Baudouy, and S. Reverchon, "Antagonistic effect of CRP and KdgR in the transcription control of the Erwinia chrysanthemi pectinolysis genes," *Mol. Microbiol.*, vol. 26, no. 5, pp. 1071–1082, 1997, doi: 10.1046/j.1365-2958.1997.6472020.x.
[16] S. M. B, M. Droux, F. Deboudard, W. Nasser, S. Meyer, and S. Reverchon, "Development and validation of an HPLC method to quantify 2-Keto-3-deoxy-gluconate (KDG) a major metabolite in pectin and alginate degradation pathways," *bioRxiv*, p. 2020.07.24.220400, Jul. 2020, doi: 10.1101/2020.07.24.220400.
[17] W. Nasser, S. Reverchon, G. Condemine, and J. Robert-Baudouy, "Specific interactions of Erwinia chrysanthemi KdgR repressor with different operators of genes involved in pectinolysis," *J. Mol. Biol.*, vol. 236, no. 2, pp. 427–440, Feb. 1994, doi: 10.1006/jmbi.1994.1155.
[18] S. Reverchon, G. Muskhelisvili, and W. Nasser, "Chapter Three - Virulence Program of a Bacterial Plant Pathogen: The Dickeya Model," in *Progress in Molecular Biology and Translational Science*, vol. 142, M. San Francisco and B. San Francisco, Eds. Academic Press, 2016, pp. 51–92.
[19] Z.-A. Ouafa, S. Reverchon, T. Lautier, G. Muskhelishvili, and W. Nasser, "The nucleoid-associated proteins H-NS and FIS modulate the DNA supercoiling response of the pel genes, the major virulence factors in the plant pathogen bacterium Dickeya dadantii," *Nucleic Acids Res.*, vol. 40, no. 10, pp. 4306–4319, May 2012, doi: 10.1093/nar/gks014.
[20] T. Lautier and W. Nasser, "The DNA nucleoid-associated protein Fis co-ordinates the expression of the main virulence genes in the phytopathogenic bacterium Erwinia chrysanthemi," *Mol. Microbiol.*, vol. 66, no. 6, pp. 1474–1490, 2007, doi: 10.1111/j.1365-2958.2007.06012.x.
[21] T. Lautier, N. Blot, G. Muskhelishvili, and W. Nasser, "Integration of two essential virulence modulating signals at the Erwinia chrysanthemi pel gene promoters: a role for Fis in the growth-phase regulation," *Mol. Microbiol.*, vol. 66, no. 6, pp. 1491–1505, 2007, doi: 10.1111/j.1365-2958.2007.06010.x.
[22] T. Lautier, N. Blot, G. Muskhelishvili, and W. Nasser, "Integration of two essential virulence modulating signals at the Erwinia chrysanthemi pel gene promoters: a role for Fis in the growth-phase regulation," *Mol. Microbiol.*, vol. 66, no. 6, pp. 1491–1505, 2007, doi: 10.1111/j.1365-2958.2007.06010.x.
[23] Y. Tutar, "Syn, anti, and finally both conformations of cyclic AMP are involved in the CRP-dependent transcription initiation mechanism in E. coli lac operon," *Cell Biochem. Funct.*, vol. 26, no. 4, pp. 399–405, 2008, doi: 10.1002/cbf.1462.
[24] C. Williams, "cAMP detection methods in HTS: selecting the best from the rest," *Nat. Rev. Drug Discov.*, vol. 3, no. 2, Art. no. 2, Feb. 2004, doi: 10.1038/nrd1306.
[25] S. Börner *et al.*, "FRET measurements of intracellular cAMP concentrations and cAMP analog permeability in intact cells," *Nat. Protoc.*, vol. 6, no. 4, pp. 427–438, Apr. 2011, doi: 10.1038/nprot.2010.198.
[26] K. Werner, F. Schwede, H.-G. Genieser, J. Geiger, and E. Butt, "Quantification of cAMP and cGMP analogs in intact cells: pitfalls in enzyme immunoassays for cyclic nucleotides," *Naunyn. Schmiedebergs Arch. Pharmacol.*, vol. 384, no. 2, pp. 169–176, Aug. 2011, doi: 10.1007/s00210-011-0662-6.
[27] L. Zhang, H. Zhang, J. Wang, and C. Pan, "Determination of Trace Level of

cAMP in Locusta Migratoria Manilensis Meyen by HPLC with Fluorescence Derivation," *Int. J. Mol. Sci.*, vol. 7, no. 8, Art. no. 8, Aug. 2006, doi: 10.3390/i7080266.
[28] S. Berthoumieux *et al.*, "Shared control of gene expression in bacteria by transcription factors and global physiology of the cell," *Mol. Syst. Biol.*, vol. 9, no. 1, p. 634, Jan. 2013, doi: 10.1038/msb.2012.70.
[29] R. S. Makman and E. W. Sutherland, "ADENOSINE 3',5'-PHOSPHATE IN ESCHERICHIA COLI," *J. Biol. Chem.*, vol. 240, pp. 1309–1314, Mar. 1965.
[30] K. C. Kao, Y.-L. Yang, R. Boscolo, C. Sabatti, V. Roychowdhury, and J. C. Liao, "Transcriptome-based determination of multiple transcription regulator activities in Escherichia coli by using network component analysis," *Proc. Natl. Acad. Sci.*, vol. 101, no. 2, pp. 641–646, Jan. 2004, doi: 10.1073/pnas.0305287101.
[31] L. Bintu *et al.*, "Transcriptional regulation by the numbers: models," *Curr. Opin. Genet. Dev.*, vol. 15, no. 2, pp. 116–124, Apr. 2005, doi: 10.1016/j.gde.2005.02.007.
[32] F. Tardy, W. Nasser, J. Robert-Baudouy, and N. Hugouvieux-Cotte-Pattat, "Comparative analysis of the five major Erwinia chrysanthemi pectate lyases: enzyme characteristics and potential inhibitors," *J. Bacteriol.*, vol. 179, no. 8, pp. 2503–2511, Apr. 1997, doi: 10.1128/jb.179.8.2503-2511.1997.
[33] M. Boccara, A. Diolez, M. Rouve, and A. Kotoujansky, "The role of individual pectate lyases of Erwinia chrysanthemi strain 3937 in pathogenicity on saintpaulia plants," *Physiol. Mol. Plant Pathol.*, vol. 33, no. 1, pp. 95–104, Jul. 1988, doi: 10.1016/0885-5765(88)90046-X.
[34] A. Duprey, G. Muskhelishvili, S. Reverchon, and W. Nasser, "Temporal control of Dickeya dadantii main virulence gene expression by growth phase-dependent alteration of regulatory nucleoprotein complexes," *Biochim. Biophys. Acta*, vol. 1859, no. 11, pp. 1470–1480, 2016, doi: 10.1016/j.bbagrm.2016.08.001.
[35] B. El Houdaigui *et al.*, "Bacterial genome architecture shapes global transcriptional regulation by DNA supercoiling," *Nucleic Acids Res.*, vol. 47, no. 11, pp. 5648–5657, Jun. 2019, doi: 10.1093/nar/gkz300.
[36] J.-A. Sepulchre, S. Reverchon, J.-L. Gouzé, and W. Nasser, "Modeling the bioconversion of polysaccharides in a continuous reactor: A case study of the production of oligogalacturonates by Dickeya dadantii," *J. Biol. Chem.*, vol. 294, no. 5, pp. 1753–1762, 01 2019, doi: 10.1074/jbc.RA118.004615.
[37] J. Monod, "The Growth of Bacterial Cultures," p. 25.
[38] B. Magasanik, "Catabolite Repression," *Cold Spring Harb. Symp. Quant. Biol.*, vol. 26, pp. 249–256, Jan. 1961, doi: 10.1101/SQB.1961.026.01.031.
[39] B. Müller-Hill, *The Lac Operon: A Short History of a Genetic Paradigm*. Walter de Gruyter, 1996.
[40] F. Moran, S. Nasuno, and M. P. Starr, "Extracellular and intracellular polygalacturonic acid trans-eliminases of Erwinia carotovora," *Arch. Biochem. Biophys.*, vol. 123, no. 2, pp. 298–306, Feb. 1968, doi: 10.1016/0003-9861(68)90138-0.
[41] B. Anderson, B. Schneider, L. Perlman, and I. Pastan, "Purification of and Properties of the Cyclic Adenosine 3',5'- Monophosphate Receptor Protein which Mediates Cyclic Adenosine 3', S-Monophosphate-dependent Gene Transcription in Escherichia coli," vol. 246, no. 19, p. 10, 1971.

Supplementary Information

1. Transcriptional regulatory model for pel expression

The transcriptional regulatory model for *pel* expression and PGA degradation is defined by the set of following ordinary differential equations, based on a previous model of the KDG/KdgR pathway (Kepseu *et al*, 2010):

$$\frac{d\rho}{dt} = \sigma * \rho * \left(1 - \frac{\rho}{\rho_s}\right) \quad (1)$$

$$\frac{dy}{dt} = \beta_2 - \left(\alpha_2 + \frac{\rho^o}{\rho}\right) * y \quad (2)$$

$$\frac{dX_E}{dt} = \beta_{1E} * \frac{\rho}{1-\rho} * \frac{K_{d6E}}{K_{d6E}+y_d} * \frac{K_{aE}+CAR*f_E}{K_{aE}+CAR} - \left(\alpha_1 * \frac{K_m}{K_m+s} * X_E\right) \quad (3)$$

$$\frac{dX_D}{dt} = \beta_{1D} * \frac{\rho}{1-\rho} * \frac{K_{d6D}}{K_{d6D}+y_d} * \frac{K_{aD}+CAR*f_D}{K_{aD}+CAR} - \left(\alpha_1 * \frac{K_m}{K_m+s} * X_D\right) \quad (4)$$

$$\frac{ds}{dt} = -\epsilon * k_{cat} * (X_E + X_D) * \frac{s}{s+K_m} \quad (5)$$

$$\frac{dz}{dt} = 2\epsilon * k_{cat} * (X_E + X_D) * \frac{s}{s+K_m} - \gamma * \rho * z \quad (6)$$

$$\frac{dw}{dt} = \gamma * (1-\rho) * z - k_5 * \left(\frac{w}{w+k_w}\right) + \frac{\rho^o}{\rho} * w \quad (7)$$

$$y = 2y_d * \left(1 + \frac{w^2}{k_{d3}^2}\right) \quad (8)$$

$$[CAP] = \frac{[cAMP]}{[cAMP]+K_{cAMP}} * \frac{[CRP]}{2} \quad (9)$$

Supplementary Table 1 : List of variables in the dynamical system

Variables	Description
ρ	Bacterial volume fraction
y	Total intracellular concentration of KdgR monomer
X_E	Concentration of extracellular PelE enzyme
X_D	Concentration of extracellular PelD enzyme
s	Concentration of extracellular PGA substrate
z	Concentration of extracellular UGA (end product of PGA degradation)
w	Concentration of KDG
y_d	Concentration of free KdgR dimer
[CAP]	Concentration of CRP.cAMP complex, inferred from the measured values of cAMP concentration (Fig. 1C)

Note that each KdgR monomer is able to bind a KDG molecule (Kepseu *et al*, 2010), whereas we consider only the dominant form of the cAMP-CRP complex involving a CRP dimer and one molecule of cAMP (Eq. 9).

Supplementary Table 2 : Parameters of the model

Parameter	Description	Value	Unit	Source
α_1	Degradation rate of PelE protein	0.0494	h^{-1}	Kepseu, 2010
α_2	Degradation rate of KdgR protein	0.0347	h^{-1}	Kepseu, 2010
β_2	Maximum rate of synthesis of KdgR	0.3	$\mu M.h^{-1}$	Kepseu, 2010
ε	Ratio of the enzymatic rate constant of Pel reactions in the culture medium and buffer	1/3		Kepseu, 2010
K_{d6E}	Equilibrium dissociation constant of the KdgR dimer at the *pelE* promoter	0.01	μM	Nasser 1997, Lautier 2007

K_{d6D}	Equilibrium dissociation constant of the KdgR dimer at the *pelD* promoter	0.0009	µM	Nasser, 1997
K_{aE}	Equilibrium dissociation constant of the cAMP-CRP complex at the *pelE* promoter	0.0006	µM	Nasser, 1997
K_{aD}	Equilibrium dissociation constant of the cAMP-CRP complex at the *pelD* promoter	0.0003	µM	Nasser, 1997
K_m	Michaelis-Menten constant for Pel.PGA complex	6800	µM	Kepseu, 2010
k_{cat}	Rate constant of the enzymatic Pel reactions in purified conditions	3600000	h^{-1}	Kepseu, 2010
[CRP]	Effective intracellular concentration of CRP monomer	0.005	µM	arb
K_{d3}	Equilibrium dissociation constant of the KdgR-KDG complex	400	µM	Nasser, 1994
f_E	*pelE* activation factor by cAMP-CRP complex	60		Nasser 1997
f_D	*pelD* activation factor by cAMP-CRP complex	180		Nasser 1997
K_{cAMP}	Equilibrium dissociation constant of cAMP and CRP dimer	10	µM	Anderson 71

Supplementary Table 3 : Unknown parameters subjected to optimization

Parameter	Description	Value	Unit
γ	Cell membrane import rate	700	h^{-1}
β_{1E}	Maximum rate of synthesis of PelE	20.0	$\mu M h^{-1}$
β_{1D}	Maximum rate of synthesis of PelD	15.0	$\mu M h^{-1}$
k_5	Phosphorylation rate of KDG	3.0	h^{-1}
k_w	Phosphorylation constant of KDG	5000	µM

Supplementary Table 4: growth parameters

Parameter	Description
ρ^o	Initial bacterial volume fraction
ρ_s	Maximum bacterial volume fraction
s_0	Initial substrate concentration
σ	Bacterial growth rate

In this study, the medium contains a fixed quantity of glucose, and the effect of varying PGA is studied. The dependence of parameters on bacterial growth on the initial amount of PGA substrate s_0 was obtained to be linear (Kepseu) and given by

$$\sigma = 0.59 + 0.56\, s_o \tag{10}$$

$$\rho_s = (1.75 + \ 1.71\, s_o) \times 10^{-3} \tag{11}$$

2. Kinetic model of cAMP import/export

As expected from earlier studies, the cAMP concentrations directly estimated from the cell pellets were not reproducible and differed from the expected increase at the transition to stationary phase (Fig. S2B). We inferred the intracellular concentrations from the reproducible values obtained in the extracellular medium (Fig. S2A) using the kinetic import/export model from (Hans). The effective equation relating the intracellular/extracellular concentration is:

$$cAMP_{in} = \frac{C_1}{OD(t)} * \frac{d(cAMP_{out})}{dt} + C_2 * cAMP_{out} \tag{12}$$

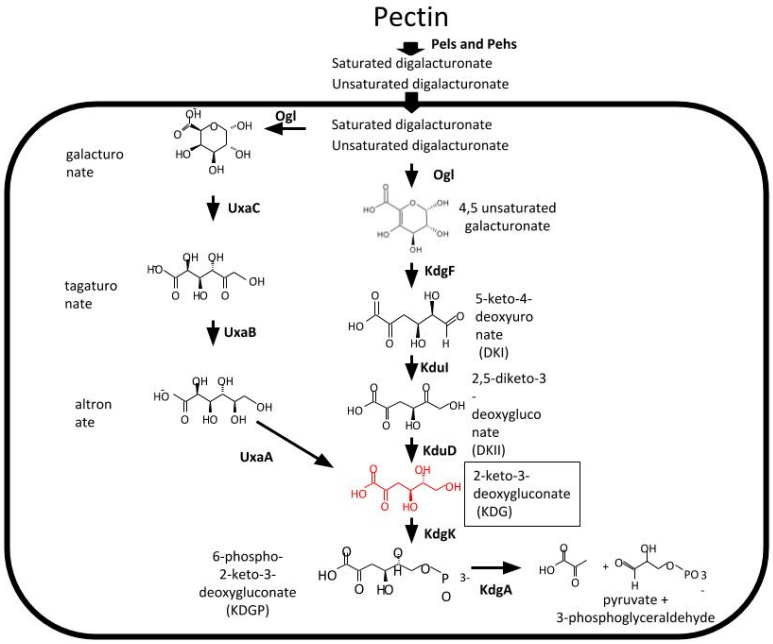
where C_1 and C_2 are two constants related to the total culture volume, cell volume, and cAMP import and export rates (Hans). The extracellular cAMP concentrations measured in our conditions are significantly higher than those measured in the original experiment with *E. coli* in microplates, but the *in vitro* regulatory properties of the purified CRP proteins from both species are indistinguishable on *pel* genes (Nasser 1997). We assumed that the kinetic import/export parameters could differ because of differences in species or growth conditions (batch culture vs microplates), and we adjusted their values based on those estimated in *E. coli*, so that the peak magnitude in internal cAMP concentration approximately matches the affinity of the metabolite for CRP (10 µM, dashed horizontal line in Fig. 1C), and thus triggers the regulatory action of the latter, as

expected for this signaling molecule. As a result, the intracellular concentrations are given by:

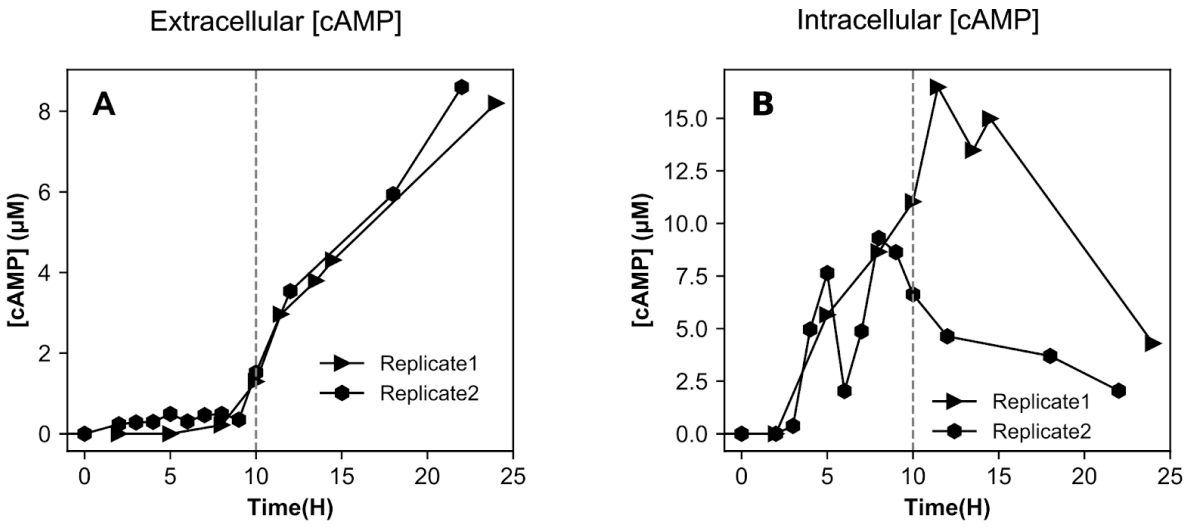
$$cAMP_{in} = \frac{22}{OD} * \frac{d(cAMP_{out})}{dt} + 0.20 * cAMP_{out} \quad (13)$$

and are computed after a spline interpolation of the extracellular measures to reduce noise. We note that the absolute magnitude (but not the timing) of the intracellular concentration peak depends on the adjusted import/export parameters and thus remains imprecise, but these quantitative variations are buffered in the resulting cAMP-CRP concentration pattern, and do not change the timing of cAMP activation.

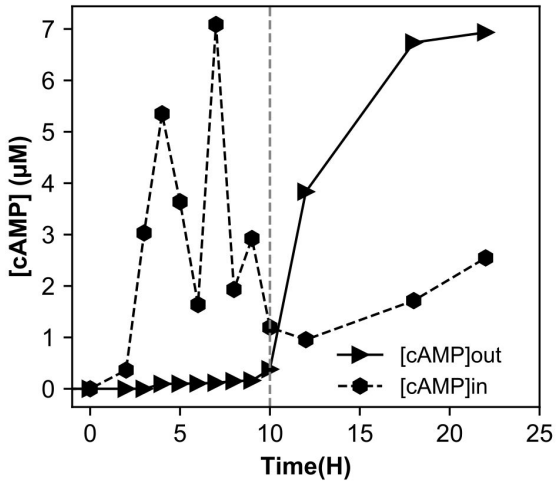
3. Supplementary Figures



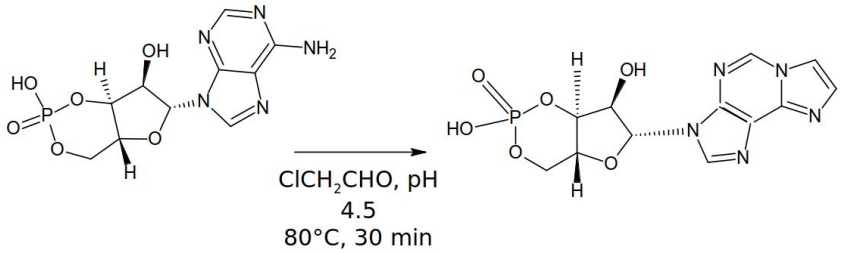
Supplementary Figure S1: Detailed pectin degradation pathway



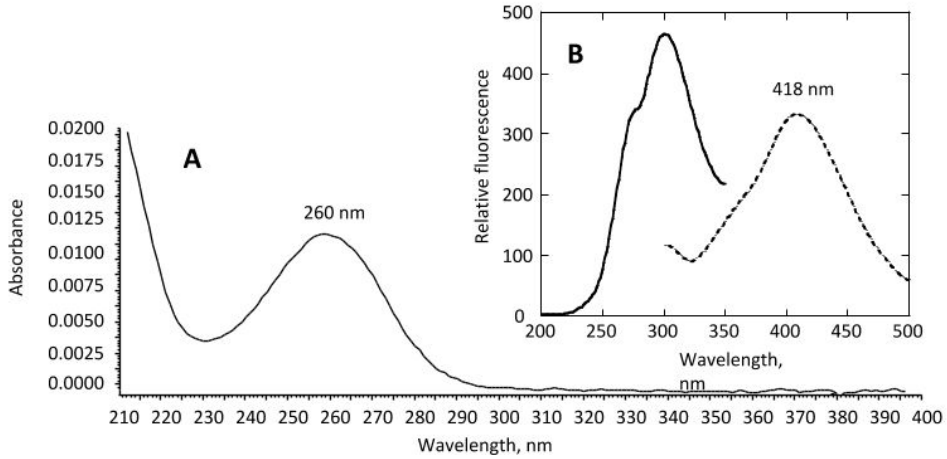
Supplementary Figure S2 : (A) Raw datapoints of extracellular cAMP concentrations in M63 Glucose PGA medium. (B) Intracellular cAMP concentrations estimated from the direct measurements of cell pellet samples deviate from the expected values and are not reproducible, as previously observed (Berthoumieux et al., 2013).



Supplementary Figure S3: Raw data points of cAMP quantifications in M63 Glucose medium, either in the extracellular medium (triangles) or in the cells as estimated from the cell pellet samples (dots).

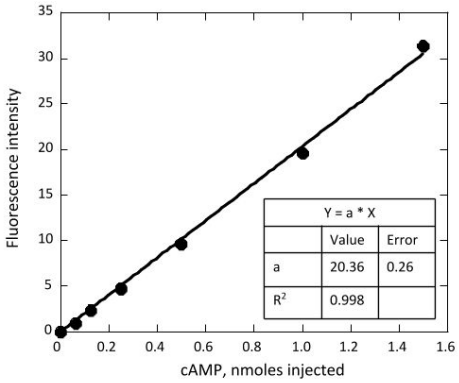


Supplementary Figure S4.a: Derivatization of cAMP with 2-chloroacetaldehyde.



A) Characteristic of the absorbance spectrum of 1 nanomole of the cAMP-chloroacetaldehyde molecule adduct. **B**) Excitation and emission peaks of the cAMP-chloroacetaldehyde molecule adduct. Spectra were performed using a SFM25 spectrofluorimeter (Kontron). The maximum of emission was of 418 nm upon excitation at 275 nm.

Supplementary Figure S4.b: cAMP absorbance/emission spectrum.



Standard curves for the pure cAMP metabolite adduct. Fluorescence intensity at 418 nm upon excitation at 275 nm. Inserts reported the value of the linear slope (a) with R^2 being the regression coefficient.

Supplementary Figure S4.C: The standard curve of the cAMP quantification on HPLC. The value of each point is the area under the cAMP peak from the corresponding standard sample.

Bibliography

Kepseu WD, Sepulchre J-A, Reverchon S & Nasser W (2010) Toward a Quantitative Modeling of the Synthesis of the Pectate Lyases, Essential Virulence Factors in Dickeya dadantii. *J. Biol. Chem.* **285:** 28565–28576

4.Conclusions

The above papers describe a combination of metabolic quantification, computational model and experimental validation. The developed HPLC quantification technique by far complements the developed transcriptional quantitative model very well. It gives the possibility of quantitative gene regulation modeling. The model has reproduced the differential expression of paralogs *pelE and pelD* and given insight into their regulatory mechanism. The measure PGA degradation kinetics agrees with the KDG accumulation quantified which resolves the drawbacks of the previous model and adds in the activation mechanism by CRP-cAMP pointing out the CCR mechanism observed in pectin degradation in *Dickeya dadantii.*

The expression pattern of *pels* gives us certain insights into the different roles of *pelE and pelD* in the virulence mechanism of *Dickeya dadantii*. The occurrence of peak expression of *pels* even after the complete degradation of PGA reflects on the pectin anticipatory property of *Dickeya dadantii* in plant environment. The expression of pectin dependent *pelD* might vary in the real plant environment depending on the pattern and concentration of pectin availability during the course of pathogenesis. This could be further looked into future perspective by conducting experiments in real plant environments and measuring *pelD* expression and *pel* activity analysis. The growth of model organisms in plants can be monitored by inoculating a known number of bacteria into the host leaf. Various stages of infection can be studied and the corresponding level of expression *pelD* and *pelE* can be quantified with qRT-PCR experiments. As discussed earlier, *pelD* expression pattern is predicted to be different in plant environment with more progressive availability of pectin unlike in the batch growth cognition in conical flask. The strong dependency of pectin should be facilitating higher expression of *pelD* in the plant environment defining it's virulence role during infection in breaking down the cell wall. In addition to the *pel* expression, we predict a higher *pel* activity observed in a real host environment due to increased stimulus by the increased pectin availability.

The CCR mechanism can be explored to cellular level with modification in the level of cAMP and measuring its effect on *pel* expression. Since *Dickeya dadantii* can simultaneously consume glucose and pectin, intermittent glucose addition to the medium could also reveal further insights into the catabolite repression mechanism. Addition of glucose at a time point when the existing glucose in the medium is depleted and at late transitional phase to observe the variation in the level of *pel* expression. Addition of glucose at the transitional phase should be reducing the boosting effect of CRP activator since presence of glucose would suppress the availability of cAMP to be bound to the CRP activator. In the same context, variation of level of cAMP is expected

to vary the gene expression level. Addition of cAMP at the beginning exponential phase should be able to provide the boost to the *pel* genes by binding to the CRP activator. Would this be observed in the experiment or will the presence of glucose suppress the effect of added cAMP. Model predicts the activation obtained in *pel* expression is proportional to the concentration of CRP-cAMP complex binding to the DNA. It would be interesting to verify this theory. How would a sudden increase in cAMP at a late exponential phase affect the expression of genes?

In the chromatogram from the cAMP quantification from bacterial cell samples, we observe multiple peaks adjacent to the identified cAMP peak, this could possibly indicate the presence of eluted purines similar to cAMP. Since the derivatization compound CAA is known to act on purine compounds, we assume the adjacent peaks being similar purine compounds like cGMP, ppGpp, etc. Thus, this quantification technique could be extended and adapted to identify possibly present purine derivatives.

5.References:

[1] J.-A. Sepulchre, S. Reverchon, and W. Nasser, “Modeling the onset of virulence in a pectinolytic bacterium,” *J. Theor. Biol.*, vol. 244, no. 2, pp. 239–257, Jan. 2007, doi: 10.1016/j.jtbi.2006.08.010.

[2] M. Boccara, A. Diolez, M. Rouve, and A. Kotoujansky, “The role of individual pectate lyases of Erwinia chrysanthemi strain 3937 in pathogenicity on saintpaulia plants,” *Physiol. Mol. Plant Pathol.*, vol. 33, no. 1, pp. 95–104, Jul. 1988, doi: 10.1016/0885-5765(88)90046-X.

[3] N. Hugouvieux-Cotte-Pattat and J. Robert-Baudouy, “Analysis of the regulation of the pelBC genes in Erwinia chrysanthemi 3937,” *Mol. Microbiol.*, vol. 6, no. 16, pp. 2363–2376, 1992, doi: 10.1111/j.1365-2958.1992.tb01411.x.

[4] H. Schmidt and M. Hensel, “Pathogenicity Islands in Bacterial Pathogenesis,” *Clin. Microbiol. Rev.*, vol. 17, no. 1, pp. 14–56, Jan. 2004, doi: 10.1128/CMR.17.1.14-56.2004.

[5] S. Reverchon, W. Nasser, and J. Robert-Baudouy, “Characterization of kdgR, a gene of Erwinia chrysanthemi that regulates pectin degradation,” *Mol. Microbiol.*, vol. 5, no. 9, pp. 2203–2216, Sep. 1991, doi: 10.1111/j.1365-2958.1991.tb02150.x.

[6] W. Nasser, J. Robert-Baudouy, and S. Reverchon, “Antagonistic effect of CRP and KdgR in the transcription control of the Erwinia chrysanthemi pectinolysis genes,” *Mol. Microbiol.*, vol. 26, no. 5, pp. 1071–1082, Dec. 1997, doi: 10.1046/j.1365-2958.1997.6472020.x.

[7] S. Reverchon, M. L. Bouillant, G. Salmond, and W. Nasser, “Integration of the quorum-sensing system in the regulatory networks controlling virulence factor synthesis in *Erwinia chrysanthemi*,” *Mol. Microbiol.*, vol. 29, no. 6, pp. 1407–1418, Sep. 1998, doi: 10.1046/j.1365-2958.1998.01023.x.

[8] T. Lautier and W. Nasser, “The DNA nucleoid-associated protein Fis co-ordinates the expression of the main virulence genes in the phytopathogenic bacterium Erwinia chrysanthemi,” *Mol. Microbiol.*, vol. 66, no. 6, pp. 1474–1490, 2007, doi: 10.1111/j.1365-2958.2007.06012.x.

[9] T. Lautier, N. Blot, G. Muskhelishvili, and W. Nasser, “Integration of two essential virulence modulating signals at the Erwinia chrysanthemi pel gene promoters: a role for Fis in the growth-phase regulation,” *Mol. Microbiol.*, vol. 66, no. 6, pp. 1491–1505, 2007, doi: 10.1111/j.1365-2958.2007.06010.x.

[10] W. Nasser and S. Reverchon, “H-NS-dependent activation of pectate lyases synthesis in the phytopathogenic bacterium Erwinia chrysanthemi is mediated by the PecT repressor,” *Mol. Microbiol.*, vol. 43, no. 3, pp. 733–748, Feb. 2002, doi: 10.1046/j.1365-2958.2002.02782.x.

[11] W. Nasser, G. Condemine, R. Plantier, D. Anker, and J. Robert-Baudouy, “Inducing properties of analogs of 2-keto-3-deoxygluconate on the expression of pectinase genes of Erwinia chrysanthemi,” *FEMS Microbiol. Lett.*, vol. 81, no. 1, pp. 73–78, Jun. 1991, doi: 10.1111/j.1574-6968.1991.tb04715.x.

[12] W. Nasser, S. Reverchon, and J. Robert-Baudouy, “Purification and functional characterization of the KdgR protein, a major repressor of pectinolysis genes of Erwinia chrysanthemi,” *Mol. Microbiol.*, vol. 6, no. 2, pp. 257–265, 1992, doi: 10.1111/j.1365-2958.1992.tb02007.x.

[13] W. Nasser, S. Reverchon, G. Condemine, and J. Robert-Baudouy, “Specific Interactions of Erwinia chrysanthemi KdgR Repressor with Different Operators of Genes Involved in Pectinolysis,” *J. Mol. Biol.*, vol. 236, no. 2, pp. 427–440, Feb. 1994, doi: 10.1006/jmbi.1994.1155.

[14] S. Reverchon, D. Expert, J. Robert-Baudouy, and W. Nasser, “The cyclic AMP

receptor protein is the main activator of pectinolysis genes in Erwinia chrysanthemi," *J. Bacteriol.*, vol. 179, no. 11, pp. 3500–3508, Jun. 1997, doi: 10.1128/jb.179.11.3500-3508.1997.
[15] W. D. Kepseu, J.-A. Sepulchre, S. Reverchon, and W. Nasser, "Toward a quantitative modeling of the synthesis of the pectate lyases, essential virulence factors in Dickeya dadantii," *J. Biol. Chem.*, vol. 285, no. 37, pp. 28565–28576, Sep. 2010, doi: 10.1074/jbc.M110.114710.
[16] L. Zhang, H. Zhang, J. Wang, and C. Pan, "Determination of Trace Level of cAMP in Locusta Migratoria Manilensis Meyen by HPLC with Fluorescence Derivation," *Int. J. Mol. Sci.*, vol. 7, no. 8, pp. 266–273, Aug. 2006, doi: 10.3390/i7080266.
[17] A. Duprey, W. Nasser, S. Léonard, C. Brochier-Armanet, and S. Reverchon, "Transcriptional start site turnover in the evolution of bacterial paralogous genes - the pelE-pelD virulence genes in Dickeya," *FEBS J.*, vol. 283, no. 22, pp. 4192–4207, 2016, doi: 10.1111/febs.13921.

Chapter III: Studying the effect of DNA supercoiling on the virulence genes

1.DNA Supercoiling: an Ancestral Regulator of Gene Expression in Pathogenic Bacteria?

Shiny Martis B*, Raphaël Forquet*, Sylvie Reverchon, William Nasser and Sam Meyer°

Université de Lyon, INSA Lyon, Université Claude Bernard Lyon 1, CNRS UMR5240, Laboratoire de Microbiologie, Adaptation et Pathogénie, 11 avenue Jean Capelle, 69621 Villeurbanne, France

* Equal first authors

° Corresponding author. E-mail: sam.meyer@insa-lyon.fr. Tel: +33 (0)4 72 43 85 16.

Abstract

DNA supercoiling acts as a global and ancestral regulator of bacterial gene expression. In this review, we advocate that it plays a pivotal role in host-pathogen interactions by transducing environmental signals to the bacterial chromosome and coordinating its transcriptional response. We present available evidence that DNA supercoiling is modulated by environmental stress conditions relevant to the infection process according to ancestral mechanisms, in zoopathogens as well as phytopathogens. We review the results of transcriptomics studies obtained in widely distant bacterial species, showing that such structural transitions of the chromosome are associated to a complex transcriptional response affecting a large fraction of the genome. Mechanisms and computational models of the transcriptional regulation by DNA supercoiling are then discussed, involving both basal interactions of RNA Polymerase with promoter DNA, and more specific interactions with regulatory proteins. A final part is specifically focused on the regulation of virulence genes within pathogenicity islands of several pathogenic bacterial species.

Keywords: DNA supercoiling; transcription; pathogenesis; genetic regulation

Introduction

DNA supercoiling (SC) has received considerable attention in recent years as a global and ancestral actor in genetic regulation. This is especially conspicuous in bacteria [1–3], where the chromosome is maintained at an out-of-equilibrium level of negative SC by a finely controlled balance of topoisomerase activity. And yet, in contrast to classical regulation based on transcription factors, quantitative models of the regulatory mechanisms by SC are essentially lacking. A possible explanation for this shortcoming is that SC affects transcription at several stages of the process, and can also be involved in various and complex interactions with regulatory proteins. As a result, virtually every investigated promoter exhibits a distinct SC response, making it difficult to dissect and model the underlying mechanisms. In this review, we wish to summarise existing evidence and models suggesting a widespread role of SC in bacterial genetic regulation, and more specifically in bacterial virulence. This topic has already been addressed in previous extensive reviews focused either on its role in bacterial growth [4] or on specific promoters that were analysed in detail [5]. Here, we propose a complementary focus on proposed mechanistic and computational models of transcriptional regulation by SC as well as accumulating information obtained from transcriptomic data, which together underline the broad relevance of the investigated phenomenon in bacterial virulence and call for a combined experimental-theoretical research effort.

1. DNA supercoiling: a global regulator of bacterial gene expression

1.1. DNA supercoiling : a relay of environmental signals to the bacterial chromosome

As observed immediately following the discovery of the double-helical structure of DNA, virtually all DNA transactions face substantial topological constraints [6]. In mechanical terms, the latter give rise to an ubiquitous torsional stress, which in turn results in DNA supercoiling (SC), i.e., the deformation of the molecule either by rotation around its helical axis (over- or under-twisting) or by the winding of this helical axis itself (writhing), as illustrated in Fig. 1 [4,7].

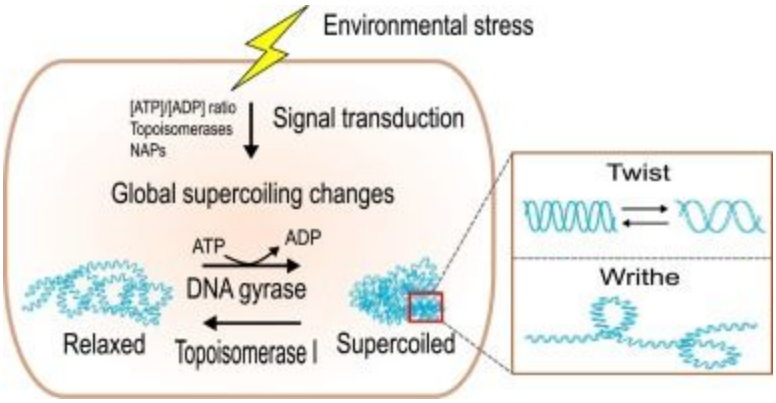


Fig. 1. DNA supercoiling acts as a sensor of environmental stress in the bacterial chromosome. Environmental cues are transduced by different mechanisms into global, stress-specific variations of the SC level. At a smaller scale, SC is distributed as twist and writhe deformations, which directly affect the transcriptional activity.

Topoisomerases are the global regulators of SC and more generally, the solvers of topological problems associated with DNA transactions [8]. In bacteria, the two main topoisomerases are topoisomerase I (topo I) and DNA gyrase. The latter maintains the chromosomal DNA in an underwound state by introducing negative supercoils in an ATP-dependent manner, while conversely, topo I relaxes the DNA (i.e. removes negative supercoils) without any ATP requirement. The global negative SC level of the chromosome is thus primarily determined by the dynamic equilibrium between these two enzymes (Fig. 1). Additional actors play a more specific role: the ATP-consuming topoisomerase IV is primarily involved in solving topological problems associated with DNA replication and cell division [9], and abundant nucleoid-associated proteins (NAPs) contribute in distributing SC along the bacterial chromosome [4,10].

The negative SC level of chromosome is finely controlled by the cell in response to environmental conditions, since almost all types of environmental challenges have been associated with SC variations, and in particular those most commonly encountered by pathogens during infection (Tab. 1). Depending on the applied stress, the chromosomal DNA experiences either a partial relaxation (+) or an increase in negative SC (-), which is usually rapid and transient. In spite of strong differences in terms of phylogeny or lifestyle, response to each specific stress is qualitatively similar in all investigated species, although these exhibit quantitatively different SC levels in standard growth conditions [3]. This observation suggests that SC is used in a wide range of bacteria to quickly transduce environmental signals toward the chromosome, with ancestral control mechanisms. Interestingly, environmental stresses have also been correlated with SC changes in archaeal species [11], suggesting that this notion could be extended to an even wider range of microorganisms.

Shock	Phylum	Species	SC change	Mechanism	Ref
Heat	P	*Escherichia coli*	Rel (+)	Gyrase and topoI activities	[12]
		Yersinia enterocolitica		Gyrase activity decrease	[13]
		Dickeya dadantii			[14]
	F	*Bacillus subtilis*			[15]
Cold	P	*Escherichia coli*	Hyp (-)	Gyrase activity increase / HU	[16]
	F	*Bacillus subtilis*			[15]
Acidic	P	*Escherichia coli*	Rel (+)		[17,18]
		Salmonella typhimurium		Gyrase activity decrease	[18,19]
		Dickeya dadantii			[20]
Osmotic	P	*Escherichia coli*	Hyp (-)	[ATP]/[ADP] increase	[21]
		Salmonella typhimurium			[22]
		Dickeya dadantii			[23]
	F	*Bacillus subtilis*			[15]
		Staphylococcus aureus			[24]
	A	*Streptomyces lividans*			[25]
Oxidative	P	*Escherichia coli*	Rel (+)	TopA activation by Fis	[26]
		Dickeya dadantii			[23]

Anaerobic	P	*Escherichia coli*	Rel (+)	[ATP]/[ADP] decrease	[27]
		Salmonella typhimurium			[28]
	F	*Bacillus subtilis*			[15]

Tab. 1. Chromosomal supercoiling response to environmental stress conditions is conserved in distant bacterial species. **Phyla:** P: Proteobacteria, F: Firmicutes, A: Actinobacteria. **SC variations:** Rel (+): relaxation, Hyp (-): hyper-supercoiling

What are the underlying mechanisms? The most clearly described pathway involves the modulation of gyrase activity by the energy charge of the cell through the [ATP]/[ADP] ratio [21,27,29]. When the latter is increased, gyrase introduces supercoils in the chromosomal DNA more actively and its negative SC level increases. Environmental stresses usually alter cellular metabolic fluxes and the energy charge in various ways [30], and this relatively simple, quick and general mechanism is indeed involved in the chromosomal response to a variety of conditions (Tab. 1). Other invoked mechanisms include (1) the action of NAPs (e.g. HU, FIS) either by direct interaction with DNA or as modulators of topoisomerase activity; (2) the regulation of topoisomerase expression (*topA*, *gyrA* and *gyrB* genes encoding topoI and gyrase subunits respectively); (3) the modulation of topoisomerase activity through post-translational modifications [31].

It should be underlined that most bacteria from Tab. 1 are pathogenic or have pathogenic serotypes. Since the infection process can be assimilated to successive environmental changes to which the pathogen needs to adapt quickly, SC appears as a general candidate for the efficient transmission of stage-specific environmental signals toward the bacterial chromosome and thus also for its transcriptional response, which in turn is critical for the subsequent steps of the infection.

1.2. Global transcriptional response to variations in DNA supercoiling

The transcriptional response to various stress conditions can be readily analysed from transcriptomic data, but it is then difficult to disentangle (1) the generic effect of the stress-induced SC variation on transcription and (2) the transcriptional effect of stress-specific pathways. The former contribution can however be analysed separately in transcriptomes obtained with gyrase inhibiting antibiotics. These are either aminocoumarins (novobiocin, coumermycin) which block the ATPase activity of topoisomerases II (gyrase and topoisomerase IV), or quinolones (ciprofloxacin, norfloxacin, nalidixic acid, oxolinic acid) which block their ligase activity, the latter resulting in many double-strand breaks and triggering a SOS response of the cell with pleiotropic effects [32]. When applied at a sublethal dosage, these drugs induce a sudden global relaxation of the chromosomal DNA, and the transcriptional response is then measured after a short time (usually 5-30 minutes), assuming

that the latter then mostly reflects the direct effect of SC, rather than indirect effects influenced by the cell's response. Such data were obtained in many organisms (Tab. 2). Note that the reported number of affected genes is strongly variable; this variability might partly reflect actual differences between organisms, but is strongly affected by the experimental conditions and methods (relaxation level, transcriptomics technology, statistical analysis). Altogether, these data consistently demonstrate a very broad response to chromosomal relaxation, with a significant effect on more than one quarter of the genes. This response is complex, with some genes being upregulated and others downregulated. These affected genes are functionally diverse, including genes involved in essential functions (e.g. DNA replication, cell division), stress responses and metabolic pathways (e.g. stringent response, DNA repair pathway), as well as virulence. They are also usually scattered throughout the chromosome, highlighting that SC-mediated regulation acts in a global way, but follows a spatial organisation pattern involving large-scale responsive domains related to structural properties of DNA [20,33].

Since this global transcriptional response is observed in a wide range of species from different phyla (Tab. 2), SC might be considered as an ancestral and widespread mode of regulation in bacteria. This notion can be related to the fundamental and highly conserved character of topoisomerase enzymes themselves [8], and may even be extended to eukaryotes, albeit with different rules [7,34]. It does not mean however that the mechanism is identical in all bacteria. The longest-running evolution experiment [35] emphasized that mutations affecting SC constitutes a "quick and efficient" way to modify the global expression pattern and gain substantial fitness, in this case by a mutation reducing topo I efficiency in less than 2000 generations [36]. It is therefore no surprise that fluctuations in topoisomerase structure and SC level were pointed in the close relatives *E. coli* and *S. typhimurium* [3] which have different lifestyles, and this is probably also frequent in different strains of the same species [19]. Changing the chromosomal SC level is thus a fundamental and generic way by which bacteria adapt to new environmental challenges, according to common ancestral rules. This extends in particular to genome-reduced bacteria almost devoid of TFs such as *Mycoplasma* or *Buchnera* [37–39], where transcriptional regulation remains poorly understood.

Phylum	Species	SC change	Method	Genes significantly affected (% genome)	Technology	Ref
P	*Escherichia coli*	Rel (+)	Norfloxacin	613 (15%)	M	[40]
		Rel (+)	Novobiocin / pefloxacin	1957 (48%)	M	[41]
		Rel (+)	Genetic engineering	740 (18%)	M	[42]
		Rel (+)	Genetic engineering / norfloxacin / novobiocin	306 (7%)	M	[43]
	Salmonella typhimurium	Rel (+)	Genetic engineering	499 (10%)	M	[44]
	Dickeya dadantii	Rel (+)	Novobiocin	1461 (32%)	M	[20]
		Rel (+)		1212 (27%)	S	[45]
	Haemophilus influenzae	Rel (+)	Novobiocin / ciprofloxacin	640 (37%)	M	[46]
F	*Streptococcus pneumoniae*	Rel (+)	Novobiocin	290 (14.2%)	M	[47]
		Hyp (-)	Seconeolitsin	545 (27%)	S	[48]
	Staphylococcus aureus	Rel (+)	Novobiocin	280 (11 %)	M	[49]
	Bacillus subtilis	Rel (+)		1075 (24 %)	M	[50]
A	*Streptomyces coelicolor*	Rel (+)	Novobiocin	121 (1.5 %)	S	[51]
	Mycobacterium tuberculosis	Rel (+)		Not provided	M	[52]
T	*Mycoplasma pneumoniae*	Rel (+)	Novobiocin	469 (43 %)	S	[37]

C	*Synechocystis*	Rel (+)	Novobiocin	Several genes	M	[53]
	Synechococcus elongatus	Rel (+)		Not provided	M	[54]

Tab. 2. Transcriptomic response to variations of DNA supercoiling in bacteria. **Phyla:** P: Proteobacteria, F: Firmicutes, A: Actinobacteria, T: Tenericutes, C: Cyanobacteria. **SC variations**: Rel (+): relaxation, Hyp (-): hypersupercoiling. **Transcriptomics technology:** M: DNA Microarray, S: RNA Sequencing

2. Mechanisms and models of transcriptional regulation by DNA supercoiling

The abovementioned transcriptional responses induced by SC variations differ qualitatively from those induced by classical transcriptional factors (TFs). The latter recognize, bind and regulate a specific subset of the genome defined by a well-defined (although often degenerate) target sequence motif, and their action can be modelled using classical thermodynamic models of activation or repression [55]. In contrast, as noted above, SC affects the transcriptome of all investigated species globally, without any identified promoter sequence determinant. And yet strikingly, regulatory models comparable to those involving TFs are essentially lacking. There are two reasons for this: first, experimentally, SC regulates gene activity in a continuous “more or less” manner as opposed to the stronger “on or off” mode of regulation by TFs [56]; second, SC can modulate transcription in a variety of ways, making them difficult to decipher. In the following, we discuss such mechanisms, most of which will be illustrated on the promoter of *pelE*, one of the major virulence genes of the phytopathogen *Dickeya dadantii* where SC-mediated regulation was studied extensively.

2.1. Basal regulation of the RNA Polymerase - DNA interaction

The ancestral and global mode of regulation by SC results, in the first instance, from it affecting the interaction between DNA and RNA Polymerase (RNAP) itself, independently from any additional regulatory protein. But this basal regulation already involves distinct mechanisms occurring at successive steps of the complex process of bacterial transcription: closed-complex formation, open-complex formation, promoter clearance, and elongation [57].

The most clearly identified – and possibly strongest - effect of SC on transcription initiation results from the requirement for RNAP to open the DNA strands and stabilize a “transcription bubble”, in order to gain access to the DNA bases in the template strand. In torsionally unconstrained DNA, this melting transition represents a substantial free energy cost of around 10 k_BT (6 kcal/mol), which in eukaryotes is provided through ATP hydrolysis by the basal transcription factor TFIIH [34]. Crucially, this cost reduces drastically when the double helix is destabilised by negative torsion at SC levels physiologically relevant in bacteria (Fig. 2A, upper panel), thus providing the physical basis for the bacterial transcription process that does not require any external energy consumption [57]. Interestingly, the same can become

true of eukaryotic RNA Polymerase II when operating on a comparably negatively supercoiled template [34], whereas temperature may replace SC as the source of melting energy in thermophilic archaea relying on reverse gyrase [58,59]. This strong regulatory effect of SC can be observed in well-controlled *in vitro* transcription assays (Fig. 2A, lower panel), where plasmids carrying a model promoter are prepared at different supercoiling levels, resulting in a drastic variation of expression strength without any modification of the DNA sequence or regulator concentration. This activation curve is quantitatively reproduced [45] using thermodynamic models of DNA opening [60–62] relying on knowledge-based enthalpic and entropic parameters for base-pairing and stacking interactions of all base sequences.

How may this mechanism lead to transcriptional regulation, i.e., the selective activation of a subset of promoters by global SC variations? The opening curve is strongly dependent on the promoter base sequence (mostly though its GC content) in the -10 region and that immediately downstream referred to as the “discriminator” [4,63–65]. Remarkably, although the latter region does not harbour any consensus binding signal for RNAP, mutation studies showed that it plays a predominant role in the SC-sensitivity of promoters: GC-rich discriminators are typically more activated by negative SC than AT-rich ones, the canonical example being those of stable RNAs strongly induced during exponential growth [4,64,65]. And yet, a systematic model of this regulation mechanism is still lacking. A possible obstacle is that the thermodynamic description used above might be insufficient, if the expression level of investigated promoters is limited not by the rate of promoter opening but of promoter escape. As an example, some promoters might attract RNAP and easily form an open-complex, but still exhibit low expression levels if the latter is too stable, resulting in abortive rather than processive transcription [4,66,67]. Since the influence of SC on these subtle kinetic steps was not dissected in detail, little can be predicted from such a scenario that may explain why promoters respond differently to SC variations [68], and some even in opposite ways [69–71]. Yet we still note that many *in vitro* investigated promoters do follow the behaviour expected from thermodynamic modelling (Fig. 2A, lower panel), which might thus account for at least a significant fraction of transcriptomic responses to SC variations.

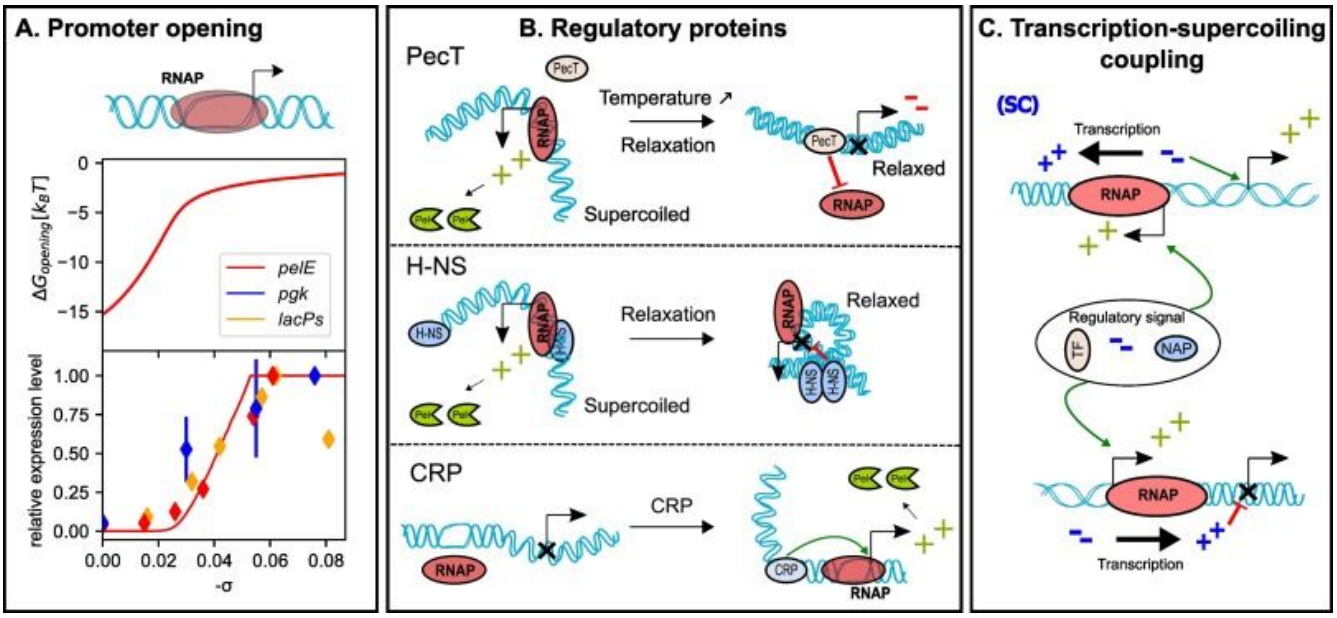


Fig. 2. Mechanisms of transcriptional regulation by DNA supercoiling. **(A)** Opening free energy of the *pelE* promoter of *D. dadantii* (upper panel) and expression level of three promoters from different species (lower panel, solid line for the model prediction) [23,45]; **(B)** Various mechanisms involving SC and NAPs/TFs at the *pelE* promoter [14,23,72]; **(C)** Local SC-mediated regulatory interaction: transcription from a given gene (left) can activate (top) or repress (bottom) its neighbour depending on their relative orientation [45].

A regulatory action of SC was proposed at least at two other transcription steps. RNAP binds promoters by recognising the -10 and -35 elements, which are separated by a spacer of variable size (15-20 nt, with an optimum at 17 nt). Because of the helical structure of DNA, these variations are associated with different relative angular positions of the recognition sites around the helical axis, possibly placing them out-of-phase for RNAP binding and closed-complex formation. A suitable (spacer length-dependent) level of torsional stress could then modulate this binding rate by untwisting the spacer DNA toward a favourable orientation [73,74]. This scenario was invoked to explain the response of several promoters to SC variations, including ribosomal RNAs which usually exhibit a suboptimal spacer length of 16 nt [73].

SC may also affect transcription at the elongation stage, where positive torsion hinders the progress of RNAP, as demonstrated from single-molecule experiments [75]. Such positive SC levels are not observed *in vivo* by usual techniques involving plasmids, but might be transiently generated locally downstream of transcribed genes (see paragraph 2.3 below). However, most *E. coli* genes were found to be transcribed at a comparable speed [76], suggesting that positive supercoils are readily eliminated by topoisomerases and do not play any regulatory role during elongation of these moderately expressed genes, whereas mechanical stalling of RNAP might occur at the most strongly expressed genes such as those

of ribosomal RNAs [77].

2.2. Specific regulation involving DNA regulatory proteins

SC affects transcription via RNAP itself, but also via promoter-specific regulatory proteins (TFs). In classical thermodynamic regulation models [55], the latter act (as activators or repressors, depending on their action on RNAP activity) at distinct binding sites characterised by their affinity, the latter being inscribed in the genomic sequence and usually described by sequence motifs. In this simplistic view, the cell has limited regulatory freedom, since the only tunable parameter is the TF concentration. In the classical example of the global regulator CRP which binds (and usually activates) hundreds of promoters in *E. coli*, increasing its concentration (or rather that of its cofactor cAMP) may activate many of these target promoters, but in this model, the *relative* amount of CRP bound at these promoters, and thus also their relative activities, is out of the control of the cell. This view is increasingly challenged since additional layers of complexity were identified, including epigenetic modifications of the regulators or promoter DNA and post-transcriptional regulation. However, we note that many regulatory proteins recognise not only the base sequence, but also the DNA shape, a mechanism often referred to as indirect readout [78]. In that case, the activity of classical regulators is intrinsically modulated, not only by chemical (epigenetic) modifications requiring dedicated enzymes, but also by the ubiquitous mechanical deformations. A computational estimation of this recognition mode suggests that, in the case of CRP, the latter may in fact be the strongest determinant of its loose sequence selectivity [79]. Crucially, in contrast to the classical “static” sequence model, this selectivity is now dependent on the mechanical state of chromosomal DNA, and thus subjected to cellular control through SC. Considering that the SC distribution itself is non-uniform and locally affected by NAPs, this additional mechanical dimension is probably a key contributor in the complexity of the binding selectivity of regulatory proteins.

This mechanism has not been investigated experimentally in many systems. One example is the thermoregulation of *pelE* by the repressor PecT, which is achieved not by a change in regulator concentration but rather by an increase of its binding affinity for the promoter resulting (at least partly) from temperature-induced chromosomal DNA relaxation (Fig. 2B, top) [14,80]. PecT belongs to the LysR-like family of TFs relying on indirect sequence readout, which is the largest TF family in enterobacteria [81], and similar mechanisms are used by other families [81–83]. Altogether, the role of DNA mechanics in TF binding selectivity is thus likely much more important than generally considered, although the precise mechanisms remain to be identified and modelled. Because most dimeric TFs are smaller than RNAP, they are probably less sensitive to the untwisting of bound DNA than the latter [79], but this sensitivity might be augmented substantially in the case of several binding sites arranged in helical phase. For example, H-NS was found to repress *pelE* on relaxed DNA (as most of its targets), but it is not the case on negatively supercoiled DNA [23]. A proposed explanation is that H-NS traps RNAP in a small loop at this promoter by bridging its two

binding sites, which in the latter condition is prevented by their unfavourable helical phasing (Fig. 2B, middle). On the other hand, regulatory proteins also exert twist deformations themselves; a spectacular example is MerR, which in the presence of Hg_{2+} binds and untwists the 19-bp spacer of the mercury resistance operon of Tn501 (a transposable element isolated from *Pseudomonas aeruginosa*), thus enabling RNAP binding [84].

In our opinion, the interplay between regulatory proteins and local SC may mostly rely on 3D modifications of the promoter conformation, i.e., writhe rather than twist. On the one hand, writhe could facilitate the formation of loops required for many regulatory interactions. The latter are favoured by the distinct mechanical properties of promoter DNA sequences [85], and were already included in regulatory models based on the transfer matrix formalism [86]. SC-dependent reduction of looping free energies can thus strongly modify the binding landscape of regulatory proteins, as already described [87,88]. Conversely, many regulatory proteins induce strong bends in DNA [83], e.g., CRP (~ 90°), LexA (~ 35°), as well as the NAPs FIS (~ 45°) and IHF (~ 180°). Since such deformations drastically reduce the energetic cost of DNA loops [89,90], they are expected to displace the twist-writhe equilibrium within the bound region in favour of the latter, and may thus induce local topological changes similar to those induced by SC [91]. In summary, just like the NAPs are involved in a complex double-sided interaction with SC that shapes the global structure of the chromosome, a comparable effect probably occurs with many more regulatory proteins at the more local scale of gene promoters, with direct consequences for local transcription.

A final important ingredient is the widespread occurrence of structural transitions in genomic DNA. The latter can switch from double-stranded B-DNA to, among others, denaturated, Z-DNA, G-quadruplex or cruciform states. The rates of these different transitions can be computed by thermodynamic modelling [92], and depend not only on the DNA sequence, but also on torsional stress that destabilises B-DNA and can be accommodated more favourably in alternate states. It was also shown that these transitions occur predominantly at bacterial gene promoters [60] which they regulate according to various mechanisms, some involving TFs. Denaturated AT-rich regions located 50-200 bp upstream of the TSS can act as “sinks” for negative SC and impede the proper opening of the promoter by thermodynamic competition. This was shown to occur for *pelE*, which *in vitro* is not expressed in absence of CRP due to such upstream strand opening; when present, CRP not only favours the correct binding of RNAP, but also “closes” the upstream AT-rich tract, possibly by bending DNA (Fig. 2B, bottom) [23]. Since denaturation bubbles are extremely flexible, they may also strongly facilitate the formation of loops [89] required by TFs [86]. Additionally, some regulatory proteins may selectively bind non B-DNA regions, as occurs at the mammalian oncogene *cMYC* where negative SC triggering DNA melting is provided by adjacent transcription [60,93]. Finally, since rho-independent termination of transcription involves RNA hairpin structures, SC might also favour structural transitions in the DNA template itself at the transcription termination site, which could then modulate the termination rate, as already

observed for the B-Z transition [94]. Since transcriptional read-through was highlighted as a widespread feature in bacterial genomes in recent years [95], an additional underestimated layer of regulation might thus also occur at this later stage of the transcription process.

2.3. Spatial heterogeneities of DNA supercoiling : the Transcription-Supercoiling Coupling

The intimate relationship between SC and transcription is not single-sided. In the elongation step, the helical structure of DNA imposes a fast rotation of the bulky RNAP relative to it (around two turns per second), but this movement is strongly hindered by the viscosity and crowding of the surrounding medium, resulting in an asymmetric accumulation of torsional stress from back to front, as recognised more than 30 years ago [96]. This phenomenon thus leads to an intrinsic dynamical coupling between SC and transcription highly dependent on gene orientations (Fig. 2C). This coupling regained significant interest in recent years since it was shown to underpin transcriptional bursting in bacteria [97], i.e., the nonlinear auto-induction of a promoter that can typically give rise to phenotypical heterogeneities among isogenic populations of cells [98]. Several theoretical models were proposed [45,77,99–101], most of them focusing on biophysical properties of transcription. A strong obstacle for their application in genomic regulation is the lack of experimental knowledge of the distribution of SC along a bacterial chromosome. A promising method involving the intercalating agent psoralen was developed in eukaryotes [7,102] but did not yet provide high-resolution data in bacteria [1]. Recently, indirect information was provided from binding distributions of topoisomerases obtained by ChIP-Seq at different resolutions [2,103]. These data, together with a systematic analysis of bacterial transcriptomes, confirmed that the distribution of SC along the chromosome is highly heterogeneous and strongly affected by gene orientations [104], leading to a fine-tuned and ancestral regulation of promoters depending on their genomic context [45]. In summary, the *local* level of SC experienced by a given promoter can strongly differ from the *global* (average) level of the chromosome, and depends on the orientation and activity of adjacent genes, providing a strong mechanistic basis for the co-regulation of co-localised operons [105]. Direct evidence for these effects was obtained from experiments involving supercoiling-sensitive promoters inserted on the chromosome in different artificial configurations [70,104]. Comparable evidence is more difficult to establish for native promoters, but was highlighted in at least two examples of divergently organised operons: the *ilv* promoters of *E. coli* [106] and the *leu-500* promoter of *S. typhimurium* [107]. Interestingly, in these two examples, the local nature of SC (generated by the divergent genes) is combined with a complex regulatory mechanism involving DNA binding proteins. In the first example, the activation of $ilvP_G$ is prevented by the denaturation of an upstream AT-rich tract, except when the NAP IHF closes that region and favours the opening of the promoter (like CRP at the *pelE* promoter, Fig. 2B). The pattern is similar for *leu*-500, although the repression is here achieved by H-NS binding at an AT-rich tract, and relieved by the TF LeuO in presence of locally generated negative SC [5,108]. Since

divergent genes involved in the same function and simultaneously expressed are commonly found in bacteria, including among those involved in pathogenicity (see below), these examples may only be the first of a large unexplored class.

The high density of bacterial genomes implies that the interaction between neighbour could in fact give rise to a collective behaviour along larger distances, forming “topological domains”. Indeed, when promoters were displaced over the chromosome, their expression and supercoiling sensitivity were found to change depending on their location and neighbouring activity [109–112]. These domains, shaped by transcription and architectural proteins, remain a poorly defined notion in bacteria. Proposed lengthscales vary from 10-20 kb [109], to 50 kb [112] and up to hundreds of kilobases [20]; while the former may underpin an extension of the notion of operons [105,113], the latter probably reflect a higher order folding of bacterial chromatin involving different actors, and this hierarchical organisation remains to be characterised.

3. DNA supercoiling and the coordination of virulence programs

3.1. An argument for DNA supercoiling being an important actor in virulence genetic regulation

Most pathogenic bacteria exhibit close genomic proximity to nonpathogenic strains, with differences located at well-defined genomic regions (of a few up to hundreds of kilobases in size) called pathogenicity islands (PAIs), which contain the virulence genes involved in pathogenesis. These regions are harboured either on the chromosome or on plasmids, and are usually acquired by horizontal gene transfer (transformation, conjugation or phage-mediated transduction). As a result, different strains of a single species can present a remarkable diversity of pathogenic phenotypes (more than 10 for *Escherichia coli*), whereas a given virulence factor can be shared between different species [114]. This mechanism explains the rapid evolution of bacterial pathogens, but also raises the question as to how the transferred genes are properly expressed after their integration into the distinct transcriptional regulatory network of the recipient cell. This problem is particularly acute for the bacterium, since any error in the expression time or strength of virulence factors immediately leads to the recognition and, ultimately, to the destruction of the invader by the host defence system [115,116].

At first glance, such drastic regulation of a few specific promoters seems to deviate from the global and non-specific regulation mode characteristic of SC. However, it appears equally incompatible with the sole action of strongly sequence-specific TFs, which would then be highly unstable during horizontal transfers between species, where these TFs are often evolutionarily distant [117]. As a matter of fact, many TFs involved in virulence indeed exhibit a weak sequence-specificity and are sensitive to the mechanical state of DNA [78], owing to an original regulatory mechanism affecting PAIs. Like other horizontally transferred regions, these usually exhibit a lower GC-content than the chromosomal average, and are therefore normally repressed by extensive binding of the NAP H-NS. Regulators can then activate the

genes without any specific contact with RNAP, by competing with H-NS for promoter binding [114,118], which can be strongly dependent on the topological state of the region.

Phylum	Family	Species	Tissue	Lifestyle	Stress encountered	Gene involved in virulence	Relax response	Ref
P	Enterobacteriaceae	*Salmonella enterica*	gastrointestinal tracts	FIP	acid	*hilD, hilC,ssrAB*	+	[91]
	Enterobacteriaceae	*Shigella flexneri*	intestinal epithelium	FIP	temperature	*virF*	+	[119]
	Enterobacteriaceae	*E. coli (EHEC)*	intestinal epithelium	FIP	temperature	*espADB*	-	[120]
	Pectobacteriaceae	*Dickeya dadantii*	plant apoplast	EP	acid,oxidative	*pelE*	-	[23]
	Pseudomanadaceae	*Pseudomonas syringae*	plant apoplast	EP	oxidative	*avrPphB*	+	[121]
	Vibrionaceae	*Vibrio cholerae*	small intestine	EP	acid	*acfA, acfD*	-	[122, 123]
	Alcaligenaceae	*Bordetella pertussis*	lung epithelial cells.	IP	temperature	*ptx*	-	[124]
	Campylobactericeae	*Campylobacter jejuni*	digestive tract	IP	temperature, pH	*momp*	-	[125]
F	Staphylococcaceae	*Staphylococcus aureus*	respiratory tract, skin	EP	osmolarity	*spa, eta*	+	[49, 126]
	Streptococcaceae	*Streptococcus pneumoniae*	respiratory tract, skin	EP	oxidative	*fatD*	-	[127]
A	Mycobacteriaceae	*Mycobacterium tuberculosis*	respiratory tract, skin	IP	oxidative	*virR, sodC*	-	[128]

Tab. 3. Virulence genes are regulated by DNA supercoiling in various pathogenic species. **Phyla**: Proteobacteria (P), firmicutes (F), actinobacteria (A). **Lifestyles**: F= facultative, I=intracellular, E=extracellular, P=pathogen. **Response** to chromosomal DNA **Relax**ation by antibiotics: repressed (-) or activated (+), with corresponding reference(s).

In most investigated species, the key signals triggering a quick activation or repression of virulence genes are precisely those environmental stress conditions that were shown to modulate the chromosome topology in various species (Tab. 3), e.g., a sharp acidity variation when *S. enterica* is transferred from the stomach to the intestine, or oxidative stress when *D. dadantii* leaves the plant apoplast. It is therefore no surprise that virulence genes from an increasing number of zoopathogenic or phytopathogenic species were shown to be directly regulated by SC, as summarized in Tab. 3. Does this mechanism play a role during the infection process, as these data suggest? In our opinion, based on the complex regulatory mechanisms illustrated above, SC is a good candidate to play the role of a basal and robust coordinator of virulence gene expression, by (1) modulating the simultaneous action of many (more specific) regulators at virulence promoters, such complexity being a characteristic feature of the latter, (2) co-regulating the adjacent genes of a PAI through the evolutionarily conserved transcription-supercoiling coupling. We present below some examples supporting this hypothesis, keeping in mind that existing results mostly concern individual genes, whereas the topological organisation of entire PAIs and its effect on their expression remains poorly understood [129].

3.2. Widespread evidence for a regulatory role of DNA supercoiling in virulence

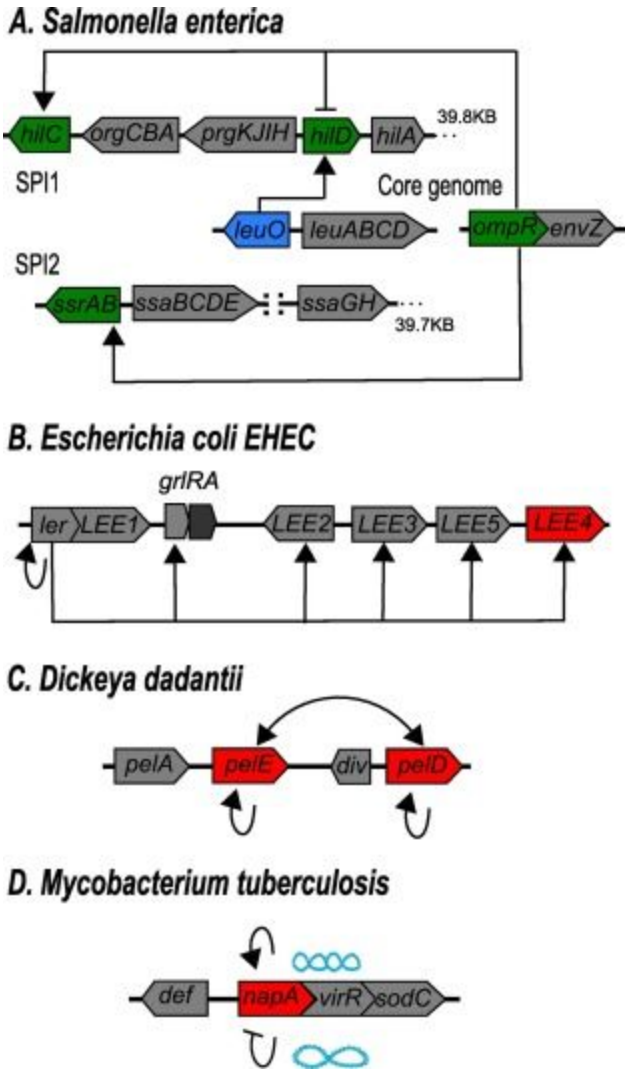


Fig. 3. Regulatory networks within several pathogenicity islands. In all species, key virulence genes are either relaxation-activated (green), relaxation-repressed (red), or regulated via the transcription-SC coupling (blue). Arrows indicate an activating (arrow) and bars a repressing effect. For *M. tuberculosis*, this effect depends on the SC level (repression on a relaxed template, activation on a supercoiled template). In *D. dadantii*, *pel* genes are self- and mutually-inductive via pectin degradation relieving the repression of both genes by the TF KdgR

Salmonella enterica is one of the most studied pathogens, and this is also true of the regulation of its virulence system by SC (mostly in C. Dorman's laboratory). Several key virulence genes within its two largest PAIs (SPI-1 and SPI-2) are supercoiling-sensitive, as well as those of the central virulence regulators OmpR and LeuO located in the core genome (Fig. 3A). In this representative example, key virulence functions were thus integrated into the

pre-existing transcription regulatory network, whereby OmpR and LeuO, primarily devoted to other functions (the former is an abundant NAP-like protein), are specifically recruited at virulence promoters in a SC-dependent manner [5,91]. Interestingly, these regulators mostly target divergently oriented operons, and the same pattern is observed in *Vibrio cholerae* (*tcpPH*) [78] and *Shigella flexneri* (*icsA*, *icsB*) [117]. While gene orientations are indifferent to classical regulatory models, divergent promoters are the ones most sensitive to the transcription-supercoiling coupling [45], which is also involved in LeuO recruitment at its own (divergent) promoter [108]. Gene orientations within the PAIs are thus likely not accidental, but rather reflect the evolved infrastructure of a local coordinated and SC-coupled gene expression.

The virulence of *E. coli* EHEC (causing severe diarrhoea) depends on a secretion system encoded in the LEE operons (Fig. 3B), including many genes shared with other *E. coli* pathogenic strains, e.g. EPEC [114]. These operons are globally repressed by H-NS, whose binding is antagonised by the activator Ler encoded in LEE1 [118]. Ler itself belongs to the H-NS family and this competition may very well be affected by the topological state of the domain. Although such an effect was not investigated in detail, the expression of LEE4 was indeed found to be SC-sensitive [120], and the same could be true of the other operons.

In the phytopathogen *D. dadantii*, pectinolytic enzymes are the main virulence factors, responsible for the soft rot symptoms [115]. These are encoded by *pel* genes scattered in several PAIs along the chromosome [20]. *pelE* and *pelD*, the two major members of this family, are paralogous genes that evolved from a unique ancestor but exhibit different expression patterns [130]. They are regulated by several NAPs (H-NS, FIS, IHF, Lrp) as well as many TFs (e.g., CRP, KdgR, PecT, PecS) but are also among the most SC-sensitive genes in the chromosome. We illustrated above how different regulatory proteins act in a complex combination, with SC modulating their relative affinities (Fig. 2B). Although no effect of locally generated SC was directly shown here, a divergent non-coding transcript (*div*) was recently identified upstream of the *pelD* promoter, which “feeds” the latter with RNAP with a strong dependence on the 3D conformation of the promoter [130], possibly involving a local transmission of SC (Fig. 3C).

Finally, since topoisomerases are conserved among all bacteria, we may expect SC to play a role in the virulence of widely distant species. This was indeed recently demonstrated in the actinobacterium *Mycobacterium tuberculosis*, where the gene of a new NAP (NapA) was identified in the same operon as major virulence factors/regulators (SodC, VirR) [128]. Interestingly, NapA autoregulates its (divergently oriented) promoter in a SC-dependent manner (Fig. 3D): in this case, global SC variations induced by environmental conditions may act as a switch, turning the operon on or off with high specificity.

Conclusion

The discussed examples show that SC plays a direct role in the regulation of virulence in many species, albeit with a remarkable variety of mechanisms involving a combination of additional regulatory actors, and these mostly remain to be characterised. A final and striking example is the remarkably small and poorly characterised tenericute *Mycoplasma pneumoniae*, which is able to infect the human respiratory tract despite being almost devoid of TFs; the ancestral regulatory action of SC is thought to play an even more cardinal role in this case [37]. The increasing interest in SC in the genetic regulation community already results in new experimental techniques facilitating the mapping of supercoil distributions at higher resolution [1,2] as well as the development of computational models at various scales of detail [45,77,101,131], which together will help elucidating the pivotal regulatory action of SC in bacterial genetic regulation advocated here.

Acknowledgements

This work has been funded by the Agence Nationale de la Recherche [ANR-18-CE45-0006-01 grant to S.M.], CNRS, INSA Lyon and Université de Lyon. We thank Georgi Muskhelishvili for his critical reading of the manuscript.

Bibliography

[1]Lal A, Dhar A, Trostel A, Kouzine F, Seshasayee ASN, Adhya S. Genome scale patterns of supercoiling in a bacterial chromosome. Nat Commun 2016;7:11055.

[2]Sutormin D, Rubanova N, Logacheva M, Ghilarov D, Severinov K. Single-nucleotide-resolution mapping of DNA gyrase cleavage sites across the Escherichia coli genome. Nucleic Acids Res 2019;47:1373–88.

[3]Rovinskiy NS, Agbleke AA, Chesnokova ON, Higgins NP. Supercoil Levels in E. coli and Salmonella Chromosomes Are Regulated by the C-Terminal 35–38 Amino Acids of GyrA. Microorganisms 2019;7:81.

[4]Travers A, Muskhelishvili G. DNA supercoiling - a global transcriptional regulator for enterobacterial growth? Nat Rev Microbiol 2005;3:157–169.

[5]Dorman CJ, Dorman MJ. DNA supercoiling is a fundamental regulatory principle in the control of bacterial gene expression. Biophys Rev 2016;8:209–20.

[6]Watson JD, Crick FH. The structure of DNA. Cold Spring Harb Symp Quant Biol 1953;18:123–31.

[7]Gilbert N, Allan J. Supercoiling in DNA and chromatin. Curr Op Gen Dev 2014;25:15–21.

[8]Forterre P, Gadelle D. Phylogenomics of DNA topoisomerases: their origin and putative roles in the emergence of modern organisms. Nucleic Acids Res 2009;37:679–92.

[9]Zawadzki P, Stracy M, Ginda K, Zawadzka K, Lesterlin C, Kapanidis AN, et al. The Localization and Action of Topoisomerase IV in Escherichia coli Chromosome

Segregation Is Coordinated by the SMC Complex, MukBEF. Cell Rep 2015;13:2587–96.
[10] Dillon SC, Dorman CJ. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. Nat Rev Microbiol 2010;8:185–95.
[11] López-García P, Forterre P. DNA topology and the thermal stress response, a tale from mesophiles and hyperthermophiles. BioEssays News Rev Mol Cell Dev Biol 2000;22:738–46.
[12] Ogata Y, Mizushima T, Kataoka K, Miki T, Sekimizu K. Identification of DNA topoisomerases involved in immediate and transient DNA relaxation induced by heat shock in Escherichia coli. Mol Gen Genet MGG 1994;244:451–5.
[13] Rohde JR, Fox JM, Minnich SA. Thermoregulation in Yersinia enterocolitica is coincident with changes in DNA supercoiling. Mol Microbiol 1994;12:187–99.
[14] Hérault E, Reverchon S, Nasser W. Role of the LysR-type transcriptional regulator PecT and DNA supercoiling in the thermoregulation of pel genes, the major virulence factors in Dickeya dadantii. Environ Microbiol 2014;16:734–45.
[15] Krispin O, Allmansberger R. Changes in DNA supertwist as a response of Bacillus subtilis towards different kinds of stress. FEMS Microbiol Lett 1995;134:129–35.
[16] Mizushima T, Kataoka K, Ogata Y, Inoue R, Sekimizu K. Increase in negative supercoiling of plasmid DNA in Escherichia coli exposed to cold shock. Mol Microbiol 1997;23:381–6.
[17] Karem K, Foster JW. The influence of DNA topology on the environmental regulation of a pH-regulated locus in Salmonella typhimurium. Mol Microbiol 1993;10:75–86.
[18] Quinn HJ, Cameron ADS, Dorman CJ. Bacterial Regulon Evolution: Distinct Responses and Roles for the Identical OmpR Proteins of Salmonella Typhimurium and Escherichia coli in the Acid Stress Response. PLoS Genet 2014;10.
[19] Colgan AM, Quinn HJ, Kary SC, Mitchenall LA, Maxwell A, Cameron ADS, et al. Negative supercoiling of DNA by gyrase is inhibited in Salmonella enterica serovar Typhimurium during adaptation to acid stress. Mol Microbiol 2018;107:734–46.
[20] Jiang X, Sobetzko P, Nasser W, Reverchon S, Muskhelishvili G. Chromosomal “Stress-Response” Domains Govern the Spatiotemporal Expression of the Bacterial Virulence Program. MBio 2015;6.
[21] Hsieh LS, Rouviere-Yaniv J, Drlica K. Bacterial DNA supercoiling and [ATP]/[ADP] ratio: changes associated with salt shock. J Bacteriol 1991;173:3914–7.
[22] Higgins CF, Dorman CJ, Stirling DA, Waddell L, Booth IR, May G, et al. A physiological role for DNA supercoiling in the osmotic regulation of gene expression in S. typhimurium and E. coli. Cell 1988;52:569–84.
[23] Ouafa Z-A, Reverchon S, Lautier T, Muskhelishvili G, Nasser W. The nucleoid-associated proteins H-NS and FIS modulate the DNA supercoiling response of the pel genes, the major virulence factors in the plant pathogen bacterium Dickeya dadantii. Nucl Ac Res 2012;40:4306–4319.

[24] Sheehan BJ, Foster TJ, Dorman CJ, Park S, Stewart GS. Osmotic and growth-phase dependent regulation of the eta gene of Staphylococcus aureus: a role for DNA supercoiling. Mol Gen Genet MGG 1992;232:49–57.
[25] Ali N, Herron PR, Evans MC, Dyson PJ. Osmotic regulation of the Streptomyces lividans thiostrepton-inducible promoter, ptipA. Microbiol Read Engl 2002;148:381–90.
[26] Weinstein-Fischer D, Elgrably-Weiss M, Altuvia S. Escherichia coli response to hydrogen peroxide: a role for DNA supercoiling, topoisomerase I and Fis. Mol Microbiol 2000;35:1413–20.
[27] Hsieh LS, Burger RM, Drlica K. Bacterial DNA supercoiling and [ATP]/[ADP]. Changes associated with a transition to anaerobic growth. J Mol Biol 1991;219:443–50.
[28] Cameron ADS, Stoebel DM, Dorman CJ. DNA supercoiling is differentially regulated by environmental factors and FIS in Escherichia coli and Salmonella enterica. Mol Microbiol 2011;80:85–101.
[29] Balke VL, Gralla JD. Changes in the linking number of supercoiled DNA accompany growth transitions in Escherichia coli. J Bacteriol 1987;169:4499–506.
[30] Hatfield GW, Benham CJ. DNA topology-mediated control of global gene expression in Escherichia coli. Annu Rev Gen 2002;36:175–203.
[31] Zhou Q, Gomez Hernandez ME, Fernandez-Lima F, Tse-Dinh Y-C. Biochemical Basis of E. coli Topoisomerase I Relaxation Activity Reduction by Nonenzymatic Lysine Acetylation. Int J Mol Sci 2018;19.
[32] Pommier Y, Leo E, Zhang H, Marchand C. DNA Topoisomerases and Their Poisoning by Anticancer and Antibacterial Drugs. Chem Biol 2010;17:421–33.
[33] de la Campa AG, Ferrándiz MJ, Martín-Galiano AJ, García MT, Tirado-Vélez JM. The Transcriptome of Streptococcus pneumoniae Induced by Local and Global Changes in Supercoiling. Front Microbiol 2017;8.
[34] Leblanc BP, Benham CJ, Clark DJ. An initiation element in the yeast CUP1 promoter is recognized by RNA polymerase II in the absence of TATA box-binding protein if the DNA is negatively supercoiled. Proc Natl Acad Sci U S A 2000;97:10745–10750.
[35] Tenaillon O, Barrick JE, Ribeck N, Deatherage DE, Blanchard JL, Dasgupta A, et al. Tempo and mode of genome evolution in a 50,000-generation experiment. Nature 2016;536:165–70.
[36] Crozat E, Philippe N, Lenski RE, Geiselmann J, Schneider D. Long-term experimental evolution in Escherichia coli. XII. DNA topology as a key target of selection. Genetics 2005;169:523–532.
[37] Junier I, Unal EB, Yus E, Lloréns-Rico V, Serrano L. Insights into the Mechanisms of Basal Coordination of Transcription Using a Genome-Reduced Bacterium. Cell Syst 2016;2:391–401.
[38] Dorman CJ. Regulation of transcription by DNA supercoiling in Mycoplasma genitalium: global control in the smallest known self-replicating genome. Mol Microbiol

2011;81:302–4.
[39] Brinza L, Calevro F, Charles H. Genomic analysis of the regulatory elements and links with intrinsic DNA structural properties in the shrunken genome of Buchnera. BMC Genomics 2013;14:73.
[40] Jeong KS, Xie Y, Hiasa H, Khodursky AB. Analysis of Pleiotropic Transcriptional Profiles: A Case Study of DNA Gyrase Inhibition. PLOS Genet 2006;2:e152.
[41] Cheung KJ, Badarinarayana V, Selinger DW, Janse D, Church GM. A microarray-based antibiotic screen identifies a regulatory role for supercoiling in the osmotic stress response of Escherichia coli. Genome Res 2003;13:206–15.
[42] Blot N, Mavathur R, Geertz M, Travers A, Muskhelishvili G. Homeostatic regulation of supercoiling sensitivity coordinates transcription of the bacterial genome. EMBO Rep 2006;7:710–715.
[43] Peter BJ, Arsuaga J, Breier AM, Khodursky AB, Brown PO, Cozzarelli NR. Genomic transcriptional response to loss of chromosomal supercoiling in Escherichia coli. Genome Biol 2004;5:R87.
[44] Webber MA, Ricci V, Whitehead R, Patel M, Fookes M, Ivens A, et al. Clinically relevant mutant DNA gyrase alters supercoiling, changes the transcriptome, and confers multidrug resistance. MBio 2013;4.
[45] El Houdaigui B, Forquet R, Hindré T, Schneider D, Nasser W, Reverchon S, et al. Bacterial genome architecture shapes global transcriptional regulation by DNA supercoiling. Nucleic Acids Res 2019;47:5648–57.
[46] Gmuender H, Kuratli K, Di Padova K, Gray CP, Keck W, Evers S. Gene Expression Changes Triggered by Exposure of Haemophilus influenzae to Novobiocin or Ciprofloxacin: Combined Transcription and Translation Analysis. Genome Res 2001;11:28–42.
[47] Ferrandiz M-J, Martin-Galiano AJ, Schvartzman JB, de la Campa AG. The genome of Streptococcus pneumoniae is organized in topology-reacting gene clusters. Nucleic Acids Res 2010;38:3570–81.
[48] Ferrándiz M-J, Martín-Galiano AJ, Arnanz C, Camacho-Soguero I, Tirado-Vélez J-M, de la Campa AG. An increase in negative supercoiling in bacteria reveals topology-reacting gene clusters and a homeostatic response mediated by the DNA topoisomerase I gene. Nucleic Acids Res 2016;44:7292–303.
[49] Schröder W, Bernhardt J, Marincola G, Klein-Hitpass L, Herbig A, Krupp G, et al. Altering gene expression by aminocoumarins: the role of DNA supercoiling in Staphylococcus aureus. BMC Genomics 2014;15:291.
[50] Sioud M, Boudabous A, Cekaite L. Transcriptional responses of Bacillus subtillis and thuringiensis to antibiotics and anti-tumour drugs. Int J Mol Med 2009;23:33–9.
[51] Szafran MJ, Gongerowska M, Małecki T, Elliot MA, Jakimowicz D. Transcriptional response of Streptomyces coelicolor to rapid chromosome relaxation or long-term supercoiling imbalance. BioRxiv 2019:602359.

[52] Boshoff HIM, Myers TG, Copp BR, McNeil MR, Wilson MA, Barry CE. The transcriptional responses of Mycobacterium tuberculosis to inhibitors of metabolism: novel insights into drug mechanisms of action. J Biol Chem 2004;279:40174–84.
[53] Prakash JSS, Sinetova M, Zorina A, Kupriyanova E, Suzuki I, Murata N, et al. DNA supercoiling regulates the stress-inducible expression of genes in the cyanobacterium Synechocystis. Mol Biosyst 2009;5:1904–12.
[54] Vijayan V, Zuzow R, O'Shea EK. Oscillations in supercoiling drive circadian gene expression in cyanobacteria. Proc Natl Acad Sci 2009;106:22564–22568.
[55] Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, et al. Transcriptional regulation by the numbers: models. Curr Opin Genet Dev 2005;15:116–124.
[56] Marr C, Geertz M, Hütt M-T, Muskhelishvili G. Dissecting the logical types of network control in gene expression profiles. BMC Syst Biol 2008;2:18.
[57] Borukhov S, Nudler E. RNA polymerase: the vehicle of transcription. Trends Microbiol 2008;16:126–134.
[58] Forterre P. A hot story from comparative genomics: reverse gyrase is the only hyperthermophile-specific protein. Trends Genet 2002;18:236–7.
[59] Meyer S, Jost D, Theodorakopoulos N, Peyrard M, Lavery R, Everaers R. Temperature Dependence of the DNA Double Helix at the Nanoscale: Structure, Elasticity, and Fluctuations. Biophys J 2013;105:1904–1914.
[60] Du X, Wojtowicz D, Bowers AA, Levens D, Benham CJ, Przytycka TM. The genome-wide distribution of non-B DNA motifs is shaped by operon structure and suggests the transcriptional importance of non-B DNA structures in Escherichia coli. Nucleic Acids Res 2013;41:5965–77.
[61] Jost D. Twist-DNA: computing base-pair and bubble opening probabilities in genomic superhelical DNA. Bioinformatics 2013;29:2479–81.
[62] Alexandrov BS, Gelev V, Yoo SW, Alexandrov LB, Fukuyo Y, Bishop AR, et al. DNA dynamics play a role as a basal transcription factor in the positioning and regulation of gene transcription initiation. Nucleic Acids Res 2010;38:1790–1795.
[63] Travers AA. Promoter Sequence for Stringent Control of Bacterial Ribonucleic Acid Synthesis. J Bacteriol 1980;141:973–6.
[64] Pemberton IK, Muskhelishvili G, Travers AA, Buckle M. The G+C-rich discriminator region of the tyrT promoter antagonises the formation of stable preinitiation complexes11Edited by M. Yaniv. J Mol Biol 2000;299:859–64.
[65] Auner H, Buckle M, Deufel A, Kutateladze T, Lazarus L, Mavathur R, et al. Mechanism of Transcriptional Activation by FIS: Role of Core Promoter Structure and {DNA} Topology. J Mol Biol 2003;331:331–44.
[66] Henderson KL, Felth LC, Molzahn CM, Shkel I, Wang S, Chhabra M, et al. Mechanism of transcription initiation and promoter escape by E. coli RNA polymerase. Proc Natl Acad Sci 2017;114:E3032–40.

[67] Wood DC, Lebowitz J. Effect of supercoiling on the abortive initiation kinetics of the RNA-I promoter of ColE1 plasmid DNA. J Biol Chem 1984;259:11184–7.
[68] Lim HM, Lewis DEA, Lee HJ, Liu M, Adhya S. Effect of Varying the Supercoiling of DNA on Transcription and Its Regulation. Biochemistry 2003;42:10718–25.
[69] Menzel R, Gellert M. Regulation of the genes for E. coli DNA gyrase: Homeostatic control of DNA supercoiling. Cell 1983;34:105–13.
[70] Dages S, Dages K, Zhi X, Leng F. Inhibition of the gyrA promoter by transcription-coupled DNA supercoiling in Escherichia coli. Sci Rep 2018;8:14759.
[71] Jha RK, Tare P, Nagaraja V. Regulation of the gyr operon of Mycobacterium tuberculosis by overlapping promoters, DNA topology, and reiterative transcription. Biochem Biophys Res Commun 2018;501:877–84.
[72] Lautier T, Blot N, Muskhelishvili G, Nasser W. Integration of two essential virulence modulating signals at the Erwinia chrysanthemi pel gene promoters: a role for Fis in the growth-phase regulation. Mol Microbiol 2007;66:1491–505.
[73] Wang JY, Syvanen M. DNA twist as a transcriptional sensor for environmental changes. Mol Microbiol 1992;6:1861–6.
[74] Unniraman S, Nagaraja V. Axial distortion as a sensor of supercoil changes: A molecular model for the homeostatic regulation of DNA gyrase. J Genet 2001;80:119–24.
[75] Ma J, Bai L, Wang MD. Transcription under torsion. Science 2013;340:1580–1583.
[76] Chen H, Shiroguchi K, Ge H, Xie XS. Genome-wide study of mRNA degradation and transcript elongation in Escherichia coli. Mol Syst Biol 2015;11:808.
[77] Sevier SA, Levine H. Properties of gene expression and chromatin structure with mechanically regulated elongation. Nucleic Acids Res 2018;46:5924–34.
[78] Dorman CJ, Dorman MJ. Control of virulence gene transcription by indirect readout in Vibrio cholerae and Salmonella enterica serovar Typhimurium. Environ Microbiol 2017;19:3834–45.
[79] Cevost J, Vaillant C, Meyer S. ThreaDNA: predicting DNA mechanics' contribution to sequence selectivity of proteins along whole genomes. Bioinformatics 2018;34:609–16.
[80] Goldstein E, Drlica K. Regulation of bacterial DNA supercoiling: plasmid linking numbers vary with growth temperature. Proc Natl Acad Sci U S A 1984;81:4046–50.
[81] Pérez-Rueda E, Collado-Vides J. The repertoire of DNA-binding transcriptional regulators in Escherichia coli K-12. Nucleic Acids Res 2000;28:1838–47.
[82] Hommais F, Oger-Desfeux C, Gijsegem FV, Castang S, Ligori S, Expert D, et al. PecS Is a Global Regulator of the Symptomatic Phase in the Phytopathogenic Bacterium Erwinia chrysanthemi 3937. J Bacteriol 2008;190:7508–22.
[83] Harteis S, Schneider S. Making the bend: DNA tertiary structure and protein-DNA interactions. Int J Mol Sc 2014;15:12335–12363.
[84] Parkhill J, Brown NL. Site-specific insertion and deletion mutants in the mer

promoter-operator region of Tn501; the nineteen base-pair spacer is essential for normal induction of the promoter by MerR. Nucleic Acids Res 1990;18:5157–62.
[85] Travers A, Muskhelishvili G. A common topology for bacterial and eukaryotic transcription initiation? EMBO Rep 2007;8:147–51.
[86] Teif VB. General transfer matrix formalism to calculate DNA–protein–drug binding in gene regulation: application to OR operator of phage λ. Nucleic Acids Res 2007;35:e80.
[87] Xiao B, Zhang H, Johnson RC, Marko JF. Force-driven unbinding of proteins HU and Fis from DNA quantified using a thermodynamic Maxwell relation. Nucleic Acids Res 2011;39:5568–77.
[88] Efremov AK, Yan J. Transfer-matrix calculations of the effects of tension and torque constraints on DNA-protein interactions. Nucleic Acids Res 2018;46:6504–27.
[89] Yan J, Marko JF. Localized Single-Stranded Bubble Mechanism for Cyclization of Short Double Helix DNA. Phys Rev Lett 2004;93:108108.
[90] Du Q, Smith C, Shiffeldrim N, Vologodskaia M, Vologodskii A. Cyclization of short DNA fragments and bending fluctuations of the double helix. Proc Natl Acad Sci 2005;102:5397–402.
[91] Cameron ADS, Dorman CJ. A Fundamental Regulatory Mechanism Operating through OmpR and DNA Topology Controls Expression of Salmonella Pathogenicity Islands SPI-1 and SPI-2. PLOS Genet 2012;8:e1002615.
[92] Zhabinskaya D, Benham CJ. Theoretical analysis of competing conformational transitions in superhelical DNA. PLoS Comput Biol 2012;8:e1002484.
[93] Kouzine F, Sanford S, Elisha-Feil Z, Levens D. The functional response of upstream DNA to dynamic supercoiling in vivo. Nat Struct Mol Biol 2008;15:146–154.
[94] Peck LJ, Wang JC. Transcriptional block caused by a negative supercoiling induced structural change in an alternating CG sequence. Cell 1985;40:129–37.
[95] Conway T, Creecy JP, Maddox SM, Grissom JE, Conkle TL, Shadid TM, et al. Unprecedented High-Resolution View of Bacterial Operon Architecture Revealed by RNA Sequencing. MBio 2014;5:e01442-14.
[96] Liu LF, Wang JC. Supercoiling of the DNA template during transcription. Proc Natl Acad Sci 1987;84:7024–7027.
[97] Chong S, Chen C, Ge H, Xie XS. Mechanism of Transcriptional Bursting in Bacteria. Cell 2014;158:314–26.
[98] Elowitz MB, Levine AJ, Siggia ED, Swain PS. Stochastic gene expression in a single cell. Science 2002;297:1183–1186.
[99] Meyer S, Beslon G. Torsion-Mediated Interaction between Adjacent Genes. PLoS Comput Biol 2014;10:e1003785.
[100] Brackley CA, Johnson J, Bentivoglio A, Corless S, Gilbert N, Gonnella G, et al. Stochastic Model of Supercoiling-Dependent Transcription. Phys Rev Lett 2016;117:018101.

[101] Ancona M, Bentivoglio A, Brackley CA, Gonnella G, Marenduzzo D. Transcriptional bursts in a non-equilibrium model for gene regulation by supercoiling. Biophys J 2019.
[102] Kouzine F, Gupta A, Baranello L, Wojtowicz D, Ben-Aissa K, Liu J, et al. Transcription-dependent dynamic supercoiling is a short-range genomic force. Nat Struct Mol Biol 2013;20:396–403.
[103] Ahmed W, Sala C, Hegde SR, Jha RK, Cole ST, Nagaraja V. Transcription facilitated genome-wide recruitment of topoisomerase I and DNA gyrase. PLoS Genet 2017;13:e1006754.
[104] Sobetzko P. Transcription-coupled DNA supercoiling dictates the chromosomal arrangement of bacterial genes. Nucleic Acids Res 2016;44:1514–24.
[105] Junier I, Rivoire O. Conserved Units of Co-Expression in Bacterial Genomes: An Evolutionary Insight into Transcriptional Regulation. PloS One 2016;11:e0155740.
[106] Opel ML, Hatfield G. DNA supercoiling-dependent transcriptional coupling between the divergently transcribed promoters of the ilvYC operon of Escherichia coli is proportional to promoter strengths and transcript lengths. Mol Microbiol 2001;39:191–198.
[107] Wu H-Y, Tan J, Fang M. Long-range interaction between two promoters: Activation of the leu-500 promoter by a distant upstream promoter. Cell 1995;82:445–51.
[108] Fang M, Wu H-Y. A Promoter Relay Mechanism for Sequential Gene Activation. J Bacteriol 1998;180:626–33.
[109] Postow L, Hardy CD, Arsuaga J, Cozzarelli NR. Topological domain structure of the Escherichia coli chromosome. Genes Dev 2004;18:1766–79.
[110] Gerganova V, Berger M, Zaldastanishvili E, Sobetzko P, Lafon C, Mourez M, et al. Chromosomal position shift of a regulatory gene alters the bacterial phenotype. Nucleic Acids Res 2015;43:8215–26.
[111] Brambilla E, Sclavi B. Gene Regulation by H-NS as a Function of Growth Conditions Depends on Chromosomal Position in Escherichia coli. G3 Genes Genomes Genet 2015;5:605–14.
[112] Ferrándiz M-J, Arnanz C, Martín-Galiano AJ, Rodríguez-Martín C, Campa AG de la. Role of Global and Local Topology in the Regulation of Gene Expression in Streptococcus pneumoniae. PLOS ONE 2014;9:e101574.
[113] Junier I, Frémont P, Rivoire O. Universal and idiosyncratic characteristic lengths in bacterial genomes. Phys Biol 2018;15:035001.
[114] Schmidt H, Hensel M. Pathogenicity islands in bacterial pathogenesis. Clin Microbiol Rev 2004;17:14–56.
[115] Leonard S, Hommais F, Nasser W, Reverchon S. Plant-phytopathogen interactions: bacterial responses to environmental and plant stimuli. Environ Microbiol 2017;19:1689–716.
[116] Stecher B, Hardt W-D. Mechanisms controlling pathogen colonization of the gut.

Curr Opin Microbiol 2011;14:82–91.
[117] Dorman CJ. Virulence Gene Regulation in Shigella. EcoSal Plus 2004;1.
[118] Connolly JPR, Finlay BB, Roe AJ. From ingestion to colonization: the influence of the host environment on regulation of the LEE encoded type III secretion system in enterohaemorrhagic Escherichia coli. Front Microbiol 2015;6.
[119] Dorman CJ, Bhriain NN, Higgins CF. DNA supercoiling and environmental regulation of virulence gene expression in Shigella flexneri. Nature 1990;344:789.
[120] Beltrametti F, Kresse AU, Guzmán CA. Transcriptional Regulation of the espGenes of Enterohemorrhagic Escherichia coli. J Bacteriol 1999;181:3409–18.
[121] Neale HC, Jackson RW, Preston GM, Arnold DL. Supercoiling of an excised genomic island represses effector gene expression to prevent activation of host resistance. Mol Microbiol 2018;110:444–54.
[122] Matson JS, Withey JH, DiRita VJ. Regulatory Networks Controlling Vibrio cholerae Virulence Gene Expression. Infect Immun 2007;75:5542–9.
[123] Parsot C, Mekalanos JJ. Structural analysis of the acfA and acfD genes of Vibrio cholerae: effects of DNA topology and transcriptional activators on expression. J Bacteriol 1992;174:5211–8.
[124] Graeff-Wohlleben H, Deppisch H, Gross R. Global regulatory mechanisms affect virulence gene expression in Bordetella pertussis. Mol Gen Genet MGG 1995;247:86–94.
[125] Dedieu L, Pages J-M, Bolla J-M. Environmental Regulation of Campylobacter jejuni Major Outer Membrane Protein Porin Expression in Escherichia coli Monitored by Using Green Fluorescent Protein. Appl Environ Microbiol 2002;68:4209–15.
[126] Fournier B. Protein A gene expression is regulated by DNA supercoiling which is modified by the ArlS-ArlR two-component system of Staphylococcus aureus. Microbiology 2004;150:3807–19.
[127] Ferrándiz M-J, de la Campa AG. The Fluoroquinolone Levofloxacin Triggers the Transcriptional Activation of Iron Transport Genes That Contribute to Cell Death in Streptococcus pneumoniae. Antimicrob Agents Chemother 2014;58:247–57.
[128] Datta C, Jha RK, Ganguly S, Nagaraja V. NapA (Rv0430), a Novel Nucleoid-Associated Protein that Regulates a Virulence Operon in Mycobacterium tuberculosis in a Supercoiling-Dependent Manner. J Mol Biol 2019;431:1576–91.
[129] Martín-Galiano AJ, Ferrándiz MJ, de la Campa AG. Bridging Chromosomal Architecture and Pathophysiology of Streptococcus pneumoniae. Genome Biol Evol 2017;9:350–61.
[130] Duprey A, Nasser W, Léonard S, Brochier-Armanet C, Reverchon S. Transcriptional start site turnover in the evolution of bacterial paralogous genes - the pelE-pelD virulence genes in Dickeya. FEBS J 2016.
[131] Lepage T, Junier I. A polymer model of bacterial supercoiled DNA including structural transitions of the double helix. Phys A 2019:121196.

2.Characterizing Seconeolitsin and studying its effect on virulence genes

a) Introduction

Looking into the aspects of expression of virulence genes, the transcription factors are already studied in the previous chapters. Further, questioning the very process of virulence gene expression, we learn that their expression should be a very quick reaction to fight the host defense barrier and establish pathogenesis. The virulence action in bacteria has to be rapid and energy efficient. Thus, we advocate DNA supercoiling to be a suitable candidate to be a factor responsible in regulating virulence genes globally in response to stress caused by host defense mechanisms [1].

DNA supercoiling is the over or under-winding of DNA under the effect of torsional stress. Supercoiling is a fundamental property of DNA. DNA in its relaxed state adopts a right-handed helical coiled conformation, the detailed structure of which is dependent on the localized sequence [2]. Double-stranded DNA typically consists of a helix with a pitch of one turn per 10.5 base pairs. The number of turns varies with respect to the level of supercoiling observed in the double-stranded DNA. Winding DNA around its axis introduces supercoils, winding in the same direction as the helix introduces positive supercoiling whereas winding in the opposite direction generates negative supercoiling [3]. Supercoiling is modulated by topoisomerase activities and NAPs (Nucleoid Associated Proteins). Changes in the action of DNA topoisomerases cause variation in global supercoiling level whereas NAPs help in organizing local superhelical structure of DNA around its position. Thus, DNA topoisomerases play a key role on varying global level superhelicity of DNA and studying modulation of supercoiling by them will enable us to understand the effect of change of supercoiling at the global scale and its effect on transcription which results in alterations in gene expression. In order to understand DNA supercoiling and its role in virulence gene expression, it is important to study the modulation of superhelicity caused by topoisomerases.

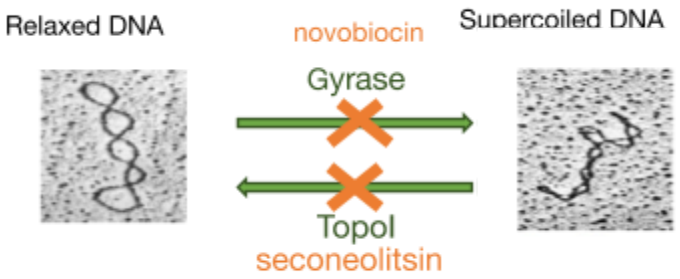


Figure 3.1.1: DNA in it's supercoiled and relaxed state. DNA Gyrase facilitates supercoiling and DNA Topoisomerases relaxes the DNA

Topoisomerases are classified into two types namely type I and type II. The type I topoisomerases cut on one of the two strands of DNA, relaxes the strand and reanneals it, whereas type II cuts both strands of the DNA helix simultaneously in order to remove tangles and supercoils. Unlike the type I topoisomerases, the type II topoisomerases derive their energy from ATP hydrolysis. In order to understand the dynamics of DNA supercoiling we choose to study the enzymes topoisomerase I, a type I topoisomerase and DNA gyrase, which is a type II topoisomerase [4]. DNA gyrase supercoils DNA by adding negative supercoils in DNA and DNA topoisomerase I relaxes DNA by removing negative supercoils. These two enzymes act in an antagonistic manner comprising the global supercoiling level of chromosome as shown in Figure 3.1.1. In order to study the effect of variation in superhelical state on gene expression, the mechanism of action of topoisomerases need to be studied. Inhibition of topoisomerase leads to variation in the level of supercoiling. Antibiotics are used to inhibit these enzymes and study their course of action. Commercially most used topoisomerase inhibiting antibiotics are Quinolones, the most studied examples are Ciprofloxacin and Norfloxacin [5]. They have DNA gyrase and topoisomerase IV inhibitory property resulting in blocking the relaxation of supercoiled DNA. Relaxation of DNA is required for transcription and replication. In bacteria they act by the rapid inhibition of bacterial DNA synthesis, leading to cell death [6]. Aminocoumarins are also studied for their topoisomerase inhibition for example Novobiocin and Coumermycin. They Inhibit DNA gyrase enzyme involved in relaxing the supercoiled bacterial DNA [7]. In this study, we use seconeolitsin, an antibiotic known to inhibit topoisomerase I enzyme. Seconeolitsin, an alkaloid compound, is known to inhibit the relaxation activity of topoisomerase I enzyme in gram-positive bacterium *Streptococcus pneumoniae* [8]. Gyrase inhibitors like novobiocin result in DNA relaxation, whereas introduction of seconeolitsin should inhibit topoisomerase I and accumulate negative supercoils, resulting in increasing the level of DNA supercoiling as shown in Figure 3.1.1. What changes are brought about in the

bacteria due to the variation in the level of supercoiling caused by topoisomerase inhibitors? Do they affect virulence gene expression as we hypothesize?

Though transcription factors are known to control the transcription, DNA supercoiling has been shown to play a vital role in transcriptional regulation in bacteria. In *Escherichia coli,* more than half of the genes are known to be sensitive to DNA supercoiling [9], advocating it to be a central regulatory factor [10]. Thus, DNA supercoiling functions as an important regulator of prokaryotic transcription [11]. To prove this, novobiocin, a gyrase inhibitor antibiotic, is usually used to study effects of variation in the level of superhelicity in *Dickeya dadantii* and its effect on gene expression. Any changes in gene expression in the presence of novobiocin is hypothesized to be the result of variation in global supercoiling levels caused by novobiocin [12]. There could also be other mechanisms playing a role in the variations caused. In transcriptomic analysis a large pool of genes were significantly activated or repressed in response to DNA relaxation. Thus, proving the sensitivity of genes to the supercoiling level and transcriptional regulation by the relaxation of the DNA.

In our study, we look at the effect of Seceoneolitsin in the model organism *Dickeya dadantii*, a gram-negative pectin degrading plant phytopathogen studied vastly and is one of the major threats to plants [13]. As stated earlier, we know seconeolitsin is expected to supercoil the DNA by inhibiting the toposomerase I enzyme. Since we are interested in studying the effect of supercoiling on virulence genes we choose the *pelE* gene to be our model gene. *pelE* is a major virulence gene in the bacteria *Dickeya dadantii* which is essential for the pathogenesis and growth of the bacteria in pectin-rich environments. *pelE* is already known to be sensitive to superhelicity based on an in-vitro study of promoter sensitivity with respect to change in superhelicity. Figure 3.1.2 demonstrates the *pelE* promoter activity sensitive to variation supercoiling level. The relative promoter activity of *pelE* gene increases with increase in superhelicity. The in-vivo novobiocin treatment studies in *Dickeya dadantii* has also proved the sensitivity of *pelE* to DNA supercoiling levels. The *pelE* expression decreases up on the novobiocin treatment, DNA relaxation has repressed the expression of virulence gene *pelE* which is coherent with the in vitro studies [12].

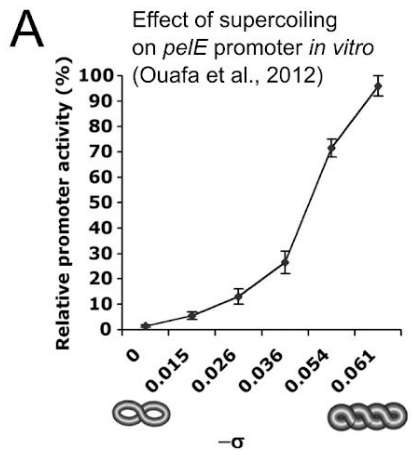


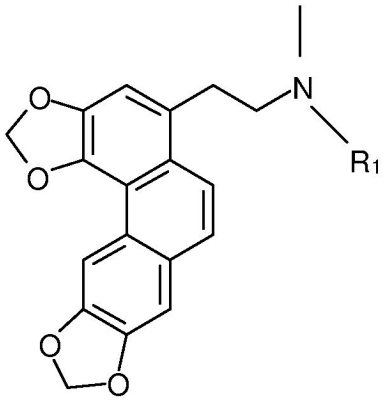
Figure 3.1.2: Effect of supercoiling on the *pelE* promoter studied in vitro. The relative promoter activity increases with increase in the supercoiling levels [12]

The effect of relaxed DNA on transcription in gram-negative bacteria has been studied using novobiocin as a gyrase inhibitor [14]. Seconeolitsin is previously studied only in gram-positive bacteria. This antibiotic is being tested on gram-negative bacteria for the first time in this work. Given the difference in the cell wall of gram-negative and gram positive bacteria, how efficiently will the antibiotic be able to get into the gram-negative cytoplasm crossing the thicker cell wall? Will the concentration of antibiotics required to inhibit DNA topoisomerases be the same as gram-positive bacteria? Upon answering these questions we can find out the effect of seconeolitsin on the supercoiling levels in bacteria. Once we are successful at modifying the supercoiling levels, we will be able to study the effect of increased supercoiling levels on transcription and expression of virulence gene *pelE.*

We know negative supercoiling favors the unwinding of the DNA double helix that is required for formation of the open complex, it is expected to increase the rate of transcription [15]. The superhelical state of DNA depends on the growth phase of cells [16]. The level of negative supercoiling is known to be decreased in the stationary phase compared to exponential growth phase [17]. Thus, the study needs to be conducted during the exponential and stationary growth phase of bacteria to study the growth phase dependencies. DNA supercoiling is also known to change transiently during shock conditions [18]. The heat stress induced in *Escherichia coli* leads to rapid relaxation of negative super-coils and then DNA topology returns back to the original state with negative supercoiling [19]. Acidic and oxidative shocks reduced the superhelical density of the DNA in growing *Dickeya dadantii* cells [12]. In this study, we have looked into the effect of DNA Topoisomerases inhibitor during growth and in

antibiotic shock state during exponential phase around 5 hours through growth and stationary phase around 10 hours during growth.

Using seconeolistin to study the effect of supercoiling on virulence gene expression



where R1 = H

Figure 3.1.3: Molecular structure of seconeolitsin (patent number WO2011073479A1) [20]

Seconeolitsin

Seconeolitsin with molecular formula $C_9H_{17}N0_4$ is a phenanthrene alkaloid compound with molecular weight 323.3g/mol produced from boldine. The melting point is found to be between 111-113°C. Seconeolitsin has been effective in inhibiting topoisomerase I as the target molecule. In order to be qualified as a potential antibiotic the compound should not only have an effective inhibition activity against the target molecule but also an effective antibacterial activity reflected as a minimal value of minimal inhibitory concentration (MIC). Seconeolitsin is synthesized as a potential candidate to inhibit DNA topoisomerase I. It has been studied in the gram-positive bacterium *Streptococcus pneumoniae* for the same and the MIC was observed to be 8-17μM. Seconeolitsin is found effective in the inhibition of both topoisomerase I activity and of cell growth at 17μM concentration [8]. Hyper-negative supercoiling was observed in an internal plasmid when treated with seconeolitsin. Seconeolitsin did not affect human cell viability. Hence, it qualifies as a good candidate for therapeutic drugs. In addition to studying the hyper-negative supercoiling caused by seconeolitsin, we also exploit the potential of seconeolitsin as an antibiotic to be used in gram-negative bacteria. The model organism studied is *Dickeya dadantii, a* plant pathogen. Since we already know about the gene responding to DNA relaxation in *Dickeya dadantii* we could also use the compound seconeolitsin to study the effect of hyper-negative supercoiling on gene expression. The combined strategy of studying the effect of probable antibiotic candidate seconeolitsin in gram-negative organisms and understanding the effect of

hyper-negative supercoiling on the gene expression level is being employed for the first time. This should be able to give us insight into using supercoiling variable antibiotics as drugs and their effect on the virulence gene expression which in turn will control pathogenesis process.

b) Materials and Methods

Process of synthesizing Seconeolitsin [20]

The synthesis of seconeolitsin was carried out in the ICBMS laboratory by the Organic and Bioorganic chemistry team. I am sincerely thankful to Lucie, Laurent and Florence for their support and help. The following methodology described in the patent for seconeolitsin was used for synthesis.

Boldine is the starting compound used for the synthesis of phenanthrenic alkaloid seconeolitsin. O-demethylation of 425mg of boldine is carried out by treating with 48% Hydrogen Bromide and glacial acetic acid under stirring condition and reflux at 130°C leading to the formation of 1,2,9,10-tetrahydroxy-N-methyl-aporphine bromide precipitates in the reaction crude. Then 850mg of this compound is dissolved in 15ml DMSO (dimethyl sulfoxide) with excess NaOH (700mg). This mixture is stirred for 90 minutes under the N_2 atmosphere at room temperature. After 90 minutes, add 0.32ml of dibromomethane and heat the mixture at 80°C for 4 hours. Then, allow it to cool and liquid/liquid extraction is carried out with dichloromethane (DCM) followed by washing with H_2O to remove traces of DMSO. Compound is purified by column chromatography using ethyl acetate as the mobile phase. Finally, the product neolitsin is obtained which crystallizes in DCM in the form of yellow needles. In order to synthesize seconeolitsin, 133mg of neolitsin is dissolved in dichloromethane and maintained at 0°C for 15 minutes. Then 1 part equivalent of α-chloroethyl chloroformate (ACE-Cl) is added, heated to reflux for 1 hour. The reaction crude is concentrated in vacuum and then dissolved in MeOH. Following is heated again for 45 min at reflux leading to the production of dry concentrated protonated form of seconeolitsin. The product is basified in order to obtain pH 10-11 base state. It is then extracted with DCM and purified by column chromatography using DCM / MeOH 9:1 mobile phase producing 80mg seconeolitsin.

The procurement of seconeolitsin

Seconeolitsin obtained in paste form initially was stored at -20°C. Aliquots of 10mM, 5mM were prepared and stored for the antibiotic treatment on *Dickeya dadantii*. Stock solution was preserved in 4°C. Encountering precipitation problems with the paste form, in order to overcome this problem seconeolitsin was procured in powdered form. Required weight of seconeolitsin is weighed and dissolved in DMSO to prepare the stock solution. Stock solution is stored at -20°C. Suspecting the possibility of any

deterioration in the activity of seconeolitsin while storing in the solution form, we decided to prepare a fresh solution of seconeolitsin as and when required for the experiments. 1.6 grams of seconeolitsin powder is dissolved in 1mL of DMSO solution to prepare 5mM stock. The stock solution is further diluted with DMSO if needed as per the requirement of seconeolitsin concentration in the experiments.

Since DMSO is toxic to the growth of bacterial cells, an attempt was made to reduce the concentration of DMSO dispensed into the medium along with seconeolitsin. 50% v/v solution of DMSO in distilled water is used to dissolve seconeolitsin. This attempt was unsuccessful since seconeolitsin is insoluble in water and would not dissolve completely in DMSO water mixture. Eventually care was taken to optimize the amount of DMSO dispensed limiting its effect on cell viability, which is explained in the coming parts.

Culture medium and bacterial strains

Dickeya dadantii wild type strain A4922 was cultured using standard conditions in M63 minimal medium supplemented with Glucose (0.2% v/v) and LB medium, when needed with the addition of PGA as the pectin source (8% v/v) and at 30°C. Gram-positive *Bacillus subtilis* strain B168 was used to test the procured seconeolitsin. NM522 is the *Escherichia coli* strain used for the study of the effect of seconeolitsin on DNA supercoiling level.

Growth kinetics monitoring with absorbance spectroscopy

The advanced absorbance microplate reader platform on the Tecan SPARK machine is used to study the growth kinetics of the bacteria. 96 well plates are used for the purpose in order to accumulate multiple samples data in the same growth conditions in order to improve the quality of data with respect to the uniformity of the growth conditions. 96 well plate plan is designed as per experimental need and the medium with required components are dispensed into these wells.16 hour old pre-culture of the organism is used to inoculate the media. The starting optical density (OD) is pre-calculated to be around 0.03 in each well. OD at 600nm wavelength and Luminescence (for *pel* expression study) is monitored every 5 min for a duration of 24 hours and recorded in the setup. The cultures are grown with agitation. The values are then obtained in the form of excel sheets. All the values are plotted with respect to time values to observe the variation in the values with respect to time and analyse the results. The experimental output data is processed to remove the noise and plots are generated by the self developed python tool. The details are discussed further in the method of data treatment section.

Modeling the growth curve to deduce the growth parameters

Bacterial growth often shows a pattern in which the specific growth rate starts at a value of zero and then accelerates to a maximal value in a certain period of time. In addition,

growth curves contain a final phase in which the rate decreases and finally reaches zero, so that an asymptote is reached, the time taken by the bacteria to initiate growth called lag time. The number of organisms is proportional to the optical density (OD) or the absorbance of the cell culture which is measured by spectrophotometer generally at 600nm wavelength. When the growth curve is defined as the logarithm of OD plotted against time, these growth rate changes result in a sigmoidal curve with a lag phase just after onset of growth. It is then followed by an exponential phase and then by a stationary phase. The Gompertz model is one of the most frequently used sigmoid models fitted to bacterial growth data [21]. The gompertz equation is rewritten with respect to growth rate, lag time and asymptote. The obtained experimental data is fit to the Gompertz equation using the least square method. This will enable us to deduce the growth parameters like maximum growth achieved (A), growth rate (μ_m) and lag time (λ). In our study, these parameters are dependent on the concentrations of antibiotic seconeolitsin used. The effect of seconeolitsin on the growth of the bacteria can be deduced by the variations on the maximum growth, growth rate and lag time.

The Gompertz equation for the bacterial growth is shown below and Figure 3.2.1 shows the parameters on the growth curve. Figure 3.2.2 is an example of the fitting done during this study to deduce the parameters.

$$y = \mathrm{A}exp\left[-exp\left[\frac{\mu_m e}{\mathrm{A}}(\lambda - t) + 1\right]\right]$$

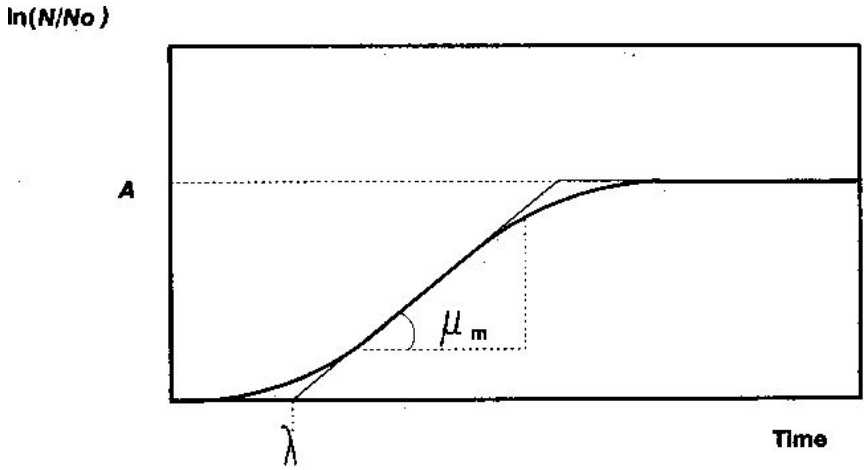


Figure 3.2.1: An example of the curve fitting using the Gompertz model of the growth curve [21]

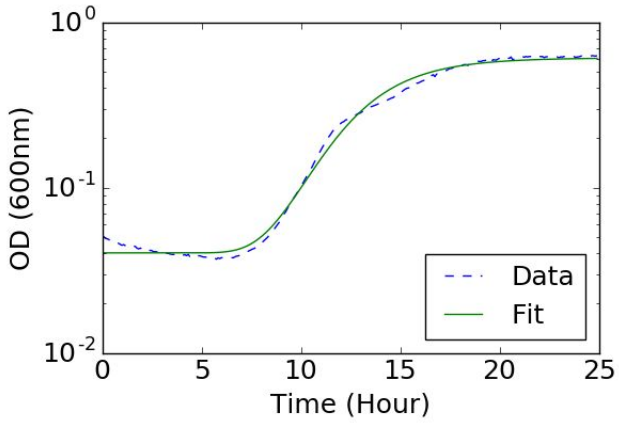


Figure 3.2.2: Experimental growth curve fit to the Gompertz growth curve model for the 200µM concentration of seconeolitsin. The dotted line corresponds to the experimental data whereas the model curve is the continuous line

Studying the effect of seconeolitsin on Supercoiling: Chloroquine gel electrophoresis

Chloroquine gel electrophoresis is a method used to monitor the superhelicity of a reporter plasmid such as pUC18. Chloroquine intercalates between base pairs of DNA [22]. Chloroquine is mainly intercalated between the nucleotide bases of the relaxed DNA which is more accessible than the supercoiled DNA as shown in Figure 3.2.3. This intercalation causes the stretching of the DNA molecule. At chloroquine concentrations of 2.5µg/mL, more negatively supercoiled molecules migrate further through the gel and more relaxed molecules migrate slower. The reporter Plasmid pUC18 is transformed into the model organism *Dickeya dadantii* and *Escherichia coli* in order to study the effect of seconeolitsin on the DNA supercoiling levels. *Dickeya dadantii* and *Escherichia coli* harboring pUC18 plasmid is treated with seconeolitsin and the plasmids are extracted run on gel containing 2.5µg/mL concentrations of chloroquine.

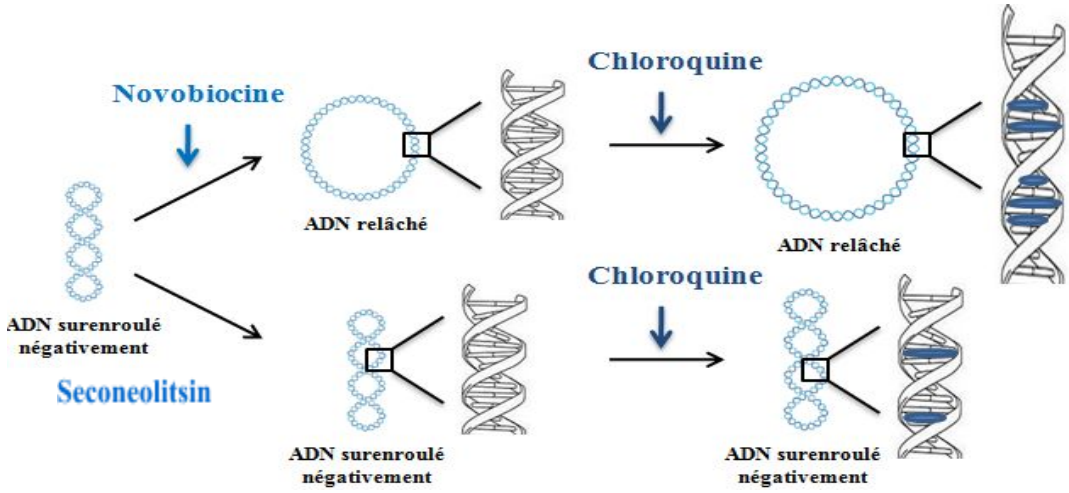


Figure 3.2.3: Intercalation of chloroquine between base pairs of DNA

The report gene system used to study the effect of Seconeolitsin on virulence gene *pelE*

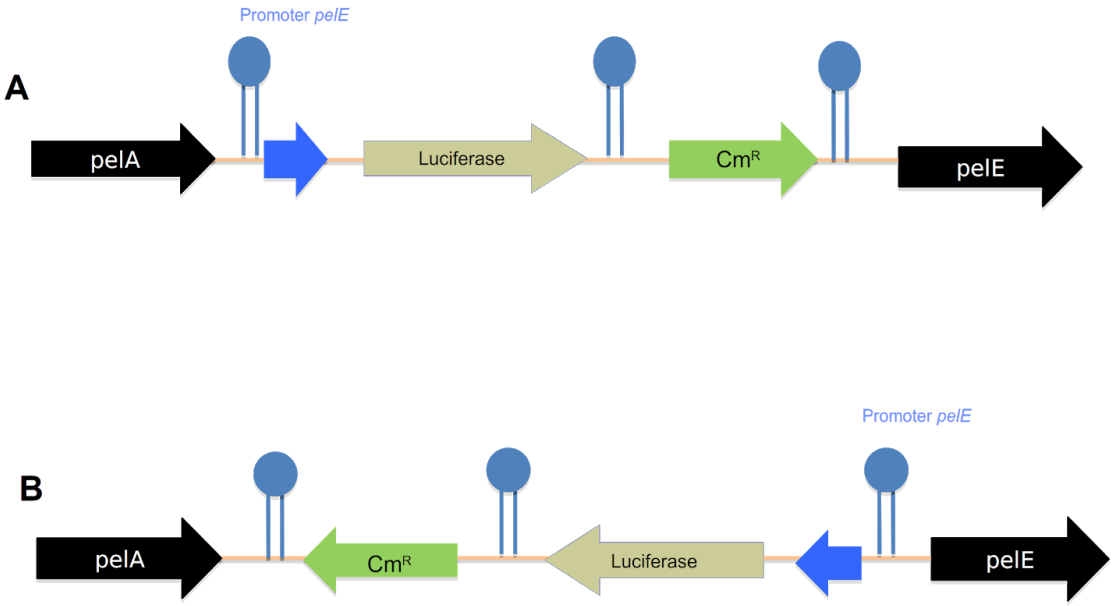


Figure 3.2.4: (A) A5720 construct with *pelE* promoter inserted in the chromosome in forward orientation fused with luciferase gene and chloramphenicol antibiotic resistance gene. (B) A5849 construct with *pelE* promoter inserted in the chromosome in reverse orientation used with luciferase gene and chloramphenicol antibiotic resistance gene

In order to study the effect of supercoiling on the genes, the *pelE* gene of the plant pathogen *Dickeya dadantii* is chosen. The *pelE* gene in *Dickeya dadantii* is studied since it is known already that it is sensitive to the changes in the supercoiling levels on the gene [12]. A genetic reporter system consists of a promoter under analysis, joined to a reporter gene in an expression vector with an antibiotic resistance gene to indicate the integrity of the reporter system. Expression of the gene under study can be determined by measuring the expression of the luciferase gene fused to the promoter under study. We have used a construct of the *pelE* promoter followed by the luciferase gene and chloramphenicol resistance gene. This construct was already available in the lab which was built for experiments earlier. Luciferase enzymes from a variety of organisms have become popular in reporter systems. They have higher utility due to their inherent sensitivity and ease of measurement. Compared to GFP or other fluorescent proteins they yield better signals since the signal from luciferase enzyme, luminescence is faster and very specific to the promoter under study [23]. The medium supplemented with luciferin emits luminescence when the gene is expressed enabling us to record the real time gene expression level under various conditions. We are at ease as the TECAN spark growth recorder also equipped with real time luminescence

recording along with the recording of growth.

Two constructs with a reporter gene on the chromosome located adjacent to the native *pelE* gene locus are used in the study. The construct A5720 is the modified wild type strain of *Dickeya dadantii*, consists of a *pelE* promoter fused with the luciferase gene and a chloramphenicol antibiotic resistance gene. Another construct A5849 with the reporter *pelE* promoter fused in the reverse orientation is also used in this study to check if the effect of change in the supercoiling levels on the gene varies with respect to the orientation of the gene. In A5720 the promoter *pelE* and following luciferase gene is inserted next to the *pelA* gene in forward direction whereas in the reverse orientation construct A5849, the promoter *pelE* and the following luciferase gene is inserted next to the natural *pelE* gene of the chromosome but in reverse orientation as shown in Figure 3.2.4. The advantage of this construct is it being on the chromosome, it reports the direct effect of change in supercoiling levels occurring in DNA along the chromosome and thus, its effect on the susceptible gene on the chromosome.

The data treatment of tecan output file

The Tecan SPARK machine records both absorbance and luminescence. The values are measured every 5 minutes and recorded. The plate consisting of the samples are shaken well before data recording to ensure uniform measurement. The data is obtained in the form of an excel data sheet. This data has to be verified for contamination and noise created by any possible interference. A Python script was written in order to automate the data analysis process. The script reads the excel data based on the wells and then looks for contamination, especially in the control wells. The negative control with just medium and antibiotic, but no inoculum is used to deduce the background noise value. Any significant increase in the negative control reading during the growth indicates a contamination problem, which needs to be analysed further before considering the results. The positive control is with medium and inoculum with no antibiotics added. The positive control sets the reference for the effect of antibiotics on the growth and gene expression. Once the data is processed, to get rid of the background noise and verified for no contamination, the base medium values are deducted from the corresponding bacterial samples. After complete processing the data for no abnormalities, the values of luminescence or optical density (OD) are plotted against the time along growth. All the values are plotted with respect to time values to observe the variation in growth and expression with respect to time and comparing them in various conditions. The effect of seconeolitsin on the reference gene expression is evaluated on the basis of fold change. The fold change refers to the increases in folds of the gene expression in the presence of antibiotic to the absence of antibiotic (positive control).

c) Results and Discussion

Seconeolitsin is known to increase the level of negative supercoiling on *Streptococcus pneumoniae* DNA. Since we are studying the effect of this antibiotic on a gram-negative organism for the first time and it is necessary to have the characterization to obtain the MIC and then proceed to study its effect on the DNA supercoiling, if so, later on the gene expression. We are thankful for the Organic and Bioorganic Chemistry team from ICBMS laboratory for synthesizing seconeolitsin and providing us with the freshly prepared seconeolitsin. It is necessary to test it first on a gram-positive organism to ensure the efficacy of the antibiotic and then proceed to characterize it on our gram-negative model organism *Dickeya dadantii*. We chose *Bacillus subtilis* as a test gram-positive organism on the basis of ease of availability in the laboratory.

Minimum Inhibitory Concentrations (MIC) are defined as the lowest concentration of an antimicrobial compound that will inhibit the visible growth of a microorganism after overnight incubation [24]. In some cases, it can be deduced from the level of inhibition of bacterial growth based on the decrease in maximum level growth reached. In other cases, the MIC is defined by the time required by the bacteria to onset the visible exponential growth. Thus, the quantification of MIC can be either in reference to the recorded maximum OD or the lag time in the presence of different concentrations of antibiotics.

In this study, we deduce the MIC on the basis of visible difference in lag time indicating the lowest concentration of drug without visible growth. We hypothesize that the bacteria overcomes the toxicity caused by the antibiotics after sometime by expelling it out or counter reacting to it when the concentrations are in the tolerable range. So in order to define the effect of antibiotics on the growth of bacteria we choose to define MIC on the basis of lag time as time taken by bacteria to counter the antibiotics. Based on the previous reference organism *Streptococcus pneumoniae,* we choose 2 hours to be the acceptable delay in course of bacterial growth to have an inhibitory effect. Figure 3.3.1B shows the growth curves of *Bacillus subtilis* in the presence of various concentrations of seconeolitsin. In order to be able to visibly assign the MIC criteria, we considered a visible delay of about 2 hours to be acceptable to have an inhibition to normal growth. On this basis, the MIC in gram-positive *Bacillus subtilis* was deduced to be below 20µM which is in agreement with the results in earlier study, *Streptococcus pneumoniae* which was found to be 17µM. The membrane outer layer of gram-negative bacteria comprises complex lipo-polysaccharide with lipid portion acting as an endotoxin. Compared to gram-positive bacteria it has an additional layer of resistance making it difficult for any drug to enter into the cytoplasm and to kill the bacteria compared to gram-positive bacteria. Thus, we expected the MIC to be possibly higher than the effective value found in gram-positive bacteria. We tested much higher concentrations of seconeolitsin initially to ensure its effectiveness in gram-negative bacteria. Growth kinetics experiments were carried out in nutrient rich LB (Liquid Broth) medium and minimal M63 media for understanding the effect of antibiotics depending on the varying growth conditions.

LB medium is nutrient rich and the Minimal M63 medium contains essential salts and limited glucose as sugar source.

A B

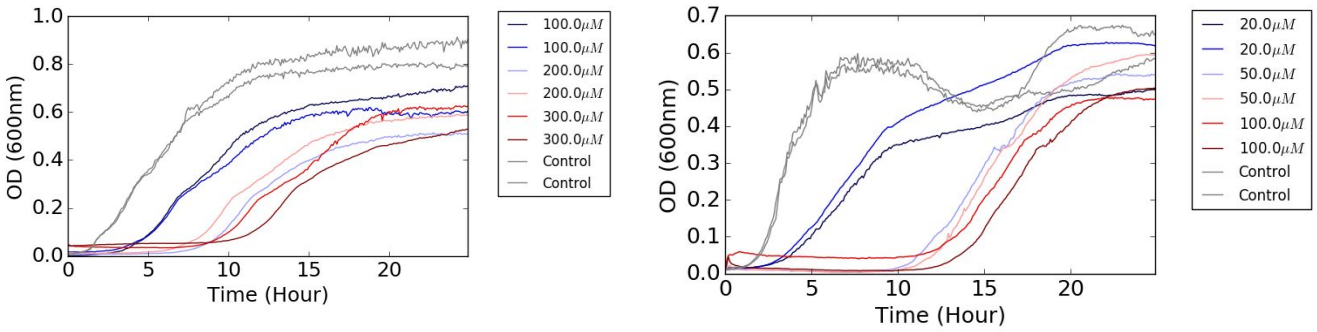


Figure 3.3.1: The growth curves of A) *Bacillus subtilis* and *B) Dickeya dadantii* grown at various concentrations of seconeolitsin in LB medium

Figure 3.3.1 shows the growth of bacteria in the presence of Seconolitsin. The MIC on gram-positive was below 20μM (Figure 3.3.1B). As expected, the gram-negative growth delay was observed at higher concentrations. For example, at 20μM about a 2 hour delay is observed in *Bacillus subtilis* whereas in *Dickeya dadantii* the delay is around 2 hours for 100μM. This confirms the hypothesis of higher MIC to be observed in gram-negative bacteria compared to gram-positive. The MIC of seconeolitsin in *Dickeya dadantii* causing visible delay of growth is deduced to be around 100μM. The MIC could vary based on the different conditions of growth the bacteria would be subjected to during the study. So for further characterization of seconeolitsin we chose 100μM, 200μM and 300μM concentrations.

We wanted to study the growth of bacteria under various conditions. In this study, we chose LB and M63 medium and also with respect to the variation in sugar source, glucose and PGA. Since the model organism *Dickeya dadantii* is a phytopathogen and flourishes in the presence of pectin source, PGA is chosen as the alternative sugar source. In order to completely visualize the effect of seconeolitsin, a known minimal medium M63 is chosen for the study. The study is individually carried out in LB with and without PGA and in Minimal M63 medium in the presence of glucose or PGA+ glucose as sugar source.

Characterization of seconeolitsin with LB medium

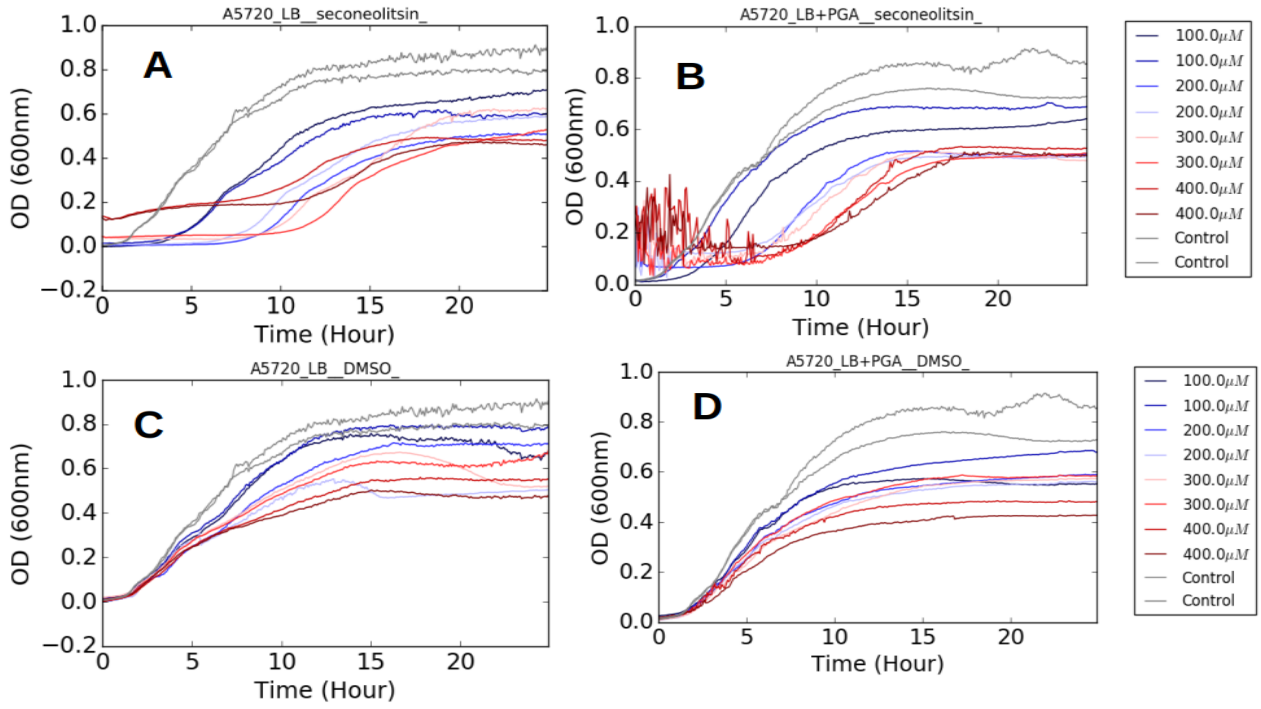


Figure 3.3.2: The growth curves of *Dickeya dadantii (*A5720) in LB media supplemented with various compounds. (A) LB medium in the presence of various concentrations of seconeolitsin. (B) LB medium supplemented with 8% v/v PGA and in the presence of various concentrations of seconeolitsin. (C) LB medium in the presence of various concentrations of DMSO corresponding to various concentrations of seconeolitsin. (D) LB medium supplemented with 8% v/v PGA and in the presence of various concentration of DMSO corresponding to various concentration of seconeolitsin

The bacterial growth experiments with seconeolitsin in LB medium is carried out in the presence and absence of PGA. Figure 3.3.2 shows the growth curve of *Dickeya dadantii* in LB medium in the presence of various supplements. Figure 3.3.2A and 3.3.2B show no difference in growth is observed with respect to presence or absence of PGA. This might be because the medium is rich in carbohydrate source in itself so the presence of PGA did not alter the growth kinetics. The presence of PGA did cause variations in the reading causing fluctuations in the curve as shown in Figure 3.3.2B. The medium with PGA was causing destabilization of absorbance measurements caused by turbidity. This did not affect the overall measurement but rather led to a technical difficulty in recording accurate absorbance value in the presence of seconeolitsin. We investigated the components of the growth culture to resolve the issue. We had the bacteria growing in LB medium along with DMSO seconeolitsin and supplemented with and without PGA.

Since DMSO is the solvent which is used to prepare seconeolitsin solution, a control experiment is carried out to check the effect of DMSO on the growth of the bacteria both in the presence and absence of PGA (Figure 3.3.2Cand 3.3.2D) in LB medium. Under 5% v/v of DMSO, volume corresponding to 100µM of seconeolitsin, the growth was not significantly altered. Presence of DMSO alters the growth of the bacteria in terms of the growth rate. This limit was considered in the next experiments as learning. The absence of perturbations in the absorbance reading in the DMSO PGA medium leads to the suspicion of seconeolitsin PGA mixture interfering with the absorbance value recording.

To conclude, in the LB medium, addition of PGA did not make a visible difference in the growth of the organism in the presence of antibiotic seconeolitsin. The limit of DMSO to be present in the culture broth was limited to 5% v/v to avoid any inhibition due to presence of DMSO on bacterial cells. Possibly the presence of viscous PGA and seconeolitsin might raise uneven shaking behavior in the medium leading to the absorbance measurement fluctuation. Though there was difficulty observed with respect to the optical density recording in the presence of PGA, the overall reading was not altered causing no real problem with the recorded value and overall analysis.

Characterization of Seconeolitsin with M63 minimal medium

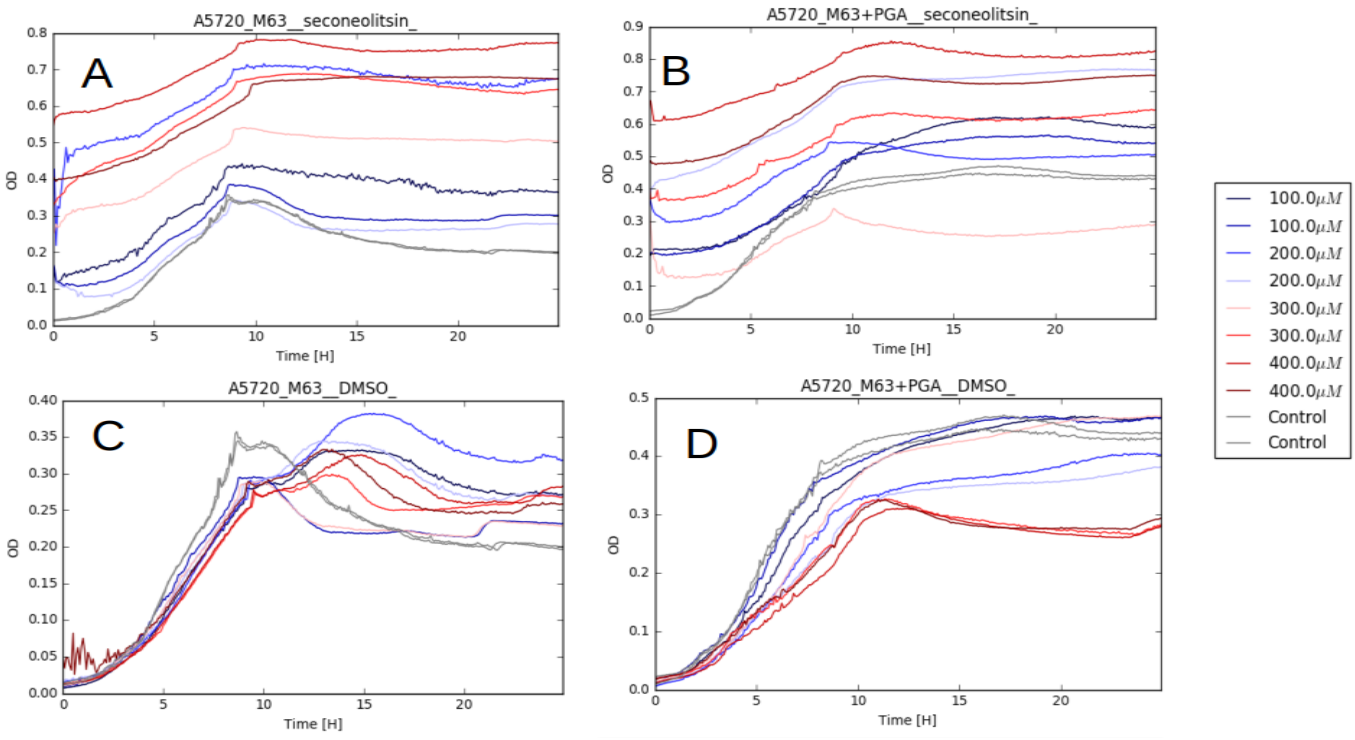


Figure 3.3.3: The growth curves of stain A5720 in M63 media supplemented with various compounds. A) M63 medium in the presence of various concentrations of seconeolitsin. B) M63 medium supplemented with 8% v/v PGA and in the presence of various concentrations of seconeolitsin. C) M63 medium in the presence of various concentrations of DMSO corresponding to various concentrations of seconeolitsin. D) M63 medium supplemented with 8% v/v PGA and in the presence of various concentrations of DMSO corresponding to various concentrations of seconeolitsin

LB medium is excellent for the growth of bacteria but all the components of LB medium are not exactly known. Since using a well studied medium with known components is better for characterizing a new antibiotic we decided to use a medium with known components. M63 medium is mainly used to study the transcriptional effect of antibiotics in our lab, hence it is well understood for this application and previous data has shown that the effect of antibiotics tend to be stronger in minimal medium. In order to minimize external interference of unknown media components and have best results, we chose M63 medium to study the transcriptional effect of seconeolitsin. So characterization was also carried out in minimum M63 media. Figure 3.3.3 shows the growth curve of *Dickeya dadantii* in M63 medium in the presence of various supplements. We encountered precipitation problems in the minimal media with seconeolitsin as observed in Figure 3.3.3. There is an up-shift in the OD curve due to the precipitation which increases with increasing concentration of seconeolitsin added as observed in 3.3.3A and B. The growth was slightly more in PGA supplemented media

compared to media with no PGA. When compared to Figure 3.3.3C and D, culture medium supplemented with DMSO and DMSO+PGA this phenomenon was not observed. Again this confirms some possible precipitation interaction of seconeolitsin. The acceptable limit of DMSO in the medium was found to be less than 5 %v/v to ensure no inhibition of cell growth. Under this limit the presence of DMSO did not affect the growth of bacteria Thus, enabling us to study the effect by seconeolitsin alone.

Minimal inhibitory concentration

MIC was deduced observing the growth of bacteria in the presence and absence of seconeolitsin. Since seconeolitsin inhibitory action creates a delay in the lag time for the bacterial growth, we deduce the MIC for this particular antibiotic based on visual growth delay. In the gram-positive *Streptococcus pneumoniae*, the MIC was 17µM where a 2 hour delay on onset of bacterial growth is observed due to antibiotic inhibition independent of growth conditions. We analyse the growth study in nutrient rich LB medium as well as the minimal M63 medium to deduce the MIC. When we observe Figure 3.3.2 and Figure 3.3.3 the onset of growth under the influence of seconeolitsin was observed to be around the same time for both medium cultures. Due to the minimal glucose level available for the bacteria, the stationary phase is reached faster in the minimal M63 medium compared to nutrient rich LB medium. In the presence of seconeolitsin, the growth rate is lesser in M63 medium compared to LB medium. In order to estimate the MIC, the growth data was analyzed to study the effect of seconeolitsin on the growth of the bacteria. The Gompertz equation is a popular mathematical model for describing bacterial growth in terms of growth rate, level of growth and lag time required for the onset of growth. Thus, the growth points from the experimental data fit to the Gompertz model of the growth curve in bacteria and in order to calculate the lag time and growth rate. Considering the definition of MIC described earlier based on lag time and the assumption of 2 hours of lag time to have visible inhibition of onset of growth the MIC was observed to be the same for both LB and M63 medium and it was deduced to be around 100µM.

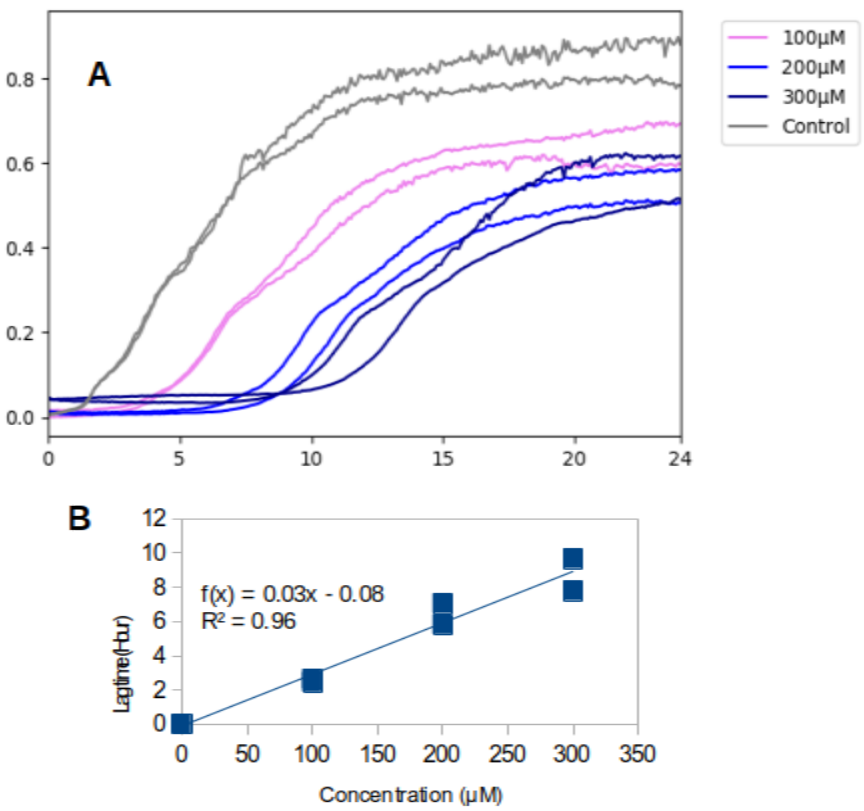


Figure 3.3.4: A) Growth curve of A5720 strain with increasing concentration of Seconeolitsin in LB medium. B) Regression of the Delay time with respect to the concentration of seconeolitsin

Figure 3.3.4A shows the growth of bacteria in the presence of 100µM, 200µM, 300µM seconeolitsin concentrations. The presence of antibiotic seconeolitsin causes a delay in the onset of bacterial growth. The following Figure 3.3.4B quantifies the growth delay observed due to the presence of seconeolitsin shown in the grown curve, a clear delay on the onset of exponential phase is seen with the increase in the concentration of seconeolitsin. The regression analysis of the delay time with respect to increase in concentration is carried out to quantify the delay time with respect to the increase in the concentration of seconeolitsin. The possible reason for the delay caused can be due to the time taken by the bacteria to overcome the inhibition of topoisomerase I by seconeolitsin, in-order to produce more topoisomerases I to facilitate the normal cause of growth. It can also be the time required by the bacteria to work on overcoming seconeolitsin with the drug efflux systems. More the seconeolitsin more the inhibition of Topoisomerases and more the time required by the bacteria to produce required Topoisomerases.

Effect of seconeolitsin on DNA supercoiling level

DNA supercoiling corresponds to the density of turns in the DNA. Mode of action of seconeolitsin is by inhibiting topoisomerase I, and the inhibition of topoisomerases causes a change of supercoiling level in the chromosome. To study the effect of seconeolitsin on the DNA supercoiling level the gel shift experiment of the plasmid in chloroquine gel is conducted. In the growth experiments seconeolitsin showed considerable inhibition in the growth of cells at higher concentration. The MIC of seconeolitsin in model organism *Dickeya dadantii* is 100μM and concentrations reasonably below 100μM can be employed to study the effect of seconeolitsin on DNA supercoiling levels to ensure no cell inhibition. We choose concentrations of seconeolitsin to be 50μM to study its effect on DNA supercoiling. Experiments are conducted in M63 minimal media to visualize the most effect of antibiotic seconeolitsin

We wanted to quantify the effect of Seconeolistin on DNA supercoiling. We employed the chloroquine gel shift experiment for the purpose. Since plasmid are isolated from bacteria subjected to seconeolitsin shock the results can be assumed true to the bacterial chromosome as well. We chose the well known *Escherichia coli* in addition to our model organism. The effect of seconeolitsin is quantified in *Escherichia coli* and *Dickeya dadantii* with chloroquine gel analysis tool with pUC18 Plasmids are extracted at exponential phase at OD=0.2 and stationary phase at OD=1.1 to study the growth phase dependencies on the effect of Seconeolitsin on DNA supercoiling.

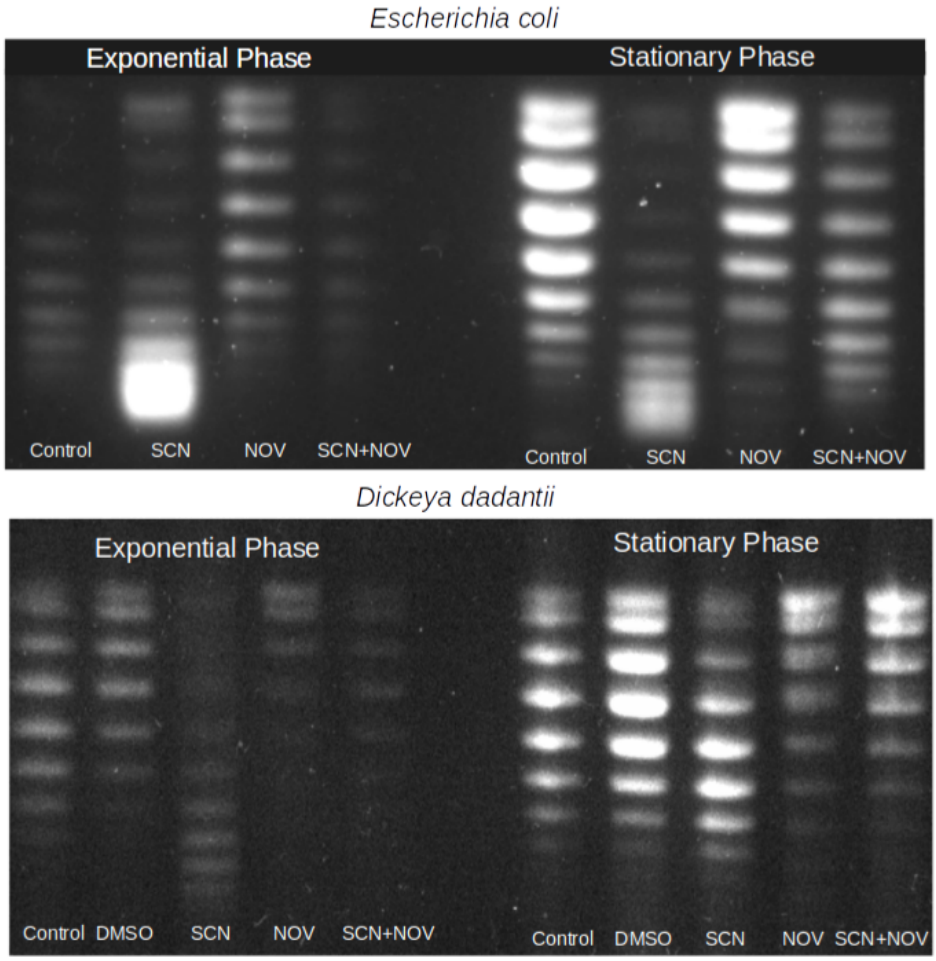


Figure 3.3.5: *Escherichia coli* (top) and *Dickeya dadantii* (bottom). The left side of the image represents the exponential phase and the right side represents the stationary phase

Figure 3.3.5 shows the gel electrophoresis of the extracted plasmid with chloroquine intercalation. Results clearly show the hyper-negative supercoiling effect of seconeolitsin. Lanes labeled Control refer to the plasmids without any treatment. The lane labeled SCN corresponds to the plasmids treated by 50μM seconeolitsin. The lanes labeled NOV refers to the novobiocin (DNA gyrase inhibitor) treated plasmids. 100μg/ml concentration of novobiocin is used. The lane labeled SCN+NOV refers to the plasmids treated with both seconeolitsin and novobiocin. In this treatment, 50μM of seconeolitsin and 50μg/ml of novobiocin is used. In *Dickeya dadantii* additional DMSO treatment is employed with treating the bacteria harbouring plasmid with the same concentration of DMSO as in the 50μM seconeolitsin sample. The less stability of exponential state plasmids compared to stationary phase plasmid can be observed in the intensities of bands in gel images.

In the chloroquine gel the supercoiled plasmids move further ahead compared to the less supercoiled plasmids. We can observe that the seconeolitsin treated plasmids have migrated further compared to the non-treated control plasmids. This proves that seconeolitsin induces hyper-negative supercoiling in the DNA. The novobiocin treated samples have migrated less indicating the relaxation induced by the antibiotic treatment. Interestingly as expected the cocktail mixture of novobiocin and seconeolitsin treated samples have found a compromise with respect to relaxation and hyper-negative supercoiling. The lanes are similar to the untreated control compound leading to the conclusion that the effect of inhibition of topoisomerase I is nullified by the inhibition of gyrase in *Escherichia coli*. In *Dickeya dadantii,* the relaxation effect of novobiocin is stronger, thus, with seconeolitsin and novobiocin together the relaxation effect of novobiocin is partially counterbalanced by seconeolitsin.

The BioRad Image Lab software is used to quantify and analyze the above gel. The bands are manually detected and quantities of the DNA present in the bands are quantified based on the intensity detection using the software. The below graph shows the result of the analysis. Relative superhelicity is calculated assuming the top most band to be topoisomerase with no added turns, with the approximation that is level 0.

Considering 10.5bp to make a turn in DNA structure, pUC18 plasmid with 2686bp size will have 255 turns at 0 level. We consider the next lanes to have consecutively added one turn and then calculate average superhelicity to be the ratio of added turn and number of turns existing in natural state.
Average superhelicity, $\Delta\sigma$ = number of turns added / 255
Relative superhelicity = relative weight of DNA in band * average superhelicity.

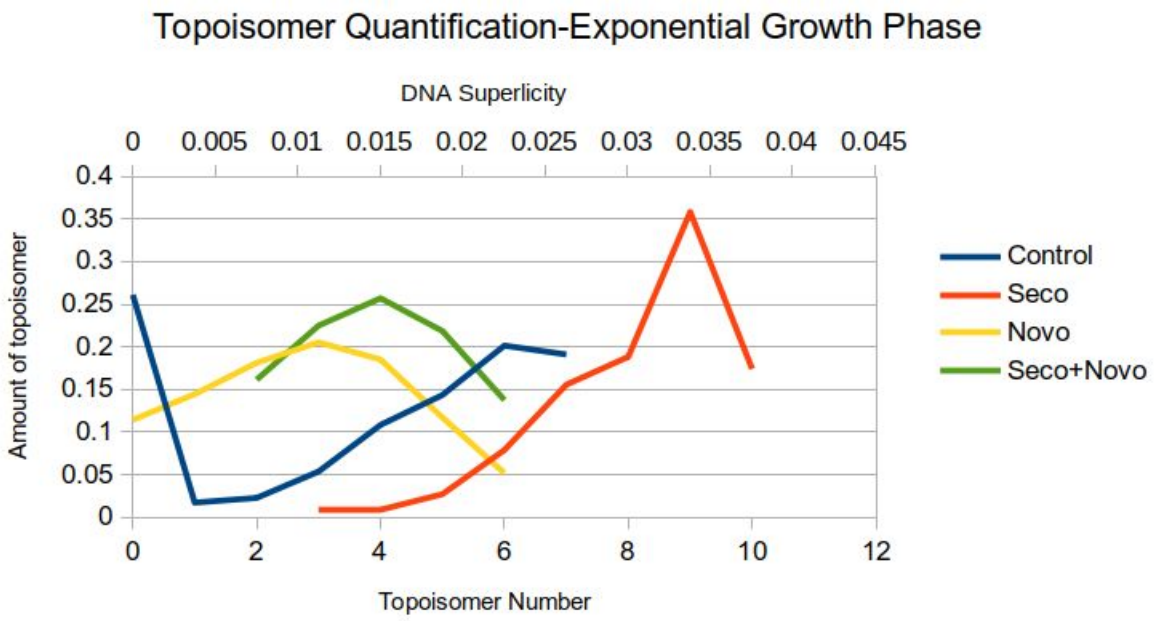


Figure 3.3.6: The result of the gel shift experiment in *Escherichia coli.* The quantity of different topoisomers is represented along the y-axis and primary x-axis represents the topoisomerase number and their numerical values of superhelicity are represented along the secondary x axis. Note that the superhelicity values are negative in this case

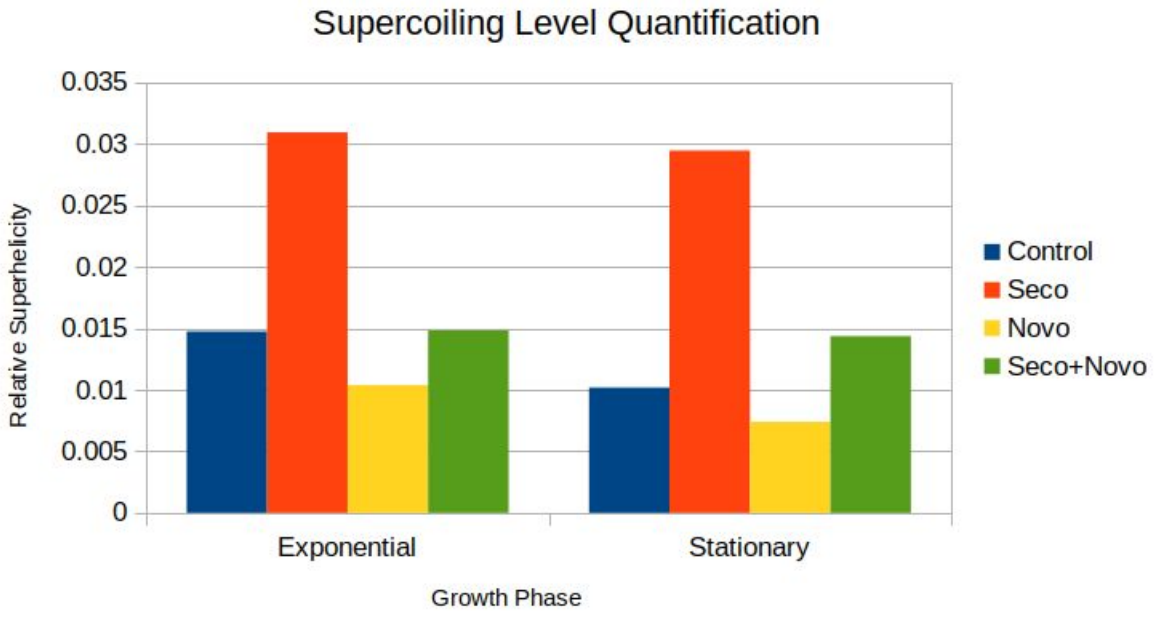


Figure 3.3.7: Chloroquine gel analysis based on the calculation of relative superhelicity by considering the intensity of the bands for *Escherichia coli*

The graphical representation of the result of the gel shift experiment is shown in Figure 3.3.6. The quantity of topisomers and superhelicity is depicted. The average superhelicity of the plasmids extracted in each condition is plotted in Figure 3.3.7. As we observe the average supercoiling is highest in seconeolitsin treated samples indicating the presence of hyper-negative supercoiling caused by the presence of antibiotics. We also observe the relaxing effect by Gyrase inhibiting antibiotic novobiocin. Interestingly, we also see a compensated effect when bacteria are treated with the mixture of seconeolitsine and novobiocin. We also observe a slight reduction in average super-helicities in the control samples in the stationary phase compared to exponential phase. The same analysis was also carried out on the model organism *Dickeya dadantii*. The results are described below.

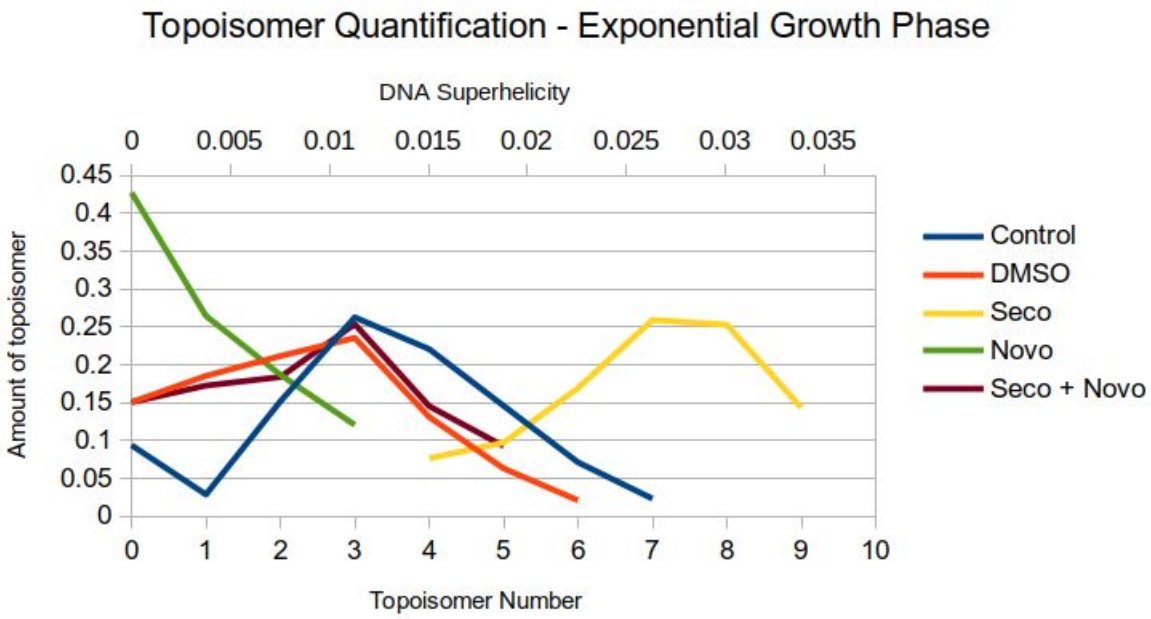


Figure 3.3.8: The result of the gel shift experiment in *Dickeya dadantii.* The quantity of different topoisomers is represented along the y axis and primary x axis represents the topoisomerase number and their numerical values of superhelicity are represented along the secondary x axis. Note that the superhelicity values are negative in this case

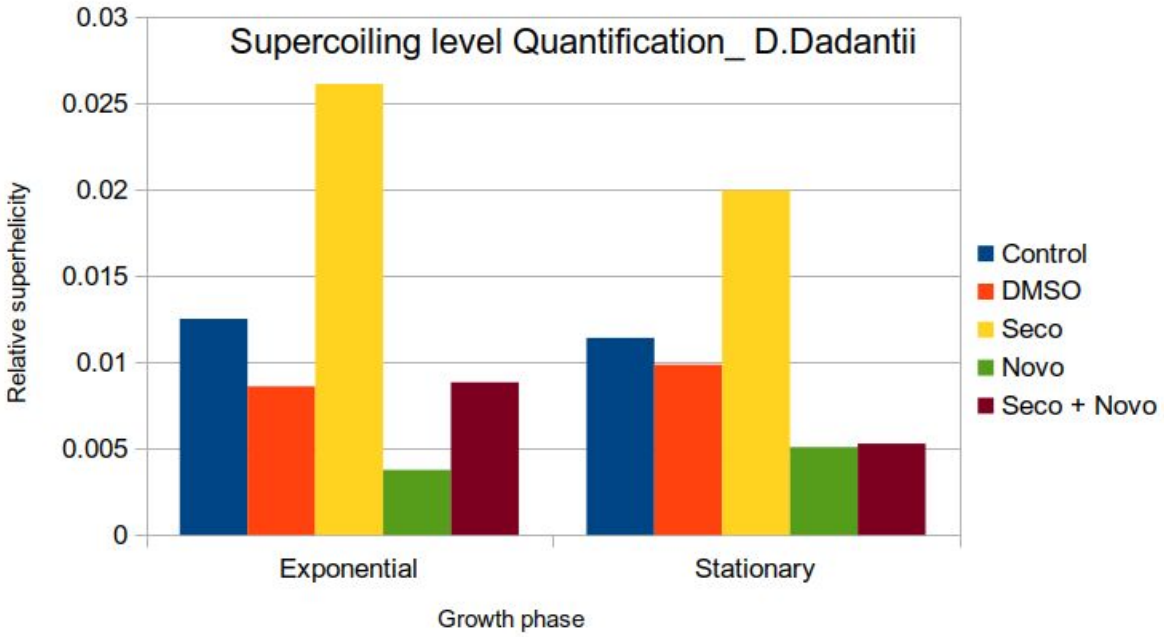


Figure 3.3.9: Chloroquine gel analysis based on the calculation of relative superhelicity by considering the intensity of the bands for *Dickeya dadantii*

The effect observed in the presence of seconeolitsin could be partially contributed by the DMSO. Since we don't have any information on the effect on DNA supercoiling by DMSO alone, in *Dickeya dadantii* an additional condition with treatment of bacterial cells with DMSO was also incorporated. This helps us to be sure of the effect of seconeolitsine alone. The topoisomerase quantification is shown in Figure 3.3.8. The average superhelicity is the highest for seconeolitsin treated plasmids and lowest for the novobiocin treated sample as expected. Compared to the non-treated samples DMSO treated samples had a slight relaxation effect observed. In any case, this emphasizes more on the supercoiling effect of seconeolitsin since with the seconeoltsin treatment, the plasmids are more supercoiled anyway. Additionally, DMSO treated samples had the superhelicity similar to the mixture of seconeolitsin and novobiocin plasmids. Compared to the control the palmids from the bacteria treated with the mixture of seconeolitsin and novobiocin had more relaxing effect in *Dickeya dadantii*. This signified that the effect of novobiocin is stronger in *Dickeya dadantii* compared to *Escherichia coli*. In order to understand the effect more quantitatively the average superhelicity in each sample is calculated and plotted in Figure 3.3.9. We do observe a difference in the effect with respect to the growth phase. Seconeolitsin is more effective in the exponential phase compared to the growth phase. This can be correlated to the change in the global supercoiling state of DNA during the growth phase.

The above analysis concludes that seconeolitsin increases the supercoiling by 40 percent in exponential phase and by 30 percent in stationary phase. Indicating that seconeolitsin is more effective in the exponential phase when compared to the stationary phase. Thus,

understanding that seconeolitsin causes a significant change in global DNA supercoiling level of bacteria, we further investigate if this effect will be observed in the level of gene expression, precisely virulent gene expression.

From the above analysis it is evident that seconeolitsin indeed has an effect on both the organisms. Compared to *Eschirichia coli*, in *Dickeya dadantii* the supercoiling effect of seconeolitsin is found during both exponential growth and stationary phase. In *Escherichia coli,* the effect is not very strong in the stationary phase. It is evident that the effect is stronger in the exponential phase in both organisms. It can be because the DNA is generally more supercoiled during the exponential growth face, indicating higher topoisomerase activity. Thus the inhibition effect of seconeolitsin is more strong during the exponential phase when the topoisomerase I activity is more prominent in maintaining the global supercoiling levels. In the transition phase, the reduced topoisomerase activity makes the supercoiling activity of seconeolistin by the inhibition of topoisomerase I less effective.

Effect of seconeolitsin on the *pel* gene expression

Seconeolitsin varied the supercoiling levels in the plasmids isolated from bacteria subjected to seconeolitsin shock. This led us to further questions if the same effect is observed on the bacterial chromosome? If yes, does the variation in the supercoiling level in chromosomes vary the gene expression pattern? SSince plasmid are isolated from bacteria subjected to seconeolitsin shock the results can be assumed true to the bacterial chromosome as well [25]. In order to study the variation of supercoiling in bacterial chromosomes and its effect on gene expression, further studies are carried out with *pelE* gene as the model promoter since its known to have a promoter activity sensitive to supercoiling [12]. The p*elE gene* is also crucial for pectin degradation which is a very important factor for the pathogenicity of the bacteria *D.dadantii* [13]. It makes the *pelE* gene an ideal choice for studying the effect of supercoiling on gene expression since it is a key virulence gene. A construct of *D.dadantii* with a fusion of *pel*E promoter with *lux* gene is used as a reporter for the gene expression (see materials and methods).

Under the application of novobiocin shock, the relaxation effect on the chromosome was dynamic. The relaxation effect of novobiocin prolonged for a duration of time and then the normal level of supercoiling of the bacteria was restored [12]. This arises a question if we use seconeolitsin in shock to study the effect on *pel* expression. Is the effect dynamic? If so, how long does the effect last? Does the expression level vary with the modification in supercoiling level caused by seconeolitsin? In order to answer these questions, knowing MIC of seconeolitsin and under MIC it has an effect on the supercoiling levels of plasmid extracted from the bacteria subjected to shock, the activity of *pelE* gene was studied under the influence of seconeolitsin.

Growth curves with different concentrations of seconeolitsin

In order to study the effect of seconeolitsin on the *pelE* gene expression, different concentrations of seconeolitsin are used. 25µM, 50µM and 100µM were the chosen concentrations and forward orientation and reverse orientation of the reporter gene is studied in minimal M63 and nutrient rich LB media. The study is conducted in both exponential and stationary growth phase in order to study the effect of growth phase on the seconeolitsin action on the *pelE* gene.

Effect of seconeolitsin on the *pelE* expression at exponential phase

The results of effect of seconeolitsin on the *pelE* expression in M63 and LB medium is described below. The effect of seconeolitsin is studied in both Shock condition and a control where bacteria is initially grown with the corresponding concentration of secoenolitsin. The effect is studied in forward (A5720) and reverse (A5489) orientation of gene promoter (as described in methodology).

The bacteria growing in corresponding media is shock treated with the antibiotic at predetermined time 5h during the exponential growth phase. The following Figure 3.3.10 shows the growth curve of the bacterium shocked by different concentrations of seconeolitsin at exponential growth phase in minimal M63 medium.

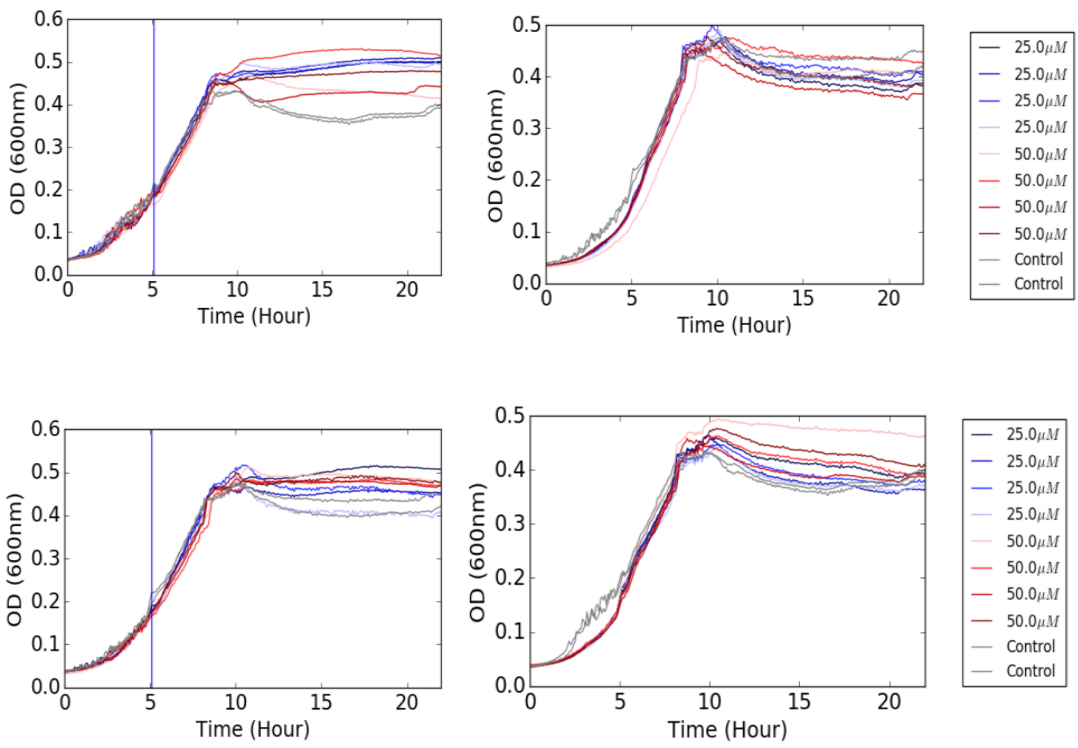


Figure 3.3.10: Growth curves corresponding to the experiments to study the effect of seconeolitsin shock of *pelE* gene expression in M63 medium. Parts A and B represent growth curve in M63 medium with 25µM, 50µM concentrations of seconeolitsin with and without shock respectively in forward orientation construct A5720 and C and D representing growth curve in M63 medium with 25µM, 50µM concentrations of seconeolitsin with and without shock respectively in reverse orientation construct A5489

Figure 3.3.10 shows the growth curve for the seconeolitsin shock experiment. Given that growth is not significantly affected by addition of antibiotics in the shock mode, the effect observed in the expression of the genes is due to the effect created by seconeolitsin. Knowing *pelE* gene is susceptible to supercoiling, the expression for the same is expected to increase soon after the shock effect is observed. The following figure shows the real time luminescence quantification corresponding to the *pelE* expression for the same bacterial growth.

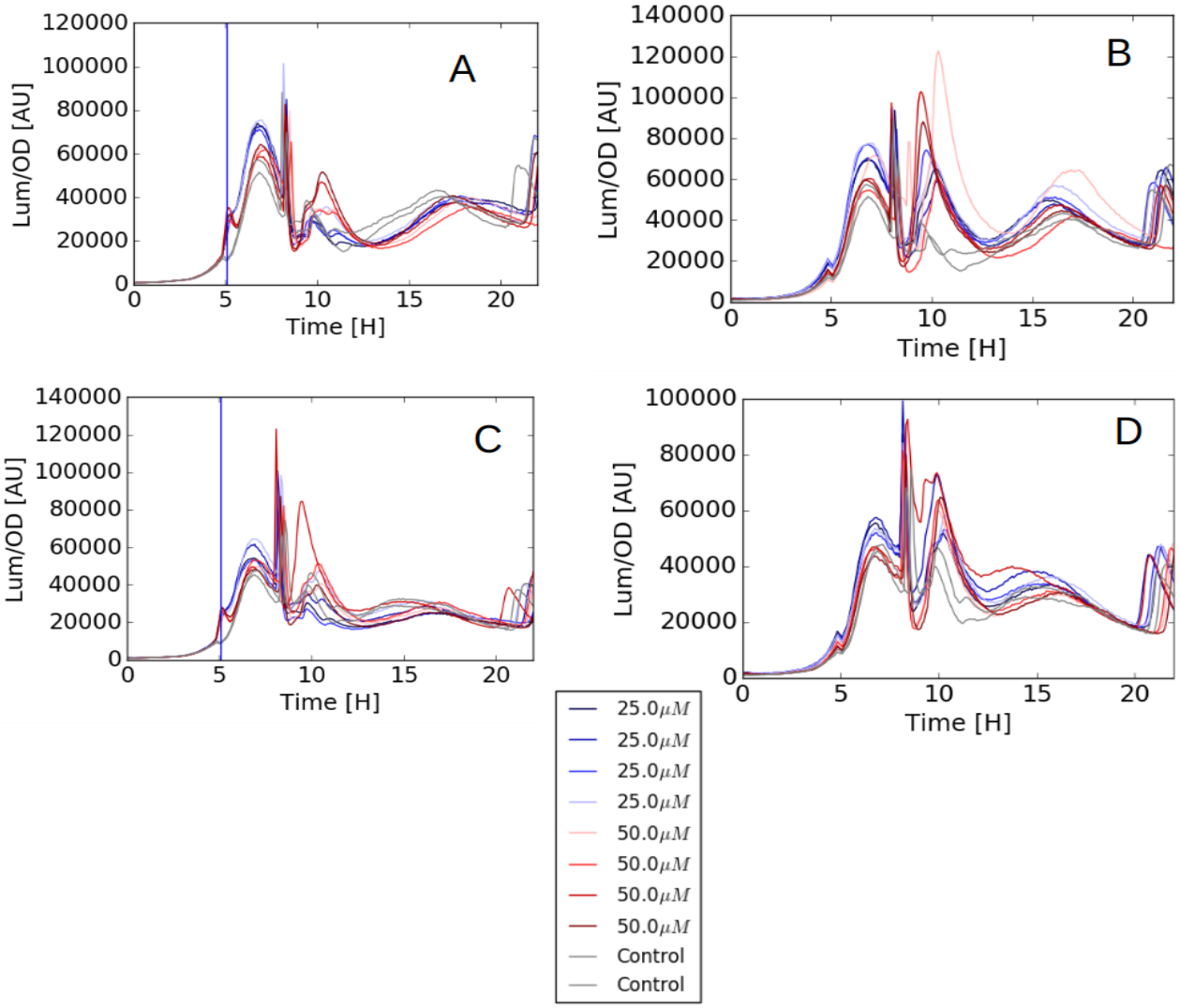


Figure 3.3.11: The real time luminescence recorded in the M63 medium with various conditions. Parts A and B represent luminescence observed in M63 medium and 25µM, 50µM concentrations of seconeolitsin with and without shock respectively in forward orientation construct A5720 and C and D representing real time luminescence recorded in M63 medium with 25µM, 50µM concentrations of seconeolitsin with and without shock respectively in reverse orientation construct A5489

Figure 3.3.11 shows the real time data of luminescence resulting from the expression of the *pelE* promoter in minimal M63 medium. In order to make sure this is an effect of the process of shock by seconeolitsin we have a control sample, where bacteria is grown in the presence of the same concentration of seconeolitsin from the onset of the growth in the same experimental set up. There is an increase in expression observed at 5h observed in these graphs proving that seconeolitsin has activated the *pelE* promoter. The experiments are carried out in both orientation of the reporter genes to study the effect of orientation on the gene expression triggered by change in supercoiling levels.

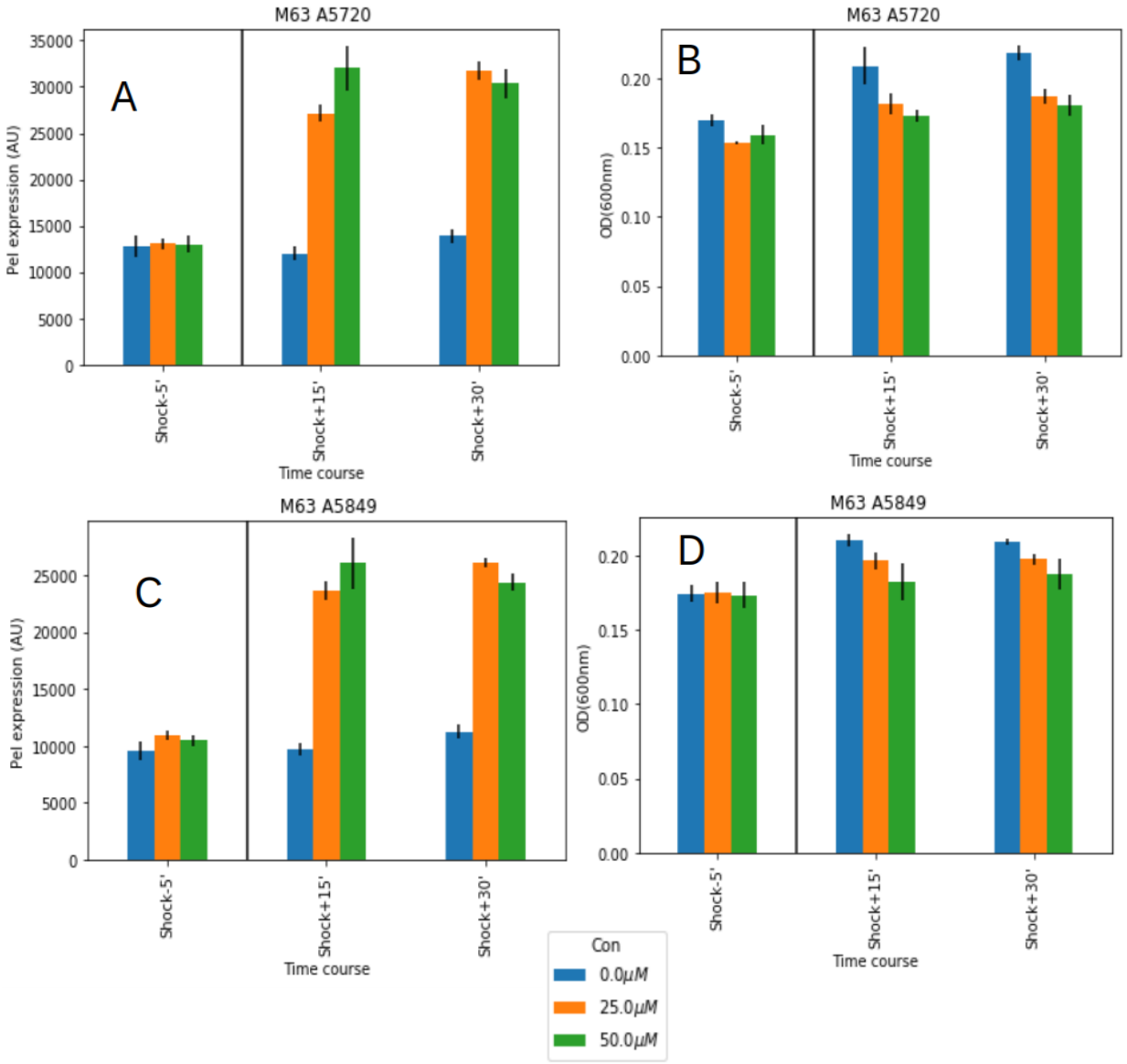


Figure 3.3.12: The bar charts representing the luminescence and OD recorded in the M63 medium just before the time of shock and 10, 15 and 30 minutes after the shock treatment with 25μM, 50μM concentrations of seconeolitsin and control with no seconeolitsin treatment. Parts A and C represent luminescence recorded in strain A5720 and A5849, respectively and B and D Show the variation in OD in strain A5720 and A5849, respectively along the shock treatment

Figure 3.3.12 shows the bar charts showing the increase in the expression level observed with the difference concentrations of seconeolistin at the time of shock and at 15 and 30 minutes after shock. In M63 medium we observe about 3 times increase in *pelE* expression level at 15 and 30 minutes after shock respectively. The same experiment also was carried out in LB medium.

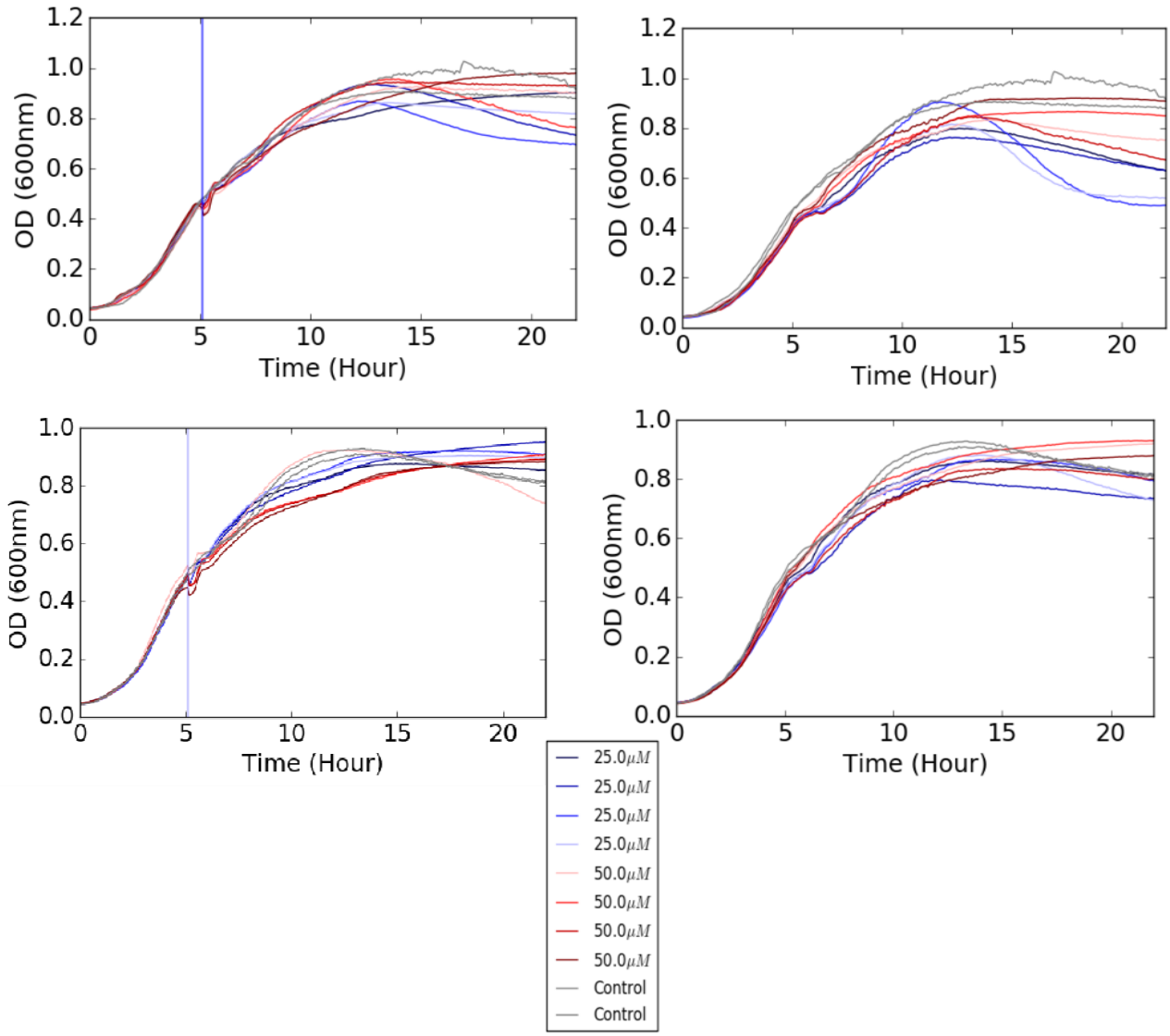


Figure 3.3.13: Growth curves corresponding to the experiments to study the effect of seconeolitsin shock of *pelE* gene expression in LB medium. Parts A and B represent growth curve in LB medium with 25μM, 50μM concentrations of seconeolitsin with and without shock respectively in forward orientation construct A5720 and C and D representing growth curve in LB medium with 25μM, 50μM concentrations of seconeolitsin with and without shock respectively in reverse orientation construct A5489

As we observe in Figure 3.3.13 growth in LB medium is not significantly affected by addition of antibiotics in the shock mode, the effect observed in the expression of the genes is due to the effect created by seconeolitsin. Knowing *pelE* gene is susceptible to supercoiling, it's expression is expected to increase soon after the shock effect is observed.

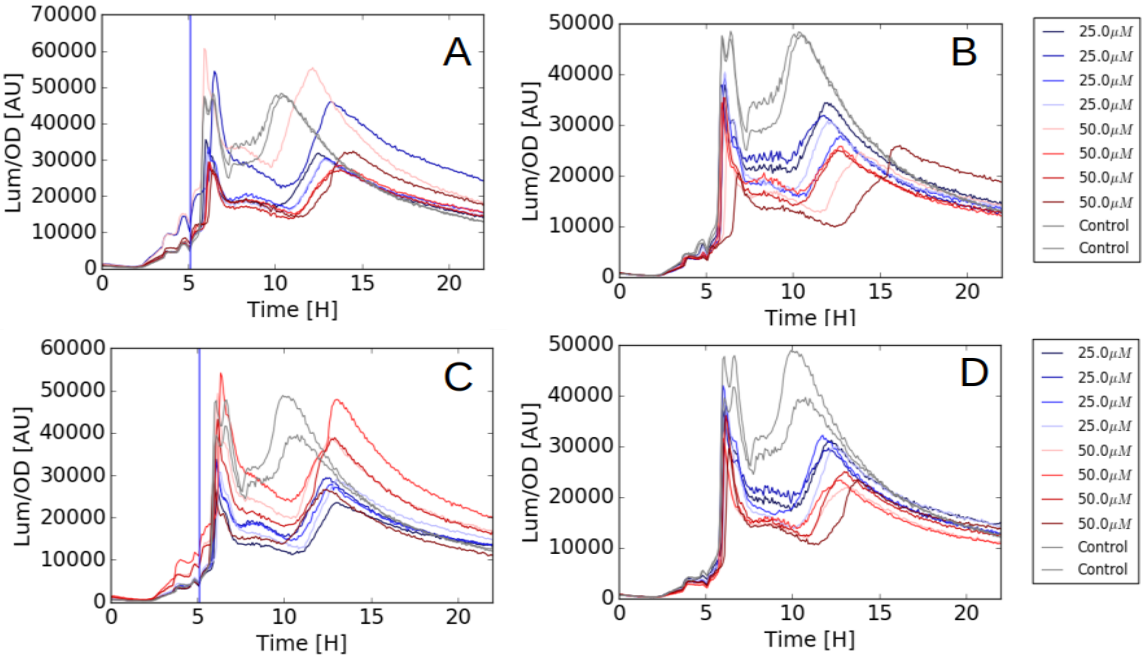


Figure 3.3.14: The real time luminescence recorded in the LB medium with various conditions. Parts A and B represent luminescence observed in LB medium and 25μM, 50μM concentrations of seconeolitsin with and without shock respectively in forward orientation construct A5720 and C and D representing real time luminescence recorded in LB medium with 25μM, 50μM concentrations of seconeolitsin with and without shock respectively in reverse orientation construct A5489

The following Figure 3.3.14 shows the real time data of luminescence resulting from the expression of the *pelE* promoter in nutrient rich LB medium. In order to make sure this is an effect of the process of shock by seconeolitsin we have a control sample, where bacteria is grown in the presence of the same concentration of seconeolitsin from the onset of the growth in the same experimental set up. There is a slight increase in expression observed at 5h observed in these graphs proving that seconeolitsin has activated the *pelE* promoter. The experiments are carried out in both orientation of the reporter genes to study the effect of orientation on the gene expression triggered by change in supercoiling levels.

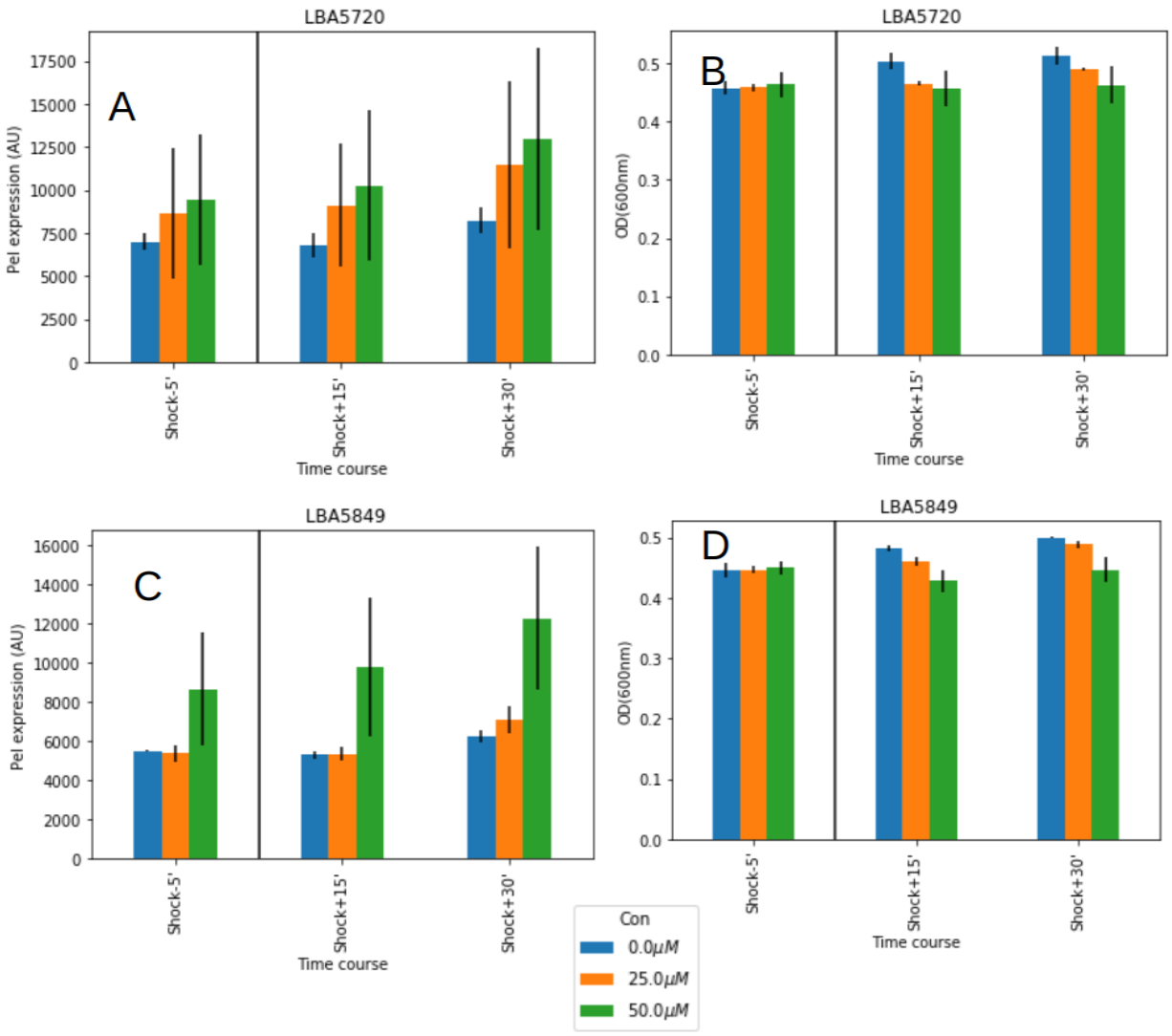


Figure 3.3.15: The bar charts representing the luminescence and OD recorded in the LB medium just before the time of shock and 10, 15 and 30 minutes after the shock treatment with 25μM, 50μM concentrations of seconeolitsin and control with no seconeolitsin treatment. Parts A and C represent luminescence recorded in strain A5720 and A5849, respectively and B and D Show the variation in OD in strain A5720 and A5849, respectively along the shock treatment

The expression levels of various concentration shock treated bacteria are recorded. Just after the shock to be time 0 and the after 10, 15 and 20 minutes respectively. Figure 3.3.15 depicts the variation observed in *pelE* expression up on seconeolitsin treatment. Compared to the control samples a slight variation in the gene expression is observed. The increase in expression level is not very statistically significant. This could be an artifact of the medium. The difference is not significant enough to be identified as an effect on expression level due to variation in supercoiling caused by the seconeolistin treatment.

A clear differentiation can be observed in minimal medium when it comes to the effect on

gene expression. The level of expression is also higher when compared to LB medium. Control experiments in M63 medium show significant changes in the *pelE* expression level when the bacterial growth is onset with seconeolitsin. In order to document the induction effect of seconeolitsin by altering the DNA supercoiling levels the activation factors are calculated as a ratio of level of expression in the non-treated to the treated samples.

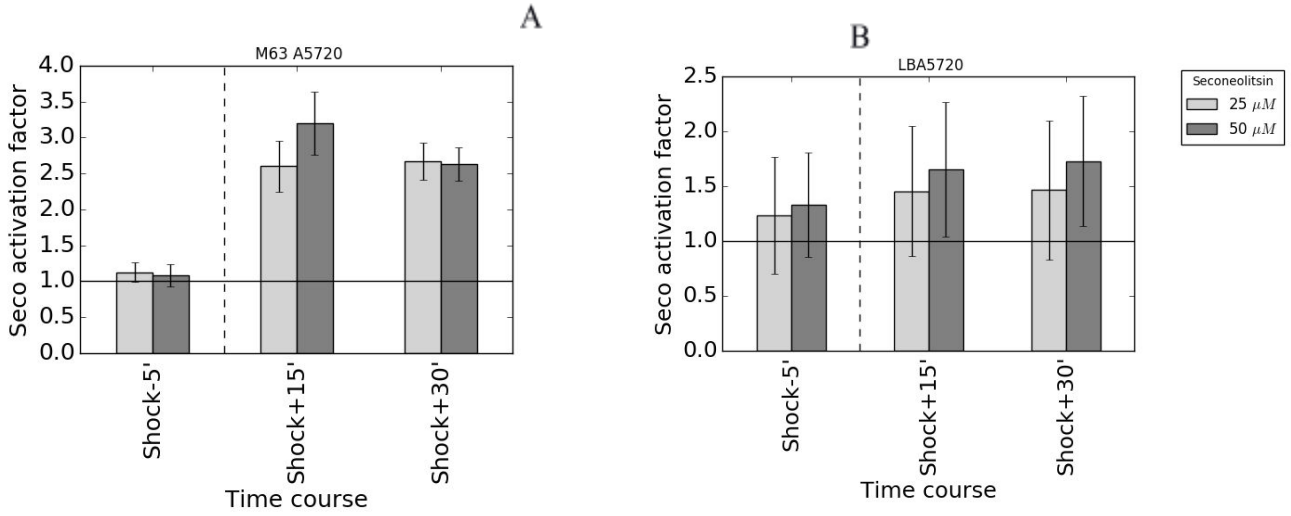


Figure 3.3.16: The *pel* activation factor with respect to time after the shock in M63 and LB medium

The Figure 3.3.16 shows the activation factors for both the mediums when the model organism is treated with the antibiotic seconeolitsin. The concentrations of seconeolitsin used for this study was 25μM and 50μM where no effect on growth was observed. The above Figure 3.3.16A depicts the seconeolitsin activation factor in the M63 medium for the concentration of 25μM and 50μM where maximum effect is observed 15 minutes aftershock and further you can infer from the expression graph that the effect reduces in an hour. The same for LB medium is shown in Figure 3.3.16B with same concentrations but here no prominent activation of the gene is observed. As expected, we observed an increase in the activity level of the *pelE* gene in the presence of seconeolitsin inhibitor in *Dickeya dadantii*. At the concentration where no effect on growth was observed, a three-fold increase (p value .00005) in the *pelE* expression level was observed within 30 minutes of the shocking process in minimal medium and in the energy rich LB medium, no significant activation was observed. The effect of antibiotic inhibition lasted for about one hour as observed in Figure 3.3.11, where the deviation in the expression due to activation by seconeolitsin shock treatment returned to the expression pattern of the non-treated control curve. This Indicates that the effect of seconeoltsin lasted for a period on *pelE* gene expression was temporary until the drug inhibition lasted.

Seconeolitsin binding to the Topoisomerases increases the level of supercoiling increasing the *pelE* expression levels, but this state is soon reversed by the additional DNA relaxation

caused in bacteria returning the *pelE* gene to the normal level of expression. It can be hypothesized that the process is easier in energy-rich media. As a result, we do not observe a strong activation factor in LB medium whereas in the minimal energy M63 medium the effect is more prominent.

Since we observed the variation in the gene activity with changes in supercoiling level of DNA, would change in orientation of the gene affect the variation caused in the gene activity? To answer this question we analysed results with respect to forward and reverse oriented *pelE* gene. We used specific constructs with onward and reverse orientation of promoter fused with luciferase gene (refer material and methods). The *pel* expression level of forward orientation strain A5720 is about 10% higher than the reverse orientation strain A5849. But the activation factor for both the orientations remains the same. The growth curve for the two strains also shows a difference in the final growth OD which explains the difference in the *pel* expression levels. Thus, this difference in the level of expression may not be caused by seconeolitsin and possibly it is the characteristic of the individual strain. With this reference, it can be hypothesized that the change in supercoiling levels induced by the antibiotic is not a local effect but has a global effect on the chromosome. Figure 3.3.17 shows the expression level of *pel* genes in forward strain A5849 and reverse strain A5720.

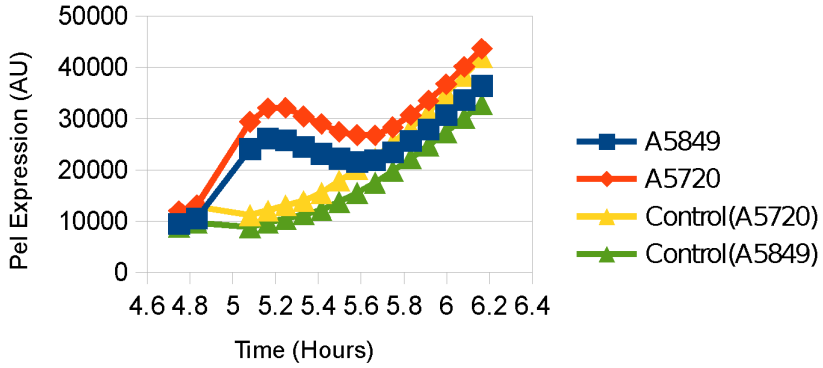


Figure 3.3.17: *pel* expression curves during the shocking in the M63 medium. A5720 and A5849 are the forward and reverse orientation *pel*E construct, respectively

The effect of seconeolitsin on *pelE* expression at stationary phase

This study is carried out by the shock treating the bacteria at a stationary phase that is around 10 hours after the growth when the optical density reaches around transition and then stationary phase is reached. Since at exponential phase the significant difference is observed only in minimal medium M63, we perform this study in M63 medium. From the characterization growth curves we observe that the bacteria reaches a stationary phase around 10 hours, so we shock the bacteria at 10 hours with three different concentrations of seconeolitsin 25µM, 50µM and 100µM.

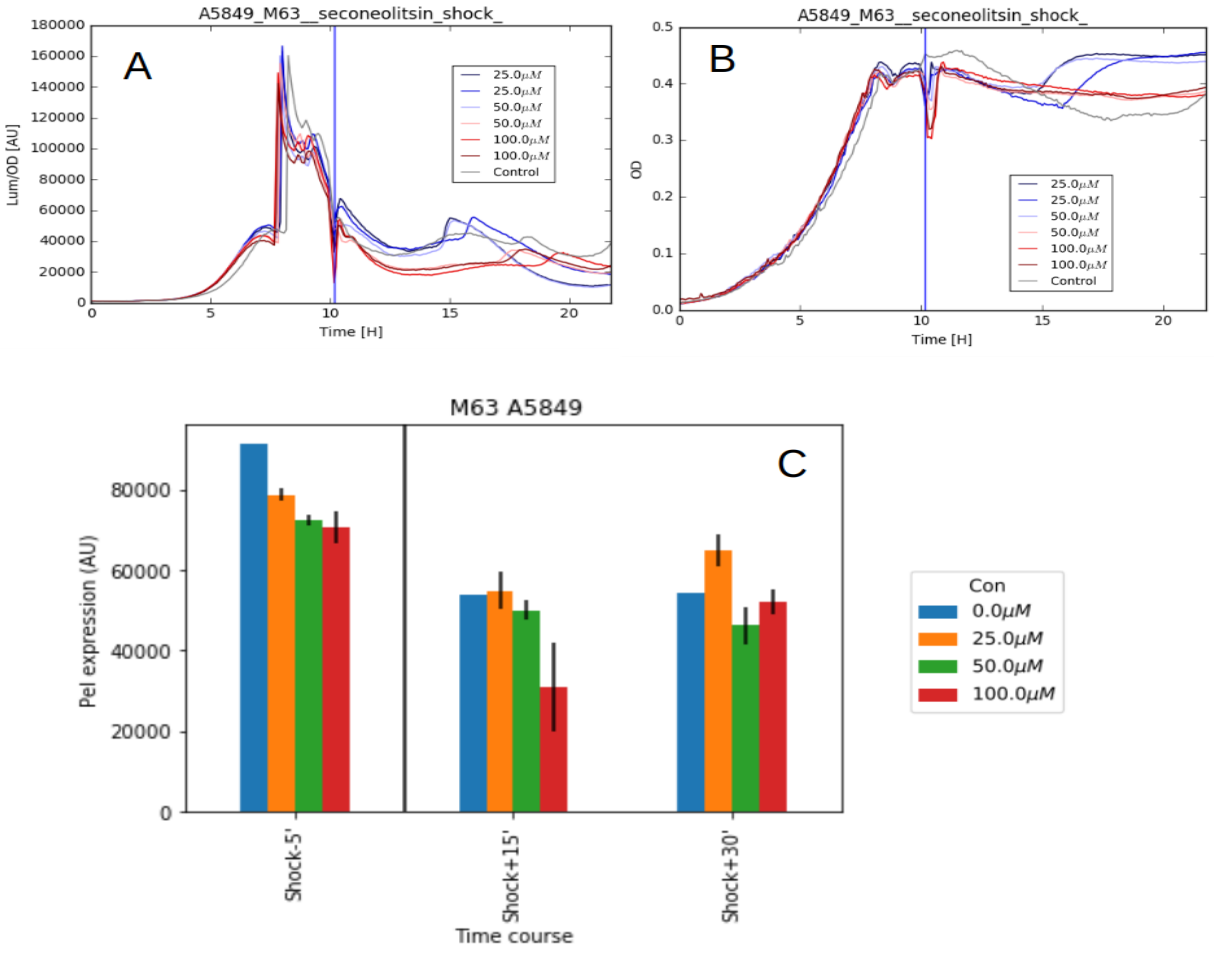


Figure 3.3.18: The effect of seconeolitsin on *pelE* expression at stationary state. A represents the luminescence recorded along the growth period represented in B. The bacteria is grown in M63 medium. C represents the luminescence recorded after seconeolitsin shock treatment

The above Figure 3.3.18 shows no significant difference observed in the levels of *pelE* expression after shock treatment. It can be concluded that since the *pelE* gene has already set expression at this period. Due to this state of the gene after about 10 hours of growth, the change in supercoiling levels created by the treatment of seconeolitsin may not have been effective in increasing level of *pelE* expression. The DNA is generally less supercoiled in stationary phase compared to exponential growth phase, this indicates that the topoisomerase activity is lesser in stationary phase. Hence, the inhibition of topoisomerase I may not be as significant as in the exponential growth phase on the expression of the gene. Looking at the expression pattern of the gene, the *pelE* gene has essentially already been expressed and triggering the expression again is not in favour of energy conservation during the stationary phase when the nutrients have been exhausted by the bacteria. In conclusion, the seconeolitsin shock treatment is found to be more effective in the exponential growth phase compared to the stationary phase.

Gene expression analysis

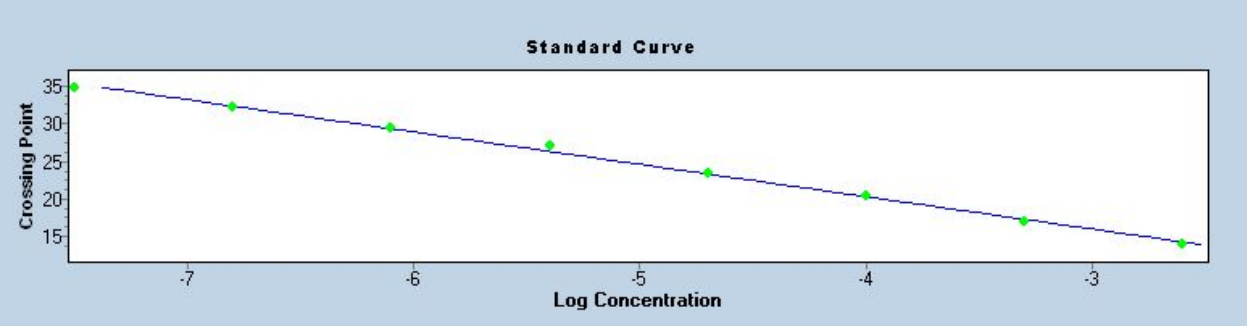
In order to study the effect of supercoiling on gene expression of *Dickeya dadantii*, RNA sequencing analysis is the widely proposed method. The antibiotic seconeolitsin is used to modify the level of DNA supercoiling. The wild type strain of *Dickeya dadantii* A4922 is grown in the minimal M63 medium supplemented with sucrose. The concentration of seconeolitsin used for shock is 50 µM as per observations recorded in chloroquine gel experiments. Bacterial culture undergoing seconeolitsin treatment procedure but not
subjected to seconeolitsin is designated as the control for the study.

The experiment is carried out in two growth phases, exponential at OD = 0.2 and stationary phase at OD = 1.1. The wild type strain A4922 is grown in the minimal M63 medium supplemented with sucrose, when the growth reaches the OD of 0.2, one portion is shocked with 50µM seconeolitsin and is grown for 15 more minutes. Same process is repeated for the culture grown until the stationary phase. All the samples are frozen and stored in -80°C until the extraction of the RNA is carried out using the cold phenol method. Extracted RNA is quantified using Nanodrop Quantification. A PCR experiment is carried out to check the quality of extracted RNA. The integrity of RNA is checked on the Electrophoresis gel Post PCR. Sample is subjected to repeated DNases to ensure purity of RNA. Purified RNA is subjected to quality control and used for cDNA synthesis to test the effect on the model genes using qRT PCR experimentation before proceeding with the sequencing analysis.

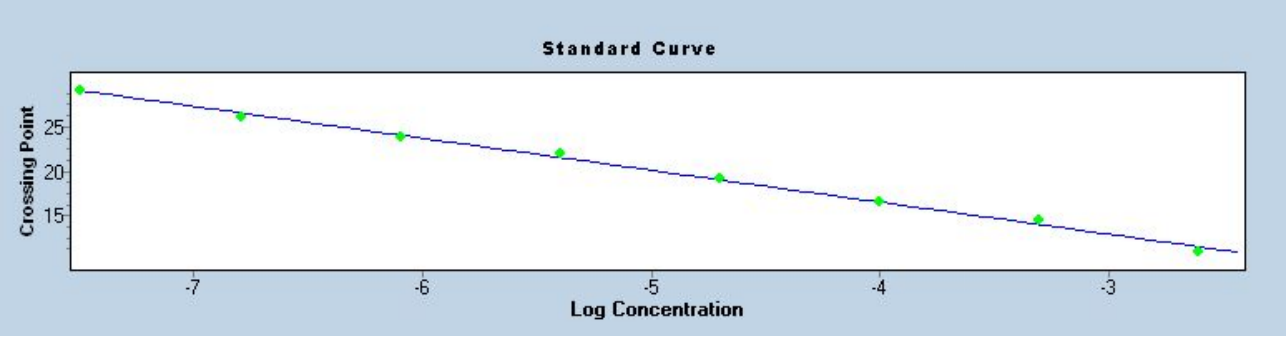
qRT-PCR analysis

The real-time recording is performed with the equipment Light cycler LC480. The experiment is carried *pAW* as external normalizers. *lpxC, hemF,* and *yafS* are used as internal normalizers. *gyrB* and *pelE* are used as model genes to know the effect of supercoiling. *pelE* expression is susceptible to variation in supercoiling as is known in earlier studies.

pAW



pelE



gyrB

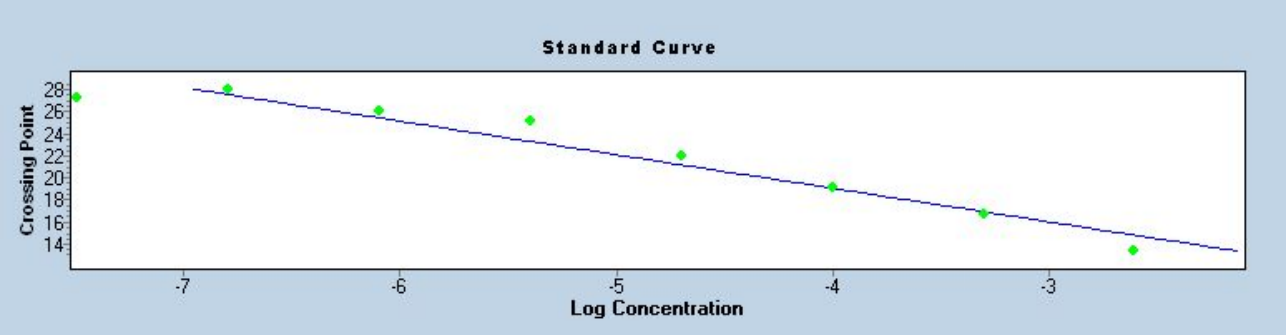


Figure 3.3.19: The standard curve for the quantification of *pAW, pelE* and *gyrB* external normalizer

Figure 3.3.19 shows the standard curve of quantification for *pAW* external normalizer, *pelE* and *gyrB* for the experiment.The external normalizer had a stable expression level indication of the quality of the experiment. Not all the internal normalizers have stable expression levels. It could be due to the variation in the level of supercoiling caused by the antibiotic. Since the antibiotic is studied for the first time in this organism, no further information on internal normalizer is known, making the interpretation of the data more difficult. Thus, the results could not be inferred quantitatively rather the external normalizers could give qualitative assurance in the experiment results.

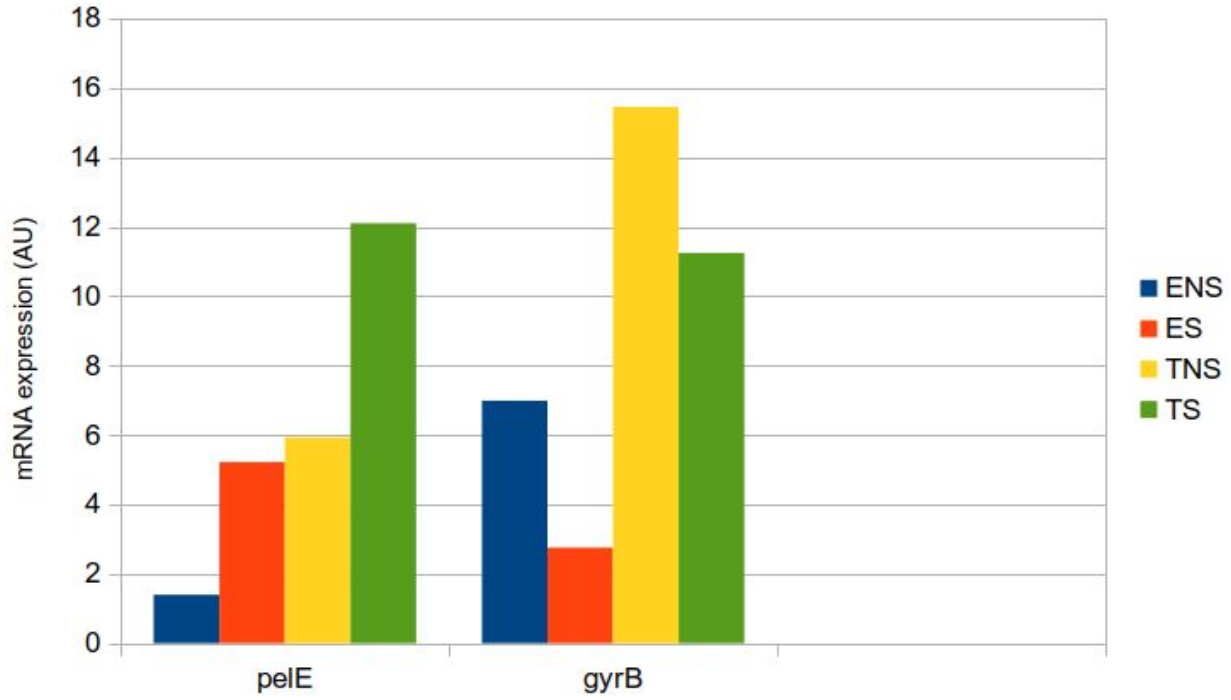


Figure 3.3.20: Quantification of the level of genes expression with qRT-PCR method. The expression of pelE, gyrB and topA genes are shown in the figure. Calculations carried out based on standard curve estimation method. The obtained result is proportional and hence multiplied by an arbitrary constant to obtain a meaningful number. The results can be compared based on their levels and are not absolute

As per out preliminary analysis shown in Figure 3.3.20, *pelE* expression was increased with Seconeolitsin treatment as expected. The *gyrB* expressions were inconclusive as the results do not follow a trend and no replicates were performed to support the results. The standard curve for quantification of *gyrB* was not very accurate. Problems were also encountered with interna normalizers questioning the validity of results. Difficulties in being able to find a stable internal normalizer leads to problems in replications of the experiment.

RNA sequence analysis was performed but there was not enough time to validate the results obtained. Currently PCR validation experiments are being repeated to validate the results. Since, there was a discrepancy in the results obtained earlier and a problem on internal normalizer the complete analysis of the transcriptomic analysis would need more time. Considering the limitation of time at disposal, and the thesis focuses on the modeling of regulation of virulence genes with respect to transcription factors and transcriptomic analysis is not very significant for the modeling part choice was made to focus on transcription factors.

d) References

[1] S. Martis B, R. Forquet, S. Reverchon, W. Nasser, and S. Meyer, “DNA Supercoiling: an Ancestral Regulator of Gene Expression in Pathogenic Bacteria?,” *Comput. Struct. Biotechnol. J.*, vol. 17, pp. 1047–1055, 2019, doi: 10.1016/j.csbj.2019.07.013.
[2] L. A. Freeman and W. T. Garrard, “DNA supercoiling in chromatin structure and gene expression,” *Crit. Rev. Eukaryot. Gene Expr.*, vol. 2, no. 2, pp. 165–209, 1992.
[3] R. R. Sinden, *DNA Structure and Function*. Burlington: Elsevier Science, 2012.
[4] R. E. Ashley, A. Dittmore, S. A. McPherson, C. L. Turnbough, K. C. Neuman, and N. Osheroff, “Activities of gyrase and topoisomerase IV on positively supercoiled DNA,” *Nucleic Acids Res.*, vol. 45, no. 16, pp. 9611–9624, Sep. 2017, doi: 10.1093/nar/gkx649.
[5] M. A. Kohanski, D. J. Dwyer, and J. J. Collins, “How antibiotics kill bacteria: from targets to networks,” *Nat. Rev. Microbiol.*, vol. 8, no. 6, pp. 423–435, Jun. 2010, doi: 10.1038/nrmicro2333.
[6] V. E. Anderson and N. Osheroff, “Type II topoisomerases as targets for quinolone antibacterials: turning Dr. Jekyll in to Mr. Hyde,” *Curr. Pharm. Des.*, vol. 7, no. 5, pp. 337–353, Mar. 2001, doi: 10.2174/1381612013398013.
[7] K. A. Hurley *et al.*, “Targeting quinolone- and aminocoumarin-resistant bacteria with new gyramide analogs that inhibit DNA gyrase,” *MedChemComm*, vol. 8, no. 5, pp. 942–951, 2017, doi: 10.1039/C7MD00012J.
[8] M.-J. Ferrándiz, A. J. Martín-Galiano, C. Arnanz, I. Camacho-Soguero, J.-M. Tirado-Vélez, and A. G. de la Campa, “An increase in negative supercoiling in bacteria reveals topology-reacting gene clusters and a homeostatic response mediated by the DNA topoisomerase I gene,” *Nucleic Acids Res.*, vol. 44, no. 15, pp. 7292–7303, Sep. 2016, doi: 10.1093/nar/gkw602.
[9] N. Blot, R. Mavathur, M. Geertz, A. Travers, and G. Muskhelishvili, “Homeostatic regulation of supercoiling sensitivity coordinates transcription of the bacterial genome,” *EMBO Rep.*, vol. 7, no. 7, pp. 710–715, Jul. 2006, doi: 10.1038/sj.embor.7400729.
[10] C. Marr, M. Geertz, M.-T. Hütt, and G. Muskhelishvili, “Dissecting the logical types of network control in gene expression profiles,” *BMC Syst. Biol.*, vol. 2, p. 18, Feb. 2008, doi: 10.1186/1752-0509-2-18.
[11] X. Jiang, P. Sobetzko, W. Nasser, S. Reverchon, and G. Muskhelishvili, “Chromosomal ‘stress-response’ domains govern the spatiotemporal expression of the bacterial virulence program,” *mBio*, vol. 6, no. 3, pp. e00353-00315, Apr. 2015, doi: 10.1128/mBio.00353-15.
[12] Z.-A. Ouafa, S. Reverchon, T. Lautier, G. Muskhelishvili, and W. Nasser, “The nucleoid-associated proteins H-NS and FIS modulate the DNA supercoiling response of the pel genes, the major virulence factors in the plant pathogen bacterium Dickeya dadantii,” *Nucleic Acids Res.*, vol. 40, no. 10, pp. 4306–4319, May 2012, doi: 10.1093/nar/gks014.
[13] X. Jiang *et al.*, “Global transcriptional response of Dickeya dadantii to environmental stimuli relevant to the plant infection,” *Environ. Microbiol.*, vol. 18, no. 11, pp. 3651–3672, 2016, doi: 10.1111/1462-2920.13267.
[14] P. Sobetzko, “Transcription-coupled DNA supercoiling dictates the chromosomal arrangement of bacterial genes,” *Nucleic Acids Res.*, vol. 44, no. 4, pp. 1514–1524, Feb. 2016, doi: 10.1093/nar/gkw007.
[15] S. Hirose and K. Matsumoto, *Possible Roles of DNA Supercoiling in Transcription*. Landes Bioscience, 2013.

[16] C. J. Dorman and C. P. Corcoran, “Bacterial DNA topology and infectious disease,” *Nucleic Acids Res.*, vol. 37, no. 3, pp. 672–678, Feb. 2009, doi: 10.1093/nar/gkn996.
[17] A. Lal, A. Dhar, A. Trostel, F. Kouzine, A. S. N. Seshasayee, and S. Adhya, “Genome scale patterns of supercoiling in a bacterial chromosome,” *Nat. Commun.*, vol. 7, Mar. 2016, doi: 10.1038/ncomms11055.
[18] C. J. Dorman, “DNA supercoiling and environmental regulation of gene expression in pathogenic bacteria.,” *Infect. Immun.*, vol. 59, no. 3, pp. 745–749, Mar. 1991.
[19] C. J. Dorman, “Flexible response: DNA supercoiling, transcription and bacterial adaptation to environmental stress,” *Trends Microbiol.*, vol. 4, no. 6, pp. 214–216, Jun. 1996, doi: 10.1016/0966-842X(96)30015-2.
[20] A. G. D. L. CAMPA, M. T. G. ESTEBAN, and M. A. B. FERRER, “Use of seconeolitsine and n-methyl-seconeolitsine for the manufacture of medicaments,” WO2011073479A1, Jun. 23, 2011.
[21] M. H. Zwietering, I. Jongenburger, F. M. Rombouts, and K. van ’t Riet, “Modeling of the Bacterial Growth Curve,” *Appl. Environ. Microbiol.*, vol. 56, no. 6, pp. 1875–1881, Jun. 1990.
[22] R. L. O’Brien, J. L. Allison, and F. E. Hahn, “Evidence for intercalation of chloroquine into DNA,” *Biochim. Biophys. Acta BBA - Nucleic Acids Protein Synth.*, vol. 129, no. 3, pp. 622–624, Dec. 1966, doi: 10.1016/0005-2787(66)90078-5.
[23] F. Fan and K. V. Wood, “Bioluminescent Assays for High-Throughput Screening,” *ASSAY Drug Dev. Technol.*, vol. 5, no. 1, pp. 127–136, Feb. 2007, doi: 10.1089/adt.2006.053.
[24] J. M. Andrews, “Determination of minimum inhibitory concentrations,” *J. Antimicrob. Chemother.*, vol. 48, no. suppl_1, pp. 5–16, Jul. 2001, doi: 10.1093/jac/48.suppl_1.5.
[25] L. L. McKay, “Functional properties of plasmids in lactic streptococci,” *Antonie Van Leeuwenhoek*, vol. 49, no. 3, pp. 259–274, 1983, doi: 10.1007/BF00399502.

Chapter IV: Conclusion and Future scope

1.Conclusion and Future Scope

We have successfully modeled the transcriptional regulation of *pelE/D* genes with respect to the KdgR regulator and CRP activator. Quantification of the metabolites KDG and cAMP proved very important for the development and validation of the model. Although KDG quantification methodology was accurate, we encountered problems in quantifications of internal cAMP concentration. The discrepancy was posed by the protocol of cell separation from medium. We employed slow centrifugation techniques for the separation of cells. Since cAMP is a very dynamic molecule, slight disturbance in the cell membrane causes an inaccurate distribution of the metabolite due to the import-export mechanism. Since the external medium is in large quantities, this dynamics of import export did not affect the quantification of external cAMP. Thus, to overcome this erroneous internal cAMP measurement problem we used a model to deduce the internal concentration using the external concentration measurement. The future scope of this work would be to obtain a methodology to accurately measure the internal concentration of cAMP. The currently employed HPLC quantification can be employed with precision if we could filter the bacterial cells from medium without disrupting the distribution dynamics of cAMP inside and outside the cell during culture harvest. Sophisticated filtration techniques can be used to achieve this cell separation from the medium [1]. All the methods have their own advantages and disadvantages. Filtration methods are most promising given the sensitivity of mobile cAMP molecules on the bacterial cell wall. Nucleopore filters are most efficient in separating cells without much disturbance to the cell structure [2]. It would be worthwhile to develop a customized methodology to separate the bacterial cells to ensure accurate measurement of internal cAMP. This could enable us to integrate directly the internal cAMP kinetics into the model intern improving the cogency of the model.

We have observed minor discrepancies in replicating the *pel* gene expression peak by 30 minutes or so when compared to the experimental peak timing. We know *pel* genes are regulated by various transcription factors as discussed in chapter I. Another future perspective would be to extend the model to other regulatory factors to further improve the credibility. We nominate the nucleoid associated protein and well known *pel* regulator FIS to be the next most probable candidate. Integration of FIS might help us reproduce the experiments to a higher extent. The concentration of FIS in bacteria is known to vary to a large extent with respect to bacterial growth stage. Since *pel* expression is also linked to the metabolic state of the cell which is determined by the growth stage, FIS could explain the peak of expression better. We already know the binding affinities of FIS to the *pel* sites obtained from the experiments conducted by our lab earlier. The binding affinities of FIS at *pelE* is estimated to be about 10nM and at *pelD* is estimated to be around 50nM [3]. We also know that the binding sites for FIS in *pelE* and overlap with the KdgR binding site [3]. This points towards a possible competition between the FIS and KdgR molecules at the respecting

gene binding site. We will have to incorporate this competitive binding in the model. We will have to measure the kinetics of available FIS molecules in the bacterial cell in order to deduce the extent of binding and the regulation effect on *pel* genes.

We have substantially studied the effect of supercoiling levels on the *pelE* gene expression. DNA supercoiling levels vary with respect to growth of the bacteria. In order to completely elucidate the effect of growth stage on virulence genes we have to understand the dependency of virulence gene expression on the DNA supercoiling levels [4]. In the future we could incorporate the supercoiling mediated regulation into the transcriptional regulation model. It will give a new dimension to the model and integrate the chromosomal architecture mediated regulation with the transcription factors mediated regulation. There has been active work understanding and modeling the relationship between DNA supercoiling levels and gene expression [5]. We could utilize this data to develop a strategy to integrate the thermodynamic model of gene expression regulation incorporating the torsional energy factors associated with supercoiling levels. New techniques are being reported to measure the levels of DNA supercoiling levels *in vivo* along the bacterial growth [6]. This could be utilized to draw a relationship between the supercoiling measured along the growth and *pel* expression.

As discussed earlier the proposed extended regulator FIS and supercoiling levels are strictly related to the bacterial growth. The DNA supercoiling levels are mediated through the NAPs like FIS. Inclusion of these two factors would reveal more information on mechanism of gene virulence regulation by NAPs organizing the chromosomal structure in relation to bacterial growth. This leads to the ultimate scope of modeling work, elucidating the link between metabolism and virulence genes regulation. We describe a transcriptional regulatory model in light of bacterial cell metabolism. Virulence genes are regulated in accordance with metabolic shifts acting as a signal. We depict the *pel* genes to be an instance of catabolite repression where the depletion of glucose causes metabolic energy depletion thus the virulence genes are boosted to enhance the uptake of alternate glucose sources to provide for the energy requirement in the cell. CCR helps bacteria derive energy in the new environments encountered during their life cycle and proves very important for pathogens to switch their energy source. CCR helps the pathogens obtain sufficient energy in the host environment and successfully infect the host organism. Metabolic switch from glucose to pectin acts as a signal activating the *pel* expression. The pectin degradation also destroys the host defense layer, i.e. the plant cell wall making it easier to infect the host. This example in the model organism strongly advocates for the regulatory link between energy requirements for the pathogen via metabolism and virulence gene regulation. Thus, pathogenesis is a process of metabolism leading to virulence gene expression overcoming the host defense and propagation of disease in the host.
The approach to tackling pathogenesis in plants is time efficient. It also enables us to utilize all the known experimental information to deduce new learning and save time in terms of running redundant experiments. Compared to the canonical approach of pathogen management involving traditional experimentation to understand the infection process and

develop measures against it, one can integrate all the information at disposal from previous studies and bibliography and develop the model to obtain meaningful inferences which can be validated with significant little experimentations. This tool can be extended to other pathogenic organisms facilitating development of disease control strategies and pathogen management. Thus, we aim to develop a sustainable approach to develop a tool to fight the pathogens with minimal time and resources.

Transcriptional regulation of virulence pel genes are modeled with respect to transcription factors based on an affinity binding thermodynamics studied in the model organism model Dickeya dadantii. This model is simple and easy to be extrapolated. It can be extended to other virulence genes in different pathogens. The model can be used to study the virulence genes in different pathogens and understand the process of infection with respect to pathogens. Having a functional and relational understanding of the virulence genes, the key strength of pathogens will help us tackle them. This model is developed to be employed to aid in developing a strategy to fight pathogens. This could help the world hunger by stopping croploss mediated by microbial pathogens. In this era of growing food insecurity around the world, I dedicate this work to the scientific community constantly trying to come up with strategies to feed the growing population.

2.References

[1] M. Almeida, A. C. García-Montero, and A. Orfao, “Cell Purification: A New Challenge for Biobanks,” *Pathobiology*, vol. 81, no. 5–6, pp. 261–275, 2014, doi: 10.1159/000358306.
[2] R. Cornell, “The use of Nucleopore filters in ultrastructural studies of cell cultures,” *Exp. Cell Res.*, vol. 56, no. 1, pp. 156–158, Jul. 1969, doi: 10.1016/0014-4827(69)90409-1.
[3] T. Lautier and W. Nasser, “The DNA nucleoid-associated protein Fis co-ordinates the expression of the main virulence genes in the phytopathogenic bacterium Erwinia chrysanthemi,” *Mol. Microbiol.*, vol. 66, no. 6, pp. 1474–1490, 2007, doi: 10.1111/j.1365-2958.2007.06012.x.
[4] S. Martis B, R. Forquet, S. Reverchon, W. Nasser, and S. Meyer, “DNA Supercoiling: an Ancestral Regulator of Gene Expression in Pathogenic Bacteria?,” *Comput. Struct. Biotechnol. J.*, vol. 17, pp. 1047–1055, 2019, doi: 10.1016/j.csbj.2019.07.013.
[5] B. El Houdaigui *et al.*, “Bacterial genome architecture shapes global transcriptional regulation by DNA supercoiling,” *Nucleic Acids Res.*, vol. 47, no. 11, pp. 5648–5657, Jun. 2019, doi: 10.1093/nar/gkz300.
[6] A. Duprey and E. A. Groisman, “FEDS: a Novel Fluorescence-Based High-Throughput Method for Measuring DNA Supercoiling In Vivo,” *mBio*, vol. 11, no. 4, Aug. 2020, doi: 10.1128/mBio.01053-20

FOLIO ADMINISTRATIF

THESE DE L'UNIVERSITE DE LYON OPEREE AU SEIN DE L'INSA LYON

NOM : MARTIS BADIADKA DATE de SOUTENANCE : 06.11.2020

Prénoms : Shiny

TITRE : Computational modeling of virulence genes expression: a study of gene regulation during plant pathogenesis

NATURE : Doctorat Numéro d'ordre : 2020LYSEI100

Ecole doctorale : Ecosystèmes Evolution Modélisation Microbiologie

Spécialité : Biomath-Bioinfo-Génomique evolutive

RESUME :

Plant pathogens are a major threat to food security all around the world. Virulence genes enable the pathogens to disease the host and pose threats. We study and computationally model the expression of virulence in pathogens. The pel genes, a major virulence factors in pectinolytic bacteria which infect plants and cause soft rot disease. The Pel enzymes released by the pathogen results in the maceration of tissue in crop and facilitate pathogenesis. We model the pel genes in our model organism *Dickeya dadantii*. We study pelD/E genes which encode proteins with similar enzymatic activities, but also share the same set of transcriptional regulators, thus raising the question of the benefit of this duplication and the specific role of these two genes. We model their expression with respect KdgR repressor and CRP activator. We develop a quantitative dynamic kinetic model of this process, that reproduces the observed behavior of the two genes and explains their specific role in the infection. The regulatory part of the model is based on experimental values, and fitted kinetic parameters are subjected to systematic evaluation. The model is robust and has multiple applications for studying the pathogenesis of pectinolytic bacteria. In addition, we explore the instance of carbon catabolite repression observed during the regulation of virulence genes *Dickeya dadantii*.
The novel antibiotic seconeolitsin is characterized in the gram-negative model organism *Dickeya dadantii*. Growth kinetics and MIC studies are carried out. The variations in the level of supercoiling in the presence of seconeolitsin and novobiocin are quantified and analyzed. The effect of supercoiling on the expression of genes, especially the *pelE* gene is studied using the novel antibiotic seconeolitsin. The study explores the role of supercoiling as a gene regulator.

MOTS-CLÉS : Transcription model, Dickeya, Virulence genes, gene regulation.

Laboratoire (s) de recherche : Microbiology Adaptation and Pathogenesis (MAP)

Directeur de thèse: Sylvie REVERCHON

Président de jury :

Composition du jury :

GENIN,	Stephane	Directeur de Recherche	CNRS	Rapporteur
de-JONG,	Hidde	Directeur de Recherche	INRIA	Rapporteur
REVERCHON,	Sylvie	Professeur	INSA, Lyon	Directrice de thèse
MEYER,	Sam	Maitre de conferences	INSA, Lyon	Co-directeur de thèse