

BIOLOGICAL NETWORK MODELLING BASED ON DIFFERENTIALLY
EXPRESSED PROTEINS IN A BACILYSIN-DEFICIENT STRAIN OF
DCEKNWU'UWDVKNKU

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MELTEM KUTNU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MOLECULAR BIOLOGY AND GENETICS

SEPTEMBER 2019

Approval of the thesis:

**BIOLOGICAL NETWORK MODELLING BASED ON DIFFERENTIALLY
EXPRESSED PROTEINS IN A BACILYSIN-DEFICIENT STRAIN OF
*DCEKNNWU'UWDVKNKU***

submitted by **MELTEM KUTNU** in partial fulfillment of the requirements for the
degree of **Master of Science in Molecular Biology and Genetics Department,**
Middle East Technical University by,

Prof. Dr. Halil Kalıpçılar

Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Ayşe Gül Gözen

Head of Department, **Biology**

Prof. Dr. Gülay Özcengiz

Supervisor, **Biology, METU**

Assoc. Prof. Dr. Nurcan Tunçbağ

Co-Supervisor, **Informatics Institute, METU**

Examining Committee Members:

Assoc. Prof. Dr. Tülin Yanık

Biology, METU

Prof. Dr. Gülay Özcengiz

Biology, METU

Assoc. Prof. Dr. Çağdaş Devrim Son

Biological Sciences, METU

Assoc. Prof. Dr. Aybar Can Acar

Informatics Institute, METU

Assoc. Prof. Dr. Nefise Akçelik

Biotechnology, Ankara University

Date: 09.09.2019

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Meltem Kutnu

Signature:

ABSTRACT

BIOLOGICAL NETWORK MODELLING BASED ON DIFFERENTIALLY EXPRESSED PROTEINS IN A BACILYSIN-DEFICIENT STRAIN OF *DCEKNNWU'UWDVKNKU*

Kutnu, Meltem

Master of Science, Molecular Biology and Genetics

Supervisor: Prof. Dr. Gülay Özcengiz

Co-Supervisor: Assoc. Prof. Dr. Nurcan Tunçbağ

September 2019, 92 pages

Bacillus subtilis is a non-pathogenic, Gram-positive organism which is known for producing a broad range of secondary metabolites with pharmacological and antimicrobial activities. The dipeptide bacilysin is one of the many antibiotics synthesized by certain strains of *B. subtilis*, and it is composed of L-alanine and the non-proteinogenic amino acid L-anticapsin. Earlier silencing studies by our group have suggested that bacilysin acts as a pleiotropic molecule on its host. Therefore, the absence or lack of bacilysin would affect the host response. In this study, data from previous proteomic studies by our group on bacilysin-deleted derivative strain *Bacillus subtilis* OGU1 have been collected. We constructed time-dependent secretome and cytosolic proteome networks by using the Prize-collecting Steiner forest (PCSF) algorithm. Functional enrichment analyses on the resulting networks have confirmed our earlier findings, as well as revealing further hidden molecules including CcpA, BglH, HxlA and YwmD, along with pathways such as amino acid metabolism, ABC transporters, sugar metabolism, and one carbon pool by folate. Biological network modeling can facilitate the identification of unknown protein interactions, as well as interaction partners in organisms such as *Bacillus subtilis* under a particular condition, in addition to guiding further experiments.

Keywords: *Deinomyces*, Bacilysin, Proteomics, Network modeling, Pathway enrichment analysis

ÖZ

***DCEKNWU'UWDVKNW*'İN BASİLİSİN ÜRETEMİYEN BİR SUŞUNDA FARKLI İFADE GÖSTEREN PROTEİNLERE DAYALI BİYOLOJİK AĞ MODELLEMESİ**

Kutnu, Meltem

Yüksek Lisans, Moleküler Biyoloji ve Genetik

Tez Danışmanı: Prof. Dr. Gülay Özcengiz

Ortak Tez Danışmanı: Doç. Dr. Nurcan Tunçbağ

Eylül 2019, 92 sayfa

Bacillus subtilis, farmakolojik ve antimikrobiyal aktiviteye sahip çok çeşitli ikincil metabolitler üreten, hastalık yapmayan, Gram-pozitif bir organizmadır. *B.subtilis*'in belli suşları tarafından üretilen bir dipeptit olan basilisin, L-alanin ve protein kaynaklı olmayan bir amino asit olan L-antikapsinden oluşmaktadır. Grubumuz tarafından yapılan önceki susturma çalışmaları, basilisinin konak hücrede pleiotropik bir molekül olarak davrandığını önermiştir. Dolayısıyla basilisin yokluğu ya da eksikliği konak hücre yanıtını etkileyecektir.

Bu çalışmada, grubumuzun basilisin susturulmuş *Bacillus subtilis* OGU1 suşu üzerine yaptığı çalışmalardan elde edilmiş veriler kullanıldı. Değişen biyolojik süreç ve yolları belirlemek için Prize-Collecting Steiner Forest (PCSF) algoritması kullanılarak zamana bağlı sekretom ve sitozolik proteom etkileşim ağları oluşturuldu. Ortaya çıkan ağlar üzerinde yapılan işlevsel zenginleştirme analizleri önceki bulgularımızı doğrulamış, ayrıca CcpA, BglH, HxlA and YwmD gibi gizli molekülleri, bununla birlikte amino asit metabolizması, ABC taşıyıcıları, şeker ve folik asit tek karbon metabolizması gibi yolları açığa çıkarmıştır.

Biyolojik ađ modellemesi, belli bir kořul altında *Bacillus subtilis* gibi organizmalarda bilinmeyen etkileřim partnerlerinin ve protein etkileřimlerinin belirlenmesini kolaylařtırmakta, ileriki deneylere kılavuzluk edebilmektedir.

Anahtar Kelimeler: *Deekmu"ımdıkkı*, Basilisin, Proteomiks, Ađ modelleme, Yolak zenginleřtirme analizi

To my beloved family...

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Prof. Dr. Gülay Özcengiz, who has accepted me as one of her students. Had it not been for her invaluable guidance, support, patience, and encouragement, it would not have been possible to complete this thesis. I am also grateful for my co-supervisor Assoc. Prof. Dr. Nurcan Tunçbağ for her constant, practical support, constructive criticism and splendid guidance, which has provided me the swiftness to bring this work into completion.

I would also like to thank Asst. Prof. Dr. Aybar Can Acar for taking his time out to help me understand the results of our whole genome shotgun (WGS) sequencing analysis and subsequent contig assembly.

I want to present my appreciation to all of my former labmates Ozan Ertekin, Caner Aktaş, İlayda Baydemir, Naz Kocabay, Sergen Akaysoy, Gözde Çelik and Nazlı Hilal Erdal for their assistance, understanding, lots of good memories and friendship. I would also like to thank H. Cansu Demirel, Gökçe Senger, Cansu Dinçer, Şeyma Ünsal Beyge, Poorya Parvizi, Güngör Budak and Etkä Yapar for their support, help and friendship. I am especially indebted to Şeyma, Güngör and Etkä for answering my most nonsense questions about coding and helping me out, as well as Özlem Durukan and Ozan Ertekin for providing me the resources and teaching me biochemistry. Additionally, I am so thankful for Özlem's amazing friendship and help in all areas of life since day one.

Last but not least, I am more than grateful to have my parents by my side through thick and thin. I want to thank my mother Tülay Kutnu and my father Alp Kutnu for their continuous support, encouragement, patience and their immense and unconditional love.

TABLE OF CONTENTS

ABSTRACT	v
ÖZ	vvii
ACKNOWLEDGEMENTS	x
TABLE OF CONTENTS	xi
LIST OF TABLES	xiii
LIST OF FIGURES	xiv
LIST OF ABBREVIATIONS	xv
CHAPTERS	
1. INTRODUCTION	1
2. LITERATURE REVIEW	3
2.1. <i>Bacillus subtilis</i> as a model organism	3
2.2. Primary metabolism in <i>Bacillus subtilis</i>	4
2.3. An overview of sporulation and quorum sensing in <i>Bacillus subtilis</i>	5
2.4. Secondary metabolites and antibiotics of <i>Bacillus subtilis</i>	11
2.5. The diverse derivative strains of <i>Bacillus subtilis</i> 168	13
2.6. Proteomics	13
2.6.1. Comparison of gel based and gel free proteomics	14
2.7. The dipeptide antibiotic bacilysin and its activity as a pleiotropic molecule ..	15
2.8. Protein-protein interaction networks (PPIN's)	20
2.9. The STRING database	21
2.10. Omics Integrator and Prize-Collecting Steiner Forest (PCSF) algorithm	23
2.11. Network analysis and visualization	24

2.12. Aim of the present study	26
3. MATERIALS AND METHODS	27
3.1. Modification of datasets.....	27
3.2. Network modelling	278
3.3. Network analysis and visualization.....	279
3.4. Functional enrichment analysis.....	30
4. RESULTS AND DISCUSSION	31
4.1. Construction of OGU1 networks	31
4.2. Functional enrichment of resulting networks.....	37
4.3. Alterations in the energy metabolism burdens cellular processes in OGU1...	38
4.4. Dynamic secretome analysis of OGU1	531
4.4.1. Bacilysin triggers carbohydrate metabolic activity and promotes exponential phase	53
4.4.2. Bacilysin deficiency causes a general stress response in the early stationary phase.....	63
4.4.3. Bacilysin regulates biofilm formation and respiration during stationary phase.....	67
5. CONCLUSION	71
REFERENCES	75
APPENDICES	89

LIST OF TABLES

TABLES

Table 2.1. Structures and functions of lantibiotics of <i>B.subtilis</i>	11
Table 2.2. Structures and functions of non-ribosomally synthesized peptide antibiotics of <i>B.subtilis</i>	12
Table 4.1. Number of proteins in each network condition.....	37
Table 4.2. Probable effectors of cellular functions under bacilysin loss in the cytosolic proteome of OGU1.....	51
Table 4.3. Probable effectors of cellular functions under bacilysin loss in the 12 th -hour secretome of OGU1.....	61
Table 4.4. Probable effectors of cellular functions under bacilysin loss in the 16 th -hour secretome of OGU1.....	66
Table 4.5. Probable effectors of cellular functions in the 24 th -hour secretome of OGU1.....	70

LIST OF FIGURES

FIGURES

Figure 1.1. Overview of the sporulation cycle in <i>B.subtilis</i>	6
Figure 1.2. Canonical bacterial quorum-sensing (QS) circuits.....	7
Figure 1.3. Rap proteins and Phr peptides in the regulation of sporulation	8
Figure 1.4. Simplified representation of the regulatory network of competence initiation and the interconnection with other networks in <i>B. subtilis</i>	10
Figure 2.1. Structure of bacilysin	16
Figure 2.2. Organization of the bacilysin gene cluster <i>bacABCDE</i>	17
Figure 2.3. Quorum sensing system and pleiotropic regulators interconnecting bacilysin biosynthesis	18
Figure 2.4. Organization of <i>bacA':lacZ::erm::bacABCDEFG</i> fusion construct	20
Figure 4.1. Cytosolic proteome network of OGU1	32
Figure 4.2. 12 th -hour secretome network of OGU1	33
Figure 4.3. 16 th -hour secretome network of OGU1	34
Figure 4.4. 24 th -hour secretome network of OGU1	35
Figure 4.5. Diagram of total number of proteins in each network	36
Figure 4.6. First-degree neighbors of AtpC	39
Figure 4.7. Overview on the regulation of genes encoding enzymes of glycolysis, the pentose phosphate shunt, and the Krebs cycle in <i>B.subtilis</i>	41
Figure 4.8. First degree neighbors of the carbon catabolite repressor CcpA	46
Figure 4.9. Enriched Gene Ontology biological process terms in the cytosolic proteome	49
Figure 4.10. Enriched KEGG pathway terms in the cytosolic proteome	50
Figure 4.11. Growth and bacilysin activity of <i>B. subtilis</i> PY79 and OGU1 in PA medium	53
Figure 4.12. Enriched GO biological process terms in the 12 th -hour secretome	58
Figure 4.13. Enriched KEGG pathway terms in the 12 th -hour secretome	59
Figure 4.14. 12 th -hour secretome case study	60

Figure 4.15. 16 th -hour secretome case study	65
Figure 4.16. Enriched GO biological process terms in the 24 th -hour secretome	69
Figure 4.17. 24 th -hour secretome case study	69

LIST OF ABBREVIATIONS

ABBREVIATIONS

STRING: Search Tool for Interactions of Genes and Proteins

2DE: 2-Dimensional Electrophoresis

MALDI-TOF: Matrix-assisted Laser Desorption Ionization Time of Flight

MS: Mass Spectrometry

NanoLC-MS/MS: In-gel tryptic digestion followed by liquid chromatography-tandem mass spectrometry

PCSF: Prize-Collecting Steiner Algorithm

CEA: Combination-based Enrichment Analysis

DAVID: Database for Annotation, Visualization and Integrated Discovery

BiNGO: Biological Networks Gene Ontology tool

FunRich: Functional Enrichment analysis tool

GO: Gene Ontology

KEGG: Kyoto Encyclopedia of Genes and Genomes

TCA: Tricarboxylic Acid Cycle

DSM: Difco Sporulation Medium

PA: Perry & Abraham Medium

CHAPTER 1

INTRODUCTION

Bacillus subtilis is a rod-shaped Gram-positive soil bacterium which produces and secretes a broad range of secondary metabolites with distinct structures and synthesis mechanisms. Bacilysin is a non-ribosomally synthesized dipeptide antibiotic which is active against fungi and some bacteria. However, recent comparative proteomic studies by our group have indicated that bacilysin is more than an antibiotic, in that it is a pleiotropic signaling molecule which influences numerous cellular processes.

The term “network modeling” implies methods, tools, and approaches which are applied to represent interactions between individuals of a group. Protein-protein interaction networks are graphs which display pairwise undirected relations between proteins. They can be used to determine previously unidentified proteins and interactions, in addition to distinguishing the steps of a signaling pathway. To infer the interactions between proteins of an organism in a given dataset, publicly available interactomes or manually curated experimental data can be utilized. A disadvantage of interactomes is that they are usually noisy and incomplete, and fail to integrate the given information to build meaningful networks. Current tools and efforts have been attempting to combine the databases in a single source and to reconstruct the signaling networks (Keskin, Tunçbağ and Gürsoy, 2016). Omics Integrator is a powerful tool for recovering hidden nodes in the remodeled network, which also incorporates weighted negative evidence by solving the Prize-collecting Steiner forest (PCSF) problem.

The purpose of this thesis is to represent and explain the interactions in bacilysin-deleted *B. subtilis* strain OGU1 via Omics Integrator. Proteomic data from comparative proteome and time-dependent secretome analyses have been collected to represent the altered interactions in OGU1. Network modeling and subsequent functional enrichment analyses have confirmed our experimental results in addition to prospective interacting proteins for each condition. Although this study cannot represent all interactions and their influence on our organism of interest, we believe our work can lead the way for future experiments.

This thesis comprises the following chapters:

Chapter 2 includes a literature review about the physiology and secondary metabolites of our organism of interest and previous research efforts. The chapter also gives information about protein-protein interaction networks, steps of network analysis, Omics Integrator and the Prize-Collecting Steiner Forest algorithm, as well as the STRING database.

Chapter 3 provides information about the experimental dataset together with the selected interactome to be used in this work. It also describes the methods and approaches in the construction of the final networks and functional enrichment analyses.

Chapter 4 consists of the networks for each condition and their functional analysis results. Important proteins which may regulate cellular functions under bacilysin loss are shown in this chapter. Chapter 4 also discusses the obtained results along with literature support.

Chapter 5 outlines the study as a whole and concludes the results and discussion chapter by commenting on the future direction of the study. It also suggests several further experimental methods such as genetic and chemical complementation.

CHAPTER 2

LITERATURE REVIEW

2.1. *Bacillus subtilis* as a model organism

Bacillus subtilis is a Gram-positive, rod-shaped motile bacterium which is usually found in soil, water sources and gastrointestinal tract of humans. It is capable of forming a stress-resistant endospore which protects the bacterium against harsh environmental conditions. Although its sibling species such as *B. anthracis* or *B. cereus* can cause anthrax or food poisoning, *B. subtilis* is non-pathogenic. Among *Bacillus* species, *Bacillus subtilis* is the most extensively studied and prominent organism (Wipat and Harwood 1999; Sonenshein et al., 2002; Lippolis et al., 2013). *Bacillus subtilis* is the second most well-understood prokaryote organism after *Escherichia coli*. Since it has evolved to survive in robust conditions such as high temperature and starvation, the genome of *B. subtilis* is considerably versatile and large, providing an elaborate investigation of this organism (Zweers et al., 2008). The genome of *B. subtilis* was first sequenced and published in 1999 by Kunst et al. and then updated in 2009 (Barbe et al., 2009). Today, two common *B. subtilis* strains, the standard strain *B. subtilis* 168 and its prototrophic derivative *B. subtilis* PY79 are commonly exploited for genetic and biochemical analyses (Spizen, 1958; Harwood and Cutting., 1990; Cui et al., 2018).

2.2. Primary metabolism in *Bacillus subtilis*

As all organisms, *Bacillus subtilis* requires energy to maintain its life cycle. It is a facultative anaerobe bacterium which was long believed to be a strict aerobe (Hoffmann et al., 1995). Studies have shown that it can grow anaerobically by using nitrite or nitrate as the terminal electron acceptor, or by fermentation (Sun et al., 1996). *B. subtilis* can therefore alter its metabolic activity according to oxygen availability.

Two-component signal transduction systems refer to complexes which consist of a membrane-bound sensor histidine kinase that perceives environmental stimuli and a response regulator that mediates the cell response, usually by changing the expression of target genes. The two-component signal transduction system ResDE is an essential regulator of anaerobic respiration in *B. subtilis*. It is composed of the sensor kinase ResE and the response regulator, ResD. Upon oxygen limitation, ResDE induces the transcription of another anaerobic transcriptional regulator, Fnr. Fnr then activates nitrate reductase genes including *narGHJI*. Unlike many anaerobes which require pyruvate formate lyase, *B. subtilis* can maintain fermentation without external electron acceptors, since pyruvate is directly metabolized by pyruvate dehydrogenase (Nakano and Zuber, 1990).

In response to changes in aerobic or anaerobic respiration, the relationship between carbon and nitrogen metabolism also becomes altered. A comparative investigation of protein expression in *B. subtilis* grown under aerobic respiration and anaerobic or nitrate fermentation conditions showed that anaerobic fermentation induced at least 44 proteins and nitrate respiration induced at least 19 proteins as compared to aerobic controls (Clements et al., 2002). Another study conducted by Choi and Saier has uncovered the connection between carbon and nitrogen metabolism by identifying a catabolite-responsive element binding site within the transcriptional regulator SigL, to which the carbon catabolite repressor protein CcpA binds (2005).

SigL regulates genes associated with amino acid metabolism such as *rocA*, *rocD*, *rocB*, *buk* and *bkdAB*, as well as carbon utilization genes *acoABCL*, and sugar specific genes *levDEFG* and *sacC*.

Furthermore, recent studies have been focusing on rewiring carbon and nitrogen metabolic networks to enhance nutrient uptake, thereby increasing protein production in *B. subtilis* as a microbial factory (Cao et al., 2018; Ma et al., 2017). These experiments represent the interrelationships between carbon and nitrogen metabolism and their effect on cellular activities.

2.3. An overview of sporulation and quorum sensing in *Bacillus subtilis*

Under nutritional starvation, *B. subtilis* stops growing and attempts to restore growth by increasing its metabolic diversity. The cells produce signals to induce of motility and chemotaxis, as well as hydrolase and antibiotic production (proteases and carbohydrases). When growth recovery fails, cells form endospores that are resistant to chemical, irradiation and desiccation. The first morphological indication of sporulation is division of the cell into a smaller forespore and a larger mother cell, each containing an entire copy of the chromosome. After mother cell engulfs the forespore and their genomes are differentially expressed with regards to regulatory pathways and developmental checkpoints, programmed death and lysis of the mother cell occurs and the mature spore is released (Stragier and Losick, 1996) (Figure 1.1). *B. subtilis* can also differentiate into a physiological state, the competent state, which allows it to undergo genetic transformation in an alternative developmental process (Solomon, Lazzazera and Grossman, 1996).

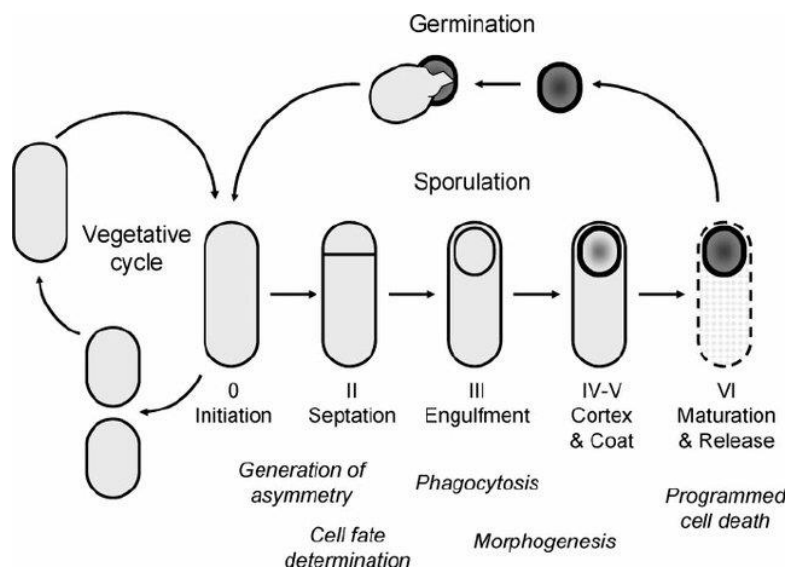


Figure 1.1. Overview of the sporulation cycle in *B. subtilis*. The vegetative cycle is favored under conditions supporting growth. Cells elongate vertically and divide from the middle to produce two identical daughter cells. Starvation induces the sporulation cycle. Stage I has been omitted and stages IV and V are combined for simplification. Released mature spores can remain dormant for an almost indefinite period before undergoing germination and outgrowth to resume the vegetative cycle (Errington, 2010).

Quorum sensing (QS) can be defined as the regulation of gene expression based on fluctuations in the cell population by small diffusible signal molecules (Miller and Bassler, 2001). There are mainly three QS mechanisms; first one is the acyl homoserine lactone (AHL) in Gram-negative bacteria. AHL is produced by a LuxI-type synthase and is perceived by a LuxR-type transcriptional activator. The second mechanism involves autoinducer peptides (AIP) in Gram-positive bacteria which are sensed by two-component regulatory systems. The final system is the autoinducer-2 (AI-2) system that is common in both Gram-negative and Gram-positive bacteria (Brackman and Coeny, 2015).

QS is an important regulatory mechanism in *B. subtilis* which partakes in a broad range of physiological processes such as competence development, sporulation, production of extracellular enzymes and antibiotics (Comella and Grossman, 2005). Like other prokaryotic organisms, signal transduction is mediated by a two-component histidine kinase system, in which a membrane-bound histidine kinase senses autoinducer peptides (AIPs) and information is channeled through a phosphorelay to a response regulator protein (Zhang and Dong, 2004). In some cases, QS signals are transported back into the cytoplasm in order to regulate gene expression by interacting with transcription factors (Rutherford and Bassler, 2012).

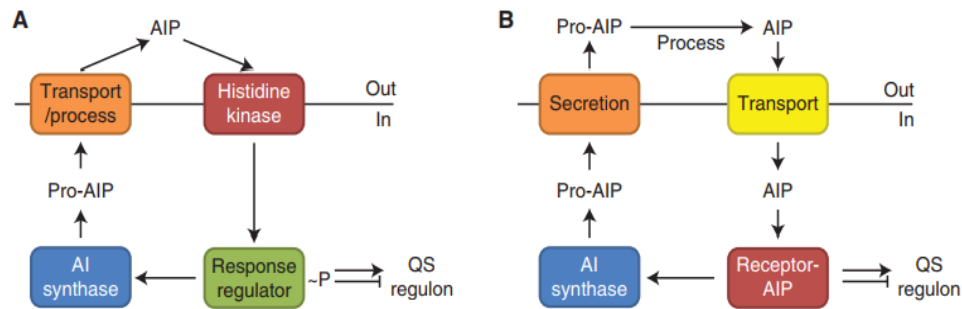


Figure 1.2. Canonical bacterial quorum-sensing (QS) circuits. Autoinducing peptide (AIP) QS in Gram-positive bacteria occurs either by (a) two-component signaling or (b) an AIP-binding transcription factor (Rutherford and Bassler, 2012).

Bacterial quorum sensing receptor family includes Rap proteins with phosphatase activity. There are 11 chromosomally encoded Rap proteins (RapA-K) and five additional plasmid encoded proteins in *B.subtilis*. Inhibitory oligopeptides of Rap proteins are called Phr peptides and each Phr is encoded in an operon with its related Rap protein (Parashar et al., 2013). The three targets of Rap proteins in *B.subtilis* are the response regulators Spo0F, ComA, and DegU.

Rap and Phr protein pairs regulate important processes during the transition to stationary growth phase. The regulation of spore formation by RapA/PhrA, extracellular protease production by RapG/PhrG and transfer of conjugative DNA elements by RapI/PhrI are some examples (Lanigan-Gerdes et al, 2007; Bendori et al., 2015) (Figure 1.3).

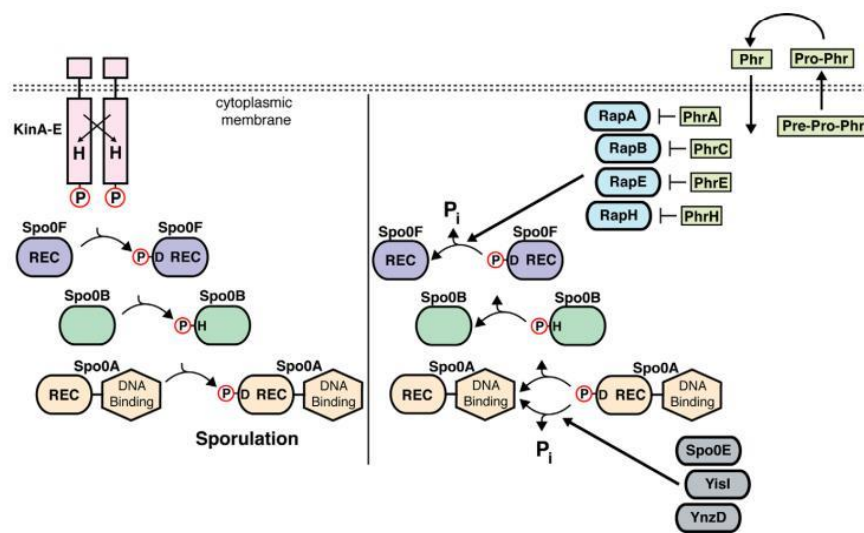


Figure 1.3. Rap proteins and Phr peptides in the regulation of *B. subtilis* sporulation. Soluble cytoplasmic receptor kinases (KinA-E) are shown as membranous proteins for simplicity. Spo0F, Spo0B, and ultimately Spo0A are consecutively phosphorylated. Sporulation commences when the phosphorylated Spo0A reaches a critical level (left). Conversely, Rap proteins dephosphorylate Spo0F and inhibit sporulation (right). Oligopeptide permease A (OppA) imports Phr peptides into the cytoplasm where they reinforce sporulation by binding to Rap proteins and inhibiting Spo0F dephosphorylation. Spo0E, YisI, and YnzD dephosphorylate Spo0A. REC: Receiver Domain; H: histidine; D: aspartate; P: phosphoryl group; P_i : inorganic phosphate (Parashar et al., 2011).

The response regulator ComA monitors more than 10% of the genome and is the main component of the ComQXPA pathway (Comella and Grossman, 2005). The pathway is activated depending on the accumulation of two pheromones: ComX (a cell driven extracellular peptide) and competence and sporulation stimulating factor (CSF). Briefly, ComX activates the membrane-bound receptor protein kinase ComP. Auto phosphorylated ComP donates the phosphate to response regulator ComA, which activates its associated promoters. Meanwhile, oligopeptide permease (Opp) imports CSF, which inhibits RapC, thereby removing its repression on ComA (Comella and Grossman, 2005). CSF inhibits the ComA controlled gene expression at high concentrations (Pottathil et al., 2008).

Other pheromones such as PhrF and PhrK also increase ComA activity. An oligopeptide permease transports these pheromones which block the inhibitory activity of their cognate Rap proteins on ComA (Pottathil et al., 2008). ComA is repressed by RapC, RapD, RapF, RapH, and RapK. However, ComA activates numerous proteins including Rap proteins and surfactin (Ogura et al., 2003). Overall, more than 20 genes directly and more than 150 genes indirectly are regulated by ComX-ComP-ComA quorum-sensing pathway (Claverys, Prudhomme and Martin, 2006).

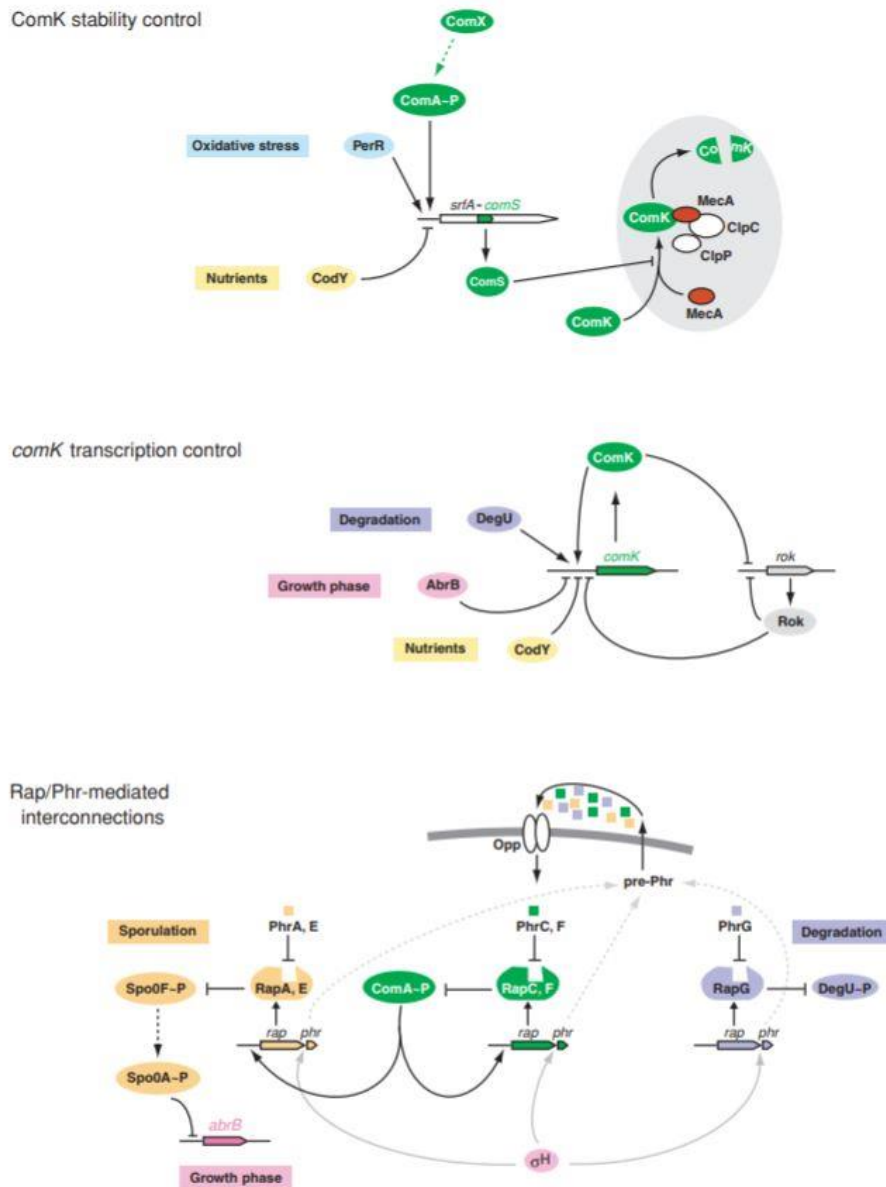


Figure 1.4. Simplified representation of the regulatory network of competence initiation and the interconnection with other networks in *B. subtilis*. Perpendiculars and arrows represent negative and positive actions, respectively. Only effectors with proven direct actions are shown). PhrA has been suggested to accumulate primarily in the periplasmic space and therefore not to function in cell density control. ComQ and ComP are not shown (Claverys, Prudhomme and Martin, 2006).

2.4. Secondary metabolites and antibiotics of *Bacillus subtilis*

The distinct metabolism of *Bacillus subtilis* is known to promote the production of a wide variety of secondary metabolites with antimicrobial and pharmacological activities. It can produce more than 20 antibiotics with diverse chemical structures. Production of peptides with antimicrobial activity in *B. subtilis* can be either ribosomal with post-translational modification (Zuber et al., 1993; Schnell et al., 1988) or non-ribosomal (Weber and Marahiel, 2001).

As numerous studies indicate, antibiotics of *B. subtilis* do not entirely act as antimicrobial molecules. For instance, non-ribosomally produced lipopeptides are associated with biofilm formation and swarming motility, lantibiotics act as pheromones in quorum-sensing, and a “killing factor” promotes programmed cell death in sister cells (Stein, 2005). Tables which show the structures and functions of lantibiotics and non-ribosomally synthesized peptide antibiotics of *B. subtilis* is shown below (Table 2.1, Table 2.2.).

Table 2.1. Structures and functions of lantibiotics of *B. subtilis*, as explained by Stein et al., 2005.

Antibiotic	Structure	Function
Subtilin	Cyclic	Pheromone-type role
Ericin	Cyclic	Similar to subtilin with weaker antibiotic activity
Mersacidin	Globular	Cell wall biosynthesis inhibition
Subtilosin A	Cyclic	Acts against some Gram (+) bacteria

Table 2.2. Structures and functions of non-ribosomally synthesized peptide antibiotics of *B.subtilis*, as explained by Stein et al., 2005. Iturin, bacillomycin, and mycosubtilin belong to the iturin family; hence, they have similar structures and functions.

Antibiotic	Structure	Function
Surfactin	Cyclic	Detergent-like action on biological membranes
Iturin	Cyclic lipopeptide	Antifungal and hemolytic activity
Bacillomycin	Cyclic lipopeptide	Antifungal and hemolytic activity
Mycosubtilin	Cyclic lipopeptide	Antifungal and hemolytic activity
Fengycin	Cyclic	Specifically acts against filamentous fungi
Bacillibactin	Cyclic	Siderophore which complexes ferric ion by its six hydroxyl residues
Bacilysin	Linear	Probable pleiotropic molecule, controlled by quorum sensing and other transcriptional mechanisms

2.5. The diverse derivative strains of *Bacillus subtilis* 168

Bacillus subtilis PY79 and *B. subtilis* 168 carry identical genomic properties with varying SNP's unique to each genome (Zeigler et al., 2008). PY79 is the tryptophan prototrophic derivative of *B.subtilis* 168, which was utilized for insertional mutagenesis analyses because of its transformation capacity (Harwood, 1990). It was constructed by introducing transposon Tn917 into the sequence of phage SP β (Youngman, Perkins and Losick, 1983; Youngman, Perkins and Losick, 1984). The tryptophan locus (*trpC*) in *B.subtilis* 168 was found to be irreversibly silenced by a three base-pair deletion, unlike PY79 which consists of a considerable amount of *trpC* islands (Zeigler et al., 2008). Therefore, the main difference between *B.subtilis* 168 and *B.subtilis* PY79 is that the former requires tryptophan as a growth factor, while the latter does not.

Moreover, our lab-adapted PY79 differs from the reference PY79, in that the former consists of 5 SNP's and one deletion in its genome. This difference is due to extensive growth and consecutive transfers in the laboratory environment for more than three decades. Further genetic studies and protein-protein interaction experiments might help elucidate how the genetic changes in lab-adapted PY79 display themselves on the phenotype. Hence, we can make an improved comparison between two strains.

2.6. Proteomics

Proteomics refers to the large-scale analysis of the proteins in a cell or biological fluids as a whole. Genomic analysis can provide very limited information with respect to an organism's adaptation mechanisms. Although DNA sequence homology can propose adaptation capabilities for that organism, investigating the protein expression profile can thus help us elucidate how an organism adjusts itself under a given condition (Cordwell et al., 2001). In addition, mRNA expression analysis can yield more information than DNA analysis.

However, it is not sufficient to represent the protein content of organism, due to the poor correlation between mRNA and protein expression in many scenarios (Liu et al., 2016; Greenbaum et al., 2003).

Splicing, polyadenylation and mRNA editing takes place after mRNA synthesis. mRNA might be subjected to regulation just before translation in eukaryotes. Moreover, proteins are also subjected to posttranslational modifications, proteolysis and compartmentalization. Thus, physiological map of an organism in particular conditions cannot be completely analyzed by solely gene or mRNA analysis.

2.6.1. Comparison of gel based and gel free proteomics

Although 2D-PAGE is an old method, it is still a convenient technique to analyze (separate and visualize) proteins in many studies (Naryzhny, 2016). However, the cost, inability to separate all proteins, insensitivity to low abundance proteins and tediousness of 2D-PAGE led to the development of gel-free proteomics techniques. Nonetheless, these shotgun proteomic approaches also have some drawbacks, based on the assumption that a protein can be identified based on the sequence of a single (or a few) tryptic peptide(s) derived from the protein.

Many proteins undergo crucial post-translational modifications, which regulate protein function and therefore will display differences in each condition. Furthermore, a given gene may result in a large number of different protein products. Discerning the isoforms of a protein then becomes a difficult task. By contrast, 2D-PAGE includes intact proteins, with changes in protein expression level, isoforms or post-translational modifications. Post-translational modifications will result in a shift of pI (in the case of phosphorylations) or relative mass (e.g. glycosylation or truncation) and show a different mobility on a 2DE gel.

Different isoforms of the same protein will therefore be visible as different spots on the 2DE gel. In gel-free based methods, Mr and pI information is lost while performing peptide analysis (Baggerman et al., 2005).

Comparative proteomics between different samples are done with 2DE-PAGE method (Bernhardt et al., 1996). However, it is restricted to first dimension pI and borders of polyacrylamide gel. Low abundant proteins cannot be observed by low resolution techniques. On the other hand, entire protein content can be analyzed by using liquid chromatography in 1D coupled LC-MS/MS technique (Old et al., 2005).

This method can also be used for comparative proteomics as a mean of low abundant protein comparison (Kuwana et al., 2002). Calculation of relative spectral counts (R_{sc}) of proteins in the gel-free approach to quantify their relative abundance differences are given below (Old et al., 2005; Beissbarth et al., 2004):

$$R_{sc} = \log_2\left(\frac{n_2+f}{n_1+f}\right) + \log_2\left(\frac{t_1-n_1+f}{t_2-n_2+f}\right)$$

For each protein, \log_2 ratio of abundance between Sample 1 and Sample 2 constitutes an R_{sc} value. n_1 and n_2 depict spectral counts for the protein in Sample 1 and Sample 2, t_1 and t_2 represent the total spectral count (sampling depth) for Samples 1 and 2; and f is the correction factor set to 1.25 instead of 0.5, as proposed by Old et al., 2005 (Demir, 2013).

2.7. The dipeptide antibiotic bacilysin and its activity as a pleiotropic molecule

The dipeptide bacilysin is among the various antibiotics produced by *B. subtilis*. It is composed of L-alanine at the N-terminus and a non-proteinogenic amino acid, L-anticapsin at the C-terminus (Walker and Abraham, 1970) (Figure 2.1).

Evidence from previous studies has confirmed that the peptide formation with L-alanine occurs in a non-ribosomal mode catalyzed by an enzyme named bacilysin synthetase (Sakajoh et al., 1987).

Bacilysin is active against some bacteria and fungi, including *Candida albicans* (Tschen, 1990; Kenig and Abraham, 1976). The antibiotic activity of bacilysin relies on its anticapsin moiety, which blocks glucosamine synthetase and therefore bacterial peptidoglycan or fungal mannoprotein biosynthesis in susceptible cells, leading to cell protoplasting and lysis (Milewski, 1993; Chmara, 1985; Chmara et al., 1982; Kenig et al., 1976; Whitney and Funderburk, 1970).

Unlike the typical non-ribosomal peptide synthesis (NRPS) mechanism, the mechanism of bacilysin production depends on a polycistronic operon named *ywfBCDEF* and a monocistronic gene *ywfG* (Inaoka et al., 2003). These genes were later renamed as *bacABCDE* and *bacF*, respectively, for convenience. The *bacABC* genes are responsible for the conversion of prephenate to anticapsin, and *bacDE* genes have roles in amino acid ligation and bacilysin immunity, respectively (Steinborn et al., 2005). The *ywfH* gene is divergently located and transcribed from the opposite DNA strand from *bacABCDE*. Later on, *ywfG* and *ywfH* genes were renamed as *bacF* and *bacG* (Parker and Walsh., 2012). The organization of *bac*-operon is shown in Figure 2.2.

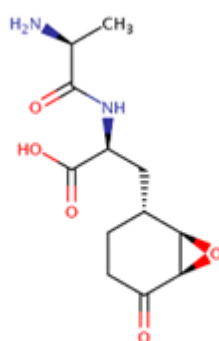


Figure 2.1. Structure of bacilysin.

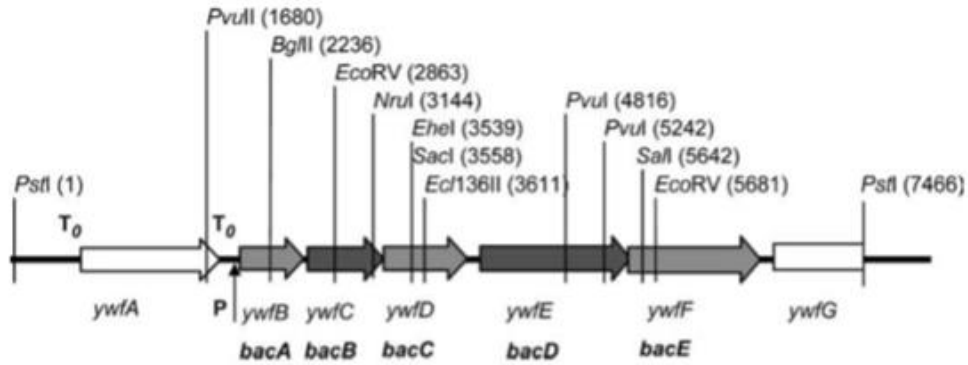


Figure 2.2. Organization of the bacilysin gene cluster *bacABCDE* relative to open reading frames *ywfABCDEFG* of *B. subtilis* 168 (Steinborn et al., 2005).

Bacilysin production is parallel to active growth in *B. subtilis* PY79 in synthetic medium and certain nutrients such as glucose, ammonium and casamino acids, as well as culture pH and temperatures above 30°C, suppress or inhibit its production (Özcengiz and Alaeddinoğlu, 1991; Özcengiz et al., 1990). Its synthesis is under stringent response and feedback regulation, as well as quorum sensing global regulation (Inaoka et al., 2003; Yazgan et al., 2001; Özcengiz and Alaeddinoğlu, 1991).

Several distinct studies regarding the control of bacilysin synthesis demonstrate that bacilysin production can be up or downregulated, depending on the conditions and the monitoring proteins (Köroğlu et al., 2011; Karataş et al., 2003). For instance, *srfA* operon, which is essential for the production of lipopeptide antibiotic surfactin, plays a direct role in bacilysin production (Karataş et al., 2003). Moreover, the expression of *bac*-operon is under the control of four transcriptional regulators: ComA, Spo0A, AbrB, and CodY. ComA and Spo0A are required for transition state induction, and ComA activity is sufficient for basal level transcription of *bac*-operon (Köroğlu et al., 2011; Karataş et al., 2003). AbrB and CodY have a simultaneous adverse effect on

bacilysin biosynthesis through accelerated suppression of *bac*-operon during exponential and early stationary-stationary phase (Köroğlu et al., 2011).

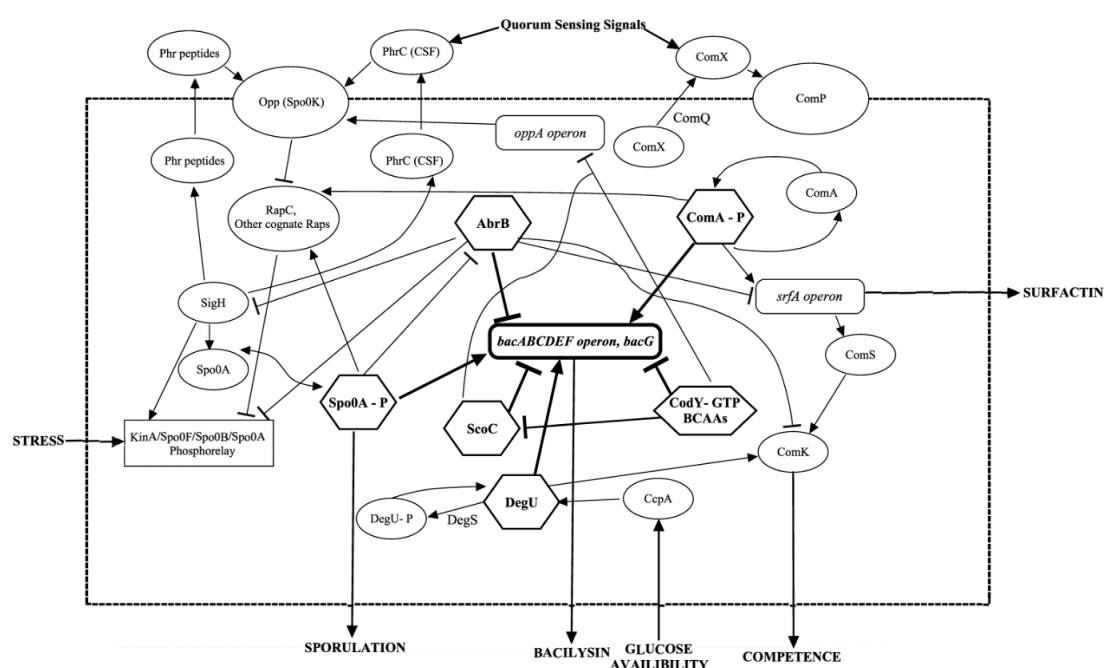


Figure 2.3. Quorum sensing system and pleiotropic regulators interconnecting bacilysin biosynthesis with surfactin production, competence development and sporulation in response to quorum sensing signals, stress and glucose availability. Arrows indicate activation; T-bars denote repression. Bold lines refer to direct interaction with P_{bacA} (adapted from Özcengiz and Ögölür, 2015).

interrelated comparative proteome studies between OGU1 and its prototrophic parental strain PY79 in our laboratory have proposed that bacilysin is more than a peptide antibiotic. In essence, the dipeptide molecule has been implicated to have multiple impacts on numerous mechanisms, which control the synthesis of bacilysin itself. Furthermore, global regulatory systems such as sporulation, quorum sensing and metabolic pathways are among other changed functions, since the absence or lack of bacilysin is recognized as a stress condition by the organism. For instance, proteome-wide 2DE MALDI-TOF MS study conducted by Aras-Taşkın (2010) has demonstrated that the lack of bacilysin induces the organism to undergo a universal adaptation process due to altered cytosolic protein expression. Similarly, the complementary nanoLC-MS study carried by Demir (2013) has shown decreased expression of proteins regarding energy metabolism, sporulation and membrane transport.

In conjunction with the studies mentioned above, the dynamic secretome analysis performed by Tekin-İşlerel (2017) which utilizes LC-MS and 2DE MALDI-TOF MS approaches has indicated that bacilysin acts on its producer as a pleiotropic molecule. Taken together, previous research has indicated that bacilysin is possibly a signaling molecule which act on different cellular functions.

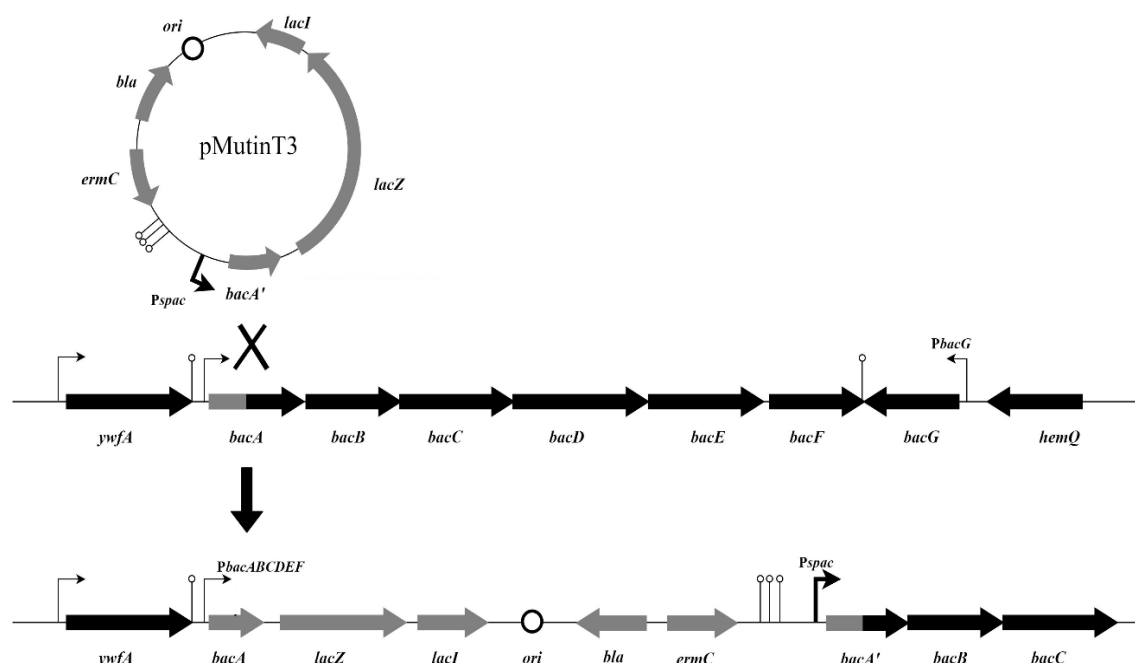


Figure 2.4. Organization of *bacA'::lacZ::erm::bacABCDEF* fusion construct in pMUTIN T3 and its location within the linearized form of plasmid.

2.8. Protein-protein interaction networks (PPIN's)

The analysis of biological networks arose from the concept of social network analysis and the application of graph theory to the social sciences. Graph theory refers to the study of mathematical structures which are used for pairwise modeling of objects. A graph consists of nodes (points) which are connected by edges (lines or arcs). Graphs can be applied to real-world problems to model relations and processes in physical, social, or information systems (Mashaghi et al., 2004). Similarly, reciprocal actions between proteins or genes can be depicted by directed or undirected graphs. Protein-protein interaction networks (PPIN's) are undirected graphs which represent the physical relations between proteins. They facilitate the prediction of biological functions of proteins and the characterization of protein complexes and pathways.

The contacts within a PPIN are specific; they occur between defined regions in the proteins and serve a biologically meaningful function.

PPIN's can display both transient and stable interactions. Transient interactions are brief interactions between proteins which transport or modify a protein. Stable interactions, however, are commonly seen in protein complexes. By representing protein-protein interactions of any kind, we can assign putative roles to proteins yet to be identified. Likewise, we can define relationships between proteins which can form complexes. Furthermore, details of the steps of a signaling pathway can be elucidated.

2.9. The STRING database

The STRING (Search Tool for the Retrieval of Interacting Genes and Proteins) database is an open source data collection which consists of known and predicted protein-protein interactions for a broad range of organisms. The database relates proteins of a particular organism based on information curated from numerous other sources. The associations in STRING include specific and biologically meaningful direct (physical) and indirect (functional) interactions.

Apart from collecting and reassessing available experimental data on protein-protein interactions, and importing known pathways and protein complexes from curated databases, interaction predictions are derived from co-expression analyses, detection of shared selective signals across genomes, automated text-mining of literature and computational transfer of interaction knowledge between organisms based on gene orthology (Szkarczyk et al, 2017).

Each protein association is scored from 0 to 1 in STRING. These protein interactions, namely edge weights, illustrate the probability that a given interaction is biologically significant, specific and reproducible. Depending on the type of evidence, interactions are scored separately in 7 channels, which are:

a) experiments, b) database, c) text mining, d) co-expression, e) neighborhood, f) fusion and g) co-occurrence.

The “experiments” channel includes evidence from actual experiments in the laboratory. “Database” consists of manually evaluated evidence from pathway databases. “Text mining” refers to mentions of protein names in literature, and frequently mentioned pairs of proteins are given a score. “Co-expression” is comprised of gene expression datasets which are retrieved from large-scale RNAseq or microarray experiments. Proteins with similar expression patterns are paired and scored accordingly. “Neighborhood”, “fusion” and “co-occurrence” are genome-based prediction channels, wherein pairs of proteins are scored regarding either their genomic neighborhood, their similarity in phylogenetic distribution or their fusion of orthologs into a single protein-coding gene. A combined score is calculated for each interaction from these channels, which is generally used as the final measure when building networks or when sorting and filtering interactions. If the combined score is high and there are multiple evidence channels which contribute to the score, it is considered a good sign of support (Szkarczyk et al, 2017).

2.10. Omics Integrator and Prize-Collecting Steiner Forest (PCSF) algorithm

Perturbed interactions in biological systems can be identified easily with “omics” methods in a precise manner. However, interpretation of experiments with different collections of data is a difficult task. Omics Integrator is a network modeling tool which recovers hidden nodes in the reconstructed network by combining different data sources together, as well as incorporating weighted negative evidence. Unless the data imply the presence of abundant hub proteins or proteins that are not expressed, Omics Integrator removes them through this negative evidence (Tunçbağ et al., 2016).

Omics Integrator consists of two modules which can be used together or individually: Garnet and Forest. Garnet aims to find transcriptional factor binding sites which may affect gene expression levels, whereas Forest generates networks which connect user-defined omic hits (i.e., transcription factors which come from Garnet or data that show a significant level of change between conditions). Forest employs the Prize-Collecting Steiner Forest (PCSF) algorithm, which takes any undirected, directed or partially directed network as input (Tunçbağ et al., 2016). The goal is to include as many terminal nodes- which are biologically significant proteins in the network- as possible, while decreasing hub dominance and the cost to generate the network (Equation 1b). To reduce possible bias which may be caused by proteins that are either conserved, highly expressed or essential, a negative weighting function for prize nodes is used.

2.11. Network analysis and visualization

The analysis of a network revolves around particular centrality metrics about the network of interest. Moreover, we can investigate a network's topology through various clustering methods which utilize the statistical parameters of the network. For instance, nodes which tend to group together when compared to other elements in the network can be identified by the approaches mentioned above. Besides, we can evaluate the importance of nodes and how they act upon one another. In the case of biological networks, the interaction patterns between proteins can be determined based on their expression values, as well as centrality measures which include but are not limited to: betweenness, degree, and closeness. The centrality of a network can be defined as the influence of a node over the network (Shao et al., 2017). In PPIN's, centrality answers the questions of which and why a protein is essential. Betweenness centrality refers to how well a node acts as a bridge in linking other nodes, therefore forming a network. Similarly, degree centrality is another measure in network analysis, which is based on how many connections a node has in a network (Kim et al., 2018).

Clustering is another common step in the analysis of networks, of whose objective is to find internally connected groups of nodes (i.e., communities), based on a weight function which specifies the influence of vertices. Depending on the weighting method, distance metric and implementation, different clusters can be generated. The modularity of a network refers to the level of connectivity across nodes. The Louvain method of clustering is an advantageous technique in that it adjusts the modularity of communities in two steps (Blondel et al., 2008). In the first step, the algorithm finds small communities by local optimization of the modularity of nodes. Next, it groups those small communities into one node and repeats the first step until no further modularity-increasing reassignments of communities are possible.

Functional enrichment analysis is another step in biological network studies. It refers to the statistical evaluation of annotation terms that significantly represent a group of functions, processes or pathways based on lists of genes or proteins in a graph. These annotation terms are provided by databases such as Gene Ontology (GO), UniProt Gene Ontology Annotation (GOA), Reactome or Kyoto Encyclopedia of Genes and Genomes (KEGG).

There are various tools and plug-ins which help interpret the experimental data such as BiNGO, ClueGO, DAVID, CEA and FunRich. Among the tools mentioned, ClueGO is a useful Cytoscape plug-in which can manage multiple annotation datasets for a broad range of organisms. It allows the inference of a network as a whole, or in parts (Mlecnik, Galon and Bindea, 2019).

Cytoscape is the most commonly used visualization tool for biological networks. It combines networks with annotation, experimental data and other biological information by utilizing numerous plug-ins which can be installed within the platform. Furthermore, Cytoscape allows customization of node and edge views of networks (Shannon et al., 2003).

2.12. Aim of the present study

Under the light of the information and the findings mentioned earlier, the present research aimed to identify the altered interactions in OGU1 compared to PY79, based on molecular interactions in *B. subtilis* 168 due to unavailable data by employing network modeling and functional enrichment analyses. We believe that this study can help elucidate novel interactions through experimental procedures, as well as confirm and expand earlier studies.

CHAPTER 3

MATERIALS AND METHODS

3.1. Modification of datasets

The interactome and protein aliases of *Bacillus subtilis* 168 were downloaded from STRING database version 10.5 (string-db.com). Protein alias file of *Bacillus subtilis* 168 consists of STRING ID's with their corresponding names and/or identification numbers. STRING ID's were next converted to locus identification numbers for each protein, as there can be multiple names for the same identification number, or vice versa. Scores from “experiments” and “database” channels were collected and a combined score for physical interactions was calculated for every edge in the interactome, as explained in the STRING website.

“Prize file” refers to a file which includes differentially expressed proteins/genes with their expression values. The expression values can be represented as either log-fold change, or negative log of the significance level describing changes between conditions. Four prize files were prepared for the data from each study, one for cytosolic proteome of OGU1, and three which were retrieved from dynamic secretome analyses at different time points (12 hours, 16 hours and 24 hours, respectively). Since different proteomic methods can identify different proteins, the results from 2DE MALDI-TOF MS, LC-MS, and nanoLC-MS techniques were combined for each dataset. Proteins which were common to either one of the methods above were selected with respect to the highest expression value. The datasets included each differentially expressed protein (for instance, 2DE MALDI-TOF and nanoLC-MS expression levels were combined in one file for OGU1/PY79 proteome log₂FC).

Cytosolic proteome data were collected and combined from the studies of previous lab members Aslı Aras Taşkın (2011) and Mustafa Demir (2013). Similarly, protein fold changes from time-dependent dynamic secretome analyses were taken from Elif Tekin İşlerel's (2017) dataset. All fold change values were \log_2 transformed. To prevent proteins in each prize file being discarded from the networks, absolute values of \log_2 fold changes were taken. Localization of data points was made via Plotly's Python graphic library and "pheatmap" library v.1.0.12 of R in version 3.5.1.

3.2. Network modeling

The Forest algorithm of Omics Integrator which is based on the Prize-collecting Steiner forest (PCSF) approach was utilized for the construction of networks for each prize file (Tunçbağ et al., 2016). The equation for the algorithm is as follows:

$$p'(v) = \beta \cdot p(v) - \mu \cdot \text{degree}(v)$$

$$f'(F) = \sum_{v \notin V_F} p'(v) + \sum_{e \in E_F} c(e) + \omega \cdot \kappa$$

Equation 1. Negative weight function and the PCSF algorithm. The input for Forest is any network $G(V, E, c(E), p'(v))$ of a node set V and edge set E . **a)** Negative weight function. **b)** The objective function to be minimized.

The input for Forest is any network $G(V, E, c(E), p'(v))$ of a node set V and edge set E . $\text{degree}(v)$ is the number of connections of a node (i.e., protein, transcription factor or a metabolite) in the interactome. β and μ parameters are for including more differentially expressed proteins in the network and less number of connections for a specific protein (i.e., hub dominance), respectively. The more β increases, the less likely a molecule (depicted as $p(v)$) is removed from the network. Yet, it has to pay the cost for every interaction it has in the network. When μ increases, it is more likely that the node is penalized for probable hub dominance, unless the data highly propose its inclusion. Negative prizes can be used to discard proteins from the network that are poorly expressed or those which are irrelevant to the study of interest (Equation 1a).

Negative weighted prize nodes are not included in the network while minimizing the objective function (Equation 1b). Edge cost for each interaction per node should also be minimized. ω and κ refer to the uniform edge cost for dummy node, which is the root of the tree (i.e., network), and the number of trees in the forest.

D , μ , β , and ω parameters were modified and randomly put in configuration files in order to include as many terminal nodes as possible. D is set between the interval [5, 10], μ is set between [0.01, 0.1], and ω is set between [6,10]. β was changed accordingly to obtain more connected networks and to maximize the number of terminal nodes. This approach resulted in separate outputs for each data set and revealed hidden intermediate proteins. For the cytosolic proteome network, the data from Aras-Taşkın (2010) and Demir (2013) were combined. Differentially expressed cytosolic proteins at 16th-hour secretome samples were also included in the total cytosolic proteome data of OGU1 (Tekin-İşlerel, 2017).

3.3. Network analysis and visualization

NetworkX package v.2.2 of Python (Schult and Swart, 2008) and iGraph package v.1.2.2 (Csardi and Nepusz, 2006) were employed for the analysis of resulting networks and community identification. Mutual nodes and edges were identified by “networkx”. The same package was used to calculate centrality measures for each graph. As stated above, degree centrality depends on the number of connections a node has in a network. For each node, it is calculated by taking the ratio of a node’s connections to all edges in the network. Likewise, the betweenness centrality is a normalization of shortest paths that pass through a given node in the network (Brandes, 2001). The formula for betweenness centrality is

$$C_B(v) = \sum_{s,t \in V} \frac{\sigma(s,t | v)}{\sigma(s,t)}$$

where s, t and v are nodes in the node set V , $\sigma(s,t)$ is the number of shortest paths between any given nodes s and t , and $\sigma(s,t | v)$ is the number of those paths passing

through node v other than s and t . To identify the most critical nodes in each network, degree centrality and betweenness centrality attributes were calculated and normalized. Final networks were visualized with Cytoscape v.3.7.0. Louvain method was used to detect communities.

3.4. Functional enrichment analysis

After clustering the networks for each condition using the Louvain method, functional enrichment analyses of each cluster have been carried out on the Cytoscape plugin ClueGO v. 2.5.4. ClueGO is a user-friendly application which allows interpretation of biological data easily. Furthermore, ClueGO also has programmatic access, and the ontology and pathway databases can be imported and updated within the app. ClueGO facilitates the enrichment analysis process by creating a term-term similarity matrix based on kappa score, which is based on shared genes between the terms (Mlecnik, Galon and Bindea, 2019). The following categories were used for functional enrichment analysis: Gene Ontology (GO) biological process, GO cellular component, GO molecular function, and KEGG pathways. Network specificity was selected as medium, GO term fusion was applied to prevent redundancy, and corresponding significance values for each term were adjusted to $p\text{-value} < 0.05$. Enrichment results of each condition were grouped according to their categories. For each condition and term, we created matrices where rows are functional annotation terms and columns are corresponding p -values.

Negative logarithm of the p -values in base 10 was taken for better interpretation and display during the matrix generation. Subsequently, the results were visualized as heatmaps using “pheatmap” library v.1.0.12 of R in version 3.5.1.

CHAPTER 4

RESULTS AND DISCUSSION

4.1. Construction of OGU1 networks

Computational prediction of protein-protein interactions provides a quick means for gaining insight into protein function, as well as verifying experimental results. By combining different types of omics data, Omics Integrator allows detection of putative interactions. Unlike previous research with Omics Integrator (Tunçbağ et al., 2016), we have attempted to model changed interactions in the prokaryotic organism *B. subtilis* in this study. As stated above, the interactome data of *B. subtilis* 168 were used to generate networks for OGU1 due to unavailable data for the prototrophic PY79. For each condition, we have prepared datasets which consist of protein expression values. We have used the PCSF algorithm to generate networks for the cytosolic proteome, 12th-hour secretome, 16th-hour secretome and 24th-hour secretome. Networks which maximize the connectivity and the number of terminal nodes were selected and merged. In addition, the mean degree of Steiner nodes were considered, since proteins coming from the interactome should not have an excessive number of connections. Proteome network for OGU1 consists of 243 nodes and 1182 edges (Figure 4.1). 12th-hour secretome network has 61 nodes and 87 edges (Figure 4.2). 16th-hour secretome network includes 36 nodes and 45 edges (Figure 4.3). Finally, 24th-hour secretome network includes 18 nodes and 19 edges (Figure 4.4). Mutual number of nodes are represented in the Venn diagram given in Figure 4.5.

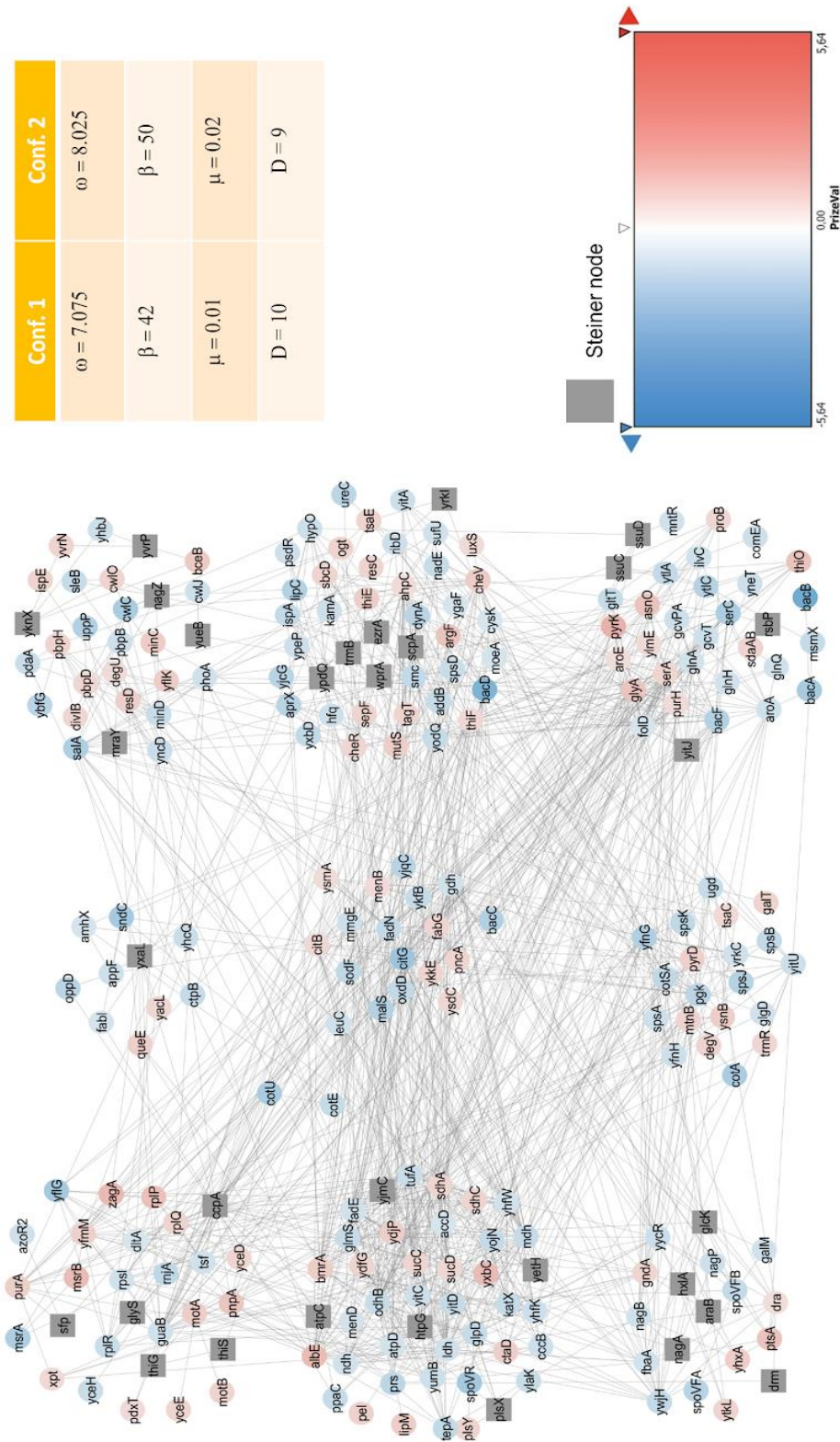


Figure 4.1. Cytosolic proteome network of OGU1. Gray squares represent intermediate proteins coming from the interactome. Red circles show increased \log_2 fold-change, blue circles show decreased \log_2 fold-change. The network consists of 243 nodes and 1183 edges, covering 10 clusters. Cluster 1 is at the top-left, Cluster 10 is in the bottom-right. Cluster 5 includes CotU and CotE.

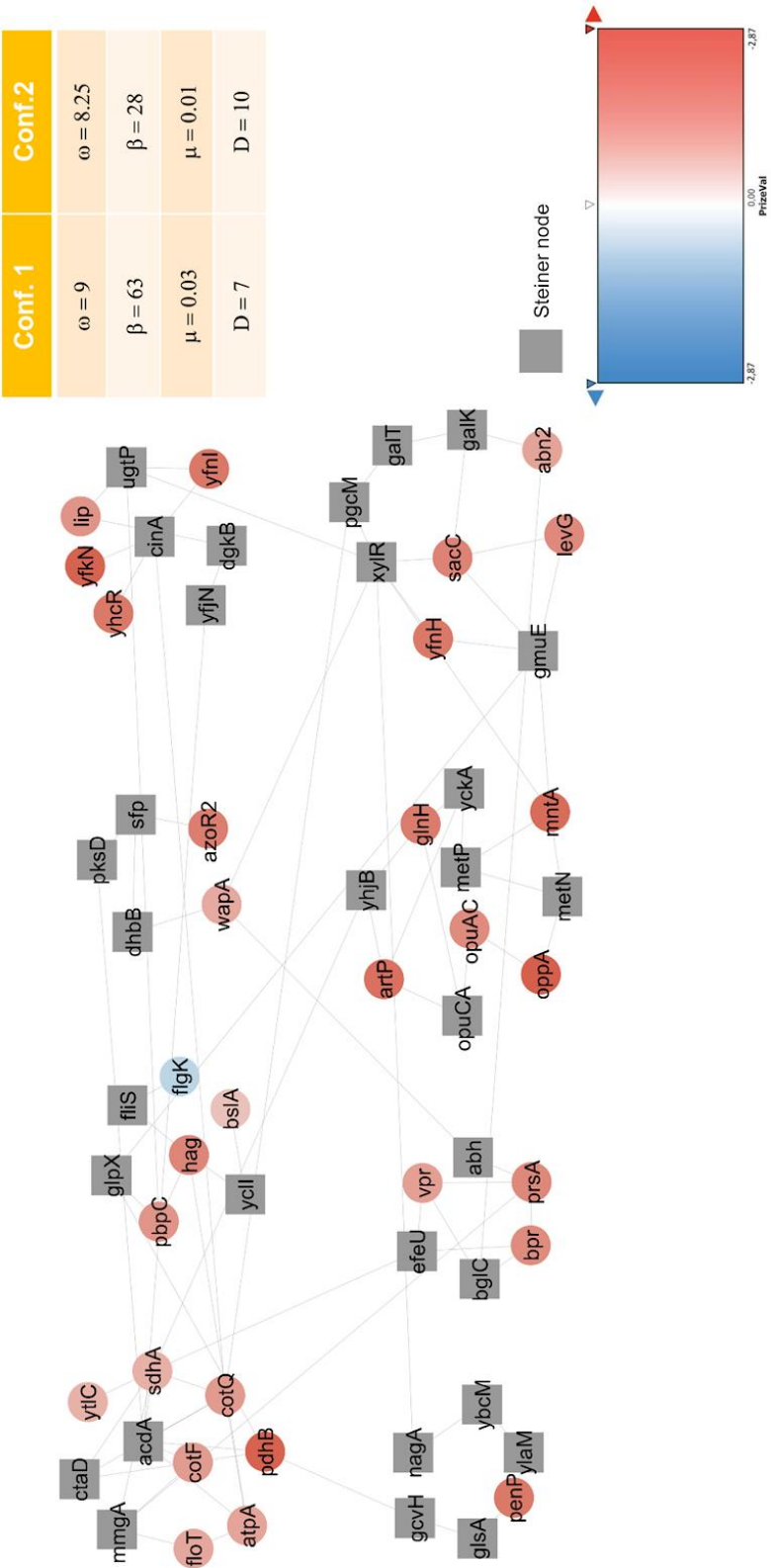


Figure 4.2. 12th-hour secretome network of OGU1. Gray squares represent intermediate proteins coming from the interactome. Red circles show increased \log_2 fold-change, blue circles show decreased \log_2 fold-change. The network consists of 61 nodes and 87 edges, and includes 8 clusters. Cluster 1 is at the top-left, Cluster 8 is in the bottom-right

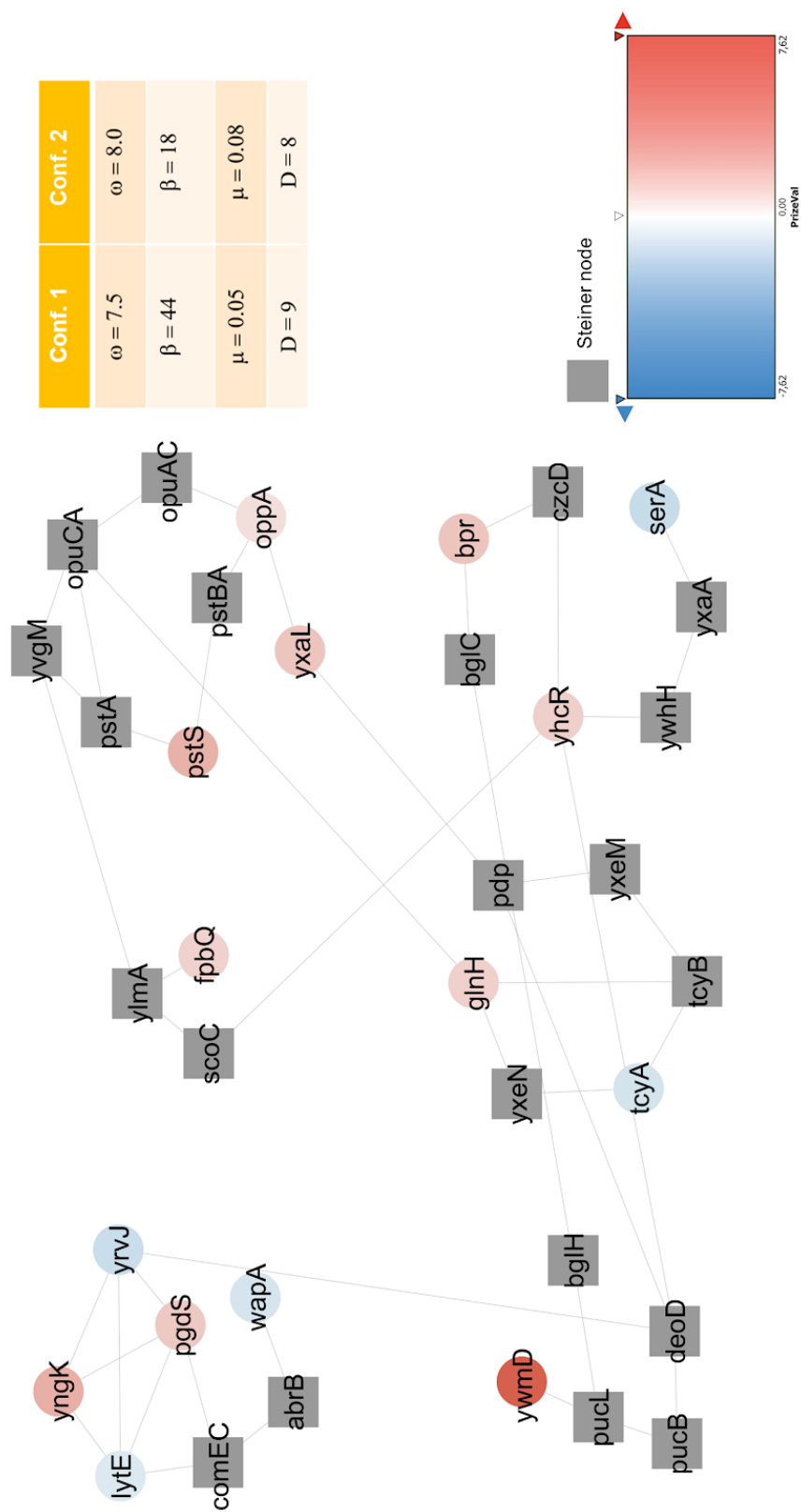


Figure 4.3. 16th-hour secretome network of OGU1. Gray squares represent intermediate proteins coming from the interactome. Red circles show increased \log_2 fold-change, blue circles show decreased \log_2 fold-change. The network consists of 36 nodes and 45 edges with 6 clusters.

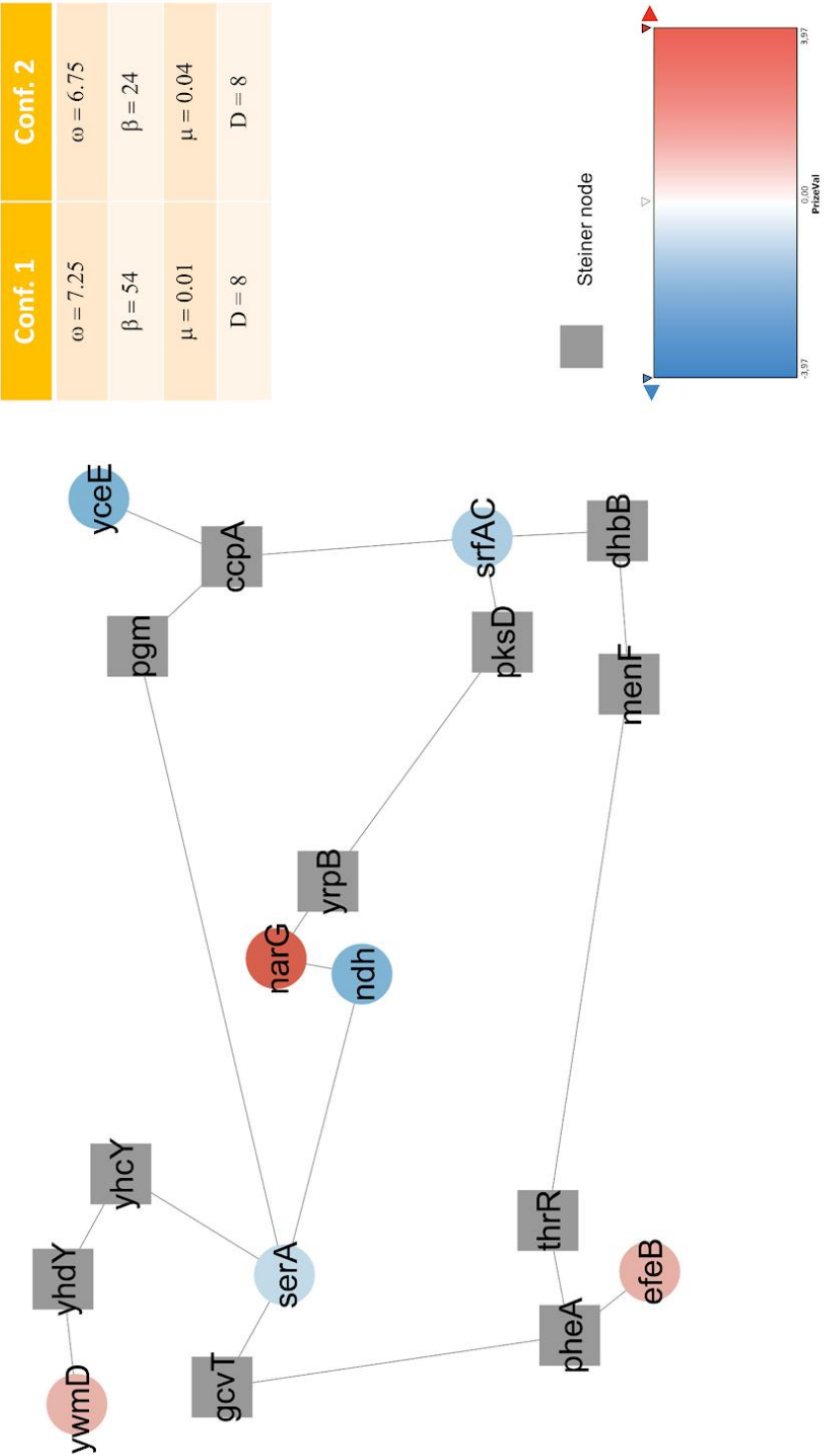


Figure 4.4. 24th-hour secretome network of OGU1. Gray squares represent intermediate proteins coming from the interactome. Red circles show increased \log_2 fold-change, blue circles show decreased \log_2 fold-change. The network consists of 36 nodes and 45 edges with 6 clusters. Cluster 1 is at the top-left, Cluster 6 is in the bottom-right.

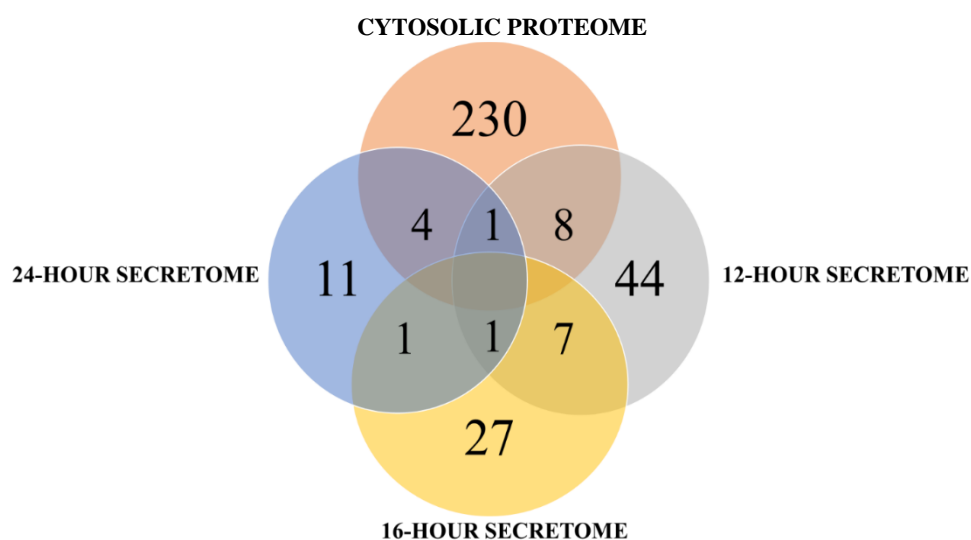


Figure 4.5. Diagram of total number of proteins found in each network. Proteome, 12-hour secretome and 16-hour secretome only have GlnH in common. Similarly, proteome, 16-hour secretome and 24-hour secretome have SerA as their mutual protein. SerA is also found to be mutually present in 16-hour secretome and 24-hour secretome networks along with YwmD. Proteome and 12-hour secretome have the most number of intersecting proteins with 8 proteins being mutual. 7 proteins are found in common for 12-hour and 16-hour secretome networks. Finally, 4 proteins are mutually present in 24-hour secretome and proteome networks.

Figure 4.1 shows that preparing a network of OGU1 based on experimental data can be used as a reference in analyzing modified interactions. Figures 4.2, 4.3 and 4.4 reveal the intercommunication of extracellular proteins. All of the figures indicate that experimental data and analysis of lists and functions differentially expressed genes are not sufficient to depict all of the interactions within a cell in a specific condition. Therefore, generating networks with the PCSF approach helps us identify the hidden proteins involved in various biological processes.

4.2. Functional enrichment of resulting networks

As described before, networks were clustered by using the Louvain method. Cytosolic proteome network of OGU1 was divided into 10 clusters. 12th-hour secretome network was partitioned into 8 clusters, 16th-hour secretome network comprised 6 clusters, and 24th-hour secretome network consisted of 5 clusters.

Table 4.1. Number of proteins in each network condition.

Network	Number of differentially expressed proteins	Number of proteins found in the interactome	Total number of proteins in the network	Number of intermediate proteins in the network	Number of edges in the network	Number of clusters
Proteome	264	212	243	31	1182	10
12-hour secretome	46	30	61	31	87	8
16-hour secretome	26	15	36	21	45	6
24-hour secretome	16	12	18	11	19	5

We then analyzed each network with ClueGO to detect specific functions, cellular locations and pathways which are enriched under each condition. GO terms and KEGG pathway enrichments were checked by applying a p-value (Benjamini-Hochberg p-value correction) cutoff of 0.05. Although ClueGO allows the enrichment analysis of multiple clusters at once, a number of clusters in each network did not have any enrichment under adjusted p-value cutoff 0.05. Some terms can be associated with more than one clusters, which makes them non-specific.

4.3. Alterations in the energy metabolism burdens cellular processes in OGU1

More than half of the clusters of OGU1 included enriched terms associated with aerobic respiration, cell morphogenesis and amino acid metabolism. Some genes which play role in the same biological process were found in different clusters, although a term can be specific for only one cluster.

Conversion of nutrients to energy is a universal process in all living organisms, including *B. subtilis*. In aerobic organisms, oxidative phosphorylation results from sugar, lipid and amino acid catabolism and culminates in the reduction of oxygen in the electron transport chain, driving ATP synthesis for a broad range of cellular activities. Our group have previously suggested that the energy demand and oxidative stress response of OGU1 may have shifted, due to the presence and upregulation of thiamine (pyruvate dehydrogenase cofactor) synthesis proteins ThiE and ThiO in the mutant strain compared to PY79. Pyruvate dehydrogenase is a connecting enzyme which catalyzes the conversion of pyruvate to acetyl-CoA in the TCA cycle by utilizing thiamine as a cofactor. Additionally, the overexpression of CtaD contributed to the argument described earlier. The significant enrichment in aerobic respiration (adjusted p-value = 4.28×10^{-7}) and its associated pathways such as glycolysis/gluconeogenesis (adjusted p-value = 2×10^{-7}), TCA cycle (adjusted p-value = 5.14×10^{-11}), oxidative phosphorylation (adjusted p-value = 9×10^{-4}) and pentose phosphate pathway (adjusted p-value = 1.93×10^{-4}) confirms our previous results, with the ATP synthase epsilon subunit AtpC as an additional interaction partner included in this study. First degree neighbors of AtpC are given in Figure 4.6. Most of its neighbors are found to be downregulated, which can be attributed to negative regulation by the stringent response.

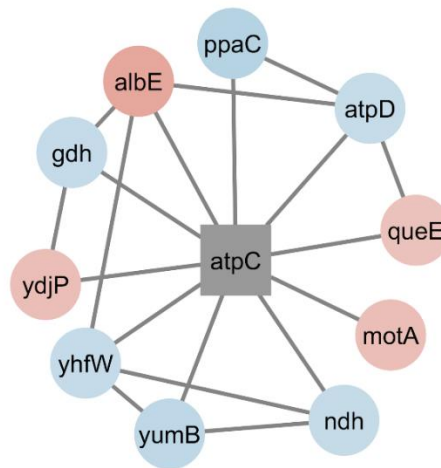


Figure 4.6. First-degree neighbors of AtpC in the cytosolic proteome network of OGU1.

During oxidative phosphorylation, NADH is oxidized back to NAD⁺ by NADH dehydrogenases, releasing H⁺. Then succinate, which was generated in the TCA cycle, is converted to fumarate by succinate dehydrogenase proteins SdhCAB (not shown). After that, oxygen is consumed and then ATP is synthesized and released by ATP synthase proteins in the F1 complex, AtpC and AtpD. During ATP synthesis, the manganese-dependent inorganic pyrophosphatase PpaC catalyzes the reaction

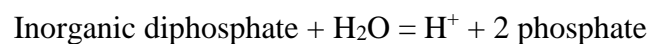


Figure 4.6 shows that NADH dehydrogenases Ndh, YhfW and YumB show decreased expression. Furthermore, AtpD and PpaC were found to be downregulated compared to PY79. However, motility protein MotA, subtilisin production protein AlbE, and tRNA modifying protein QueE are upregulated in the mutant strain. The sigma-W regulated stress protein YdjP, of whose first degree interaction partners are linked to sporulation, is also upregulated. Previous studies by our group have demonstrated decreased cell wall synthesis in OGU1.

In optimum conditions, phosphorylated GlcNAc is incorporated into glycolysis, although it is used for cell wall synthesis for the most part. AlbE is induced under oxygen limitation and nutrient starvation in OGU1 (Bertram et al., 2011). It is possible that bacilysin deficiency results in less GlcNAc utilization by Gdh, causing less ATP synthesis. The absence of bacilysin might have been sensed as a starvation signal, stimulating the expression of YdjP, AlbE and MotA. The upregulated MotA may lead to increased motility and chemotaxis for nutrient uptake.

Glycolysis and pentose phosphate (PP) pathway are interconnected in all organisms, in that glucose can either be oxidized to ribose 5-phosphate by PP pathway for amino acid and nucleotide synthesis, or to pyruvate for energy derivation (Figure 4.7). The partitioning of glucose into either one of the pathways depends on the reversible balancing of NADPH and NADP^+ by various dehydrogenases in *B.subtilis* (Rühl et al., 2012; Weckbecker and Hummel, 2005; Zamboni et al., 2004; Dauner et al., 2002).

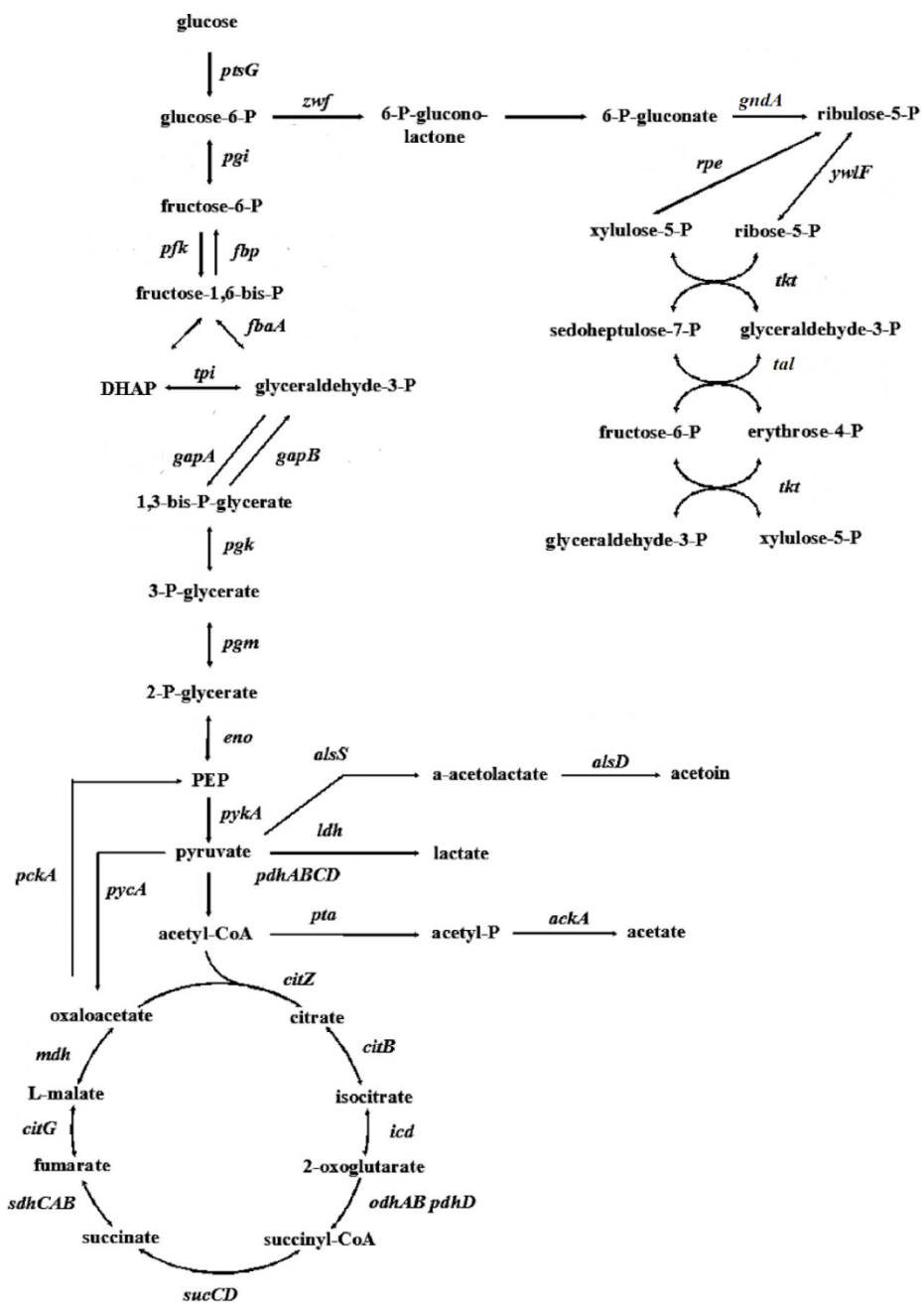
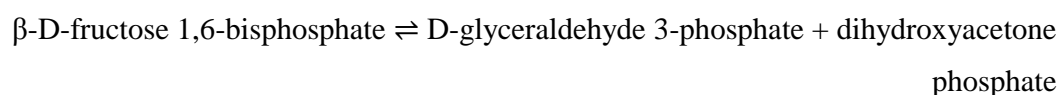


Figure 4.7. Overview on the regulation of genes encoding enzymes of glycolysis, the pentose phosphate shunt, and the Krebs cycle in *B. subtilis* (Schilling et al., 2007).

The PP pathway was found to be significantly enriched in the cytosolic proteome of OGU1. It is also linked to amino acid and nucleotide metabolism. During glycolysis in *B.subtilis*, glucose is converted to D-glutono-1,5-lactone (GDL) by glucose 1-dehydrogenase Gdh. Through several steps, Prs protein converts GDL to PRPP, which is a crucial intermediate in nucleotide and histidine synthesis. HxlA and GndA are other proteins involved in the synthesis and utilization of D-ribose 5-phosphate, which acts as a scaffold for both amino acids and nucleotides. In the meantime FbaA catalyzes the reversible reaction below to generate pyruvate (or vice versa) for downstream reactions:



It is probable that the decrease in expression levels of Prs, YwjH and FbaA indirectly triggers the oxidative phase in PP pathway, in which glucose is oxidized to D-ribose 5-phosphate by Drm. This may contribute to the oxidative stress which impairs sporulation, cell wall organization, cell division and regulation of cell shape.

One carbon metabolism is a universal process and refers to reactions in which amino acids and nucleotides are formed by the transfer of one-carbon group such as methyl (-CH₃) or formyl (-CHO). It is mediated by folate cofactors such as the versatile tetrahydrofolate (THF), which is produced in the one carbon metabolism and required for break down and *de novo* synthesis of amino acids, RNA and DNA (Bertacine Dias et al., 2018). In the one carbon pool pathway of *B.subtilis*, methyltransferases GcvT or GlyA convert folate to THF, which is then transformed to 10-formyl THF to be used in purine synthesis. 10-formyl THF is reversibly converted to 5,10 methenyl THF, which is reduced to 5,10 methylene THF by Fold. The latter molecule is then turned into 5-methyl THF by the methionine-regulated YitJ to be utilized in pyrimidine biosynthesis.

In our study, first degree neighbors of YitJ were found to mostly exhibit decreased protein expression. Among its first interacting partners are downregulated proteins SpsK, DynA, LuxS and CysK, which are related to spore coat polysaccharide synthesis, determination of cell shape, cell division, motility, quorum sensing and regulation of transcription factors.

Fold is a negative regulator of stringent response. Stringent response can be defined as the reaction of a bacterium against a stress condition such as amino acid starvation, heat shock or another limiting condition (Irving and Corrigan, 2018). The downregulation of Fold in OGU1 shows that the lack or absence of bacilysin triggers the stringent response. Furthermore, it induces a flux towards purine synthesis along with the upregulation in GlyA and YkkE. GlyA is linked to amino acid uptake, glycine production and genetic competence. It is found to be directly linked to serine phosphatase RsbP, which participates in protein modification by serine phosphorylation. GlyA is known to be phosphorylated either from a serine, threonine, or tyrosine residue. Upon phosphorylation, it is possible that RsbP activates GlyA to carry out its functions. Therefore it can be deduced that OGU1 compensates the effects of bacilysin deficiency through this indirect mechanism.

Glycine, serine and threonine metabolism is another affected pathway in OGU1 (adjusted p-value = 1.2×10^{-2}). In this pathway, SerA converts the intermediate molecule 3-phospho-D-glycerate from glycolysis to 3-phosphate-hydroxy-pyruvate, which is then converted to phosphoserine by SerC. SerB then hydrolyzes phosphoserine to serine. Then GlyA, which requires THF as a cofactor, catalyzes the reversible conversion of serine to glycine. Serine can also be integrated into pyruvate metabolism (i.e., gluconeogenesis) and TCA cycle for amino acid and fatty acid synthesis by L-serine deaminase heterodimer molecules SdaAA and SdaAB. The increase in SerA and SdaAB, as well as the slight decrease in SerC suggest integration of serine either through conversion to pyruvate by SdaAB, or from 3-phospho-D-

glycerate in glycolysis. Upregulation of GlyA also hints to glycine production, which has been shown to favor sporulation at the end of growth phase along with L-serine (Sugae and Freese, 1970). Greenwich et al. have shown that serine levels decrease in stationary phase under optimum conditions, triggering biofilm formation (2019). These findings demonstrate that although sporulation frequencies of OGU1 and PY79 are similar, it is possible that elevated serine levels inhibit biofilm production, and bacilysin acts as a regulator of biofilm repressor SinR or other histidine kinases.

Obviously, one of the key pathways afflicted by bacilysin deficiency is the biosynthesis of secondary metabolites (adjusted p-value = 3×10^{-4}). The connections of *bac*-operon genes to other proteins with diverse functions in the network indicate that the loss of bacilysin from the cell has a crucial impact on general cellular activity.

Arginine biosynthesis occurs in multiple ways within the urea cycle in *B. subtilis*. In the first route, it is produced by the degradation of glutamine to NH_3 , which is contributed into glutamate structure. Probable N-acetylornithine deacetylase YodQ then converts glutamate to ornithine as an intermediate molecule in the urea cycle. In the urea cycle, ArgF transfers a carbamoyl group to ornithine, which produces another intermediate molecule, citrulline. Alternatively, NH_3 can directly be used to generate citrulline by ArgF. Arginine is then produced through either L-argininosuccinate which releases fumarate as a byproduct, or directly by NosA, which reduces citrulline to arginine. The second route for arginine synthesis occurs from the TCA cycle intermediate 2-oxoglutarate, which produces glutamate and then ornithine to join the urea cycle and yield arginine. Arginine can be degraded to CO_2 and urea by UreC, which can be used as an alternative nitrogen source.

Arginine plays an essential role in the physiology of *B. subtilis*, in that its phosphorylation on multiple sites at different proteins regulates many critical cellular processes, such as protein degradation, motility, competence, and stringent and stress responses (Elsholz et al., 2012). Twin arginine motif in the N-terminus of passenger proteins provide transportation of folded proteins across the plasma membrane (Barnett et al., 2007). Accumulation of positively charged amino acids arginine, ornithine and citrulline has been shown to substitute for potassium, which acts as a counterion for glutamate and controls gene expression, as well as enzyme activation (Gundlach et al., 2017). A previous study by Cruz-Ramos et al. has demonstrated partial growth defect in minimal medium, and suggested that ornithine degradation produces sufficient nitrogen catabolites to permit rapid growth and reduce urease expression (1997). Indeed, UreC expression was found to be $\sim 2.5 \log_2$ -fold decreased in OGU1, indicating a flux of ornithine to citrulline. Additionally, the slight decrease in GlnA does not change the conversion of glutamine to citrulline by ArgF, of whose expression is upregulated by 2.49-fold in logarithm base 2. Downregulation of YodQ also proposes a fluctuation towards the direct conversion of N-acetylornithine to citrulline. The expression of YodQ is regulated by mother cell-specific transcription factor σ_E (sigma-E) (Feucht, Evans and Errington, 2003).

Upon initiation to sporulation, the decline in YodQ expression proposes a dysregulation in the sporulation process in the absence of bacilysin. Furthermore when looked up for the first neighbors of YodQ, two secreted proteins, WprA and YxaL were present. WprA is a serine-type protease which is associated with protein quality control and cell wall growth during exponential phase (Tjalsma et al., 2004; Margot and Karamata, 1996), whereas YxaL is a kinase linked to DNA repair along with PcrA (Noirot-Gros et al., 2002). It is probable that the downregulation of YodQ results in reduced production of cell wall-associated proteins.

CcpA is a global transcriptional regulator which binds to catabolite repressor elements in the promoters of target genes in *B. subtilis* (Cao et al., 2018; Singh et al., 2008; Choi and Saier, 2005). In our study, CcpA was found to neighbor 43 proteins with different functions in TCA cycle, competence, flagellar activity, cell wall synthesis, and resistance with varying expression levels (Figure 4.8).

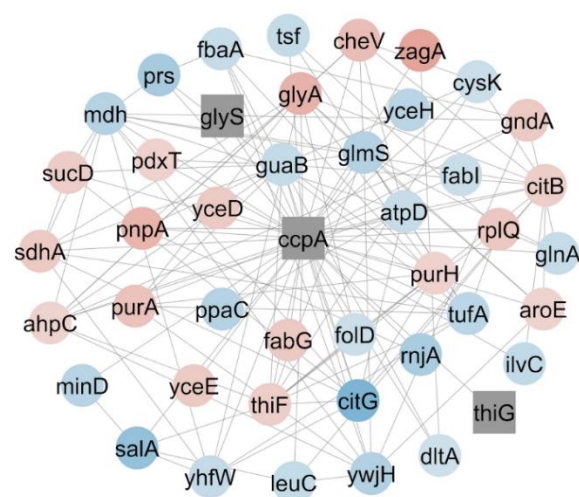


Figure 4.8. First-degree neighbors of the carbon catabolite repressor CcpA. Absence of bacilysin triggers stringent response, negatively regulating proteins Prs, Tsf, TufA, cell wall synthesis proteins GlmS and DltA, as well as competence protein PnpA. This results in the activation of chemotaxis and biofilm formation protein CheV to acquire nutrients. Additionally, fatty acid synthesis proteins have a slight countereffect on cell wall synthesis. Aromatic amino acid synthesis proteins such as AroE, GlyA, PurA and PurH, as well as TCA cycle and PP pathway proteins such as SdhA and GndA then become activated. This triggers oxidative phosphorylation, which can be explained by the upregulation of thiamine synthesis protein ThiF and oxidative stress resistance protein AhpC.

Diaminopimelic acid (DAP) is a characteristic component of some bacterial cell walls. In the early steps of cell wall synthesis in *B. subtilis*, uridine nucleotide becomes attached to and activates the cell wall components N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (Mur2Ac). Later in the pathway, Mur2Ac is integrated to

DAP-based peptidoglycan cell wall by penicillin-binding proteins PbpD, PbpH and PbpB. Transpeptidase PbpD increased approximately 1.6 log₂-fold, PbpH which facilitates formation of a cell wall increased 1.78 log₂-fold, and cell division protein PbpB decreased by 1.75 log₂-fold in OGU1. Under normal conditions during stationary phase, expression of PbpB and PbpH reaches a peak, whereas PbpD levels drop. However, bacilysin deficiency reverses the effects mentioned, and sporulation and cell wall formation is dysregulated.

Another enriched biological process in OGU1 is cell septum assembly of cluster 7 (adjusted p-value = 5.18×10^{-2}), which includes proteins EzrA, DynA and SepF. During cell division, EzrA stimulates the recruitment of PbpA to the septum, which shuttles between cell poles and septum, and required for peptidoglycan synthesis. The dynamic behavior of PbpA requires various cell division proteins including FtsZ, DynA and EzrA (Pompeo et al., 2015). EzrA's neighbors in the cytosolic proteome network are SepF and Smc. SepF protein which was found to be elevated ~1.8 log₂-fold in OGU1 recruits FtsZ to cell membrane for cytokinesis. A recent study by Gao et al. has indicated that overproduction of SepF disrupts the assembly of late division proteins responsible for septum synthesis, in addition to suppressing the two-component system WalRK that stimulates the expression of FtsZ (2017). Apart from co-localizing with FtsZ during cell division, DynA induces local lipid bilayer fusion and sealing membrane gaps upon membrane damage (Sawant et al., 2015; Dempwolff et al., 2012). DynA was downregulated 2.55 log₂-fold in OGU1. Therefore, decreased levels of DynA also increases membrane stress along with decreased cell division.

Smc is a part of the condensin complex, and is also involved in the segregation of replication origins along with ScpA and ScpB, which bind to the head domain of Smc (Krepel et al., 2018). The three proteins are repressed by the stringent response in *B. subtilis*. Reduced levels of Smc imply a decline in cell division activity, due to repression by the stringent response as a result of bacilysin insufficiency.

The ATP-dependent genetic competence and DNA repair protein AddB recognizes the chi sequence 5'- AGCGG-3' and transforms into a helicase from a destructive endonuclease by changing its conformation (Gilhooly et al., 2016). AddB was found to be decreased ~2.1 log₂-fold in OGU1. AddB is also an interaction partner of ScpA, which was mentioned earlier. Therefore, the decline in cell division can be associated with diminished genetic competence and DNA repair. These findings demonstrate the interrelations of proteins in OGU1. The enriched pathways of OGU1 are given in Figure 4.6. and some of the most central nodes with crucial functions are listed according to their ranks in Table 4.2. A list of enriched biological functions are given in Figure 4.7.

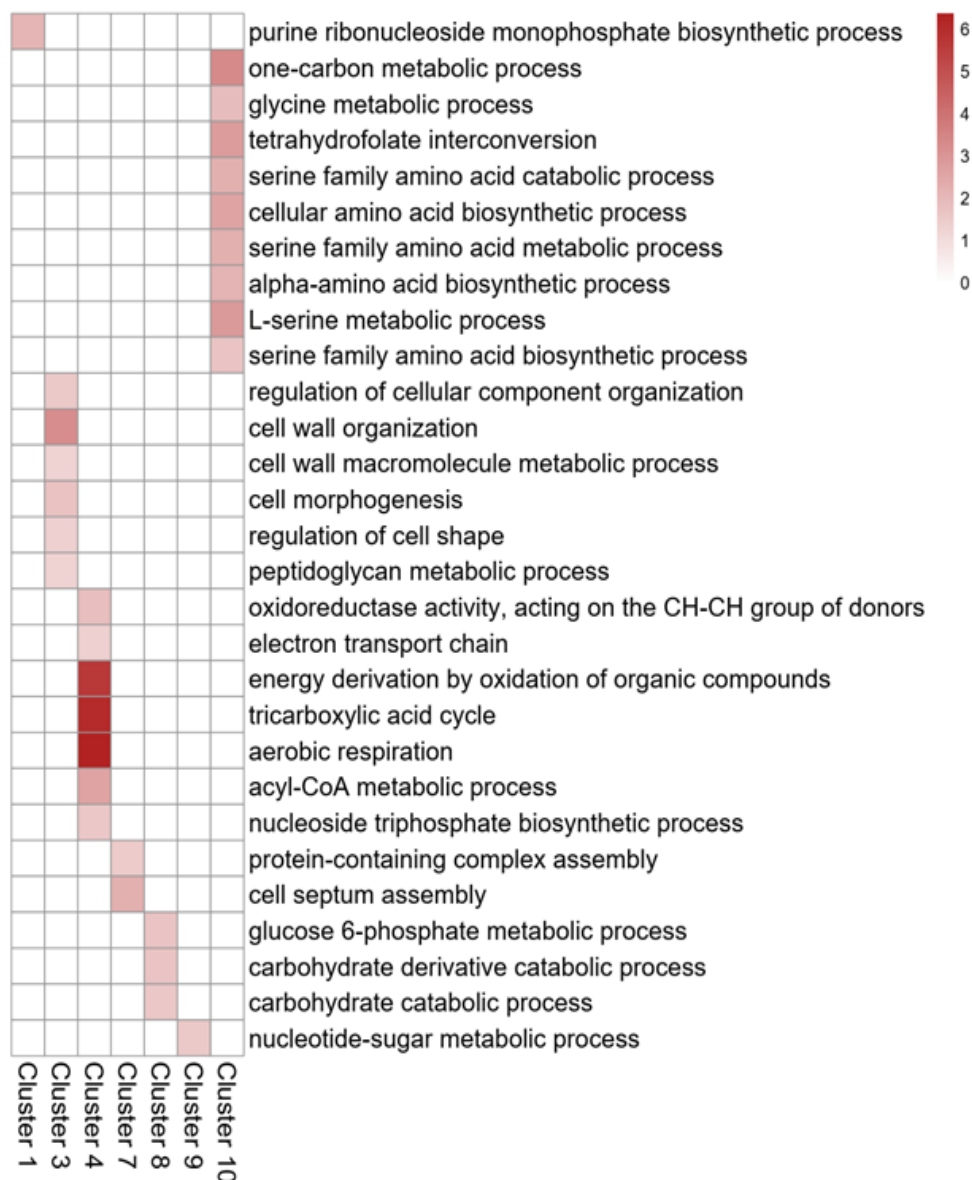


Figure 4.9. Enriched Gene Ontology biological process terms for clusters in the cytosolic proteome network of OGU1. Adjusted p-value cutoff was set to 0.05 for the enrichment results and the adjusted p-values were transformed into negative logarithm base 10.

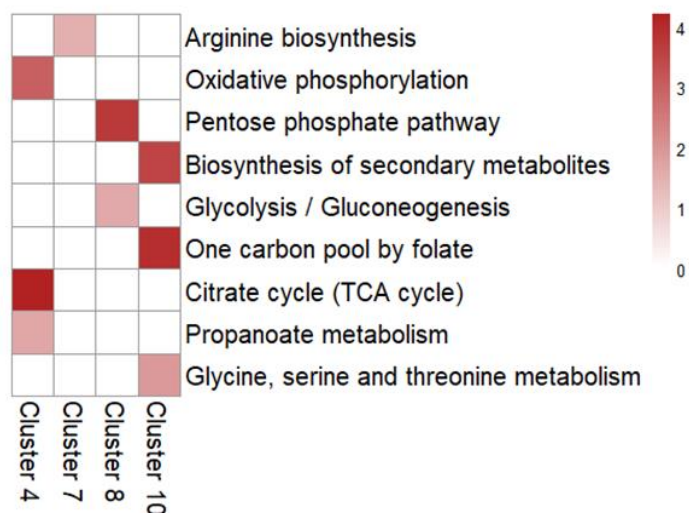


Figure 4.10. Enriched KEGG pathway terms for clusters in the cytosolic proteome network of OGU1. Adjusted p-value cutoff was set to 0.05 for the enrichment results and the adjusted p-values were transformed into negative logarithm base 10.

Table 4.2. Probable effectors of cellular functions under bacilysin loss in the cytosolic proteome network of OGU1.

Protein	Function	Operon/Regulon	Type of node in the network	Degree centrality	Betweenness centrality
CcpA	Transcriptional regulator, carbon catabolite repressor	SigA, CcpA	Intermediate	0.1777	0.0955
GlyA	Glycine biosynthesis, genetic competence	PurR, T-box, SigA	Proteomic	0.1653	0.0613
HtpG	Class III heat-shock protein (molecular chaperone)	SigA	Intermediate	0.1281	0.0232
GlmS	Cell wall synthesis	SigA	Proteomic	0.124	0.0213
AlbE	Antilisterial bacteriocin (subtilisin) production	Rok, ResD, AbrB, SigA	Proteomic	0.1157	0.0192
YitJ	Methionine biosynthesis, tetrahydrofolate interconversion	S-box	Intermediate	0.0579	0.0143
WprA	Cell wall, protein quality control	SigA, CcpA, YvrHB	Intermediate	0.0413	0.0114
NagZ	Cell wall recycling, utilization of cell wall components	MurR	Intermediate	0.0331	0.0389
DegU	Regulation of degradative enzymes, genetic competence and other adaptive responses	SigA, DegU, TnrA, CcpA, SinR	Proteomic	0.0331	0.0035
YxaL	Unknown, increases the processivity of the PcrA helicase	Rok, DnaA, AbrB	Intermediate	0.0289	0.0179
GlnH	Glutamine uptake	TnrA, SigE	Proteomic	0.0248	0.0043
NagA	N-acetylglucosamine utilization	SigA, CcpA, NagR	Intermediate	0.0165	0.0082
Sfp	Biofilm formation, biosynthesis of antibacterial compounds	SrfA	Intermediate	0.0124	0.0082
MotA	Motility and chemotaxis	SigD	Proteomic	0.0124	0.0083
EzrA	Control of FtsZ ring formation	SigA	Intermediate	0.0083	0.0082

4.4. Dynamic secretome analysis of OGU1

The secretome refers to the subset of proteome which comprises secreted proteins, as well as their components in a cell (Tjalsma et al., 2004; Greenbaum et al., 2001). The secretome studies help us better understand the survival mechanisms of many microbes which have the ability to secrete considerable amounts of valuable substances with special properties (Greenbaum et al., 2001). Thus, secretome is highly dynamic and responsive to various environmental stimuli and pathologies (Caccia et al., 2013).

B. subtilis secretome changes at the onset of the sporulation to include new membrane- and wall-binding proteins as well as those secreted outside, such as new degradative enzymes that may provide the cells with new nutrients and those engaged in detoxification, communication, defense and many more. The absence of an outer membrane in *Bacillus subtilis* simplifies protein secretion pathways and allow the organism to secrete high levels of extracellular proteins.

As mentioned above, our group have elucidated the effect of bacilysin in a time-dependent secretome analysis. Furthermore, the expression of bacilysin is relatively constant during the exponential growth, but increases during the transition between exponential and stationary phases and reaches to its maximal level upon entry into stationary phase (16th-hour) (Figure 4.11). Time-dependent interactions and their results on OGU1 are described in the following sections.

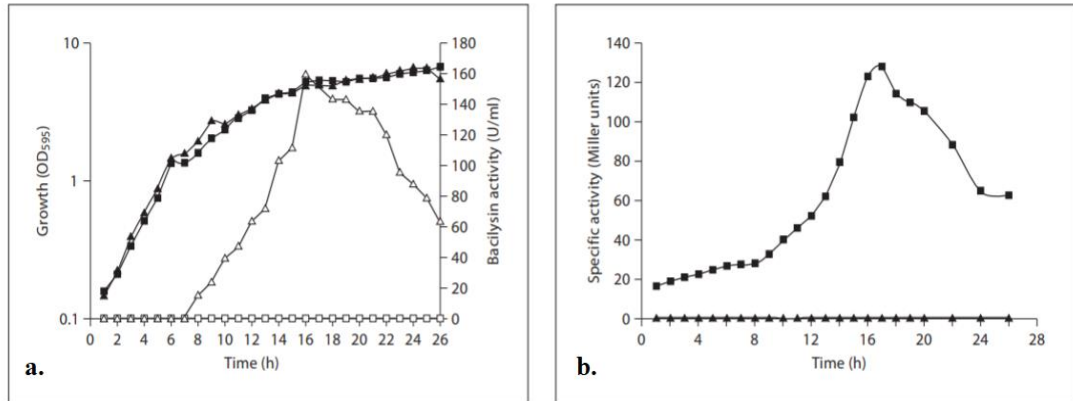


Figure 4.11. a. Growth (closed symbols) and bacilysin activity (open symbols) of *B. subtilis* PY79 (triangles) and OGU1 (squares) in PA medium. **b.** Expression of transcriptional *bacA-lacZ* fusion in PA medium. Specific activity of *B. subtilis* PY79 (black triangles) and OGU1 (black squares) (Köroğlu et al., 2011).

4.4.1. Bacilysin triggers carbohydrate metabolic activity and promotes exponential phase

As seen in the graphs above, cell growth of both PY79 and OGU1 increases in the 12th-hour, indicating higher levels of protein expression. Dynamic secretome analysis on the 12th-hour samples has shown increased protein expression of extracellular proteins.

5 clusters are found to include pathway terms which are specific to the 12th-hour secretome samples of OGU1. Most of the pathways are associated with carbohydrate and amino acid transport. Cluster 7 is enriched in ABC transporters (adjusted p-value = 8.26×10^{-4}), and cluster 8 is enriched in amino sugar and nucleotide sugar metabolism (adjusted p-value = 1.41×10^{-4}), as well as fructose and mannose metabolism (adjusted p-value = 7.58×10^{-4}).

ABC transporters refer to membrane proteins which either export or import their specific substrates. Like other bacteria, ABC transporters in *B. subtilis* can transfer minerals, lipids, amino acids, sugars and peptides across the plasma membrane. In OGU1, these upregulated membrane protein complexes are mainly associated with amino acid transport, specifically glycine, methionine, arginine, glutamine, and cystine (adjusted p-value = 9.44×10^{-5}). The proteins associated with amino acid transport are ArtP, GlnH, MetN, MetP, OpuAC, OpuCA, MntA, OppA and YckA.

OppA is regulated by ScoC during exponential growth (Koide, Perego and Hoch, 1999), and involved in the import of competence and sporulation stimulating factor CSF (Zafra et al., 2012). In OGU1, OppA is directly neighbored by OpuAC and MetN. MetN and MetP permit the entry of methionine by binding S-adenosylmethionine (AdoMet), another one carbon group cofactor, to their S-box motif (Hullo et al., 2004). Methionine may act as the sole sulfur source during exponential growth. OpuAC is an osmotic stress related protein that protects the cell against rupture during elongation (Hoffmann and Bremer, 2017). Additionally, OpuAC and OpuCA are targets of cyclic diAMP, which is essential for growth in many Gram-positive bacterial species (Fahmi, Port and Cho, 2017). OpuCA By interacting with OppA, OpuAC (and OpuCA) and MetN allow glycine and methionine transport, while protecting the cell during growth in bacilysin deficiency.

ArtP and GlnH transport arginine and glutamine which are essential for the physiology of *B. subtilis*, as discussed earlier. MntA encodes manganese-binding lipoprotein A and is important for manganese ion transport. Manganese uptake in *B. subtilis* is active in the exponential growth and tightly regulated on transcription level (Que and Helmann, 2000). Manganese uptake also occurs during initiation of sporulation. As mentioned before, sporulation frequencies of OGU1 and its parental strain were similar.

It is possible that MntA benefited for sporulation initiation in OGU1 during exponential growth and contributed to comparable sporulation frequency with PY79 (Tekin-İşlerel, 2017).

Amino sugar and nucleotide sugar metabolism (adjusted p-value = 1.41×10^{-4}) is interconnected to fructose and mannose metabolism (adjusted p-value = 7.58×10^{-4}) in OGU1. Like all living organisms, this mutant strain utilizes carbohydrates efficiently for its growth, even in the absence of bacilysin. The first step of amino-sugar nucleotide synthesis involves the hydrolysis of the N-acetyl group of N-acetylglucosamine-6-phosphate by NagA to yield glucosamine 6-phosphate (GlcN-6P) and acetate. GlcN-6P then participates in various metabolic pathways, including cell wall synthesis, pentose phosphate pathway, and starch and sucrose metabolism. YbcM is a similar enzyme to GlmS, which is triggered by the binding of GlcN-6P as a cofactor so that the gene is only expressed in the absence of GlcN-6P (Collins et al., 2007). The increase in NagA and GlcN-6P probably blocks YbcM activity in the 12th-hour.

GmuE is a fructokinase which adds a phosphate group to fructose, producing fructose-6-phosphate (Fru-6P) which is also an intermediate in GlcN-6P synthesis and glycolysis. GmuE's first degree neighbors are MntA, levanase SacC, fructose uptake protein LevG and glucose-1-phosphate cytidyltransferase YfnH. During exponential growth, SacC catalyzes the degradation of polysaccharide levan to fructose, and LevG transfers fructose across the membrane by using phosphoenolpyruvate (PEP) as a cofactor to be utilized in other carbohydrate-associated processes. The presence of elevated GmuE along with LevG and SacC in the 12th-hour secretome network confirms the cell's demand for energy during growth. GmuE is also related to starch and sucrose metabolism (adjusted p-value = 3.58×10^{-3}), in that phosphorylated fructose is converted to D-glucose-phosphate by PgcM as a final step. D-glucose-phosphate can then be integrated into energy production, biofilm formation, secondary

metabolite synthesis and two-component systems. The cellulose degrading protein BglC also contributes to D-glucose-phosphate formation during glycolysis.

Another enriched pathway in the 12th-hour secretome of OGU1 is alanine, aspartate and glutamate metabolism (adjusted p-value = 9.64×10^{-3}). Alanine plays a crucial role in bacterial growth and viability, the L-isomer of this amino acid is one of the building blocks for protein synthesis, and the assimilation of D-isomer results in the incorporation of D-alanine into the cell wall (Sidiq et al., 2019). Similarly, aspartate limitation in the DSM results in cell lysis and depletion of peptidoglycan precursor DAP due to glutamate activity as a competitive inhibitor of aspartate uptake (Zhao, Roistacher and Helmann, 2018). Proteins associated with this pathway in the 12th-hour network are GlsA, YlaM and YbcM. GlsA and YlaM have PenP as their common neighbor, which is a narrow spectrum β -lactamase with cell wall and biofilm resistance activity (Bucher et al., 2019). It was shown to increase almost 2.4 log₂-fold in OGU1. GlsA and YlaM are both glutamine degradation proteins. It can be therefore deduced that GlsA and YlaM are downregulated and PenP is upregulated to protect the cell from lysis in OGU1 during the exponential phase, and glutamine is either incorporated into cell wall structure by GlmS or used for amino acid biosynthesis.

As expected, flagellar assembly and motility is found to be significant during exponential growth (adjusted p-value = 9×10^{-3}). During flagellar assembly, flagellin synthesizing protein Hag produces flagellin monomers, which are bound to and delivered by the chaperone protein FliS to the secretion apparatus. After binding to the extracellular chaperone FliD, flagellin folding and assembly is catalyzed. FliD serves as the filament cap, and polymerizes atop two structural transition proteins called FlgK and FlgL that bridge the flagellar hook and filament (Mukherjee and Kearns, 2014). Upregulation of Hag protein in OGU1 provides a faster flagellum assembly, thereby accelerating motility and chemotaxis. In our predicted network, Hag is directly linked to PbpC and YclI. Loss of penicillin binding proteins has been shown to antagonize flagellum assembly in *E.coli* (Evans et al., 2013). In another

study, BacA and BacB have been shown to colocalize with PbpC, which is also involved in stalk and flagellin formation (Andari, Altegoer, Bange and Graumann, 2015). Silencing bacilysin operon then probably results in shorter membrane extensions, while increasing flagellar activity and movement towards nutrients in OGU1.

MmgA, which is related to fatty acid β -oxidation, is regulated by σ_E (sigma-E) regulon. Sigma-E regulon sustains appropriate metabolic conditions in the mother cell in order to maintain proper spore morphogenesis (Steil, Serrano, Henriques and Völker, 2005). However, Mmg operon is subjected to catabolite repression by CcpA during sporulation under normal conditions (Bryan, Beall and Moran Jr, 1996). The increment in protein expression of MmgA along with FloT, CotF, SdhA and PrsA demonstrate a contrary effect under bacilysin loss, leading to a probable repression of *ccpA*-operon. The transport protein PrsA is required only in the presence of a cell wall (Wahlstroem, Vitikainen, Kontinen and Sarvas, 2003). SdhA probably reversibly catalyzes the conversion of succinate to fumarate, which is subsequently converted to acetyl-CoA to be integrated to fatty acid synthesis for cell wall and spore coat formation.

There is a link between TCA cycle function and sporulation. TCA cycle enzymes are maximally induced just before the onset of sporulation, and the absence of these activities results in a sporulation deficiency (Gao, Jiang, Pogliano and Aronson, 2002). PdhB catalyzes the conversion of pyruvate to acetyl-CoA, linking glycolysis and TCA cycle. The 2.8 log₂-fold upregulated protein neighbors ATP synthesis protein AtpA, fatty acid degradation enzyme AcdA, spore coat proteins CotF and CotQ, and succinate dehydrogenase SdhA. The presence of AcdA suggests a contribution to alanine, aspartate and glutamate metabolism for cell wall production and growth.

To sum up, these findings strongly emphasize that bacilysin is directly or indirectly associated with spore coat formation and protection, as well as swarming motility. Augmented protein expression in the absence of bacilysin during exponential phase proposes a coping mechanism for the cell to perform its vital functions properly. Enriched biological processes of 12th-hour secretome are given in Figure 4.7. Some of the crucial proteins in the constructed 12th-hour secretome network of OGU1 are listed in Table 4.3.

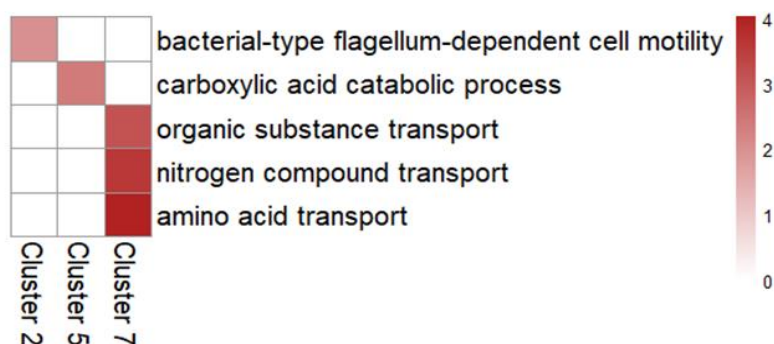


Figure 4.12. Enriched GO biological process terms for clusters in the 12th-hour secretome network of OGU1. Adjusted p-value cutoff was set to 0.05 for the enrichment results and the adjusted p-values were transformed into negative logarithm base 10.

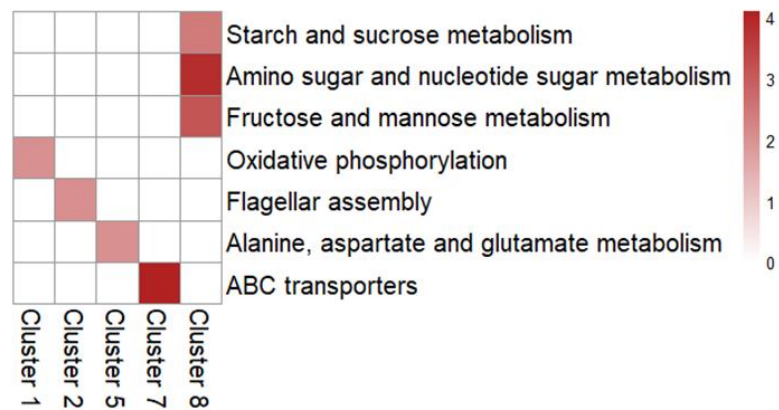


Figure 4.13. Enriched KEGG pathway terms for clusters in the 12th-hour secretome network of OGU1. Adjusted p-value cutoff was set to 0.05 for the enrichment results and the adjusted p-values were transformed into negative logarithm base 10.

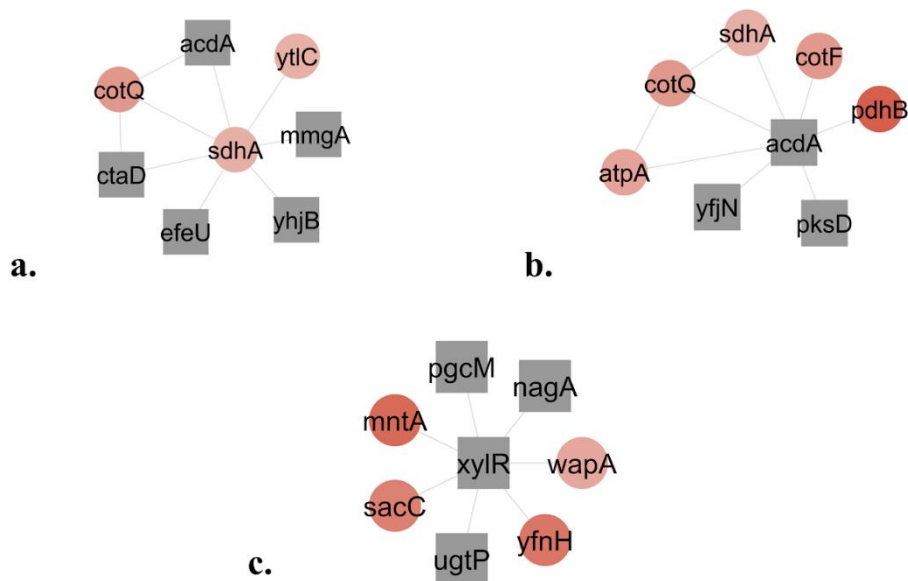


Figure 4.14. 12th-hour secretome case study: First degree neighbors of **a.** SdhA, **b.** AcdA and **c.** XylR based on their degree centrality measures. Proteins of primary metabolism such as SdhA are upregulated in the exponential phase of OGU1. MmgA, SacC, NagA and CtaD are repressed by CcpA under normal conditions. However, it is probable that bacilysin deficiency could block CcpA activity in the 12th-hour of growth in OGU1, leading to a compensation mechanism to prevent cell wall and spore coat stress by increasing the expression of proteins such as WapA, CotQ and YtlC. The presence of carbohydrate utilizing protein PgcM, probable glucose utilization protein YhjB (Werne, Seidel and Kuipers, 2015), early to mid-sporulation protein PksD (Fawcett et al., 2000) and iron transporter EfeU suggest a counteraction in the absence of bacilysin to sustain cellular functions.

Table 4.3. Probable effectors of cellular functions under bacilysin loss in the 12th-hour secretome network of OGU1.

Protein	Function	Operon/Regulon	Type of node in the network	Degree centrality	Betweenness centrality
SdhA	TCA cycle	SigA, FsrA	Proteomic	0.1167	0.2408
AcdA	Fatty acid degradation	FadR, SigA	Intermediate	0.1167	0.2101
Hag	Motility and chemotaxis	CodY, ScoC, SigD, CsrA	Proteomic	0.0667	0.1387
WapA	Cell wall synthesis	DegU, YvrHB, WalR, SigA	Proteomic	0.05	0.0865
GmuE	Glucomannan utilization	GmuR, CcpA, AbrB, SigA	Intermediate	0.0833	0.0748
NagA	N-acetylglucosamine utilization, cell wall turnover	NagR, CcpA, SigA	Intermediate	0.0333	0.0678
PksD	Polyketide synthesis	AbrB, CodY	Intermediate	0.0333	0.0499
MmgA	Cell wall recycling, utilization of cell wall components	CcpA, SigE, SigV	Intermediate	0.0667	0.0485
Sfp	4-phosphopantetheinyl transferase (surfactin synthetase-activating enzyme), required for production of surfactin, plipastatin and bacillaene	SrfA	Intermediate	0.05	0.0476
MetP	Methionine uptake	S-box, CodY	Intermediate	0.05	0.0399
GlnH	Glutamine uptake	TnrA, SigE	Proteomic	0.05	0.0371
ArtP	Arginine uptake	YlxR	Proteomic	0.05	0.0371
YckA	Unknown, similar to ABC transporter	YckBA	Intermediate	0.05	0.0317
OpuCA	Glycine transport, hyper-osmotic stress	OpcR, RemA, YlxR	Intermediate	0.05	0.024
OppA	Initiation of sporulation, competence development	TnrA, CodY, ScoC	Proteomic	0.0333	0.0186
PenP	Resistance to β -lactam antibiotics, cell wall stress resistance (Bucher et al., 2019)	SigA	Proteomic	0.0333	0.0116
OpuAC	Glycine transport, hyper-osmotic stress	RemA, SigA	Proteomic	0.0333	0.0102

4.4.2. Bacilysin deficiency causes a general stress response in the early stationary phase

As stated earlier, bacilysin levels peak at the 16th-hour (early stationary phase) of culture under normal conditions. It is thought that a considerable number of biological functions would be affected when bacilysin is absent from the growth medium due to its signaling property. Indeed, only two pathways were found to be significant to the 16th-hour secretome samples of OGU: Purine metabolism (adjusted p-value = 1.38×10^{-2}) and ABC transporters (adjusted p-value = 3.98×10^{-10}).

Purine metabolism in *B. subtilis* is interconnected with folate synthesis, PP pathway and amino acid synthesis. Purine nucleotides can be synthesized either *de novo* or through salvage pathways. DeoD is a purine salvage and interconversion protein which recycles free purines to nucleotides. During this process, hypoxanthine is produced as an intermediate molecule by PucABCDE. Hypoxanthine has been studied in relation to its mutagenic effect during sporulation in *B. subtilis* (Ayala-García et al., 2016). Additionally, excessive amount of hypoxanthine has been attributed to wrinkle formation in colonies as a result of oxidative stress. The oxidative stress is caused by the oxidation of hypoxanthine to uric acid, forming reactive oxygen species in a similar manner to eukaryotes (Gallegos-Monterrosa et al., 2017). Uric acid is further degraded to allantoin by uric oxidase PucL. First degree neighbors of PucL are YwmD and BglH. YwmD is a putative exported protein similar to D-amino acid dehydrogenases, with a 7.6 log₂-fold upregulation in OGU1. According to STRING database v.10.5, it interacts with a number of membrane proteins such as the zinc metalloprotease YwhC, putative phosphorelay sensor kinase YwpD, heme synthesis protein HemY, and many other transporters. Based on this information, upregulation of YwmD is also involved in oxidative phosphorylation and possible sensory activity, relaying the absence of bacilysin to different proteins including PucL for recovering the cell.

BglH is a major enzyme involved in β -MUG (4-methylumbelliferyl- β -d-glucoside) hydrolysis during spore germination and outgrowth. β -MUG is taken into the cell by phosphotransferases and hydrolyzed after phosphorylation (Setlow, Cabrera-Martinez and Setlow, 2004). Previous experiments have shown that the sporulation frequency of OGU1 and PY79 are similar (Tekin-İşlerel, 2017). It is then probable that BglH may contribute to compensating the effects of bacilysin loss along with BglC of the same operon during stationary phase. Bpr, which appeared alongside BglC in the 12th-hour secretome, is found to be upregulated by 2-fold in log₂ in the stationary phase. *bpr* encodes a serine-type extracellular protease to degrade proteins from the environment (Pohl et al., 2013; Veening et al., 2008). Bpr is one of the major exoproteases in *B. subtilis* which are activated by phosphorylated DegU (Marlow et al., 2014; Tsukahara and Ogura, 2008; Verhamme et al., 2007). Although not essential for growth or sporulation, the transcription of minor extracellular proteases is thoroughly regulated in *B. subtilis* by other different regulators such as AbrB, ScoC, SinR, and SpoIID. Therefore, the expression of extracellular proteases is induced during nutrient exhaustion in stationary phase as well as other starvation conditions (Barbieri et al 2015; Allenby et al., 2005). The two-component response regulator DegU was found to increase 2.7 fold in the cytosolic proteome of OGU1 (Aras-Taşkın et al, manuscript submitted). DegU is involved in the synthesis of degradative enzymes as well as competence in *B. subtilis* (Dahl et al., 1992). Mutations in the *degS* or *degU* genes cause a more stable phosphorylation of DegU, resulting in the hyperproduction of degradative enzymes (Antelmann et al., 2001). This suggests that after exponential phase, protein degradation increases in the absence of bacilysin, as it is sensed as a type of stress condition in which the organism overproduces extracellular enzymes to use all possible sources from the environment as a coping mechanism.

The decrease in cell wall and spore coat synthesis in the 16th-hour secretome can be explained by the presence of a number of proteins in the network. YrvJ is a cell wall organization protein which hydrolyzes peptidoglycans for cell wall elongation along with LytE. One of its interaction partners is DeoD, which was addressed above. The newly identified sporulation and germination protein YngK which is found to be 3.6- \log_2 fold upregulated in OGU1 (Nicolas et al., 2012) is also a neighbor of downregulated YrvJ and LytE. Another interaction partner, PgdS, is a hydrolase which breaks down gamma-polyglutamic acid (γ -PGA). γ -PGA is a naturally occurring polymer which is synthesized at the beginning of stationary phase due to nutrient starvation (Hsueh et al., 2017; Candela, Mock and Fouet, 2005). Although cell growth proteins are downregulated in the stationary phase, the increase in PgdS and YngK indicate a considerable defense against nutrient starvation in the absence of bacilysin.

As explained before, OppA is regulated by ScoC during exponential growth (Koide, Perego and Hoch, 1999), and involved in the import of competence and sporulation stimulating factor CSF (Zafra et al., 2012). In the 16th-hour secretome network of OGU1, OppA is neighbored by cyclic diAMP targets OpuCA and OpuAC, and phosphate transporter PstBA. During exponential phase, OpuAC and OpuCA interact with overexpressed OppA to keep the cell from breaking open during elongation. In the stationary phase, however, the expression of OppA decreases to a \log_2 -fold change of 1.21. One of the interaction partners of OppA is 2.5 \log_2 -fold upregulated YxaL, which is normally repressed by AbrB (Chumsakul et al., 2011; Banse et al., 2008). Likewise, CodY represses OppA. Elevated levels of both YxaL and OppA indirectly suggest the reciprocal action of bacilysin on its own regulator proteins.

TcyA, TcyB, YxeM and YxeN of cluster 5 are all involved in cysteine uptake, which is an essential process for efficient transformation (Chilton et al., 2017; Sysoeva et al., 2015) and intracellular zinc buffering (Eide, 2014). TcyA is also a major DAP importer for cell wall synthesis (Zhao, Roistacher and Helmann, 2018).

YxeM and YxeN may contribute to cell wall synthesis together with TcyA and TcyB. Downregulation of TcyA in the 16th-hour secretome results in less DAP uptake by the cell, thereby less cell wall production in OGU1. Crucial proteins with their first-degree interacting partners are given in Figure 4.15. A few of the important proteins targeted by bacilysin deficiency are given in Table 4.4.

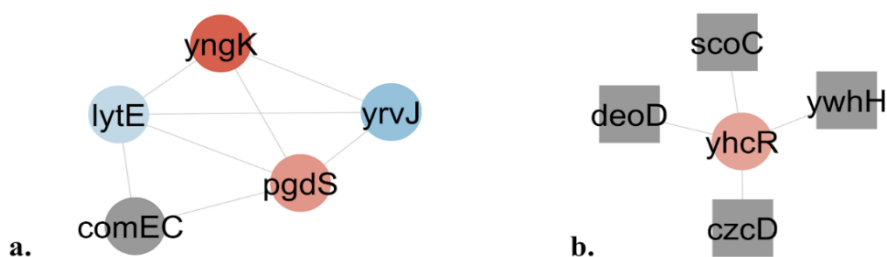


Figure 4.15. 16th-hour secretome case study: First degree neighbors of **a.** Pgds and **b.** YhcR based on their degree centrality measures. As stated earlier, YngK and Pgds are expressed in the beginning of stationary phase due to nutrient starvation. Pgds is also indirectly associated with flagella, in that flagella regulate poly- γ -glutamate gene expression (Diethmaier et al., 2017; Chan et al., 2014; Cairns et al., 2013). In the absence of bacilysin during stationary phase, cell wall organization proteins LytE and YrvJ are downregulated. This condition probably triggers the expression of motility and chemotaxis associated genes, which can be explained by the presence of competence protein ComEC in the subnetwork. Purine salvage protein DeoD is associated with similar sporulation frequencies of OGU1 and its parental strain due to overproduction of nutrient degrading extracellular enzymes, as described above. Its connections to gene regulators ScoC and YwhH support the compensatory effect of DeoD in bacilysin-deficiency during early stationary growth. CzcD, a cation transport protein, probably allows quality control during spore formation by the accumulation of cations on the outer membrane of forespores in a varying manner in OGU1 (Sirec et al., 2019).

Table 4.4. Probable effectors of cellular functions under bacilysin loss in the 16th-hour secretome network of OGU1.

Protein	Function	Operon/Regulon	Type of node in the network	Degree centrality	Betweenness centrality
PgdS	Polyglutamic acid degradation	SigD	Proteomic	0.1143	0.0782
YhcR	Utilization of nucleic acids, RNase	YhcR-YhcS	Proteomic	0.1143	0.3996
DeoD	Purine salvage and interconversion	DeoD-LdcB	Intermediate	0.1143	0.5841
TcyB	Cystine and diaminopimelate uptake	TcyABC	Intermediate	0.0857	0.1271
GlnH	Glutamine uptake	TnrA, SigE	Proteomic	0.0857	0.0851
YvgM	Unknown, molybdenum transport	YvgLM	Intermediate	0.0857	0.1289
AbrB	Regulation of gene expression during the transition from growth to stationary phase	Spo0A, AbrB, SigA	Intermediate	0.0571	0.0571
YxeM	Uptake and utilization of S-(2-succino)cysteine	CymR, SigA	Intermediate	0.0571	0.1427
YxeN	Uptake and utilization of S-(2-succino)cysteine	CymR, SigA	Intermediate	0.0571	0.0085
ScoC	Transition state regulator	AbrB, SenS, SalA, CodY, SigA, ScoC	Intermediate	0.0571	0.1804
PstBA	High-affinity phosphate uptake	SigA, PhoP	Intermediate	0.0571	0.0291
BglC	β -1,4-glucan degradation	SigA	Intermediate	0.0571	0.0172
BglH	β -glucoside utilization	SigA, CcpA, LigT	Intermediate	0.0571	0.0253
OpuAC	Glycine transport, coping with hyper-osmotic stress	RemA, SigA	Intermediate	0.0571	0.0333

4.4.3. Bacilysin regulates biofilm formation and respiration during stationary phase

B. subtilis becomes fully differentiated and stops growing at the end of stationary phase, in which it regulates its own metabolism under nutrient starvation. Glycine, serine and threonine metabolism becomes more prominent during 24th-hour (adjusted p-value = 9.66×10^{-4}). The phosphoglycerate dehydrogenase protein SerA was shown to be upregulated during the transition from early stationary to stationary phase. Yet, SerA expression in the 24th-hour is downregulated when compared to cytosolic proteome. This is presumably due to the difference in protein extraction and detection methods (Tekin-İşlerel, 2017; Demir, 2013; Aras-Taşkın, 2010). One of SerA's interaction partners is GcvT, which degrades glycine to produce serine by using THF as a cofactor. GcvT is also present in the cytosolic proteome network of OGU1. As stated earlier, the expression of SerA decreases while transiting from early stationary to stationary phase in OGU1. Serine levels decrease in stationary phase under optimum conditions, triggering biofilm formation as explained earlier. This may have caused a flux towards GcvT for serine synthesis, thereby sustaining attenuated biofilm production in OGU1. Alternatively, glycine production may have been triggered to maintain stationary growth. Other direct partners of SerA are two-component histidine kinase protein YhcY, NADH dehydrogenase Ndh, and surfactin synthesis protein SrfAC. YhcY is regulated by the LiaR regulon, which is associated with cell envelope oxidative stress (Wolf et al., 2010). YhcY was also shown to regulate growth by phosphorylating its regulator counterpart, YhcZ in *B. thuringiensis* (Jia et al., 2018). *srfAC* encodes surfactin synthetase subunit 3 which takes role in the surfactin biosynthesis (Tanovic et al., 2008). In *B. subtilis*, surfactin is a lipopeptide which is involved in the extracellular matrix production for biofilm formation via paracrine signaling (Marvasi et al., 2010). The surfactin biosynthetic operon has also role in bacilysin production since bacilysin biosynthesis is also blocked in a *srfA*-blocked mutant (Yazgan Karataş et al., 2003). In OGU1, SrfAC has been shown to be downregulated by 1.6 fold in log₂ during the 24th-hour of incubation. The NADH

dehydrogenase Ndh is present in both cytosolic proteome and 24th-hour secretome with a decreased expression from early stationary to stationary phase. Taken together, YhcY and YhcZ may have sensed the lack of bacilysin and stimulated energy conservation and biofilm formation, albeit in a diminished manner.

Phenylalanine, tyrosine and tryptophan biosynthesis (adjusted p-value = 4.83×10^{-3}) and siderophore group nonribosomal peptide synthesis (adjusted p-value = 1.44×10^{-3}) are associated with each other, in that these amino acids are incorporated into metal binding nonribosomal peptides by conversion of chorismate into the intermediate molecule isochorismate by MenF to produce menaquinone (vitamin K). During phenylalanine metabolism, chorismate mutase ThrR reversibly converts chorismate to prephenate, which is also a precursor of bacilysin (Hilton et al., 1988).

In *B. subtilis*, menaquinone synthesis is membrane associated and implicated in oxidative phosphorylation (Farrand and Taber, 1974). Reduced form of menaquinone exhibits antioxidant properties and can play a role in protecting cellular membranes from lipid oxidation (Nowicka and Kruk, 2010). Additionally, menaquinone is shown to be involved in the active transport of molecules across the cell membrane and in sporulation (Hojo et al., 2007; Farrand and Taber, 1974). Menaquinone concentration also determines aeration of proliferating biofilm structures (Mahdinia, Demirci and Berenjian, 2017; Ikeda and Doi, 1990). EfeB is an iron binding secreted protein, which is immediately induced upon iron starvation (Pi and Helmann, 2017; Baichoo et al., 2002). Similarly, the expression of isochorismatase Dhbb is strongly enhanced at the beginning of biofilm formation under iron deficiency (Pisithkul et al., 2019). Iron is essential for biofilm formation because it regulates surface motility and stabilizes the polysaccharide matrix (Ahmed and Holmström, 2014; Chhibber et al., 2013; Weinberg, 2004). EfeB is 2 log₂-fold upregulated in OGU1. In light of the information mentioned above, we can deduce that bacilysin deficiency induces iron uptake during stationary phase and promotes biofilm formation.

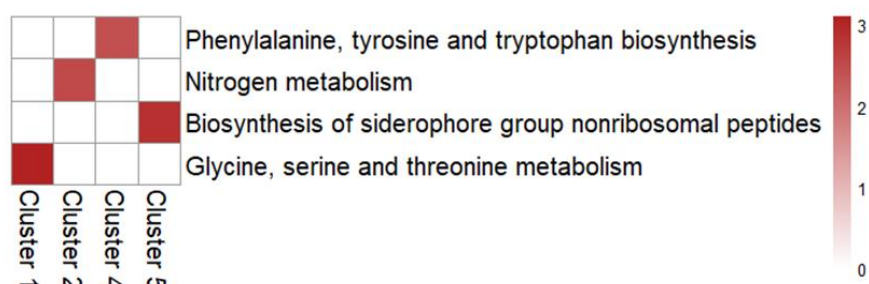


Figure 4.16. Enriched GO biological process terms for clusters in the 24th-hour secretome network of OGU1. Adjusted p-value cutoff was set to 0.05 for the enrichment results and the adjusted p-values were transformed into negative logarithm base 10.

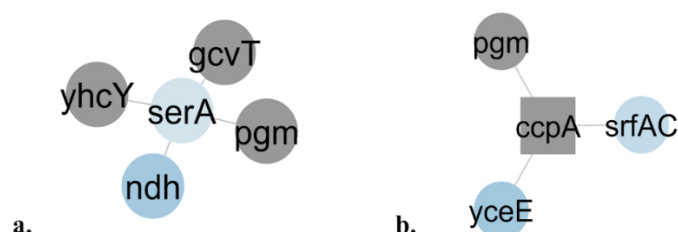


Figure 4.17. 24th-hour secretome case study: First degree neighbors of **a.** SerA and **b.** CcpA based on their degree centrality measures. SerA, Ndh, CcpA, YceE and GcvT are also present in the cytosolic proteome network. SerA, Ndh and YceE are downregulated in the stationary phase compared to 16th-hour. SrfAC stimulates biofilm formation under normal conditions, yet *bacA* disruption reduces its expression, as well. Complete differentiation of cells in the 24th-hour culture elicits maintenance of biofilm formation and energy, regardless of bacilysin production.

Table 4.5. Probable effectors of cellular functions under bacilysin loss in the 24th-hour secretome network of OGU1.

Protein	Function	Operon/Regulon	Type of node in the network	Degree centrality	Betweenness centrality
SerA	Biosynthesis of serine	CcpA	Proteomic	0.2353	0.5331
CcpA	Transcriptional regulator, carbon catabolite repressor	SigA, CcpA	Intermediate	0.1765	0.2426
PheA	Biosynthesis of phenylalanine	SigA	Intermediate	0.1765	0.2279
GcvT	Glycine utilization	Gly-box	Intermediate	0.1176	0.2353
SrfAC	Antibiotic synthesis/competence	PhoP, CodY, ComA, PerR, Spx, Abh, SigA	Proteomic	0.1765	0.2353
YhcY	Unknown/two component sensor histidine kinase	LiaR, SigA	Intermediate	0.1176	0.2206
PksD	Polyketide synthesis	AbrB, CodY	Intermediate	0.1176	0.0846
NarG	Nitrate respiration, nitrogen assimilation	Fnr, SigA	Proteomic	0.1176	0.0956
YrpB	Unknown/nitronate dehydrogenase	YrpB	Intermediate	0.1176	0.0699

CHAPTER 5

CONCLUSION

- Finding the interactions, regulations and mechanisms within a group of differentially expressed proteins or genes have become prominent. Combining multiple ‘omic’ hits in a biologically meaningful way helps us understand the molecular pathways and cellular mechanisms that are active during a perturbation. For this aim, a variety of network modeling approaches which accurately uncovers crucial pathways and molecules have been developed.
- In this work, we have applied the Prize-collecting Steiner Forest (PCSF) algorithm in an attempt to visualize and interpret altered protein interactions in *in bacABCDE*-deleted OGU1 strain. PCSF algorithm can help us infer the relations between experimental hits by integrating hidden proteins, providing a better explanation on disconcerted cellular processes.
- To model the interactions in OGU1, the interactome of *B. subtilis* 168 was selected, as there were no data available for OGU1’s parental strain, *B. subtilis* PY79. Nevertheless, PY79 itself is a prototrophic derivative of *B. subtilis* 168.
- 4 experimental datasets with differential expression values were organized for analysis: one for the cytosolic proteome network of OGU1, and three from dynamic secretome analyses at 12, 16 and 24 hours of bacterial culture. This helped us evaluate time-dependent changes in signaling at exponential and stationary growth phases, as well as determine the roles of proteins with previously unknown functions.

- Cytosolic proteome network of OGU1 consists of 243 nodes and 1182 edges, 12th-hour secretome network comprises 61 nodes and 87 edges, 16th-hour secretome network includes 36 nodes and 45 edges, and 24th-hour secretome network consists of 18 nodes and 19 edges.
- To explain the modified processes and pathways for each condition, networks were clustered by Louvain method and analyzed with ClueGO v. 2.5.4. This resulted in 10 clusters for cytosolic proteome, 8 clusters for 12th-hour secretome and 6 clusters for 16th-hour secretome.
- KEGG pathway terms and GO terms (biological process, cellular component and molecular function) were analyzed in terms of their enrichment status. We could accurately recall various pathways such as quorum sensing, oxidative phosphorylation, ABC transporters and secondary metabolite biosynthesis. We also identified additional processes such as amino acid transport and carboxylic acid catabolism that are indirectly linked to motility, regulation of cell shape and cell wall, germination and competence.
- The presence of CcpA and its related proteins in the cytosolic proteome network have indicated that carbon catabolite repression is an essential process in all constructed networks. It can be reinforced that bacilysin possibly acts as a signaling molecule which prompts transcription factors such as CcpA to maintain cellular activities.

- We could also detect the presence of sulfur metabolism and sporulation associated proteins SsuC and SsuD, stress resistance proteins HxlA, PlsX, TrmB, YknX and HtpG, stringent response protein ScpA, and probable RNase YpdQ in the same network. These could be targeted in further research regarding bacilysin and its effect on cellular stress.
- As previous research by our group has demonstrated, molecular transport is a highly affected process in both exponential and early stationary growth. Proteins such as EfeU, YhjB, UgtP and XylR indicate that in 12th-hour secretome, bacilysin deletion possibly causes the cell to uptake essential molecules to a greater extent.
- Former research by our group has demonstrated that the loss of bacilysin has a more profound effect at the 16th-hour of cell culture. Since bacilysin levels peak at this time, the lack of the dipeptide causes the cell to seriously alter its functions. Indeed, protein levels at 16-hour samples are found to be the lowest. We could analyze the course of relations between extracellular, cell wall and cell membrane proteins at the stationary growth phase of OGU1.
- Membrane transport, peptide degradation, cell wall synthesis, competence and sporulation efficiency were among modified cellular activities in OGU1. Existence of transcriptional regulators AbrB and ScoC in the 16th-hour secretome network strongly and repeatedly suggests that bacilysin acts upon several transcriptional elements as a signaling molecule to regulate multiple functions.

- Our research has shown that this kind of integrated network modeling approach can help us understand and evaluate perturbations at protein level in a prokaryotic organism under a particular condition. We could also estimate and explain the roles of several proteins which were previously considered unknown. Recent biological network studies on *Bacillus subtilis* have focused on gene regulations. In this thesis, we have attempted to represent different metabolic activities or molecular changes in a *bac* operon-deleted derivative of *Bacillus subtilis* at the protein level.
- We aim to improve and expand this study by including transcriptional data. In addition, we believe that this study can guide further experiments such as genetic or chemical complementation and induction studies, which can enhance our knowledge on our organism of interest.

REFERENCES

- Allenby, N.E., O'Connor, N., Prágai, Z., Ward, A.C., Wipat, A., and Harwood, C.R. (2005). Genome-wide transcriptional analysis of the phosphate starvation stimulon of *Bacillus subtilis*. *J Bacteriol*, 187 (23), 8063-8080.
- Antelmann, H., Tjalsma, H., Voigt, B., Ohlmeier, S., Bron, S., van Dijl, J.M., and Hecker, M. (2001). A proteomic view on genome-based signal peptide predictions. *Genome Res*, 11 (9), 1484-1502.
- Aras-Taşkın, A. (2010). Proteome-wide Analysis of the Functional Roles of Bacilysin Biosynthesis in *Bacillus subtilis* (Master's thesis). Retrieved from <http://etd.lib.metu.edu.tr/upload/12612409/index.pdf>
- Baggerman, G., Vierstraete, E., De Loof, A., and Schoofs, L. (2005). Gel-based versus gel-free proteomics: a review. *Comb Chem High T Scr*, 8 (8), 669–677.
- Barbe, V., Cruveiller, S., Kunst, F., Lenoble, P., Meurice, G., Sekowska, A. and Danchin, A. (2009). From a consortium sequence to a unified sequence: the *Bacillus subtilis* 168 reference genome a decade later. *Microbiology*, 155 (6), 1758-1775.
- Barbieri, G., Voigt, B., Albrecht, D., Hecker, M., Albertini, A.M., Sonenshein, A.L., and Belitsky, B.R. (2015). CodY regulates expression of the *Bacillus subtilis* extracellular proteases Vpr and Mpr. *J Bacteriol*, 197 (8), 1423-1432.

Beissbarth, T., Hyde, L., Smyth, G.K., Job, C., Boon, W.M., Tan, S.S., Scott, H.S. et al. (2004). Statistical modeling of sequencing errors in SAGE libraries. *Bioinformatics*, 20, 31–9.

Bernhardt, J., Völker, U., Völker, A., Antelmann, H., Schmid, R., Mach, H., Hecker, M. et al. (1996). Specific and general stress proteins in *B. subtilis* -a two-dimensional electrophoretic study. *Microbiology*, 143, 999–1017.

Bertacine Dias, M.V., Santos, J.C., Libreros-Zúñiga, G.A., Ribeiro, J.A. and Chavez-Pacheco, S.M. (2018). Folate biosynthesis pathway: mechanisms and insights into drug design for infectious diseases. *Future Medicinal Chemistry*, 10 (8), 935-959.

Blondel, V.D., Guillaume, J.L., Lambiotte, R. and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, (10), P10008.

Brandes, U. (2001). A Faster Algorithm for Betweenness Centrality. *Journal of Mathematical Sociology*, 25(2),:163-177. Taken from <http://www.inf.uni-konstanz.de/algo/publications/b-fabc-01.pdf>

Bucher, T., Keren-Paz, A., Hausser, J., Olender, T., Cytryn, E. and Kolodkin-Gal, I. (2019). An active β -lactamase is a part of an orchestrated cell wall stress resistance network of *Bacillus subtilis* and related rhizosphere species. *Environmental Microbiology*, 21 (3): 1068-1085.

Candela T., Mock M., Fouet A. (2005). CapE, a 47-amino-acid peptide, is necessary for *Bacillus anthracis* polyglutamate capsule synthesis. *J. Bacteriol*, 187, 7765–7772.

Cao, H., Villatoro-Hernandez, J., Weme, R.D.O., Frenzel E. and Kuipers, O.P. (2018). Boosting heterologous protein production yield by adjusting global nitrogen and carbon metabolic regulatory networks in *Bacillus subtilis*. *Metabolic Engineering*. DOI: <https://doi.org/10.1016/j.ymben.2018.08.001>

Chilton,, S.S., Fabel ,T.G., Hromada, S. and Burton, B.M. (2017). A Conserved Metal Binding Motif in the *Bacillus subtilis* Competence Protein ComFA Enhances Transformation. *J Bacteriol*, 199 (15), e00272-17.

Chmara, H., Milewski, S., Dzieduszycka, M., Smulkowski, M., Sawlewicz, P. and Borowski, E. (1982). Epoxypeptides- a novel group of metabolic inhibitors in prokaryotic and eukaryotic organisms. *Drugs under Experimental and Clinical Research*, 8, 11-12.

Chmara, H. (1985). Inhibition of glucosamine synthase by bacilysin and anticapsin. *Microbiology*, 131 (2), 265-271.

Choi, S.-K. and Saier Jr, M.H. (2005). Regulation of *sigL* Expression by the Catabolite Control Protein CcpA Involves a Roadblock Mechanism in *Bacillus subtilis*: Potential Connection between Carbon and Nitrogen Metabolism. *J Bacteriol*, 187 (19): 6856–6861.

Chumsakul, O., Takahashi, H., Oshima, T., Hishimoto, T., Kanaya, S., Ogasawara, N. and Ishikawa, S. (2011). Genome-wide binding profiles of the *Bacillus subtilis* transition state regulator AbrB and its homolog Abh reveals their interactive role in transcriptional regulation. *Nucleic Acids Res*, 39 (2): 414-428.

Claverys, J.P., Prudhomme, M. and Martin, M. (2006). Induction of Competence Regulons as a General Response to Stress in Gram-Positive Bacteria. *Annu Rev Microbiol*, 60: 451–75

Clements, L.D., Streips, U.N. and Miller, B.S. (2002). Differential proteomic analysis of *Bacillus subtilis* nitrate respiration and fermentation in defined medium. *Proteomics*, 2, 1724-1734.

Collins, J.A., Irnov, I., Baker, S. and Winkler, W.C. (2007). Mechanism of mRNA destabilization by the *glmS* ribozyme. *Genes Dev*, 21 (24): 3356-68.

Csardi, G. and Nepusz, T. (2006). The igraph software package for complex network research, *InterJournal, Complex Systems* 1695. URL: <http://igraph.org>

Cui, W., Han, L., Suol, F., Liu, Z., Zhou, L. and Zhou, Z. (2018). Exploitation of *Bacillus subtilis* as a robust workhorse for production of heterologous proteins and beyond. *World Journal of Microbiology and Biotechnology*, 34, 145.

Dahl MK, Msadek T, Kunst F, Rapoport G. (1992). The phosphorylation state of the DegU response regulator acts as a molecular switch allowing either degradative enzyme synthesis or expression of genetic competence in *Bacillus subtilis*. *J Biol Chem*, 267 (20), 14509-14.

Diethmaier, C., Chawla, R., Canzoneri, A., Kearns, D.B., Lele, P.P and Dubnau, D. (2017). Viscous drag on the flagellum activates *Bacillus subtilis* entry into the K-state. *Mol Microbiol*, 106 (3): 367–380.

Demir, M. (2013). Proteome-Wide Analysis of the Role of Expression of Bacilysin Operon on Idiophase Physiology of *B. subtilis* (Master's thesis). Retrieved from <http://etd.lib.metu.edu.tr/upload/12615570/index.pdf>

Eide, D.J. (2014). Bacillithiol, a new role in buffering intracellular zinc. *Mol Microbiol*, 94(4), 743–746.

Elsholz, A.K., Turgay, K., Michalik, S., Hessling, B., Gronau, K., Oertel, D., Mäder, U., Bernhardt, J., Becher, D., Hecker, M. and Gerth, U. (2012). Global impact of

protein arginine phosphorylation on the physiology of *Bacillus subtilis*. *Proc Natl Acad Sci USA*, 109 (19), 7451-7456.

Errington, J. (2010). From spores to antibiotics via the cell cycle. *Microbiology*, 156, 1-3. DOI: 10.1099/mic.0.035634-0.

Feucht, A., Evans, L. and Errington, J. (2003). Identification of sporulation genes by genome-wide analysis of the sigma-E regulon of *Bacillus subtilis*. *Microbiology*, 149 (10), 3023-34.

Gilhooly, N.S., Carrasco, C., Gollnick, B., Wilkinson, M., Wigley, D.B., Moreno-Herrero, F. and Dillingham, M.S. (2016). Chi hotspots trigger a conformational change in the helicase-like domain of AddAB to activate homologous recombination. *Nucleic Acids Res*, 44 (6), 2727-2741.

Hagberg, Schult and Swart (2008). “Exploring network structure, dynamics, and function using NetworkX”, in Proceedings of the 7th Python in Science Conference (SciPy 2008), Gäel Varoquaux, Travis Vaught, and Jarrod Millman (Eds), (Pasadena, CA USA), pp. 11–15, Aug 2008.

Harwood, C.R. and Cutting, S.M. (1990). Molecular Biological Methods for *Bacillus* (pp. 1-3). A Wiley- Interscience Publication.

Hilton, M.D., Alaeddinoğlu, N.G. and Demain, A.L. (1988). Synthesis of bacilysin by *Bacillus subtilis* branches from prephenate of the aromatic amino acid pathway. *J Bacteriol*, 170 (1), 482-484.

Hoffmann, T., Troup, B., Szabo, A., Hungerer, C. and Jahn, D. (1995). The anaerobic life of *Bacillus subtilis*: cloning of the genes encoding the respiratory nitrate reductase system. *FEMS Microbiol Lett* 131, 219-225.

Hsueh, Y.-H., Huang, K.-Y., Kunene S.C. and Lee, T.Y. (2017). Poly- γ -glutamic Acid Synthesis, Gene Regulation, Phylogenetic Relationships, and Role in Fermentation. *Int J Mol Sci* 18 (12): 2644.

Hullo, M.F., Auger, S., Dassa, E., Danchin, A. and Martin-Verstraete, I. (2004). The *metNPQ* operon of *Bacillus subtilis* encodes an ABC permease transporting methionine sulfoxide, D- and L-methionine. *Research in Microbiology*, 155 (2004), 80-86.

Inaoka, T., Takahashi, K., Ohnishi-Kameyama, M., Yoshida, M. and Ochi, K. (2003). Guanine nucleotides guanosine 5'-diphosphate 3'-diphosphate and GTP co-operatively regulate the production of an antibiotic bacilysin in *Bacillus subtilis*. *Journal of Biological Chemistry*, 278 (4), 2169-2176.

Irving, S.E. and Corrigan, R.M. (2018). Triggering the stringent response: signals responsible for activating (p)ppGpp synthesis in bacteria. *Microbiology*, 164 (3), 268-276.

Jia, L.D., Gao, T.T., Peng, Q., Lv, J., Zhang, J., Chen, M. and Song, F.P. (2018). The functions of *yhcZ* gene during *Bacillus thuringiensis* growth (Article in Chinese). *Yi Chuan*, 40(5), 415-424. DOI: 10.16288/j.ycz.18-061.

Karataş, A.Y., Çetin, S. and Özcengiz, G. (2003). The effects of insertional mutations in *comQ*, *comP*, *srfA*, *spo0H*, *spo0A* and *abrB* genes on bacilysin biosynthesis in *Bacillus subtilis*. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* 1626 (1), 51-56.

Kenig, M. and Abraham, E.P. (1976). Antimicrobial activities and antagonists of bacilysin and anticapsin. *Microbiology*, 94 (1), 37-45.

Kenig, M., Vandamme, E., and Abraham, E.P. (1976). The mode of action of bacilysin and anticapsin and biochemical properties of bacilysin-resistant mutants. *Microbiology*, 94 (1), 46-54.

Keskin, Ö., Tunçbağ, N. and Gürsoy, A. (2016). Predicting Protein–Protein Interactions from the Molecular to the Proteome Level. *Chemical Reviews*, 116 (8), 4884-4909.

Kim, Y. and Jang, S. (2018). Mapping the knowledge structure of frailty in journal articles by text network analysis. *PLoS One*, 13 (4): e0196104.

Köroğlu, T., Ögülür, İ., Mutlu, S., Yazgan-Karataş, A. and Özcengiz, G. (2011). Global regulatory systems operating in bacilysin biosynthesis in *Bacillus subtilis*. *Journal of molecular microbiology and biotechnology*, 20 (3): 144–155.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G. and Azevedo, V. (1997). The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature*, 390(6657): 249-256.

Lippolis, R., Siciliano, R., Mazzeo, M., Abbrescia, A., Gnoni, A., Sardanelli, A. and Papa, S. (2013). Comparative secretome analysis of four isogenic *Bacillus clausii* probiotic strains. *Proteome Science*, 11 (1): 28.

Ma, W., Liu, Y., Shin, H., Li, J., Chen, J., Du, G. and Liu, L. (2017). Metabolic engineering of carbon overflow metabolism of *Bacillus subtilis* for improved N-acetyl-glucosamine production. *Bioresource Technology*. DOI: 10.1016/j.biortech.2017.10.007

Margot, P. and Karamata, D. (1996). The *wprA* gene of *Bacillus subtilis* 168, expressed during exponential growth, encodes a cell-wall-associated protease. *Microbiology*, 142 (12), 3437-3444.

Mariappan, A., Makarewicz, O., Chen, X.H. and Borriss, R. (2012). Two component response regulator DegU controls the expression of bacilysin in plant-growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. *Journal of Molecular Microbiology and Biotechnology*, 22 (2), 114-125.

Marvasi, M., Visscher, P. T., and Casillas Martinez, L. (2010). Exopolymeric substances (EPS) from *Bacillus subtilis*: polymers and genes encoding their synthesis. *FEMS Microbiology Letters*, 313(1), 1-9.

Mashaghi, A., Ramezanpour, A. and Karimipour, V. (2004). Investigation of a protein complex network. *European Physical Journal, B* 41: 113–121.

Milewski, S. (1993). Chemical modification studies of the active site of glucosamine-6 phosphate synthase from baker's yeast. *Biochim. Biophys. Acta* 1161: 279-284.

Miller, M.B. and Bassler, B.L. (2001). Quorum sensing in bacteria. *Annu Rev Microbiol*, 55, 165-199.

Mlecnik, B., Galon, J. and Bindea, G. (2019). Automated exploration of gene ontology term and pathway networks with ClueGO-REST. *Bioinformatics*, 1-3. DOI: 10.1093/bioinformatics/btz163.

Nakano, M.M. and Zuber, P. (1990). Molecular biology of antibiotic production in *Bacillus*. *Crit Rev Biotechnol* 10, 223-240.

Naryzhny, S. (2016). Towards the Full Realization of 2DE Power. *Proteomes*, 4 (4), 33. DOI: 10.3390/proteomes4040033

Nicolas, P., Mäder, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N. and Becher, D. (2012). Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science*, 335 (6072), 1103-1106.

Noirot-Gros, M.F., Soultanas, P., Wigley, D., Ehrlich, S., Noirot, P., and Petit, M.A. (2002). The beta-propeller protein YxaL increases the processivity of the PcrA helicase. *Molecular Genetics and Genomics*, 267 (3), 391-400.

Ogura, M., Shimane, K., Asai, K., Ogasawara, N., and Tanaka, T. (2003). Binding of response regulator DegU to the *aprE* promoter is inhibited by RapG, which is counteracted by extracellular PhrG in *Bacillus subtilis*. *Molecular Microbiology*, 49(6), 1685-1697

Old, W.M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K.G., Mendoza, A., Sevinsky, J.R., Resing, K. et al. (2005). Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol Cell Proteomics*, 4 (10), 1487–502.

Ögölür, İ. (2008). The effects of twelve quorum-sensing gene products on the expression of *bacABCDE* operon in *Bacillus subtilis* (Master's thesis). Retrieved from <http://etd.lib.metu.edu.tr/upload/12609983/index.pdf>

Özcengiz, G. and Alaeddinoğlu, N.G. (1991). Bacilysin production by *Bacillus subtilis*: effects of bacilysin, pH and temperature. *Folia microbiologica*, 36 (6), 522-526.

Özcengiz, G. and Ögüller, İ. (2015). Biochemistry, genetics and regulation of bacilysin biosynthesis and its significance more than an antibiotic. *New Biotechnology*, 32 (6), 612-619.

Özcengiz, G., Alaeddinoğlu, N.G. and Demain, A.L. (1990). Regulation of biosynthesis of bacilysin by *Bacillus subtilis*. *J Ind Microbiol*, 6, 91-100.

Parker, J.B. and Walsh, C.T. (2012). Stereochemical Outcome at Four Stereogenic Centers During Conversion of Prephenate to Tetrahydrotyrosine by BacABGF in the Bacilysin Pathway. *Biochemistry* 51 (28), 5622-5632.

Plotly Technologies Inc. (2015). Collaborative data science. Plotly Technologies Inc. Montréal, QC. URL: <https://plot.ly>.

Pohl, S., Bhavsar, G., Hulme, J., Bloor, A.E., Misirli, G., Leckenby, M.W. and Harwood, C.R. (2013). Proteomic analysis of *Bacillus subtilis* strains engineered for improved production of heterologous proteins. *Proteomics*, 13 (22), 3298-3308.

Rutherford, S.T. and Bassler, B.L. (2012). Bacterial quorum sensing: its role in virulence and possibilities for its control. *Cold Spring Harb Perspect Med*, 2 (11).

Sakajoh, M., Solomon, N.A. and Demain, A.L. (1987). Cell-free synthesis of the dipeptide antibiotic bacilysin. *Journal of Industrial Microbiology & Biotechnology*, 2 (4), 201-208.

Schnell, N., Entan, K.D., Schneider, U., Gotz, F., Zahner, H., Kellner, R. and Jung, G. (1988). Pre-peptide sequence of epidermine, a ribosomally synthesized antibiotic with four sulphide-rings. *Nature*, 133, 276-278.

Setlow, B., Cabrera-Martinez, R.M. and Setlow, P. (2004). Mechanism of the hydrolysis of 4-methylumbelliferyl- β -d-glucoside by germinating and outgrowing spores of *Bacillus* species. *J Appl Microbiol*, 96 (6), 1245-1255.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* 13 (11), 2498-2504.

Shao, H., Mesbahi, M., Li, D. and Xi, Y. (2017). Inferring Centrality from Network Snapshots. *Scientific Reports*, 7, 40642.

Solomon, J.M., Lazazzera, B.A. and Grossman, A.D. (1996). "Purification and characterization of an extracellular peptide factor that affects two different developmental pathways in *Bacillus subtilis*". *Genes Dev*, 10, 2014–2024.

Sonenshein, A.L., Hoch, J.A. and Losick, R. (2002). *Bacillus subtilis*: from cells to genes and genes to cells. In AL Sonenshein, JA Hoch & R Losick, *Bacillus subtilis and its closest relatives: from genes to cells* (pp.3–5). Washington DC: American Society for Microbiology Press.

Spizen, J. (1958). Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. *Proc Natl Acad Sci USA*, 44, 1072-1078.

Stein, T. (2005). *Bacillus subtilis* antibiotics: structures, syntheses, and specific functions. *Mol Microbiol*, 56 (4), 845-857.

Steinborn, G., Hajirezaei, M.R. and Hofemeister, J. (2005). *bac* genes for recombinant bacilysin and anticapsin production in *Bacillus* host strains. *Archives of microbiology*, 183 (2), 71-79.

Stragier, P. and Losick, R. (1996). Molecular genetics of sporulation in *Bacillus subtilis*. *Annu Rev Genet* 30, 297-341.

Sugae, K. and Freese, E. (1970). Requirement for Acetate and Glycine (or Serine) for Sporulation Without Growth of *Bacillus subtilis*. *J Bacteriol* 104 (3), 1074-1085.

Sun, G., Birkey, S.M. and Hulett, F.M. (1996). Three two-component signal transduction systems interact for Pho regulation in *Bacillus subtilis*. *Mol Microbiol* 19, 941-948.

Sysoeva, T.A., Bane, L.B., Xiao, D.Y., Bose, B., Chilton, S.S., Gaudette, R. and Burton, B.M. (2015). Structural characterization of the late competence protein ComFB from *Bacillus subtilis*. *Biosci Rep*, 35 (2), e00183.

Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P., Jensen, L.J., von Mering, C. (2017). The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Research*, 45 (D1): D362-D368.

Tanovic, A., Samel, S.A., Essen, L.O., and Marahiel, M.A. (2008). Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science*, 321 (5889), 659-663.

Tekin-İşlerel, E. (2017). Dynamic Analysis of Secretome Alterations in a Bacilysin Knock-out Mutant of *Bacillus subtilis* (Doctoral dissertation). Retrieved from <http://etd.lib.metu.edu.tr/upload/12621675/index.pdf>

Tjalsma, H., Antelmann, H., Jongbloed, J.D.H., Braun, P.G., Darmon, E., Dorenbos, R., Dubois, J.Y.F., Westers, H., Zanen, G., Quax, W.J., Kuipers, O.P., Bron, S., Hecker, M. and van Dijl, J.M. (2004). Proteomics of Protein Secretion by *Bacillus subtilis*: Separating the “Secrets” of the Secretome. *Microbiol Mol Biol Rev*, 68 (2): 207-233.

Tschen, M. (1990). Gegen Pilze wirksame Antibiotika der *Bacillus subtilis*-Gruppe (Antifungal antibiotics of the *Bacillus subtilis*-group). *Forum Mikrobiology*, 3(90):156-163.

Tsukahara, K. and Ogura, M. (2008). Characterization of DegU-dependent expression of *bpr* in *Bacillus subtilis*. *FEMS microbiology letters*, 280 (1): 8-13.

Tunçbağ, N., Gosline, S.J.C., Kedaigle, A., Soltis, A.R., Gitter, A. and Fraenkel, E. (2016). Network-Based Interpretation of Diverse High-Throughput Datasets through the Omics Integrator Software Package. *PLoS Comput Biol* 12(4): e1004879.

Tunçbağ, N., Milani, P., Pokorný, J.L., Johnson, H., Sio, T.T., Dalin, S., Iyekegbe, D.O., White, F., Sarkaria, J.N. and Fraenkel, E. (2016). Network Modeling Identifies Patient-specific Pathways in Glioblastoma, *Scientific Reports*, 6: 28668.

Veening, J.W., Hamoen, L.W. and Kuipers, O.P. (2005). Phosphatases modulate the bistable sporulation expression pattern in *Bacillus subtilis*. *Molecular Microbiology*, 56 (6), 1481-1494.

Verhamme, D.T., Kiley, T.B., and Stanley-Wall, N.R. (2007). DegU coordinates multicellular behaviour exhibited by *Bacillus subtilis*. *Molecular Microbiology*, 65 (2), 554-568

Walker, J.E. and Abraham, E.P. (1970). The structure of bacilysin and other products of *Bacillus subtilis*. *Biochemical Journal*, 118 (4): 563-570.

Weber, T. and Marahiel, M.A. (2001). Exploring the domain structure of modular nonribosomal peptide synthetases. *Structure*, 9: R3-R9.

Whitney, J.G. and Funderburk, S.S. (1970). Anticapsin, a new biologically active metabolite. IV. Mechanism of action. In: Abstracts for the X international congress for microbiology, Mexico City, p 101.

Wipat, A. and Harwood, C.R. (1999). The *Bacillus subtilis* genome sequence: The molecular blueprint of a soil bacterium. *FEMS Microbiology Ecology*, 28 (1): 1–9.

Yazgan, A., Özcengiz, G. and Marahiel, M.A. (2001). Tn10 insertional mutations of *Bacillus subtilis* that block the biosynthesis of bacilysin. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1518 (1): 87-94.

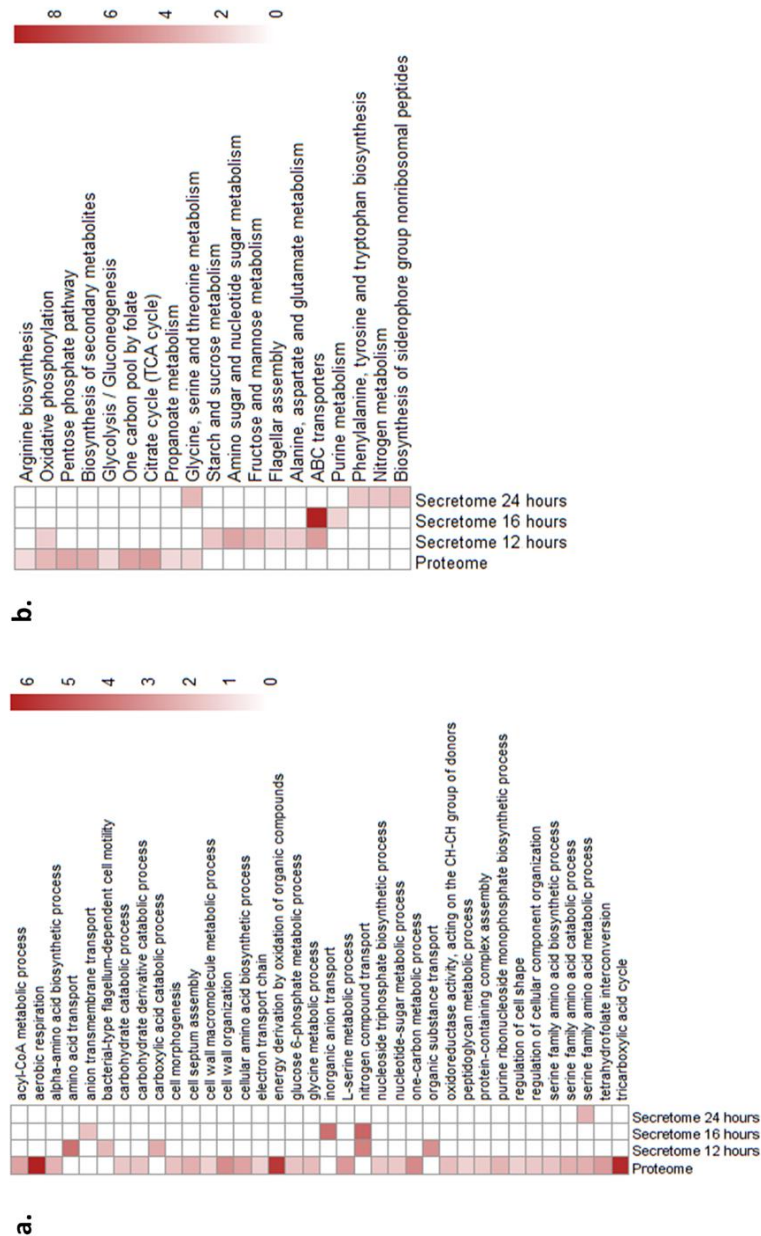
Zhang, L.H., and Dong, Y.H. (2004). Quorum sensing and signal interference: diverse implications. *Molecular microbiology*, 53 (6), 1563-1571.

Zuber, P., Nakano, M.M. and Marahiel, M.A. (1993). Peptide antibiotics in *Bacillus subtilis* and other Gram-positive bacteria (pp. 897-916). *American Society of Microbiology*.

Zweers, J.C., Barak, I., Becher, D., Driessen, A.J.M., Hecker, M., Kontinen, V.P., Saller, M.J., Vavrova, L. and van Dijl, J.M. (2008). Towards the development of *Bacillus subtilis* as a cell factory for membrane proteins and protein complexes. *Microb Cell Fact*, 7:10.

APPENDICES

A. Common biological processes (a) and pathways (b) in each condition



B. Number of clusters and their corresponding proteins in the cytosolic proteome network

Cluster	Number of proteins	Proteins
Cluster 1	28	AzoR2, CcpA, DltA, GlyS, GuaB, MotA, MotB, MsrA, MsrB, PdxT, PnpA, PurA, RnjA, RplP, RplQ, RplR, RpsI, Sfp, ThiG, ThiS, Tsf, Xpt, YceD, YceE, YceH, YciC, YflG, YfmM
Cluster 2	10	AmhX, AppF, CtpB, FabI, OppD, QueE, SndC, YacL, YhcQ, YxaL
Cluster 3	28	BceB, CwlC, CwlJ, CwlO, DegU, DivIB, IspE, MinC, MinD, MraY, NagZ, PbpB, PbpD, PbpH, PdaA, PhoA, ResD, SalA, SleB, UppP, YbfG, YflK, YhbJ, YknX, YncD, YueB, YvrN, YvrP
Cluster 4	42	AccD, AlbE, AtpC, AtpD, BmrA, CccB, CtaD, FadE, GlmS, GlpD, HtpG, KatX, Ldh, LipM, Mdh, MenD, Ndh, OdhB, Pel, PlsX, PlsY, PpaC, Prs, SdhA, SdhC, SpoVR, SucC, SucD, TepA, TufA, YdfG, YdjP, YetH, YhfK, YhfW, YitC, YitD, YjmC, YlaK, YojN, YumB, YxbC
Cluster 5	2	CotE, CotU
Cluster 6	18	BacC, CitB, CitG, FabG, FadN, Gdh, LeuC, MalS, MenB, MmgE, OxdD, PncA, SodF, YjqC, YkfB, YkkeE, YsdC, YsmA
Cluster 7	44	AddB, AhpC, AprX, ArgF, BacD, CheR, CheV, CysK, DynA, EzrA, Hfq, HypO, IspA, Kama, LipC, LuxS, MoeA, MutS, NadE, Ogt, PsdR, ResC, RibD, SbcD, ScpA, SepF, Smc, SpsD, SufU, TagT, ThiE, ThiF, TrmB, TsaE, UreC, WprA, YgaF, YitA, YjcG, YodQ, YpdQ, YpeP, YrkI, YxbD
Cluster 8	18	AraB, Dra, Drm, FbaA, GalM, GlcK, GndA, HxlA, HagA, HagB, NagP, PtsA, SpoVFA, SpoVFB, YhxA, YtkL, YwjH, YycR
Cluster 9	20	CotA, CotSA, DegV, GalT, GlgD, MtnB, Pkg, PyrD, SpsA, SpsB, SpsJ, SpsK, TrmR, TsaC, Ugd, YfnG, YfnH, YitU, YrkC, YsnB
Cluster 10	33	AroA, AroE, AsnO, BacA, BacB, BacF, ComEA, FldD, GcvPA, GcvT, GlnA, GlnH, GlnQ, GltT, GlyA, IlvC, MntR, MsmX, ProB, PurH, PyrK, RsbP, SdaAB, SerA, SerC, SsuC, SsuD, ThiO, YitJ, YlmE, YneT, YtlA, YtlC

C. Number of clusters and their corresponding proteins in the 12th-hour secretome network

Cluster	Number of proteins	Proteins
Cluster 1	10	AtpA, YtlC, YvdP, YuaG, CotQ, SdhA, CtaD, PdhB, MmgA, AcdA
Cluster 2	7	FlgK, FliS, PbpC, Hag, YuaB, YclI, GlpX
Cluster 3	5	WapA, AzoR2, DhbB, Sfp, PksD
Cluster4	8	YhcR, YfkN, CinA, YfnI, UgtP, Lip, YfjN, DgkB
Cluster 5	6	PenP, NagA, YlaM, YbcM, GcvH, GlcA
Cluster 6	6	Vpr, Bpr, PrsA, EfeU, BglC, Abh
Cluster 7	10	MntA, MetN, MetP, GlnH, ArtP, YckA, OppA, OpuAC, YhjB, OpuCA
Cluster 8	9	Abn2, GalT, GalK, LevG, PgcM, SacC, XylR, YfnH, GmuE

D. Number of clusters and their corresponding proteins in the 16th-hour secretome network

Cluster	Number of proteins	Proteins
Cluster 1	7	AbrB, ComEC, WapA, LytE, YngK, YrvJ, PgdS
Cluster 2	3	FpbQ, YlmA, ScoC
Cluster 3	8	OppA, PstBA, YxaL, PstS, OpuAC, OpuCA, YvgM, PstA
Cluster 4	5	DeoD, PucB, PucL, YwmD, BglH
Cluster 5	6	TcyB, TcyA, GlnH, YxeM, Pdp, YxeN
Cluster 6	7	YhcR, CzcD, YwhH, Bpr, SerA, YxaA, BglC

E. Number of clusters and their corresponding proteins in the 24th-hour secretome network

Cluster	Number of proteins	Proteins
Cluster 1	5	YhcY, YhdY, YwmD, GcvT, SerA
Cluster 2	3	Ndh, NarG, YrpB
Cluster 3	3	YceE, CcpA, Pgm
Cluster 4	3	PheA, ThrR, EfeB
Cluster 5	4	SrfAC, PksD, DhbB, MenF