PROTEOMIC ANALYSIS OF TWO CEPHAMYCIN C OVERPRODUCER AND AN INDUSTRIAL CLAVULANIC ACID OVERPRODUCER STRAINS OF STREPTOMYCES CLAVULIGERUS IN COMPARISON WITH THE STANDARD STRAIN NRRL 3585

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

BY

ESER ÜNSALDI

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
BIOLOGY

SEPTEMBER 2016

Approval of the thesis:

PROTEOMIC ANALYSIS OF TWO CEPHAMYCIN C OVERPRODUCER AND AN INDUSTRIAL CLAVULANIC ACID OVERPRODUCER STRAINS OF *STREPTOMYCES CLAVULIGERUS* IN COMPARISON WITH THE STANDARD STRAIN NRRL 3585

submitted by **ESER ÜNSALDI** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Biology Department, Middle East Technical University** by,

Prof. Dr. Gülbin Dural Ünver _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Orhan Adalı _____
Head of Department, **Biology**

Prof. Dr. Gülay Özcengiz _____
Supervisor, **Biology Dept., METU**

Examining Committee Members:

Prof. Dr. Ufuk Gündüz _____
Biology Dept., METU

Prof. Dr. Gülay Özcengiz _____
Biology Dept., METU

Prof. Dr. Mustafa Akçelik _____
Biology Dept., Ankara University

Assoc. Prof. Dr. Bülent İçgen _____
Environmental Eng. Dept., METU

Asst. Prof. Dr. Aslıhan Kurt Kızıldoğan _____
Agricultural Biotech. Dept., Ondokuz Mayıs University

Date: *23.09.16*

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, Last name: Eser Ünsaldı

Signature:

ABSTRACT

PROTEOMIC ANALYSIS OF TWO CEPHAMYCIN C OVERPRODUCER AND AN INDUSTRIAL CLAVULANIC ACID OVERPRODUCER STRAINS OF *STREPTOMYCES CLAVULIGERUS* IN COMPARISON WITH THE STANDARD STRAIN NRRL 3585

Ünsaldı, Eser

Ph.D., Department of Biology

Supervisor: Prof. Dr. Gülay Özcengiz

September 2016, 347 pages

In this study, two cephamycin C (CC) overproducers, *Streptomyces clavuligerus* AK39 and TB3585 strains, and a clavulanic acid (CA) overproducer *S. clavuligerus* DEPA were subjected to proteomic analysis to elucidate their differential protein expression profiles when compared to that of the standard strain *S. clavuligerus* NRRL 3585. Two proteomics approaches were employed: 2DE technique coupled with MALDI-TOF/MS and LC-MS/MS. By using two techniques, a total of 40 proteins were identified as upregulated while 70 proteins were downregulated in AK39. For TB3585 strain, a total of 44 proteins were overrepresented while 50 proteins were underrepresented. For both AK39 and TB3585, "Hypothetical/Unknown Proteins" category harbored higher numbers of upregulated proteins which was followed by "Secondary Metabolism". Both CC overproducers specifically overexpressed enzymes related with CC biosynthesis and underexpressed proteins involved in the biosynthesis of diverse secondary metabolites like clavams and polyketides. As for the DEPA strain, a total of 80 proteins were upregulated, and 174 proteins were downregulated. Upregulated protein categories in DEPA were ranked as

"Hypothetical/Unknown", "Others/General Function", "DNA Replication, Recombination, Repair, Transcription" and "Secondary Metabolism", respectively. Three CA biosynthetic enzymes as well as CcaR were overrepresented whereas enzymes of two other secondary metabolites were underrepresented along with two important global regulators. A decrease in amino acid metabolism, particularly in methionine biosynthesis, appeared as one of the prominent mechanisms of success of DEPA strain as a prolific producer of CA. Results described herein provide useful information for further improvement of CC and CA production in *S. clavuligerus* with rational approaches.

Keywords: *Streptomyces clavuligerus*, cephamycin C, clavulanic acid, proteomics

ÖZ

STREPTOMYCES CLAVULIGERUS'UN İKİ AYRI SEFAMİSİN C AŞIRI ÜRETİCİ SUŞU VE BİR ENDÜSTRİYEL KLAVULANİK ASİT AŞIRI ÜRETİCİSİNİN STANDARD SUŞ NRRL 3585 İLE KARŞILAŞTIRMALI PROTEOMİK ANALİZİ

Ünsaldı, Eser

Ph.D., Biyoloji Bölümü

Tez Yöneticisi: Prof. Dr. Gülay Özcengiz

Eylül 2016, 347 sayfa

Bu çalışmada, iki sefamisin C (CC) aşırı üreticisi suşu, *Streptomyces clavuligerus* AK39 ve TB3585, ve bir klavulanik asit (CA) aşırı üreticisi olan *S. clavuligerus* DEPA suşunun standard suş *S. clavuligerus* NRRL 3585 ile karşılaştırmalı proteomik analizleri yapılarak protein ifadelerindeki farklılıklar ortaya konulmuştur. İki ayrı proteomik yaklaşım kullanılmıştır: 2DE yöntemini takiben MALDI-TOF/MS analizi ve LC-MS/MS metodu. Bu iki tekniği kullanarak, AK39 suşunda toplam 40 proteinin seviyesinde artış, 70 proteinin seviyesinde ise azalma belirlenmiştir. TB3585 suşunda ise, her iki yöntem sonucu, toplam 44 proteinin aşırı ifade edildiği, 50 proteinin ifadesinin azaldığı belirlenmiştir. AK39 ve TB3585 suşlarında, "Hipotetik/Bilinmeyen Proteinler" kategorisi sayıca en fazla proteini içermekte ve onu takiben "İkincil Metabolizma" gelmektedir. Her iki CC aşırı üreticisi, özellikle CC biyosentezi ile ilgili olan proteinleri aşırı üretmekte ve klavam ve poliketid gibi çeşitli ikincil metabolitlerin üretiminde görev alan proteinleri daha az üretmektedir. DEPA suşunda ise, toplamda 80 proteinin seviyesi artarken 174 proteinin ifadesi azalmıştır. Aşırı üretilen

proteinlerin bulunduğu kategorileri sırasıyla “Diğerleri/Genel Fonksiyon”, “DNA Replikasyonu, Rekombinasyon, Onarım, Transkripsyon” ve “İkincil Metabolizma” oluşturmaktadır. Üç CA biyosentetik enzimi ve CcaR proteininin seviyesinde artış gözlemlenirken diğer iki ikincil metabolitin üretiminde görev alan enzimler ile iki önemli global regülatörün ifadesinde azalış belirlenmiştir. Amino asit metabolizmasında, özellikle metiyonin biyosentezinde, görev alan enzimlerde belirlenen genel düşüş, DEPA suşunun oldukça verimli bir CA üreticisi oluşuna açıklık getirmiştir. Bu araştırmanın sonuçları, *S. clavuligerus*’da CC and CA üretiminin rasyonel yaklaşımlarla daha da arttırılması yönünde yeni hedefler gösteren bilgiler temin etmiştir.

Anahtar kelimeler: *Streptomyces clavuligerus*, sefamisin C, klavulanik asit, proteomik

To My Beloved Family

ACKNOWLEDGMENTS

There are some stages that can be distinguished by certain borders in one's life-time. One of those stages for me has been the term that covered my Ph.D studies. It has not only contributed to me scientifically, but also touched my life deeply with sometimes a sense of humor and excitement and sometimes with awe and dread. It has given me an invaluable life experience as if to make me ready to the life that I have wanted to build for myself. There were moments of fulfillment, pleasure, disappointment and despair. However, I was bestowed by people who shared those pleasant as well as dreadful moments of mine with their sincere support and comfort whenever I needed.

First of all, I would like to thank my advisor, Prof. Dr. Gülay Özcengiz, for her never-ending support and guidance throughout my studies. She has always been so considerate for me and her students that she never withheld any support and help she could provide for us in any aspect of our life.

I would also like to thank Assoc. Prof. Dr. Servet Özcan from GENKÖK in Erciyes University for his support, encouragement and invaluable advices. I have learned a lot from him through his genius discussions.

I am also indebted to Prof. Dr. Dörte Becher and Dr. Birgit Voigt in Greifswald University for their kind help, valuable advices and hospitality in Germany.

I would like to express my gratitude to Dr. Volkan Yıldırım who always patiently helped me with his generous advices and knowledge.

I am also indebted to Asst. Prof. Dr. Aslıhan Kurt Kızıldoğan for her warm and understanding support during the tough times I had been through.

My sincere thanks go to my fellow labmates Çiğdem Yılmaz, Elif Tekin, İsmail Cem Yılmaz and Zeynep Eran Ak and former labmates Ayça Çırçır Hatıl, Mustafa Çiçek whom we shared so many memories with and supported one another all the time. I

would also like to express my gratitude to former fellow labmates Dr. Orhan Özcan, Dr. Burcu Emine Tefon and Dr. Sezer Okay for their generous help and advices.

I would like to thank Proteomics Group in GENKÖK under the supervision of Assoc. Prof. Dr. Servet Özcan, namely Ayşegül Murat, Mustafa Burak Açar, Ahmet Bakırbaş, Eda Mert, Ayşe Hale Güçkır and Fatih Ömerli, for their kind hospitality and help in Erciyes University in Kayseri.

My precious friend Bahar Değirmenci, who never left my side and supported me with her sincerity and always believed in me deserves my all heartful thanks. Even though we have been miles apart and hours away, I have always felt her support and faith in me.

I also appreciate my dear friends and the loveliest couple that I have ever seen, Arezoo Hosseini and Navid Hosseini, for their support and sincere friendship.

I would also like to thank Giray Bulut for his warm friendship and generous help whenever I needed.

My precious, dear family also deserves to be appreciated. I have a loving family with all the warmness and affection that one can ever see. My beloved lovely sisters, Esin Ünsaldı and Müyesser Ünsaldı, never stopped to support me and were always by my side with their heartful support. My dear brother, Fethi Ahmet Ünsaldı, always believed in me and showed me how lucky I am to deserve such a love. My lovely mother, Mehtap Ünsaldı, and always-caring father, Sami Ünsaldı, wished me the best all the time and never got tired of supporting me.

I acknowledge the financial support granted by TUBITAK KBAG (Project no: 114Z648) and METU Research Fund [BAP-07-02-2014-007-050]. I also fully appreciate the grant endowed by TUBITAK 2211-C Doctorate Scholorship Programme.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

TSB	Tryptic Soya Broth
SA	Starch-Asparagine
CA	Clavulanic Acid
CC	Cephamycin C
SARP	*Streptomyces* Antibiotic Regulatory Proteins
TCS	Two Component System
MALDI-TOF	Matrix Assisted Laser Desorption/Ionization Time of Flight
MS	Mass Spectrometry
LC	Liquid Chromatography
1DE	One Dimensional Gel Electrophoresis
2DE	Two Dimensional Gel Electrophoresis
PAGE	Polyacrylamide Gel Electrohoresis
1D PAGE	One Dimensional Polyacrylamide Gel Electrohoresis
2D PAGE	Two Dimensional Polyacrylamide Gel Electrohoresis
SDS	Sodium Dodecyl Sulfate
ACN	Acetonitrile
FASP	Filter-Aided Sample Preparation
TFA	Trifluoroacetic Acid
Mw	Molecular Weight

Da	Dalton
PTM	Postranslational Modification

CHAPTER 1

INTRODUCTION

1.1. The Genus *Streptomyces*

Streptomyces is the largest genus identified in the phylum of *Actinobacteria*, which resemble filmentous fungi in such aspects as forming hypha and spores during morphological development (Flärdh and Buttner, 2009; Nett et al., 2009). *Streptomyces* are well-known for having high G+C content in their DNA (Hopwood, 2006).

Streptomyces are usually soil inhabitants and recyclers of organic materials.Some of them also dwell in aquatic environments and can also be plant or animal pathogens (Flärdh and Buttner, 2009). Another conspicuous property of them is that some members have linear chromosomes, and these linear chromosomes possess discrete telomeres with unique terminal-inverted repeats binding terminal proteins (Bao and Cohen 2001). Furthermore, large linear plasmids as well as abundant circular plasmids are another distinguishing property of them (Paradkar et al., 2003).

To date genomes of 17 *Streptomyces* species have been completely sequenced (Kim et al., 2015a). *Streptomyces* have large linear chromosomes up to 10 Mb (Hopwood, 2006) and more than 20 secondary metabolic gene clusters, which encode the biosynthesis of polyketides by polyketide synthases (PKSs) (Staunton and Weissman, 2001), peptides by nonribosomal peptide synthetases (NRPSs) (Marahiel and Essen, 2009), bacteriocins (Moore, 2008), terpenoids, shikimate-derived metabolites, aminoglycosides, and other natural products (Nett et al., 2009), thus they are notorious for their high capacity for secondary metabolite production. The linear chromosomes (8-10 Mb) usually contain “core” regions of 5–6 Mb and “arm” regions (Hopwood, 2006). Genes that are essential for the survival of the microorganism are located in the

core regions (Medema et al., 2010) while arm regions contain mostly non-essential genes that carry secondary metabolite gene clusters (Nett et al., 2009). Moreover, some *Streptomyces* spp. also possess linear plasmids with their own telomere-like structures (Chater and Kinashi, 2007; Chen, 2007).

Another distinctive property of *Streptomyces* is their complex developmental life cycle which involves differentiation of cells both physiologically and morphologically and is well correlated with the secondary metabolite production (Figure 1.1) (McCormick and Flärdh, 2012). Mainly, its life cycle consists of three steps: vegetative (substrate) mycelium, aerial (reproductive) mycelium formation and spore formation. Spores germinate to give rise to vegetative mycelium which grows in and on the top of the medium. Thin layered septa divide the vegetative mycelium, composed of filamentous multinucleoidal hypha, into large compartments (Ohnishi et al., 2002; Manteca et al., 2008). As a result of a lack of nutrient sources, there comes a stage where macromolecular synthesis is decreased and the aerial mycelium formation begins (Manteca et al., 2008). Aerial mycelium is formed as upward extensions from the tips of vegetative mycelium (Wildermuth, 1970; Miguélez et al., 1999; Manteca et al., 2007). As the aerial mycelium develops, *Streptomyces* colonies produce diverse secondary metabolites and synthesize antibiotics (Elliot et al., 2007). After a while, extensive compartmentalization at the tips of the aerial hypha occurs, which then forms the spore chains wrapped at the tip of the aerial hypha. Upon the release and germination of the spore chains, life cycle of *Streptomyces* begins (Kwak and Kendrick, 1996; Ohnishi et al., 2002; Manteca et al., 2008).

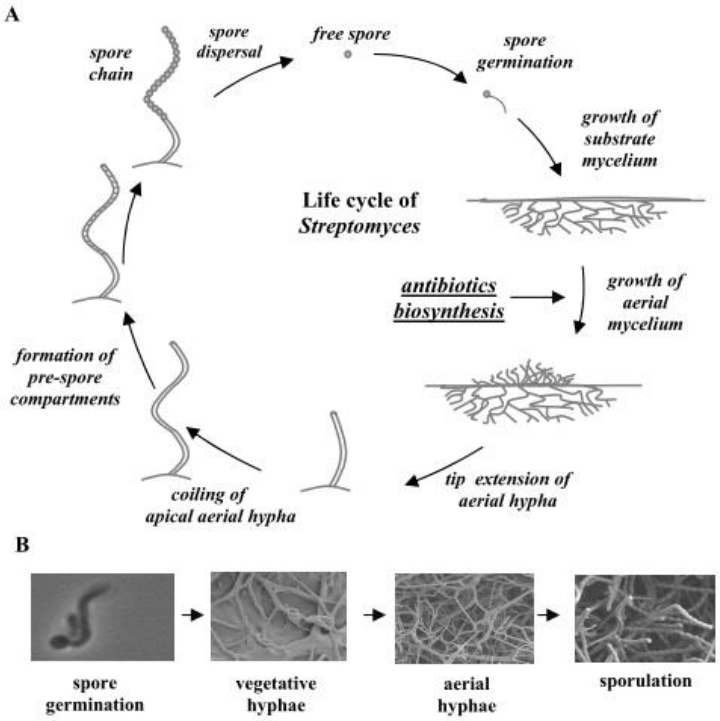


Figure 1. 1. Life cycle of *Streptomyces coelicolor* A3(2). A) Schematic representation. B) Images of different stages of life cycle of *S. coelicolor* from scanning electron micrograph (Swiatek, 2012).

There are diverse stresses such as physical, chemical or biological coming from soil where *Streptomyces* mostly inhabit. Given that streptomycetes are inmotile, their adaptation towards production of a large variety of secondary metabolites to combat the unfavorable conditions is quite plausible (Bentley et al., 2002). Among the secondary metabolites that *Streptomyces* can produce are antibacterial, anticancer, immunosuppressive, antihelmintic and antifungal agents (van Wezel and McDowell, 2011). Moreover, diverse chemical inhibitors of cellular processes such as antibiotics, fungicides, cytostatics, modulators of the immune response, and effectors of plant growth were found to be synthesized by *Streptomyces* (Ōmura, 1992; Hopwood, 2007). Of all the discovered antibiotics, more than half is produced by *Streptomyces* (Berdy 1995; Challis and Hopwood 2003; Medema et al., 2010). The production of secondary metabolites by *Streptomyces* occurs just before or during the aerial hypha

formation, which is represented as the stationary phase in liquid grown cultures (Bibb, 2005).

1.2. *Streptomyces clavuligerus* as a Prolific Secondary Metabolite Producer

Streptomyces clavuligerus was first isolated from a South American soil sample and named by Higgens and Kastner (1971) as "clavuligerus" meaning "bearing little cubes" (*clavula* little cube; *-igerus* bearing) due to the fact that this species was composed of branched aerial mycelia that eventually form spore chains (Higgens and Kastner, 1971).

Over the years, *Streptomyces* proved to be one of the most industrially important microorganism by producing several natural products, some of which are used in the clinic. *Streptomyces clavuligerus*, for example, produces two very important β-lactam antibiotics (cephamycin C and clavulanic acid) (Medema et al., 2010). Cephamycin C is a broad spectrum β-lactam antibiotic while clavulanic acid is a β-lactamase inhibitor as well as a weak β-lactam antibiotic (Mellado et al., 2002; Bussari et al., 2009). This species also produces several clavam antibiotics (clavam-2-carboxylate, 2-formyloxymethylclavam, 2-hydroxymethylclavam, hydroxyethylclavam and alanyl-clavam) (Liras et al., 2008), the pyrrothine class antibiotic holomycin (Kenig and Reading, 1979; Oliva et al., 2001) shown to have antitumor activity (Liras et al, 2008), a tacrolimus-like macrolide (Kim and Park, 2008), and a compound related to the nucleoside antibiotic tunicamycin (Kenig and Reading, 1979).

Medema et al. (2010) sequenced *S. clavuligerus* genome and described the presence of 6.8 Mb chromosome and 1.8 Mb linear plasmid in the genome. They showed 23 putative secondary metabolite gene clusters resided in 6.8 Mb chromosome while 1.8 Mb linear plasmid (pSCL4) harbored 25 such clusters. They identified 7,281 putative protein-encoding genes on these two replicons, six rRNA operons, 72 tRNA genes, and 14 pseudogenes (Medema et al., 2010). Another study conducted on *S. clavuligerus* genome by Song et al. (2010a) confirmed the results represented by Medema et al. (2010). They showed the presence of one linear chromosome (6,736,475 bp, 72.69% GC) and four linear plasmids, pSCL1 (10,266 bp, 71.96% GC), pSCL2 (149,326 bp, 70.07% GC), pSCL3 (442,792 bp, 70.77% GC), and pSCL4

(1,796,117 bp, 71.85% GC) in *S.clavuligerus* genome. While they also confirmed the results of previous studies (Netolitzky et al., 1995; Wu and Roy 1993; Wu et al. 2006). Moreover, though it possessed the conserved core region in the genome like other streptomyces, it lacked the large chromosomal arms. The supercluster responsible for the cephamycin C and clavulanic acid biosynthesis as well as one of the clavam clusters (Tahlan et al. 2007) was shown to be located on the main chromosome while the alanylclavam cluster (Zelyas et al. 2008) resided on the 1.8 Mb plasmid (pSCL4). Furthermore, pSCL4 harbored a distant homolog of clavaminate synthase gene, whose function could not be identified but was suggested to be another β-lactam antibiotic that has yet to be described in *S. clavuligerus*. (Medema et al., 2010). Through their study, Song et al. (2010a) also found gene clusters for nonribosomal peptide synthetases, polyketide synthases, and the hybrids on pSCL3, pSCL4, and the chromosome. They also identified two gene clusters coding for enediyne-containing compounds which are potent antitumor agents (Liu et al., 2003) and many others for secondary metabolism such as staurosporine (Salas and Méndez, 2009), moenomycin (Ostash et al., 2009), terpenes, pentalenenes, phytoenes, siderophores, and lantibiotics (Song et al., 2010a).

1.3. β-Lactam antibiotics, Cephamycin C and Clavulanic acid

1.3.1. β-Lactam Antibiotics

β-lactam antibiotics are one of the oldest and most commonly used antibacterial compounds in the history. They are modified peptides that have bactericidal effect on bacteria (Liras, 1999). They all share a common chemical structure consisting of a four-membered, heterocyclic ring called β-lactam ring (Figure 1. 2). β-lactam rings are fused to a second ring thus β-lactam antibiotics form a bicyclic structure. The ring structure of this second ring determines the type of β-lactams. If it is five membered, β-lactams are called penam/penem antibiotics or if six membered, as cepham/cephem antibiotics (Figure 1. 2) (MacKenzie, 2007).

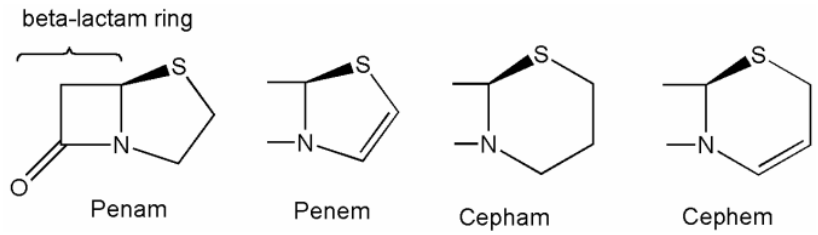


Figure 1. 2.Common motif of the β-lactam antibiotics. Complete β-ring structure is shown at the very left of the images. Nomenclature of β-lactam antibiotics are also shown depending on the fused (2^{nd}) ring to the β-ring (MacKenzie, 2007).

β-lactam antibiotics interfere with the synthesis of bacterial cell wall and they have bactericidal effect on the microorganisms. The main component of the bacterial cell wall is peptidoglycan layer which is composed of glycan layers and short peptide linkages that connect the glycan layers.

Peptidoglycan biosynthesis mainly consists of three stages (Figure 1. 3). First stage is initiated in the cytoplasm and involves the biosynthesis of UDP-N-acetylmuramyl (UDP-MurNAc)-pentapeptide and UDP-N-acetylglucosamine (UDP-GlcNAc) precursors. Second stage is marked by the linkage of UDP-MurNAc-pentapeptide to the transport lipid (undecaprenyl pyrophosphate), resulting in the formation of lipid I. GlcNAc derived from UDP-GlcNAc precursor is added to the lipid I, thereby converting lipid I to lipid II, to which a peptide moiety is attached at the third amino acid of the pentapeptide. Lipid II is transported through the membrane to the exterior side. At the final stage, transpeptidation and transglycolsylation reactions catalyzed by penicillin-binding proteins (PBPs) and transglycosylases, respectively, allow the moiety carried by lipid II to be incorporated into the newly synthesized peptidoglycan (Kohanski et al., 2010; Pinho et al., 2013).

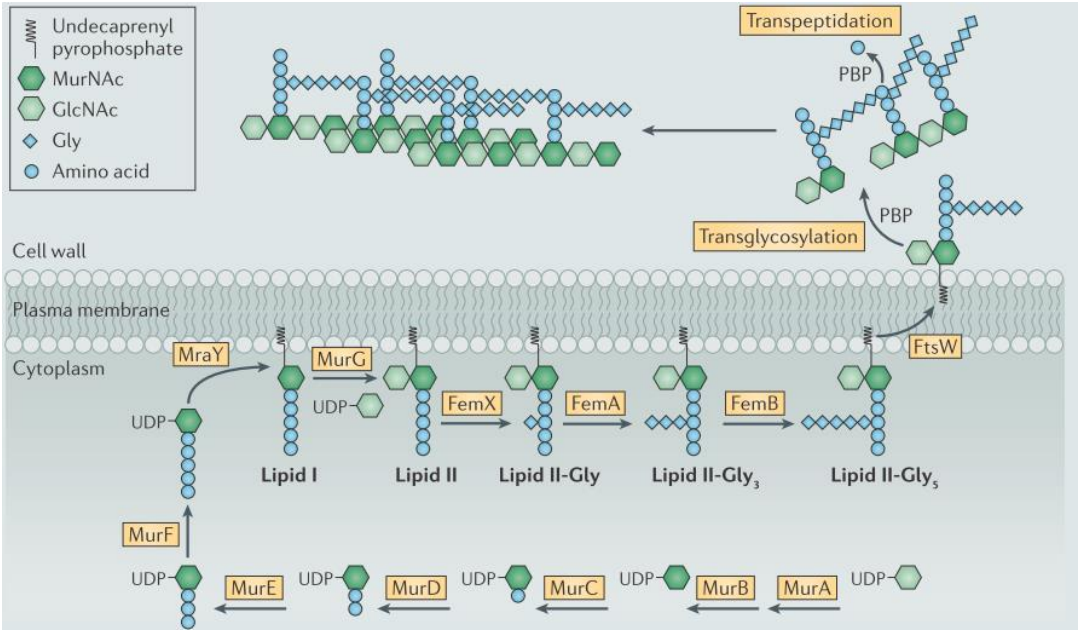


Figure 1. 3. Peptidoglycan biosynthesis of the bacterial cell wall structure (Pinho et al., 2013).

This step is where the β-lactam antibiotics come into the play. β-lactams act like D-alanyl-D-alanine moieties within the pentapeptide structure through which transpeptidation occurs and hence can bind to the PBPs. Upon binding the active site of the enzyme, their ring structure is hydrolyzed and then they acylate the serine residue in the active site forming a stable acyl-enzyme complex (Figure 1. 4) (MacKenzie, 2007). As a result of the inhibition of PBPs, peptide bridges cannot be introduced for the complete peptidoglycan layer formation. Formerly, it was thought that due to the lack of cross-linkages among the glycan layers, cell walls cannot resist the osmotic pressure produced inside of the cell and thus lysis occurs (Tipper and Strominger, 1965; MacKenzie, 2007). However, the lysis process was shown to be more complex than the simple build-up of osmotic pressure. Autolysin enzymes, which are peptidoglycan hydrolases and also used during cell division and elongation, were shown to take part in the hydrolysis of the peptidoglycan structure after the β-lactam antibiotics block the activity of PBPs (Kohanski et al., 2010; Pinho et al., 2013).

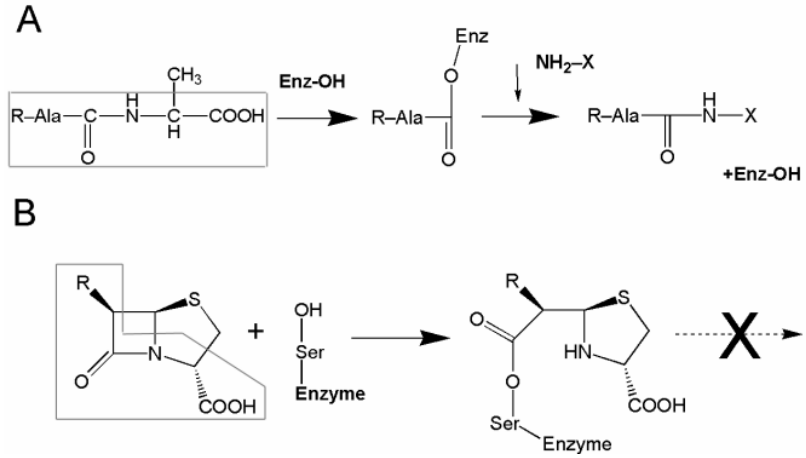


Figure 1. 4. Activity of PBPs and β-lactam antibiotics (A) Transpeptidation reaction catalyzed by PBPs. (B) Inhibition of PBPs by β-lactam antibiotics (MacKenzie, 2007).

1.3.2. Cephamycin C

Discovery of cephamycins dates back to 1970s during a screening program at Merck and Lilly Co. They are composed of a β-lactam ring fused to a six-membered cephem ring and had α-aminoadipoyl in the 7β position just like cephalosporin antibiotics but can be differentiated from them by having a methoxy substituent at the 7α position of the β-lactam ring. (MacKenzie, 2007; Page, 2012). Cephamycin A and B show broad spectrum of activity against both Gram-positive and Gram-negative bacteria but their stability is quite low. On the other hand, cephamycin C shows very similar spectrum of activity like cephamycin A and B but is more effective against Gram-negative bacteria. All cephamycins share the property of being resistant to β-lactamases and are more stable against serine β-lactamases than cephalosporins. The resistance comes from the methoxy substitution at the C7-position (Figure 1. 5) (MacKenzie, 2007, Bussari et al., 2009).

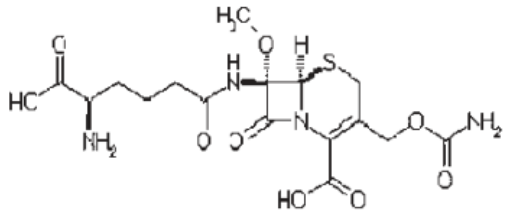


Figure 1. 5. Chemical structure of cephamycin C (Bussari et al., 2009).

Cephamycin C is produced by *Streptomyces cattleya*, *Streptomyces clavuligerus* and *Nocardia lactamdurans* (Bussari et al., 2009). However, cephamycin C is not directly used in the clinic, instead semi-synthetic antibiotics such as cefoxitin, cefotetan, cefametazole, and temocillin derived from cephamycin C are used commercially (Figure 1. 6) (Kagliwal et al., 2009; Bussari et al., 2009). Semi-synthetic cephamycins show higher resistance against class A extended-spectrum β-lactamases while they were shown not to be so stable against the class C β-lactamases, and a number of the plasmid-encoded class C enzymes (e.g., FOX-1 enzyme) (Leiza et al., 1994; Page, 2012).

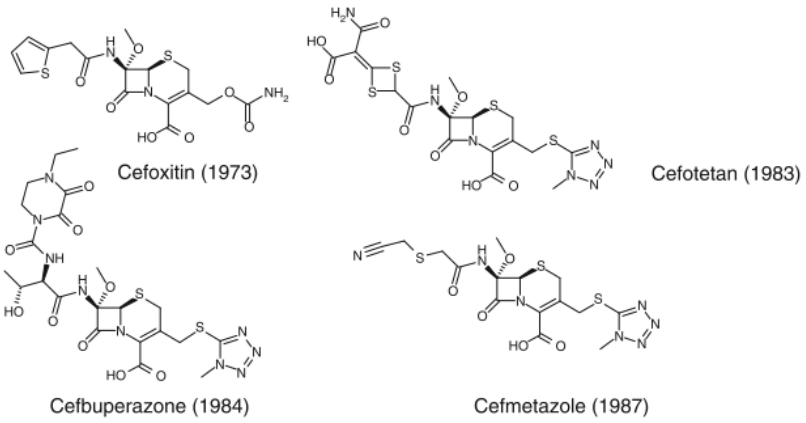


Figure 1. 6. Semi-synthetic cephamycins (Page, 2012).

1.3.3. Clavulanic Acid

Clavulanic acid (CA) is a broad spectrum β-lactam antibiotic active against both Gram-positive and Gram-negative bacteria but shows quite a weak antibacterial activity compared to other broad spectrum antibiotics (Demain and Elander, 1999; Saudagar et al., 2008). Though having weak activity, Brown et al. (1976) showed that it has a strong β-lactamase inhibitory activity. It was first isolated from *S. clavuligerus* and after its characterization as a β-lactamase inhibitor, it has been widely used in clinic combined with other β-lactam antibiotics like amoxicillin or ticarcillin, under the trade names of Augmentin ® and Timentin®, respectively (Song et al., 2010b; Page, 2012). Clavulanic acid is categorized as a clavam metabolite and has a bicyclic structure with 3R, 5R stereochemistry consisting of a β-lactam ring fused to a second five-membered oxazolidine ring (MacKenzie, 2007; Song et al., 2010b). 3R, 5R stereochemistry renders the clavulanic acid β-lactamase inhibitory activity (Figure 1. 7) (Baggaley et al., 1997).

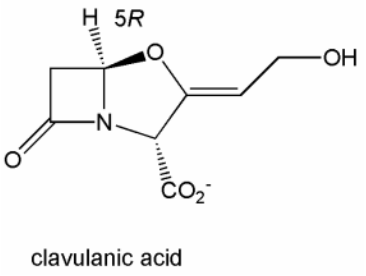


Figure 1. 7. The structure of (3R,5R)-clavulanic acid (MacKenzie, 2007).

Clavulanic acid is a potent inhibitor of class A serine β-lactamases which inhibit penicillins, it can also inhibit some of the class D serine β-lactamases that selectively hydrolyze oxacillins, but it does not show considerable activity against Class C serine β-lactamases which usually target cephalosporins and have no activity against class B metallo-β-lactamases (Baggaley et al., 1997; MacKenzie, 2010; Page, 2012).

Clavulanic acid acts as a substrate analogue for the targeted β-lactamase and remains bound to the active site of the enzyme long enough for β-lactam antibiotic co-administered with CA to evade being degraded by the β-lactamase (MacKenzie, 2007).

1.4. Biosynthesis of Cephamycin C and Clavulanic Acid

Biosynthesis of clavulanic acid and cephamycin C occurs in parallel in *S. clavuligerus*; moreover, the genes responsible for the coding of both antibiotics form separate clusters on the genome but these clusters are close to each other forming an arrangement of a supercluster of approximately 60 kb (Liras et al., 2008).

Accessory genes as well as genes for the synthesis of β-lactam compounds are found in the cluster. Accessory genes are not required directly for the biosynthesis of the antibiotics but for processes such as regulation of the genes within the supercluster, export of the compounds, and self-resistance. (MacKenzie, 2007). Moreover, a paralog cluster for the clavulanic acid synthesis was shown to be located on pSCL4 plasmid of *S. clavuligerus* (Tahlan et al., 2004a; Song et al., 2010a)

1.4.1. Biosynthesis of Cephamycin C

There are three amino acids used as precursors for the cephamycin C biosynthesis: L-valine (L-Val), L-cysteine (L-Cys) and L- α -aminoadipic acid (L-α-AAA). These three precursors are necessary for the synthesis of all β-lactam antibiotics. Of these, L-valine and L-cysteine are common amino acids while L-α-aminoadipic acid is a non-proteogenic amino acid that is to be synthesized by the microorganism before antibiotic production (Liras and Martín, 2006). Lysine and α-aminoadipic acid are produced through aspartate pathway which leads to the cephamycin C production (Figure 1. 8). Aspartate is the precursor of the pathway and aspartokinase initiates the first reaction by converting aspartate to β-aspartyl phosphate, which is then converted to aspartate β-semialdehyde by aspartate semialdehyde dehydrogenase. At this point, pathway is divided to two branches, one leading to the formation of threonine, methionine and isoleucine amino acids through homoserine dehydrogenase activity, the other, also called lysine branch, leads to the cephamycin C production (Tunca et al., 2004).

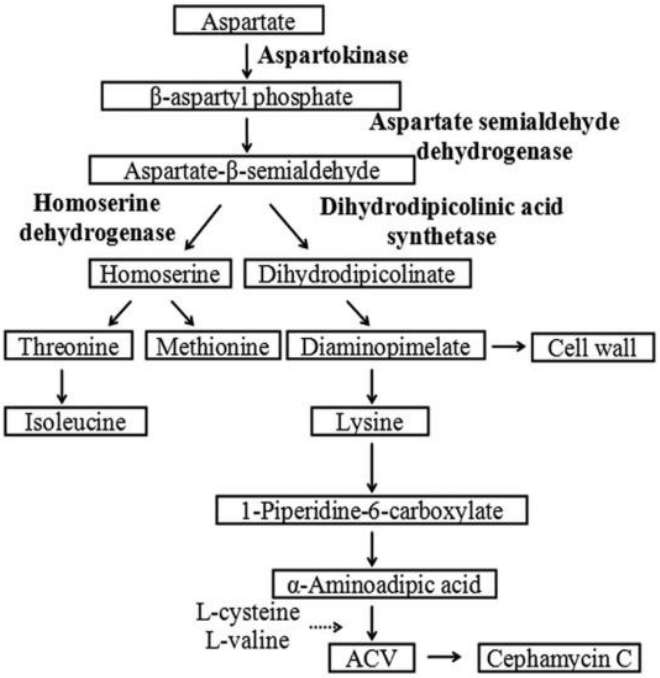


Figure 1. 8. The aspartate pathway in *S. clavuligerus* (Özcengiz et al., 2010).

Lysine produced through aspartate pathway is converted to α-aminoadipic acid semialdehyde by the lysine-6-aminotransferase (LAT) enzyme coded by *lat* gene (Coque et al. 1991; Liras, 1999). α-aminoadipic acid semialdehyde is spontaneously cyclized to piperideine-6-carboxylic acid (P6C) which is the cyclic form of a-aminoadipic semialdehyde (Martín, 1998) and then piperideine-6-carboxylate dehydrogenase (P6C-DH) coded by *pcd* gene converts P6C to α-aminoadipic acid (Figure 1. 9) (Martín et al., 2010; Özcengiz and Demain, 2013). However, it was a wonder that *pcd* mutants could produce cephamycin at levels of 30% to 70% when compared to wild type (Alexander et al., 2007).

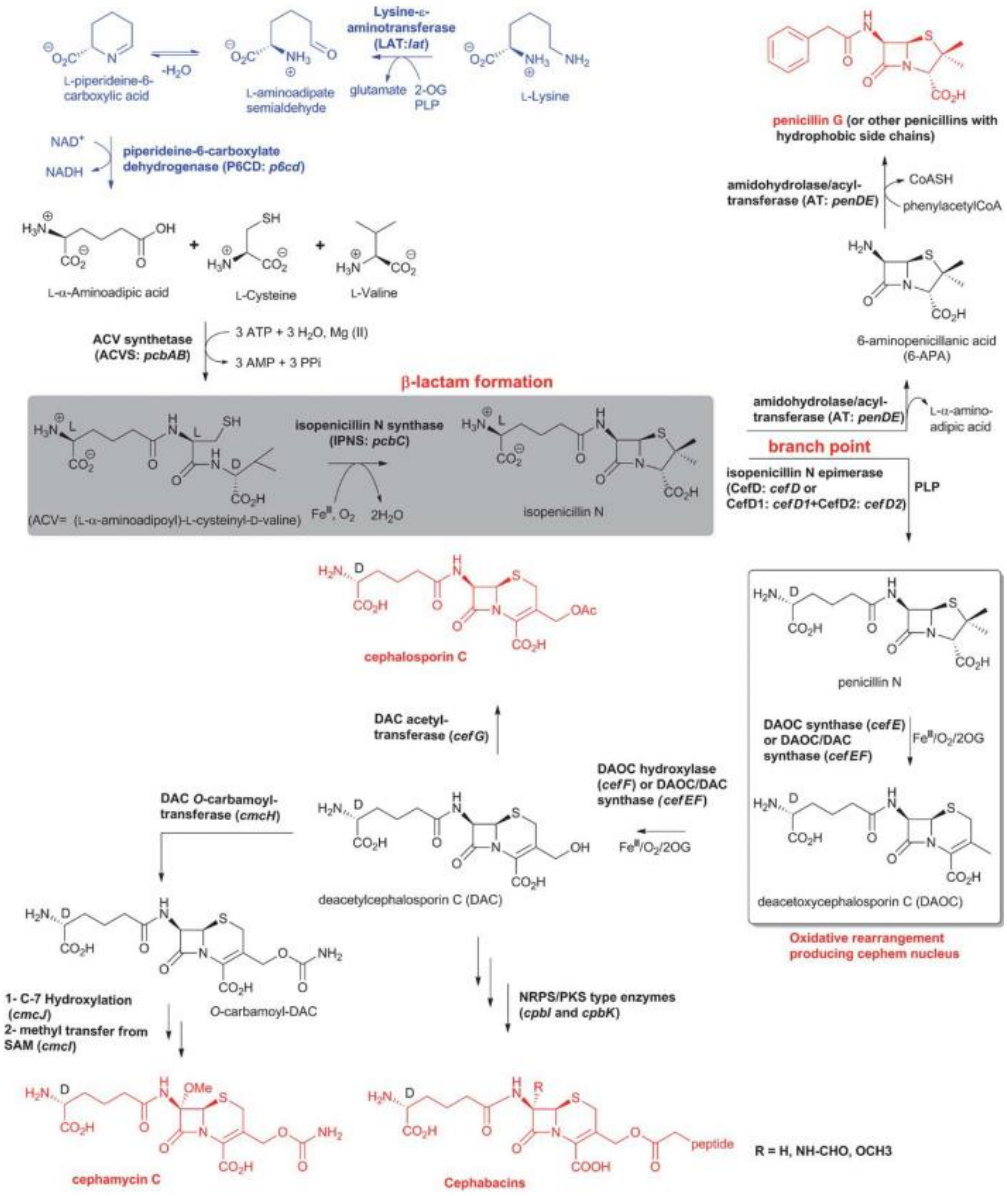


Figure 1. 9. Penicillin and cephalosporin biosynthetic pathways in fungi and bacteria (Hamed et al., 2013).

δ -(L- α -aminoadipyl)-L-cysteinyl-D-valine (ACV) synthetase (ACVS) is a non-ribosomal peptide synthetase which is responsible to form the peptide L- δ (α -aminoadipyl)-L-cysteinyl-D-valine from the precursors (ACV). ACVS coded by *pcbAB* gene has three conserved domains which specifically binds each amino acid

precursor to activate them by using ATP. It first forms aminoacyl-adenylates after which they are bound to the enzyme as thioesters and then epimerizes L-valine to the D-valine finally linking three amino acids so that the peptide ACV is formed (Martín et al., 2000; Liras and Martín, 2006). Reduced ACV tripeptide is converted to isopenicillin N by isopenicillin N synthase (IPNS) also known as ACV cyclase coded by *pcbC* gene (Liras and Martín, 2006; Özcengiz and Demain, 2013). The action mechanism of isopenicillin N is by the formation of the β-lactam ring first and the closure of the thiozolidine ring thereafter, thereby forming the first bicyclic structure in the pathway (Liras and Martin, 2006; Özcengiz and Demain, 2013). The *pcbAB* coding for the ACVS and *pcbC* genes coding for the IPNS are adjacent to each other in cephamycin C gene cluster suggesting a probable co-evolutionary process (Hamed et al., 2013). Isopenicillin N epimerase encoded by *cefD* is a pyridoxal phosphate-dependent enzyme and catalyzes the reversible epimerization of L-α-aminoadipoyl side chain in isopenicillin N to the D -α-aminoadipoyl side chain in penicillin N, i.e., catalyzes the conversion of isopenicillin N to penicillin N (Liras and Martín, 2006; Hamed et al., 2013). The next step in the cephamycin biosynthesis is the formation of deacetoxycephalosporin C, DAOC during which five-membered thiazolidine ring of penicillin N is expanded to six-membered dihydrothiazine ring by deacetoxycephalosporin C synthase (DAOCS), also known as expandase, coded by *cefE* gene. By using molecular oxygen, Fe^{2+}, and α-ketoglutarate, DAOCS produces DAOC and succinic acid. DAOC intermediate is converted to deacetylcephalosporin C (DAC) by deacetylcephalosporin C synthase (DACS) also known as hydroxylase coded by *cefF* gene and the DAOC is hydroxylated at C-3' position to form DAC (Liras and Martín, 2006). In fungi, DAOCS and DACS form a single bifunctional enzyme while they function as separate enzymes in *S. clavuligerus*. However, these proteins coded by *S. clavuligerus* show approximately 70% identity in amino acid level to each other (Kovacevic and Miller, 1991) and 60% identity to the protein coded by *cefEF* in fungi (Liras and Martín, 2006). Morover, despite to a low level, DAOCS enzyme shows DACS activity and DACS shows DAOCS activity (Baker et al., 1991) suggesting that gene duplication during the evolution of the microorganism (Kovacevic and Miller, 1991) and also horizontal gene transfer between bacteria and

fungi (Liras and Martín, 2006; Hamed et al., 2013). For the final step of cephamycin C production, there are three gene products involved: CmcI, CmcJ, and CmcH. DAC is carbamoylated at C-3' position by the gene product CmcH and methoxylated at C-7 position by the CmcI-CmcJ complex although it is not known which modification comes first; it has been suggested that at first C-3' carbamoylation occurs and then methoxylation at the 7α position takes place, as a result of which cephamycin C is produced (Öster et al., 2006). The carbamoyl transferase (*cmcH*) carries out an ATP-dependent reaction and uses carbamyl phosphate as the carbamoyl donor (Özcengiz and Demain, 2013). C-7 methoxylation occurs via two steps; first hydroxylation occurs at C-7 position, the hydroxyl group of which is derived from molecular oxygen and then C-7 methylation of the hydroxylated intermediate, the methyl group of which is derived from methionine (Öster et al., 2006). C-7 methoxylation process is proposed to be mediated by CmCI-CmCJ complex but the actual mechanism could not be revealed yet; however, with immunoaffinity chromatgoraphy, it was observed that these two proteins co-purified as complex and when isolated separately, CmcI and CmcJ proteins could not bind deacetylcephalosporin C (Enguita et al., 1996; Liras and Martín, 2006).

1.4.2. Biosynthesis of Clavulanic Acid

Clavulanic acid (CA) and 5S clavams (clavam-2-carboxylate, 2-hydroxymethylclavam, 2-formyloxym-ethylclavam, and alanylclavam) are clavam metabolites and are produced by *Streptomyces* spp. (Jensen, 2012).

Clavulanic acid was first discovered to be produced by *Streptomyces clavuligerus* in 1976 and then by several other *Streptomyces* spp. such as *S. jumonjinensis*, *S. katsurahamanus* (Brown et al., 1976; Jensen and Paradkar, 1999). Through genome sequencing projects, it was suggested that a wider spectrum of microorganisms such as *S. flavogriseus* ATCC 33331, *Saccharomonospora viridis* DSM 43017 might have the biosynthetic ability to produce CA. Nonetheless, of the all clavam producers, *S. clavuligerus* is the only species who could produce both CA and 5S clavams. Other species could either synthesize 5S clavams or CA (Jensen, 2012).

Gene cluster related to the CA biosynthesis are found adjacent to cephamycin C biosynthesis genes on the chromosome. CA cluster contains the genes encoding the enzymes involved in both CA and 5S clavam biosynthesis in the early steps as well as the genes coding for the enzymes which are only involved in CA biosynthesis in late steps (Figure 1. 10) (Jensen, 2012).

Biosynthesis of CA can be divided in two main parts; first is the synthesis of (3S, 5S)-clavaminic acid and the second is the conversion of clavaminic acid to CA (Hamed et al., 2013). Clavaminic acid is the common precursor for the synthesis of both clavulanic acid and clavams (Egan et al., 1997). Enzymes and steps leading to the clavaminic acid biosynthesis could be identified successfully whereas the rest of the steps has not yet been understood very well (Jensen, 2012).

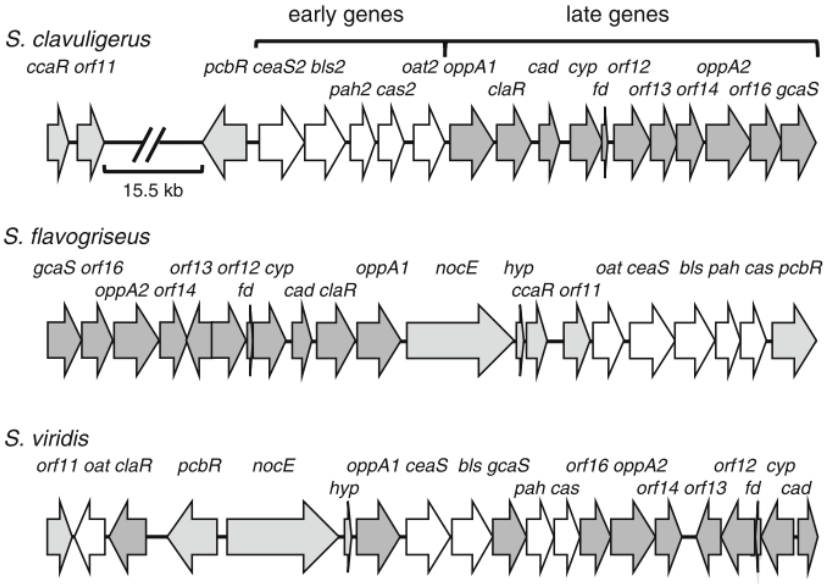


Figure 1. 10. CA gene clusters from known and presumed clavulanic acid producer species. White arrows represent the early genes, dark gray arrows represent the late genes. Light gray arrows indicate genes with regulatory or unknown functions (Jensen, 2012).

The first reaction in the synthesis of CA is the condensation of L-arginine with glyceraldehyde-3-phosphate and carried out by carboxyethylarginine synthase (Ceas1/Ceas2) producing N^2 -(2-carboxyethyl)arginine as the final product. Ceas requires thiamine pyrophosphate for its biological activity (Figure 1. 11) (Jensen, 2012). Although arginine is the precursor for CA production, ornithine was shown to be a much stronger stimulator of CA production than arginine and can suppress cephamycin C production (Chen et al., 2003; Özcengiz and Demain, 2013). This might be a result of the ornithine being the precursor for the production of arginine. OAT coded by *orf6* which is located in clavulanic acid biosynthetic gene cluster transfers acetyl group to L-glutamate from L -N-acetylornithine to form L-ornithine and L-N-acetylglutamate, the process which initiates arginine biosynthesis steps (Kershaw et al., 2002).The second step in CA biosynthesis is catalyzed by β-lactam synthetase (Bls1/Bls2) that is in charge of introducing the monocyclic β-lactam ring by converting N^2-(2-carboxyethyl)arginine to deoxyguanidinoproclavaminate (DGPC) (Bachmann et al., 1998). However, the production of β-lactam ring in penicillin or cephalosporins is quite different, which involved the cyclization of a non-ribosomally generated δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine tripeptide by isopenicillin N synthase. The cyclization reaction in CA production, on the other hand, proceeds by the help of an oxoanion intermediate after the adenylation of N^2-(2-carboxyethyl) arginine β-carboxylate (Miller et al., 2002). After the formation of deoxyguanidinoproclavaminate, there remain 4 steps to produce clavaminic acid. Clavaminic acid synthase (CAS) is responsible for the catalysis of three out of four reactions while proclavaminic acid amidino hydrolase (PAH) enzyme is responsible for one reaction. However, what is unusual in this 4-step reaction is the fact that the first step is catalyzed by CAS in which DGPC is hydroxylated at C3 position to give (2S,3R)-guanidino-proclavaminic acid (GPC) and the second reaction is catalyzed by PAH which hydrolyzes the guanidine group of GPC in order to produce proclavaminic acid (PCA) and then last two steps are carried out by the same CAS enzyme having catalyzed the first step. It is quite rare to observe the action of two enzymes catalyzing subsequent reactions in which one enzyme switches in and out pattern. At the third step, CAS enzyme forms dihydroclavaminic acid by removing β-lactam 4'-*pro*-S

hydrogen and the hydrogen of the hydroxyl group at C-3 of PCA so that desaturative ring closure takes place meaning cyclization occurs. Finally, CAS produces clavaminic acid from dihydroclavaminic acid by removing hydrogens leading the desaturation. In other words, CAS-catalyzed reactions are hydroxylation, cyclization and desaturation, respectively (Jensen, 2012; Hamed et al., 2013). The above-mentioned reactions represent the early steps of CA biosynthesis. Late steps begin with the conversion of clavaminic acid to N-glycylclavaminic acid by glycylclavaminic acid synthase (GCAS) coded by the gene *orf17* (Arulanantham et al., 2006). The last step for the production of clavulanic acid through the reduction of clavaldehyde is carried out by clavaldehyde dehydrogenase (CAD), encoded by *cad* (Nicholson et al., 1994; Pérez-Redondo et al., 1998). Intermediate steps or reactions that lead the conversion of N-glycylclavaminic acid to clavaldehyde have not been completely revealed yet, especially the inversion of 3S, 5S stereochemistry of N-glycylclavaminic into the 3R, 5R stereochemistry of clavaldehyde, which confers the CA β-lactamase inhibitory property in contrast to other clavams in 3S, 5S configuration (Jensen, 2012). Fulston et al. (2001) could identify the presence of an unstable intermediate 3R,5R-clavulanate-9-aldehyde during CA production and this molecule requires the conversion of C-9 amino acid to a C-9 hydroxyl and the inversion of the 3S,5S stereochemistry to the 3R, 5R stereochemistry to be produced (Fulston et al., 2001).

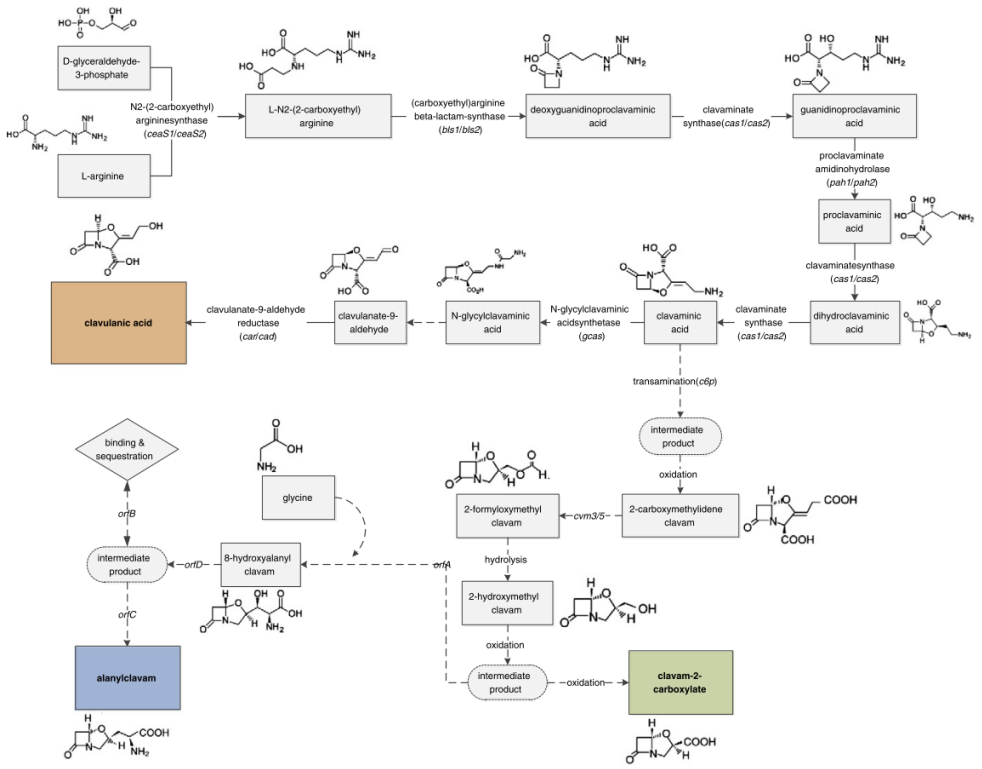


Figure 1. 11. Biosynthetic pathways for clavulanic acid and 5S clavams in *Streptomyces clavuligerus*. Solid lines represent known steps and dashed arrows indicate unknown steps (Özcengiz and Demain, 2013).

1.5. Regulation of the Secondary Metabolism in *Streptomyces*

Secondary metabolism of *Streptomyces* is regulated by several factors mainly by pleiotropic (global) factors and pathway-specific factors (Liras et al., 2008). Pathway-specific regulators usually located in the antibiotic gene clusters are controlled by pleiotropic regulators that do not only affect the antibiotic production but many other physiological properties in the microorganism. Moreover, the control of whole regulatory mechanism is further encompassed by several environmental factors such as nutrient limitation, growth rate and γ-butyrolactones (Paradkar et al., 1998).

1.5.1. Pleiotropic (Global) Factors

Pleiotropic factors affect the organism in a wide perspective from morphological differentiation to secondary metabolism (Liras et al., 2008). *bld* genes are the most prominent examples for the pleiotropic factors found in *Streptomyces*. For instance, *bldG* is an anti-anti sigma factor that affect both morphological differentiation and antibiotic production in *S. clavuligerus*. *bldA,* on the other hand, affects only morphological differentiation in *S. clavuligerus*, whereas it was shown to be involved in both antibiotic production and aerial hypha formation in *S. coelicolor*, *S. lividans*, and *S. griseus*. Thus, it is possible to observe some differences in the functions of the same genes in different *Streptomyces* spp. (Bignell et al., 2005)

In *S. coelicolor*, downstream of the anti-anti sigma factor, *bldG* is *apgA* gene encoding an anti-sigma factor. It was shown that BldG can be transcribed as both monocistronic and polycistronic transcripts together with ApgA (also known as ORF3). The presence of this "anti-anti sigma/anti-sigma pair" might suggest that they may either regulate the activity of a single sigma factor under selective environmental conditions or more than one sigma factor. Furthermore, they can act differentially on individual sigma factors so that while one regulates the activity of antibiotic production the other can regulate differentiation; alternatively the pair can act on sigma factor(s) to regulate both antibiotic production and differentiation together (Bignell et al., 2000). These suggestions have been partially clarified by the study of Sevcikova et al. (2010) who showed that BldG not only interacts with ApgA but also with another anti-sigma factor UshX which is specific for σ^H sigma factor related with morphological differentiation as well as osmotic stress response in *S. coelicolor* A3(2). They also suggested a "partner-switching-like mechanism" in which that BldG protein is dephosphorylated by an unknown mechanism under osmotic stress or some other stimulus, and thus binds both ApgA and UshX anti-sigma factors making them release the sigma factors that they are bound to. The sigma factors that are released are complexed with the RNA polymerase machinery responsible for the expression of the genes to be stimulated by the related sigma factors (Figure 1. 12) (Sevcikova et al., 2010).

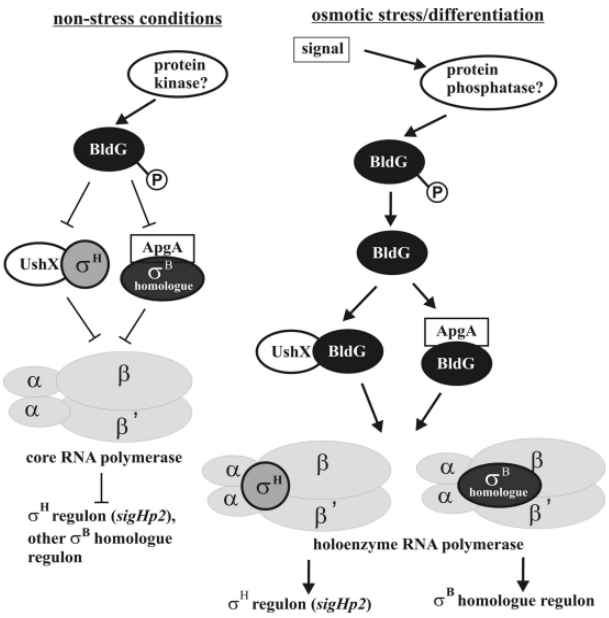


Figure 1. 12. Model of activation of σ^H activity by BldG in *S. coelicolor* (Sevcikova et al., 2010).

Another well-known example is the *bldA* gene which regulates both antibiotic production and differentiation in *S. coelicolor*. It encodes a tRNA that can translate UUA corresponding to a rare leucine codon in *Streptomyces*. For example, *S. coelicolor* has TTA codon only in 2 % of its genes and the translation of these genes depends on the translation of the *bldA* gene (Leskiw et al., 1991; Chandra and Chater, 2008). However, since essential genes do not contain this specific codon but only the ones related with antibiotic production and morphological differentiation, *bldA*-null mutants can be grown without differentiation and antibiotic production. Main actor in morphological differentiation induced by *bldA* gene is AdpA protein which is also a pleiotropic factor containing a TTA codon in its genetic make-up (den Hengst et al., 2010). *S. clavuligerus*, on the other hand, did not show any defect in antibiotic production though *ccaR* gene, regulator of cephamycin and clavulanic acid production, contains TTA codon (Trepanier et al., 2002). The fact that *ccaR* gene product can be

produced in *bldA* mutants despite having a TTA codon, was attributed to the mistranslation during the protein synthesis (Higo et al., 2011).

AdpA is another pleiotropic factor, the presence of which has been identified in several *Streptomyces* species such as *S. coelicolo*r A3(2) (Nguyen et al., 2003; Takano et al., 2003), *S. lividans* (Zhu et al., 2005), *S. antibioticus* (Zhu et al., 2005), *S. ansochromogenes* (Pan et al., 2009), *S. clavuligerus* (López-García et al., 2010) and *S. avermitilis* (Komatsu et al., 2010). Upon expression of *adpA*, the transcription of many genes related to morphological differentiation and secondary metabolism are activated. *adpA* also contains a TTA codon which requires the transcription product of *bldA* gene, and its expression is activated when A-factor builds up in the cell by time (Trepanier et al., 2002; Kato et al., 2007) and reaches a certain level so that it binds to the A-factor-specific receptor (ArpA) in *Streptomyces griseus*, which remains attached to the promoter of *adpA* (Section 1.5.3). Upon binding of A-factor, ArpA is released from the promoter and transcription of *adpA* gene is initiated (Ohnishi et al., 1999; Higo et al., 2011).

1.5.2. Pathway-Specific Regulators

Pathway-specific transcriptional regulators are found in antibiotic biosynthesis gene clusters and they can affect antibiotic production in either a negative or positive way (Paradkar et al., 1998).

One of the most common pathway-specific regulators found in *Streptomyces* species is in the group of SARP (*Streptomyces* antibiotic regulatory proteins) family proteins with a molecular weight of around 25 kDa. They possess a DNA-binding domain and a transcriptional activation domain. ActII-ORF4, CdaR and RedD proteins that control the production of actinorrhodin, undecylprodigiosin and CDA (calcium-dependent-antibiotic) in *S. coelicolor*; CcaR protein that controls the production of cephamycin C and clavulanic acid in *S. clavuligerus*; StrR protein that controls the production of streptomycin in *S. griseus* belong to SARP family proteins (Martín and Liras, 2010).

CcaR is a transcriptional regulator that positively affects the cephamycin and clavulanic acid production in *S. clavuligerus* and as expected from a pathway-specific

regulator, it enhances the production by regulating the transcription of related biosynthetic genes (Liras et al., 2008). It is also an autoregulatory protein regulating its own expression. Furthermore, lysine aminotransferase (LAT), isopenicillin-N synthase (IPNS), isopenicillin-N epimerase (IPNE) and desacetoxycephalosporin-C synthase (DAOCS) proteins involved in cephamycin C biosynthesis are regulated by CcaR protein (Alexander and Jensen, 1998). *cefD-cmcI* bidirectional promoter located in cephamycin C gene cluster produces large transcript spanning from *cefD* to *pcd* (*cefDE-pcd*) and the fact that CcaR also binds to this bidirectional *cefD-cmcI* promoter (Pérez-Llarena et al., 1998; Santamarta et al., 2002) further proves that it regulates the transcription of early, middle and late cephamycin biosynthetic genes in the cluster (Liras et al., 2008). However, regulation by CcaR protein is not only limited to cephamycin C gene cluster but also extended to clavulanic acid gene cluster. CcaR binds *ceaS2* gene promoter which encodes the carboxyethylarginine synthase that initiates the first step in the clavulanic acid production. It also binds *claR* promoter that controls the late steps in clavulanic acid biosynthesis (Santamarta et al., 2011; Kurt et al., 2013). In other words, CcaR not only controls the cephamycin C biosynthesis but also the early and late steps of the clavulanic acid biosynthesis.

ClaR belonging to the LysR type transcriptional regulator family is another pathway-specific regulator required for the expression of late genes in the clavulanic acid gene cluster that acts after clavaminic acid formation to produce clavulanic acid (Paradkar et al., 1998; Özcengiz and Demain, 2013). ClaR is a 47 kDa protein having two DNA binding domains in N- and C-termini and is expressed as a monocistronic transcript of 1.7 kb in size. *claR* mutants cannot produce clavulanic acid and accumulates clavaminic acid while cephamycin C production continues even at higher levels (Liras and Rodriguez-Garcia, 2000).

1.5.3. γ-Butyrolactones

γ-butyrolactones, also referred to as “microbial hormones”, are small extracellular autoregulators and function as important signaling molecules in secondary metabolism and sometimes in morphological differentiation of *Streptomyces* (Takano, 2006; Liras et al., 2008). They usually have a 2,3-di-substituted γ-butyrolactone skeleton and can

be generally classified into three categories according to their chemical structures (Figure 1. 13): (1) A-factor type, possessing a 1'-keto group; (2) virginiae butanolide (VB) type, possessing a 1'-α-hydroxyl group; and (3) IM-2 type, possessing a 1'-β-hydroxyl group (Choi et al., 2003). They can diffuse in and out of the cells being efficacious even at nanomolar concentrations (Hsiao et al. 2009) and when their cognate receptors are present, they act as signaling molecules (Choi et al., 2003). Butyrolactone receptors are dimeric DNA binding proteins which are bound to specific DNA sequences, so called ARE boxes, in the absence of γ-butyrolactones (Choi et al., 2003; Santamarta et al., 2005). Generally, when these autoregulators bind to the receptors which usually act as repressors, receptors dissociate from the target DNA sequences making them available for transcriptional machinery (Choi et al., 2003; Özcengiz and Demain, 2013). For example, Brp protein, γ-butyrolactone receptor, was identified in *S .clavuligerus* and shown to have inhibitory effect on the cephamycin and clavulanic acid production (Kim et al., 2004, Santamarta et al., 2005). There are two different ARE boxes that Brp could bind in *S. clavuligerus* (Santamarta et al., 2005). ARE boxes are palindromic inverted repeats of 22-26 bp located upstream of the genes to be regulated (Santamarta et al., 2005; Liras et al., 2008). One is located in its own promoter region indicating an autoregulatory control and the second one is located 815 nt upstream of the *ccaR* transcription start point, that is, 75 nt upstream of the *ccaR* ATG codon (Wang et al., 2004). Furthermore, it was also shown that there was another protein that binds the same ARE sequence upstream the *ccaR* gene (Santamarta et al., 2007). This protein is called AreB encoded by *areB* gene and belongs to IclR family of regulatory proteins. IclR family proteins can be repressors or activators and can bind the target sequences on DNA as dimers or tetramers (Santamarta et al., 2007). *areB* gene is expressed from a divergent promotor, in the opposite site of which resides *leuCD* cluster responsible for leucine biosynthesis. AreB has autoregulatory effect on its own transcription and also regulates the transcription of *leuCD* gene cluster. Moreover, AreB requires the presence of a small molecule in order to bind the ARE sequences upstream of the *ccaR* gene but does not need it for the interaction with *areB-leuCD* intergenic region. This shows the differential effect

of AreB on antibiotic production. Moreover, AreB is necessary for leucine metabolism as well as for fatty acid catabolism as carbon source (Liras et al., 2008).

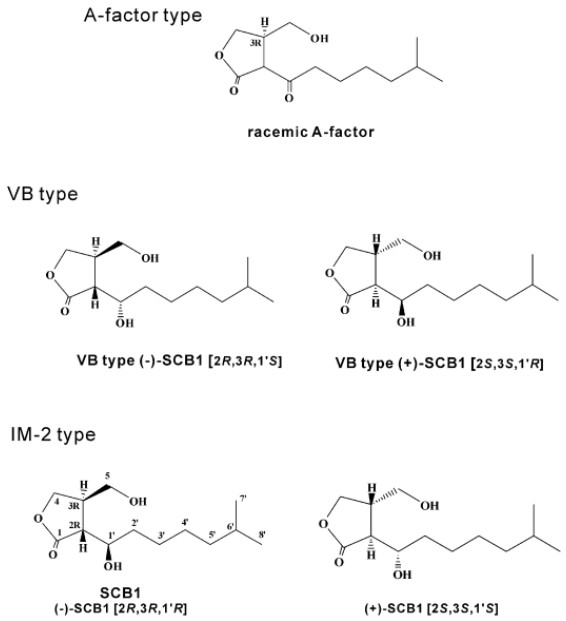


Figure 1. 13. Types of γ-butyrolactones (Choi et al., 2003).

A-factor (2-isocaproyl-3R-hydroxymethyl-γ-butyrolactone) is a well-known γ-butyrolactone in regulating the streptomycin production and cell differentiation in *S. griseus* (Ohnishi et al., 1999). It is assumed to act as a signaling molecule not only for the separate cells in a single hypha but also between the neighboring hyphae inducing antibiotic production and cell differentiation within a short time (Figure 1. 14), thus the "microbial hormone" term best fits to the function of A-factor (Ohnishi et al., 2005; Horinouchi and Beppu, 2007). Concentration of A-factor increases in the cell in a growth-dependent manner (Ohnishi et al., 2005). *afsA* gene encodes an A-factor biosynthetic enzyme and after A-factor reaches the threshold value in nanomolar concentrations, it binds A-factor receptor ArpA (Ohnishi et al., 2005; Horinouchi and

Beppu, 2007). ArpA is bound to ARE boxes, as mentioned earlier. Upon binding of A-factor, ArpA dissociates from the DNA so that *adpA* is freed to be transcribed. *adpA* encodes a transcriptional activator protein AdpA which binds operators of several genes as such forming a regulon. AdpA regulon regulates the transcription of several genes ranging from *strR,* which is a transcriptional activator activating the streptomycin biosynthesis gene cluster, to the genes involved in cellular differentiation (Horinouchi and Beppu, 2007).

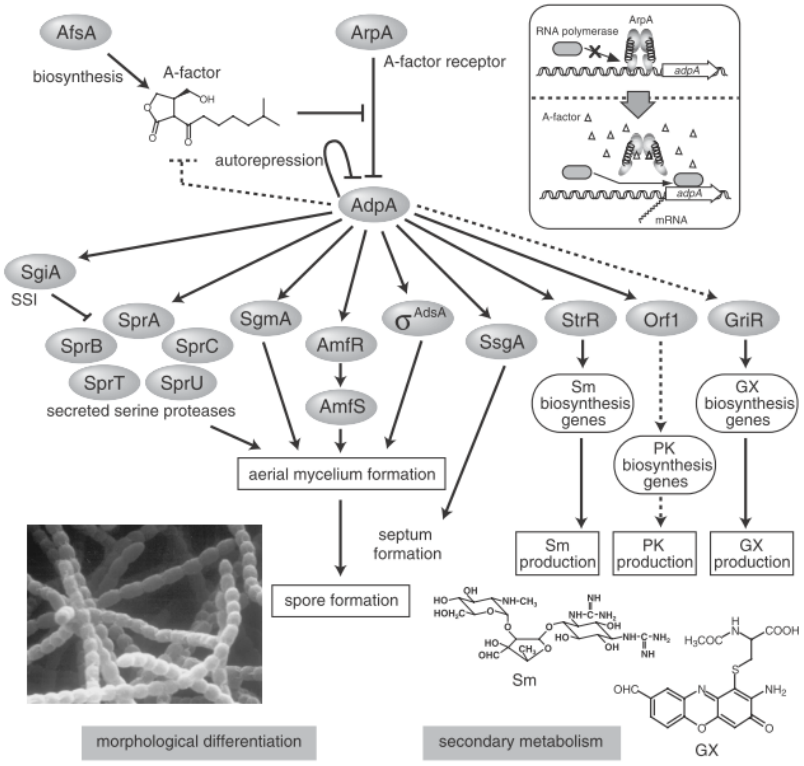


Figure 1. 14. A-factor-induced regulatory pathways that lead to secondary and morphological development (Horinouchi and Beppu, 2007).

1.5.4. Stringent Response

Under severe nutrient limitation, especially amino acid starvation, bacterial cells resort to a type of protection system called stringent response, which includes highly phosphorylated guanosine nucleotides ppGpp and pppGpp as the main actors (Cashel and Kabalcher, 1970; Haseltine et al., 1972). Stringent response mechanism has been first observed in enterobacteria and depended on two proteins: RelA protein, a (p)ppGpp synthetase encoded by *relA* gene and SpoT protein showing (p)ppGpp 3′-pyrophosphohydrolase activity and encoded by *spoT* gene (Figure 1. 15) (Sy, 1977; Gomez-Escribano et al., 2008). Under nutrient starvation, level of (p)ppGpp (stringent factor) molecules increase in the cell (Jain et al., 2006). RelA protein is a ribosome associated protein and is activated when uncharged tRNAs enter the A-site of the ribosomes (Cashel et al., 1996). As a result, (p)ppGpp is produced by the transfer of a pyrophosphate moiety from ATP to the 3' hydroxyl position of the ribose of GDP/GTP by the act of RelA (Figure 1. 15) (Jain et al., 2006). SpoT protein has hydrolase activity and turns (p)ppGpp to GTP or GDP, as such both proteins adjust the intracellular levels of the stringent factors under stress conditions (Jain et al., 2006). Production of (p)ppGpp results in this factor's binding the RNA polymerases and re-directing the gene transcription such that biosynthesis of proteins from many genes involved in growth of the bacteria decreases whereas transcription through genes encoding proteins related to stationary phase or physiological stress response are enhanced (Magnusson et al., 2005; Gomez-Escribano et al., 2008).

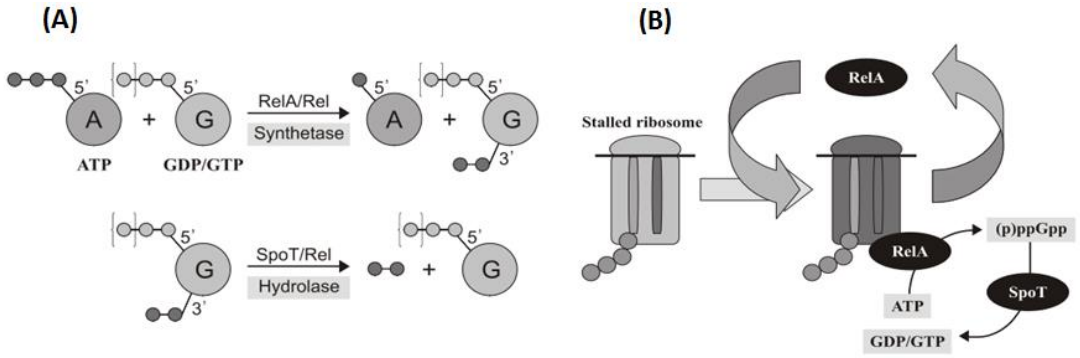


Figure 1. 15. (A). Synthesis and hydrolysis reactions of (p)ppGpp by Rel or SpoT/Rel. (B) Action mechanism of Rel or RelA and SpoT. When an uncharged tRNA enters during translation, Rel or RelA binds to the ribosome, resulting in the synthesis of (p)ppGpp. SpoT catalyzes the degradation of (p)ppGpp when stringent conditions are relieved (Jain et al., 2006).

In *Streptomyces*, the first stringent response was identified in *Streptomyces hygroscopicus* (Riesenberg et al., 1984). While in enterobacteria, there are two related but functionally different proteins (RelA and SpoT) responsible for stringent response, there is one bifunctional RelA/SpoT protein in Gram-postive bacteria and *Streptomyces* (Martínez-Costa et al., 1996; Wendrich and Marahiel, 1997). *relA*-null mutants in *S. coelicolor* showed impaired actinorhodin production and undecylprodigiosin production under nitrogen limitation (Martínez-Costa et al., 1996; Chakraburtty and Bibb, 1997). On the other hand, in case of *S. clavuligerus*, there has been some contradicting reports. Jin et al. (2004) constructed a *S. clavuligerus relA*::*hyg* strain by inserting hygromycin resistance genes into *relA* gene while Gomez-Escribano et al. (2008) produced two mutant strains (*S. clavuligerus relA*::*neo* or *S. clavuligerus ΔrelA*), both of which cannot synthesize ppGpp. Gomez-Escribano et al. (2008) showed an increase in clavulanic acid and cephamycin C production in these mutants compared to wild type strain while results of Jin et al. (2004) revealed that *S. clavuligerus relA::hyg* could produce neither of these antibiotics. Subculturing techniques as well as the intergenic sequence differences upstream of *relA* in the parental strains that each group used were thought to account for the conflict (Liras et al., 2008).

1.5.5. Two Component System Regulators

Another fundamental mechanism that controls both the secondary metabolism and differentiation in *Streptomyces* is two component system (TCS). Under unfavorable environmental conditions such as nutritional changes or other stress factors, these organisms trigger two component signal transduction pathways that will result in TCS regulators' binding the pathway-specific regulatory gene promoters to activate them or TCSs affecting other regulatory pathways indirectly (Martín and Liras, 2010; Rodríguez et al, 2013). However, only a few of TCSs were shown to bind directly to pathway specific regulator gene promoters in *S. coelicolor* (Rodríguez et al., 2013). Hutchings et al. (2004) showed the presence of 67 paired system of TCSs consisting of a response regulator (RR) and a sensor kinase (SK) acting in concert while there were 13 orphan RR genes and 17 orphan SK genes which comprise the orphan two component system proteins within *S. coelicolor* genome sequence. Classical TCS (paired system) consists of a membrane-bound sensor kinase (SK) induced by the environmental cues and a cognate regulator; i.e., response regulator (RR) that is stimulated by the SK to mediate the cellular response usually by regulating the transcriptional activation of specific genes (Mascher et al., 2006; Hutchings et al., 2004). However, in case of orphan TCS proteins, an SK is not accompanied by an RR or RR is the one that is not accompanied by a SK. In contrast to classical TCS, orphan TCS system can use quite unique mechanisms to change the cellular responses against environmental stimuli (Raghavan and Groisman, 2010). TCS protein coding genes usually reside in the core regions of the *Streptomyces* chromosome, not in the instable arm regions (Martín and Liras, 2010).

The TCS PhoP/R is one of the major signal transduction system in *Streptomyces* coming into play under phosphate limitation in the environment. It consists of PhoR protein being the sensor kinase and PhoP protein being the response regulator. Mainly, under inorganic phosphate limitation, PhoR phosphorylates PhoP, and then activated PhoP binds to specific sequences called PHO boxes. One of these PHO boxes is located in the promoter of the polymerase omega factor gene *rpoZ* which is required for both the undecylprodigiosin and actinorhodin production and the morphological

differentiation in *S. coelicolor* (Santos-Beneit et al., 2011a). Furthermore, PhoP was also shown to bind to the promoter region of *afsS* gene, exerting transcriptional repression on it, thus negatively affecting the actinorhodin and undecylprodigiosin production in *S coelicolor* (Figure 1. 16) (Horinouchi, 2003; Rodríguez et al., 2013).

Another system being TCS AfsQ1/Q2 is activated under nitrogen starvation conditions. In the system, AfsQ2 is the sensor kinase and AfsQ1 is the response regulator. AfsQ1 regulator was found to have binding sequences on the promoter regions of *actII-ORF4*, *cdaR*, and *redZ,* thus can directly activate the actinorhodin, calcium-dependent antibiotic and undecylprodigiosin production, respectively. *S. coelicolor* (Shu et al., 2009; Wang et al., 2013). Moreover, AfsQ1 activates the transcription of *sigQ* which is a negative regulator of antibiotic production, most probably by binding directly to the intergenic region between *afsQ1* and *sigQ* thereby controlling the antibiotic regulation by the help of both positive and negative regulators (Wang et al., 2013). Furthermore, it was also shown that AfsQ1/Q2 regulates the transcription of genes related with nitrogen assimilation and phosphorus and carbon uptake as well as morphological differentiation (Figure 1. 16) (Martín et al., 2011).

There are also TCSs that can act as repressors of the antibiotic production. One of them is DraR/K system. *ΔdraRK*, *ΔdraR*, *ΔdraK* mutants of *S. coelicolor* showed increased levels of undecylprodigiosin (RED) in minimal medium (MM) supplied with high concentration of nitrogen sources but decreased levels of actinorhodin (ACT) under the same conditions indicating that TCSs can also have repressive role in antibiotic production under some circumstances (Figure 1. 16) (Yu et al., 2012).

Although it is possible to divide the control mechanisms of secondary metabolism in *Streptomyces* into different categories, there is no way to separate them with certain boundaries; all are somewhat interconnected and intertwined with each other that the regulation of the secondary metabolism is the collective and synchronous work of many regulatory mechanisms.

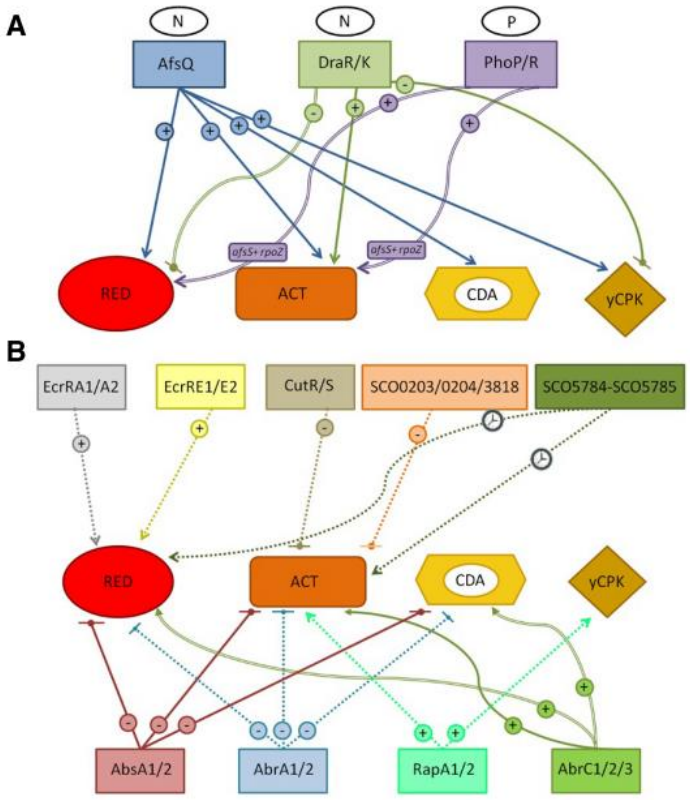


Figure 1. 16. Overview showing how the antibiotic production is regulated by TCSs with nitrogen (N) or phosphate (P) signals in *S. coelicolor*. (Rodríguez et al. 2013).

1.6. Proteomics

Proteomics, as originally coined by Wilkins et al. in 1996, refers to the large-scale study of proteins in a biological compartment, cell, tissue, or even whole organism at a particular time and under certain conditions (Beranova-Giorgianni, 2003; Abdallah et al., 2012). Proteins are the actual players inside the cell. While DNA is a stable set of codes that carry the whole genetic information and thus, it is far from showing the dynamic intracellular changes. mRNA levels, on the other hand, are not much informative about the fate of a protein after translation; for example, whether it undergoes any posttranslational modifications or what kind of protein-protein interactions it performs to regulate cellular responses against certain stimuli, etc. cannot be answered (Beranova-Giorgianni, 2003). However, the proteins harbor essential information that determines the dynamics and the state of the cell such as

abundance, co- or post-translational modification, type of modification, subcellular localization, 3D structure of the proteins and the protein-protein interactions or protein-biomolecule interactions (Griffin and Aebersold, 2001). Therefore, information obtained from proteomics can be applied to many different areas of basic and applied biosciences and medicine. In microbiology, host-pathogen interactions, microbial communities, resistomes and microbial biotechnology are only some of the areas that the proteomics analyses can successfully be applied (Otto et al., 2014).

1.7. Top-Down, Middle-Down and Bottom-Up Proteomics

There are two general approaches used in proteomics for protein characterization (Figure 1. 17 and Figure 1. 18). Most commonly used one is "bottom-up" proteomics (Zhang et al., 2014). It literally means that proteins are fragmented into their corresponding peptides by proteolytic cleavage and then analyzed with MS for the characterization of the proteins (Zhou and Veenstra, 2008; Zhang et al., 2013a). For this, the mixture of the proteins are firstly digested by enzymes such as trypsin, subsequently mass of the peptides are determined by either electrospray ionization (ESI) or matrix-assisted laser desorption/ionization (MALDI) and then these peptides are further fragmented in CID (Collision-Induced Dissociation) chamber as a result of collisions with inert gases like helium or argon to find out their sequence or modifications (Chait, 2006; Zhou and Veenstra, 2008). The bottom-up proteomics can also be referred to as shotgun proteomics, a term coined by Yates laboratory due to the approach being analogous to shotgun genomic sequencing (Yates, 1998). Since bottom-up proteomics necessitates the protein digestion prior to identification and there is the probability that some of the peptide fragments can go undetected, it is hard to identify protein isoforms or modified proteins with this approach. However, 2 DE can be used as a savior in this respect. When it is used as the preferred method in bottom-up proteomics, protein isoforms or modified proteins can be visualized on the gel regarding their pI and mass changes before protein digestion and subsequent identification (Rogowska-Wrzesinska et al., 2013). Therefore, bottom-up strategy have many applications such as proteome profiling, protein quantification, protein modification, and protein–protein interaction (Zhang et al., 2013a).

Top-down proteomics, on the other hand, includes the characterization of intact proteins in contrast to bottom-up proteomics exploiting the peptide fragments for characterization (Zhang et al., 2013a). Intact proteins are ionized and by being subjected to the gas phase they are fragmented in the MS; thereby, masses of both the protein and fragmented peptide ions are determined with this approach (Chait, 2006). Top-down proteomics offers several advantages like the ease of PTM determination over bottom-up proteomics and also makes it possible to reveal the complete primary structure of the protein (Zhang et al., 2013a); however, since it is the intact proteins that are handled, the protein ionization and protein fragmentation in the gas phase have been proven to be difficult (Chait, 2006). Intact proteins with large sizes constitutes another obstruction in characterizing the proteins. Han et al. (2006) exploited a technique that makes it possible to identify proteins with molecular weights of slightly over 200 kDa. However, working with the intact proteins still poses problem especially while handling complex protein mixtures which contain proteins that have highly different physicochemical properties, which makes it challenging to separate them efficiently (Chait, 2006).

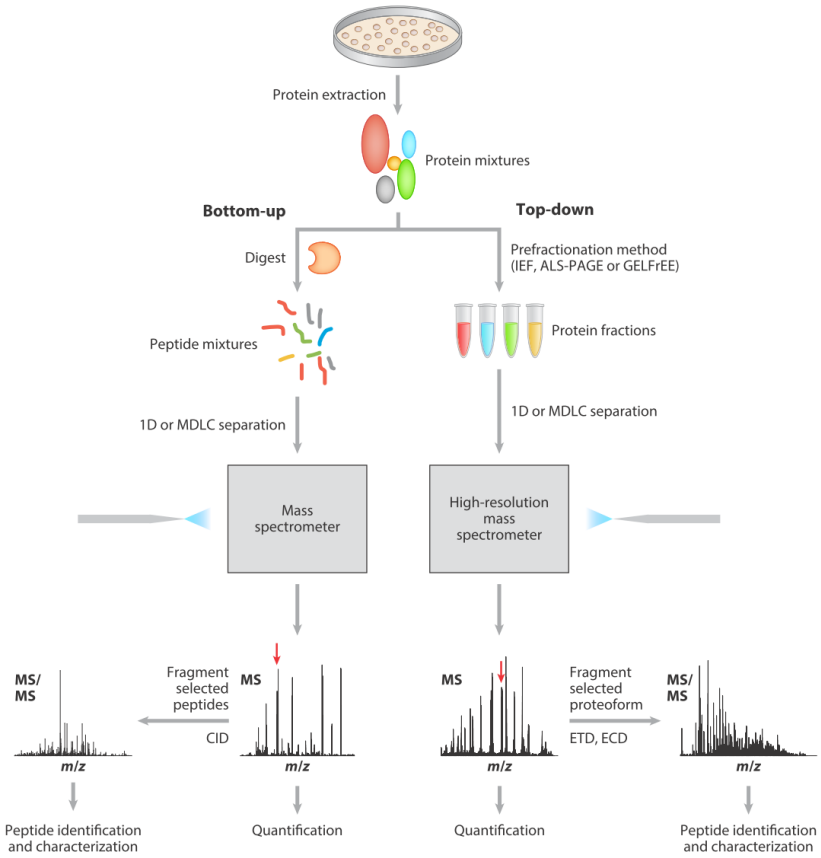


Figure 1. 17. General workflow for MS-based high-throughput bottom-up and top-down proteomics. Abbreviations: ALS-PAGE, acid-labile surfactant–polyacrylamide gel electrophoresis; CID, collision-induced dissociation; ECD, electron capture dissociation; ETD, electron transfer dissociation; GELFrEE, gel-eluted liquid fraction entrapment electrophoresis; IEF, isoelectric focusing; LC, liquid chromatography; MDLC, multidimensional liquid chromatography separation; MS, mass spectrometry. (Zhang et al., 2014).

Recently, another approach gaining more and more popularity has come into light and it is somewhere between the top-down and bottom-up proteomics. It was used by Garcia et al. (2007) for better observation of PTM localization on histone proteins and was named as middle-down proteomics (Figure 1. 18) (Garcia et al., 2007). This approach is based on producing larger peptide fragments by restricted digestion using usually endoproteases like Lys-C or Glu-C and hence ending up with peptide

fragments of 5-15 kDa in size when compared to the proteins of 50 kDa or more analyzed in top-down proteomics (Han et al., 2008; Armirotti and Damonte, 2010). It has proved to have several advantages over top-down proteomics. The intact protein ionization and fragmentation problem is overcome by this approach as well as the sensitivity is better than the top-down being much closer to that of the bottom-up approach (Moradian et al., 2014). Both the top down and middle-down proteomics were shown to provide successful characterization of different PTMs on the proteins (Han et al., 2008).

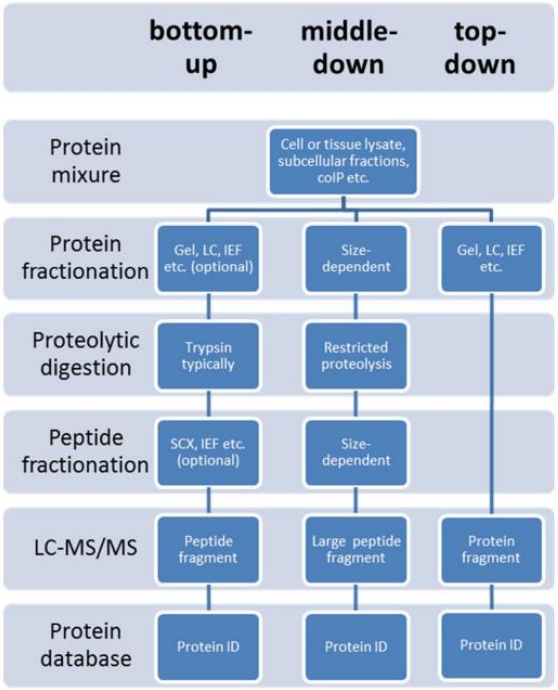


Figure 1. 18. Proteomic strategies: bottom-up vs top-down vs middle-down. (Zhang et al., 2013a).

Quantitation of proteomics is another challenging area and has several options to offer, each having its own advantages and disadvantages. Quantitative proteomics can be grouped as gel-based or gel-free and as label-free and label-based methods. Label-

based methods have sub-categories depending on the labelling strategy such as chemical or metabolic labeling. (Abdallah et al., 2012).

1.8. Gel-Based Proteomics

In gel-based proteomics, a mixture of proteins or peptides extracted from a sample (tissue, cell, body fluid, microorganism, etc.) is loaded into a gel matrix and separated according to their electrophoretic mobility. Electrophoretic mobility depends on the physicochemical properties of the proteins such as surface charge and molecular size plus the electrophoretic properties such as pore size of the matrix and applied voltage. With gel-based methods, it is possible to retain the several properties of the proteins so it gives us crucial information about the properties of them. For example pI, Mw, PTMs, primary structure and even isoforms of the protein are maintained (Wöhlbrand et al., 2013). Polyacrylamide gel electrophoresis (PAGE) has been the most commonly used technique for the separation of the proteins since 1959 (Raymond & Weintraub, 1959).

1.8.1. SDS-PAGE Gel Electrophoresis (1DE)

SDS (sodium dodecyl sulphate), a strong anionic detergent, is used to solubilize proteins, especially integral membrane proteins which are considerably hard to solubilize, by masking the proteins with a constant negative charge per mass unit of 1.4 g SDS per g protein and hence disrupting the secondary and tertiary structure completely. When a certain voltage is applied to the gel system, proteins are separated according to their molecular weight (Westermeier, 2006; Wöhlbrand et al., 2013). 1-DE is usually coupled to LC (liquid chromatography) before MS. As a preparation for the LC-MS, SDS gel containing the lane with separated proteins is excised and cut into several slices, then each slice is subjected to proteolytic digestion. Henceafter, peptide fragments can be passed through LC and subsequently analyzed by MS/MS (Abdallah et al., 2012).

1.8.2. Blue Native (BN) PAGE

Membrane protein isolation can be problematic due to their propensity to form aggregates -even when they are treated with detergents- as well as the possibility of

denaturation especially if there are some labile subunits of the multiprotein complexes cleaved during the isolation procedure (Schägger, 1994). Therefore, to isolate membrane proteins, Native Blue PAGE has been the most preferred method after developed by Schägger and von Jagow (1991) (Westermeier, 2006). In this technique, proteins are first separated under native conditions using mild nonionic-detergents (Triton X-100, digitonin, etc.), thus their 3D structure is preserved as in *in vivo* (Abdallah et al., 2012; Wöhlbrand et al., 2013). Coomassie Blue G-250 in cathodal buffer is added into the vertical gel chamber and Coomassie Blue dye competes with nonionic detergents and replaces them, rendering the intact proteins negative charges (Westermeier, 2006). As a result, proteins migrate in the gel according to their Mw as well as shape. After the first dimension, the gel lane is excised and inserted just on the top of a classic SDS-PAGE in a perpendicular axis compared to the first separation axis. Now denatured, proteins and their components run according to their Mw on the second dimension.

BN-PAGE method is usually linked with hydrophobic membrane proteins in mitochondria, plant membranes and chloroplasts and also useful for determining the subunit composition of the proteins, molecular weight and native state of proteins and complexes (Schägger, 1994; Abdallah et al., 2012).

1.8.3. Two-Dimensional Polyacrylamide Gel Electrophoresis (2D SDS-PAGE)

2D SDS-PAGE (or simply 2DE) is also known as the classical proteomic method. It was first invented by O'Farrel in 1975 (O'Farrell, 1975; Otto et al., 2014) and since then it has been the main component of the gel-based proteomics especially used for the determination of differential protein expression (Abdallah et al., 2012). It consists of two dimensions, both of which exploit the physicochemical properties of the protein. The first dimension separates the proteins according to their isoelectric points (pI). Proteins are amphoteric molecules carrying negative, positive and zero net charge. Zero net charge is obtained when the sum of positive and negative charges are equal to zero at a specific pH. Thus the pH representing the zero net charge on a protein is called pI and pI values for each protein are unique (Chevalier, 2010). In the first dimension, also known as IEF (Isoelectric Focusing), a gel matrix consisting of a

variable pH range along the gel is used and this pH gradient is generated by the use of carrier ampholytes which are highly mobile, low molecular-weight organic molecules with high buffering capacity (Westermeier, 2006); however, nowadays commercial IPG (Immobilized pH gradient) strips are commonly used instead of the lab-made gel matrices using carrier ampholytes in order to increase reproducibility and resolution. In order to produce IPG strips, amphoteric acrylamido derivatives (so-called immobilines) are co-polymerized with conventional acrylamide on thin films (Görg, et al., 2004). When a certain electric field is applied to the IPG strip loaded with the protein mixture, proteins will begin to migrate along the pH gradient on the gel. When the net charge of the proteins turns to zero; in other words, they reach their isoelectric point at a certain pH, they stop migration, that is, they are focused. Thus, at the first dimension (IEF), proteins are separated according to their pI values and focused. Commercial IPG strips differ in separation distances (7-24 cm) as well as pH gradients. For example, as for pH gradients, pH 4-5 is a narrow range, pH 4-7 or pH 5-11 is a medium range and pH 3-11 represents a broad range IPG strip. Medium to broad range strips are suitable for samples of medium-level complexity like bacterial proteomes while for more complex samples like eukaryotic proteomes, several medium to narrow range strips can be used to cover or catch enough proteins in the whole proteome (Hoving et al., 2000; Görg et al., 2004; Görg et al., 2009; Wöhlbrand et al., 2013).

After the first dimension, strips are equilibrated with SDS buffer so that all the proteins are denatured and masked with the negative charges (Westermeier, 2008). Therefore, the only physicochemical property that can differentiate the proteins in the second dimension can be the molecular weight of the proteins. Thus, in the second dimension, classical SDS-PAGE is performed to separate the proteins according to their Mw. Gels can be prepared in horizontal or vertical systems (Görg et al., 1995). Vertical systems are preferred for running more than one gel in parallel (Wöhlbrand et al., 2013).

Ever after completing the first and second dimensional separation of the proteins, the procedure is far from complete since the proteins that are run on the gel are also to be visually observed and then analyzed by MS. So many methods have been exploited for the observation of these proteins, and their sensitivity, reproducibility and

quantitative accuracy are important parameters to be considered while choosing the visualization method. However, none of the methods can satisfy all these parameters at the same time (Görg et al., 2004). Common staining methods include Coomassie staining using Coomassie Brilliant Blue dye, silver staining, fluorescent dyes, such as Sypro® Ruby (Molecular Probes®) or Deep Purple TM (GE Healthcare), or radioactive labeling (Curreem et al., 2012; Dhingra et al., 2005). Although being the most common staining method, colloidal Coomassie staning has the least sensitivity, ranging from 30-100 ng of protein per spot (Neuhoff et al., 1985; Anderson et al., 1991; Westermier, 2006); nonetheless, it requires a simple staining procedure and gives high quantitation and is completely compatible with mass spectrometric analyses (Dyballa and Metzger, 2009) apart from being reproducible, providing a clear background with an average sensitivity (Wang et al., 2007). Coomassie Brilliant Blue used for Coomassie staining is an anionic triphenylmethane dye and has two modified versions Coomassie R-250 (Red tint) and Coomassie G-250 (Green tint). In acidic environment, dyes can can bind to the amino groups of the proteins by both electrostatic and hydrophobic interactions, as a result of this non-specifical binding, an unbiased quantification of the protein content on the gel is possible (Westermeier and Marouga, 2005). Neuhoff et al. (1985, 1988) established Colloidal Coomassie Blue staining method in which by decreasing the pH and increasing the colloidal form of the dye to a certain limit, they increased the sensitivity, reproducibility and also what is lacking in previous procedures a clear background. This method has been the cornerstone of nearly all CBB staining methods and is still being exercised with some modifications and improvements.

Another technique with a higher sensitivity is silver staining. It is possible to detect 2-5 ng protein/spot with this technique (Shevchenko et al., 1996; Dhingra et al., 2005). The application of silver staining in polyacrylamide gels was introduced by Merril et al. (1979) and since then, there have been around 100 modifications on the technique to increase the sensitivity, compatibility, etc. (Westermeier and Marouga, 2005). Although there are so many modifications, two types of silver staining methods are the most common to be used for 2D gels: silver nitrate and silver diamine procedures (Westermeier, 2006). Basically silver staining relies on the autocatalytic reduction of

the silver ions to elemental silver and the presence of the proteins facilitates the reduction and thus the silver precipitates forming spots from dark brown to black on a yellowish to light brown background (Westermeier and Marouga, 2005; Wölhbrand et al., 2013). However, silver staining is not MS compatible since glutaraldehyde used in the procedure cross-links with the polypeptide chains irreversibly thereby hindering the effective tryptic digestion of proteins before MS analyses (Candiano et al., 2004; Wang et al., 2007). It is possible to exclude glutaraldehyde during silver staining procedures at the expense of sensitivity, though. Despite being quite sensitive, there are some other drawbacks of silver staining such as decreased reproducibility compared to CBB and also protein to protein variability which makes it less suitable especially for quantification analyses (Dhingra et al., 2005).

1.8.3.1. Image Acquisition and Evaluation

2D gel electrophoresis is especially used for examining differential protein expression profiles between an experimental and a control group. Depending on the size and the intensity of the protein spot on the gel, it is possible to determine if there is an increased, or decreased level of protein and also disappearance or appearance of it among the groups. The identification of the protein of interest can then be carried out by MS (Görg et al., 2004). For this, stained gels are first scanned with special scanners providing high resolution and homogenous background throughout the gel (Westermeier, 2006). Secondly, the scanned images are imported to specific 2D image-analysis software programs such as DeCyder (GE Healthcare), Proteomweaver (Bio-Rad), PDQuest (Bio-Rad), Delta2D (DECODON). What all these software programs do is basically to align the gel images, match the spots, calculate the spot intensities by removing the noise, and providing normalization and background adjustment. Peak intensities coming from each spot, the boundaries of which are set mainly by the software, are used for the quantification (Abdallah et al., 2012). However, image analysis is the most challenging part of the 2D-PAGE since it is usually not fully automated and user interference is required for better aligning, determination of the spot boundaries, separation or merging of the spots. After handling the image analysis, the program gives a quantitation table showing volumes

of spot areas as well as the differential expression patterns between the spots. Most of the image-analysis softwares also include statistical tools so that the significance of the differential expression observed for a protein or spot can be further checked on the program (Westermeier, 2006; Westermeier et al., 2008).

1.8.3.2. Pros and Cons of 2DE

2D gels have the power to resolve up to 5000 proteins depending on the length and pI of the IPG strip used (Görg et al., 2004). Moreover, they give direct information about the intact proteins, making them an important top-down proteomic method. For example, with 2DE, comparative analyses can be successfully performed as in the case of treated vs untreated or control vs experimental groups to find out the differential expression of proteins. Thus, it gives direct information about the quantitative properties of the protein of interest and also makes possible to visualize the localization of the protein on the gel. Thus, experimental pI and Mw of the protein can be determined. After the MS results, theoretical pI and Mw (coming from the protein sequence) can be compared with the experimental results. This opens the way to identify post-translational modifications (PTMs) that change the molecular mass as well as pI of proteins (Rogowska-Wrzesinska et al., 2013).

Although 2DE has all its advantages, it has some drawbacks, as well. First of all, only a limited pH range can be used. Thus, highly acidic or basic proteins might not be determined (Rogowska-Wrzesinska et al., 2013). Moreover, separation of very large proteins and very small proteins are underrepresented on the gel. Not only 2D gels but also other gel-based proteomics methods suffer from low detecting power of membrane proteins. Having hydrophobic components, membrane proteins are hardly soluble in buffers prepared for IEF. Thus, only a few membrane proteins can be visualized on 2D gels. Finally, the abundance of proteins expressed in a cell differs incredibly depending on the protein. While some can be expressed as millions of copies, the others might be only in hundreds. This huge difference in expression patterns can cause underrepresentation of low-abundance proteins on 2D gels (Chevalier, 2010).

2D gel is not an end-point analysis method *per se* and neither is any other proteomics approach. Only when different methods are combined like gel-based methods complemented with gel-free methods, can it be possible to gain the comprehensive information on proteome (Westermeier, 2016).

1.9. Gel-Free Proteomics

Gel-free (MS-based) proteomics is a bottom-up (shotgun) proteomics approach in which proteins are first proteolytically digested to produce peptide fragments which will be subsequently analyzed by MS-based technologies. There are several gel-free strategies to identify the proteins which are generally categorized as label-free and label-based techniques (Wöhlbrand et al., 2013). Method of choice depends on the sample characteristics (Abdallah et al., 2012) and design of the experiment as well as the budget.

1.9.1. Label-Based Methods

Label-based approaches depend on isotopic labels that are either chemical, metabolical or enzymatical labels (Wöhlbrand et al., 2013). The most commonly used label-free methods are Isotope-Coded Affinity Tag (ICAT), ^{14}N/ ^{15}N labelling, Isobaric Tags for Relative and Absolute Quantification (iTRAQ), Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC), ^{18}O /^{16}O enzymatic labeling, Isotope Coded Protein Labeling (ICPL) (Zhu et al., 2009). For example, SILAC method exploits the mass differences coming from light and heavy amino acids incorporated into the proteins during culturing of cell populations. One of the cultures is supplied with heavy amino acids and the other with light amino acids. Thereafter, cell cultures are combined, proteins are extracted and peptide fragments are analyzed by MS. Peak intensity ratios of the peptide are used to find out the differential expression of the proteins (Mann, 2006). There are important advancements in quantitative proteomics analyses in isotope labeling methods, even determination of absolute molar concentration of proteins of interest has been possible by using stable isotope-labeled internal standard peptide sequences (Barnidge et al., 2003; Gerber et al., 2003; Barnidge et al., 2004). However, most of the label-based proteomics has limitations. Requirement of high sample concentration, complex sample preparation, substantial time to be dedicated,

specific quantification software, expensive reagents, and sometimes incomplete labeling are potential drawbacks (Zhu et al., 2009).

1.9.2. Label-Free Methods

As an alternative to label-based approaches, label-free methods can be used. They include straightforward steps such as (i) protein extraction, reduction, alkylation and digestion, (ii) peptide separation by liquid chromatography (LC), (ii) protein preparation like reduction, alkylation and digestion and then analysis by LC/MS-MS, (iii) peptide/protein identification by MS (Zhu et al., 2009). With label-free approaches, it is also possible to study differential proteomics analyses and these analyses cost much less than those of isotope-labelled methods; furthermore, they can be applied to any biological sample and the complexity of the proteomes is lower (Schulze and Usadel, 2010).

Since there is no labeling of proteins in label-free methods, different approaches are exploited for the quantification and comparison of the expression differences between samples. There are two main measurement methods; one of them is ion intensity which is calculated using peptide peak areas or peak heights in chromatography and the second one is the spectral count of the protein identified (Zhu et al., 2009).

1.9.2.1. Quantification by Spectral Counts

Abundance of a protein in a sample correlates with peptides corresponding to that protein identified during MS/MS scanning (Dicker et al., 2010). Therefore, the information coming from the number of peptides that match a certain protein can be interpreted as tandem mass spectra, or spectral count, and spectral counts are positively associated with the protein abundance in a sample (Fu et al., 2008; Choi et al., 2008). Spectral counts especially come into use when expression of proteins in two different samples are to be compared. However, it is not enough just to use the raw data of spectral counts. For comparison purposes, they are to be normalized or modified for more plausible deduction. There are several strategies suggested to make use of spectral counts for reliable quantitation of the proteomics data. One of them is the protein abundance index (PAI) calculation. PAI is defined as the number of observed

peptides divided by the number of observable peptides for each protein (Rappsilber et al., 2002). PAI was further modified by Ishihama et al. (2005) to emPAI in which 1 is subtracted from the exponential form of PAI and the resultant value is directly correlated with the protein abundance (Ishihama et al., 2005). Another strategy is the inclusion of protein length into the calculation. Since as the length of the protein increases, the number of the peptides observed are expected to increase. Taking this into account, NSAF method has been devised. With this technique, spectral count (SpC) of each protein is first divided by the length (L) of the same protein. Then for the normalization of this value (SpC/L), (SpC/L) is divided by the sum of SpC/L values of all proteins in the sample (Equation 1. 1) (Zybailov et al., 2006; Neilson et al., 2011).

$$(\mathrm{NSAF})_k = \frac{(\mathrm{SpC}/L)_k}{\sum_{i=1}^{N}(\mathrm{SpC}/L)_i} \qquad \text{Equation 1. 1}$$

Old et al. (2005) used another technique to compare the spectral counts coming from two separate samples. It depends on Rsc value and is based on the $\log_2$ ratio to find out the abundance of proteins in two samples. This equation necessitates the spectral count for each protein and all the spectral counts coming from sample 1 and 2 (Equation 1.2) (Old et al., 2005).

$$R_{\mathrm{SC}} = \log_2[(n_2 + f)/(n_1 + f)] + \log_2[(t_1 - n_1 + f)/(t_2 - n_2 + f)] \qquad \text{Equation 1. 2}$$

There are several other strategies to quantitate the label-free proteomics results using the spectral counts, as well (Lu et al., 2007; Griffin et al., 2010).

1.10. Mass Spectrometry

Mass spectrometry (MS) is used for protein identification and is the main player of the large-scale proteomics (Yates et al., 2009). It is an analytical technique measuring the mass-to-charge (m/z) of ions produced after the proteins or peptides enter the mass spectrometer (Wysocki et al., 2005; Westermeier et al., 2008). Three main components

are what make a mass spectrometer (Figure 1. 19): an ionization source, a mass analyzer, and an ion detector (Yates, 2004). Very briefly, an ion source is responsible for converting the molecules entering the MS to gas-phase ions, a mass analyzer separates the ions according to their m/z ratio and the number of ions at a specific m/z ratio is detected by a detector (Han et al, 2008).

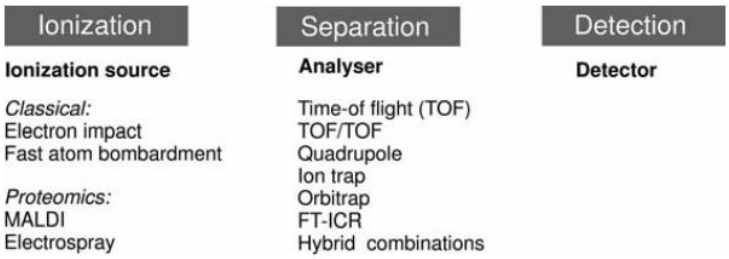


Figure 1. 19. Three components of a mass spectrometer: Ion source, mass analyzer and detector (Westermeier et al., 2008).

1.10.1. Ionization Techniques

Ionization in MS is responsible for converting molecules into ions in gaseous phase so that they can be responsive to electric and magnetic fields in which they will be manipulated (Yates, 2004). Although ionization of biomolecules dates back to 1976, it was not until 1988 when the soft ionization techniques (MALDI, ESI) were introduced into mass spectrometry world. Even though ionization techniques such as plasma desorption and fast atom bombardment (FAB) could ionize biomolecules, proteins are non-volatile, polar and thermally unstable molecules, which makes them harder to turn into gas-phase ions without extensive degradation (Westermeier et al., 2008; Yates et al., 2009). Therefore, introduction of MALDI and ESI after 1988 was the major breakthrough for the protein identification by mass spectrometry, which gave high efficiency of ionization of the proteins (Westermeier et al., 2008).

1.10.1.1 Matrix-Assisted Laser Desorption Ionization (MALDI)

MALDI, first introduced by Karas and Hillenkamp in 1988 (Karas and Hillenkamp, 1988), is still in use and just one of the two "soft ionization" techniques (Mann et al, 2001). The application of MALDI relies on the analyte molecules being mixed with an organic molecule of conjugated aromatic ring structure and hence upon irradiation, they can absorb at a wavelength of a laser (Figure 1. 20) (Mann et al., 2001; Wysocki et al., 2005). These organic molecules forming the matrices are usually composed of two molecules: α-cyano-4-hydroxycinnamic acid or dihydrobenzoic acid (DHB), and the optimal ratio of mixing an analyte with the matrix is in the range of 1:1000 to 1:10000 (analyte to matrix). Upon drying, most of the matrices form polycrystalline structures; furthermore, matrices are capable of absorbing UV light at certain wavelength, usually N_2-laser (λ=337 nm) (Kaufmann, 1995; Yates, 1998; Mann et al., 2001). When they are irradiated with a laser, the laser causes the desorption, subsequently ionizing the analyte and the matrix. However, the energy transferred from α-cyano-4-hydroxycinnamic acid or dihydrobenzoic acid (DHB) to analyte is different during desorption and ionization. For example, α-cyano-4-hydroxycinnamic is "hotter" than the latter, thus DHB is mostly preferred for large proteins, while α-cyano-4-hydroxycinnamic acid is a standard matrix with high sensitivity (Mann et al., 2001; Westermeier et al., 2008). As for the desorption/ionization process, it occurs either via proton loss or gain, even though it is possible for multiple-charged ions to be formed during the process, singly charged ions dominate in MALDI (Wysocki et al., 2005). After the formation of ions, they are accelerated to the MS analyzer (Westermeier et al., 2008).

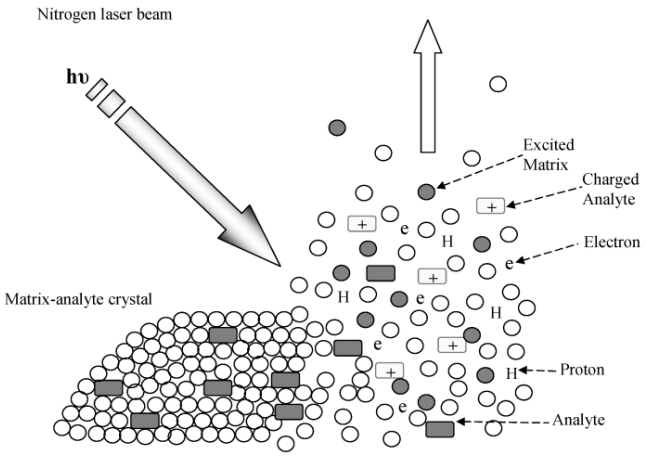


Figure 1. 20. MALDI ionization process (El-Aneed et al., 2009).

1.10.1.2. Electrospray Ionization (ESI)

Electrospray ionization (ESI) technique to be used in mass spectrometers for biological purposes was developed by Fenn et al. (1989). Working principle of ESI is very briefly based on an analyte solution passed through a needle like structure (a metal capillary) to which an electric field is applied. The charged molecules formed inside of the capillary are ejected from the tip of the needle as spray particles and directed towards the mass analyzer (Banerjee and Mazumdar, 2012). However, the mechanism is more complicated than it seems. The more detailed version is as follows: Under an atmospheric pressure, sample dissolved in a polar volatile solvent in dilute concentration (< mM) is passed through a metal capillary and between the capillary and the mass spectrometry inlet is applied an electric field of 2- 6kV (Westermeier et al., 2008; Banerjee and Mazumdar, 2012, Konermann et al., 2012). At the tip of the capillary, solution forms a Taylor cone and is sprayed as fine droplets due to the intense electric field at the tip of the capillary (Figure 1. 21) (Fenn, 2003). Taylor cone is a result of the repulsive coulombic forces observed between the like charges (Westermeier et al., 2008). Furthermore, surrounding the capillary tube is a sheath gas (dry N_2) flowing around the tube, which improves the spray generation (nebulization) and also helps the spray be directed towards the mass spectrometer (Banerjee and

Mazumdar, 2012). After the solution leaves the Taylor cone as spray droplets, they continue to evaporate along the trajectory (Konermann et al., 2012). Here come the concepts known as Rayleigh instability and Rayleigh limit (Fenn, 2003). Evaporation occurs till the droplets reach what is known as Rayleigh limit. According to this, due to the presence of like charges on the droplets, Coulomb repulsion will appear and droplets will undergo fission as a result of this repulsion. More specifically, as the evaporation occurs, the density of the charges on the surface of the droplet will increase, and at a certain critical value (also known as Rayleigh limit), they will repel each other, thus "Coulomb repulsion of the surface charges overcomes the surface tension that tries to contain the droplet liquid in a spherical shape"; as a result, this instability (Rayleigh instability) will generate more and more small-sized droplets (Figure 1. 22) (Fenn, 2003). At a certain point, the charged analytes are freed from the droplets. Usually there are heated capillaries (with a temperature of 100–300°C) lying at the interface of the atmospheric pressure chamber through which droplets are traveling and the vacuum chamber housing the mass spectrometer to complete the evaporation (desolvation) of the ions. After the heat capillary, the analyte ions in gasous phase are now directed toward the mass spectrometer (Banerjee and Mazumdar, 2012). Furthermore, in contrast to MALDI, the charges on the analytes are usually multiple. Thus, the mass range of the analyte is decreased, which facilitates the analysis of large molecular weight proteins (Westermeier et al., 2008)

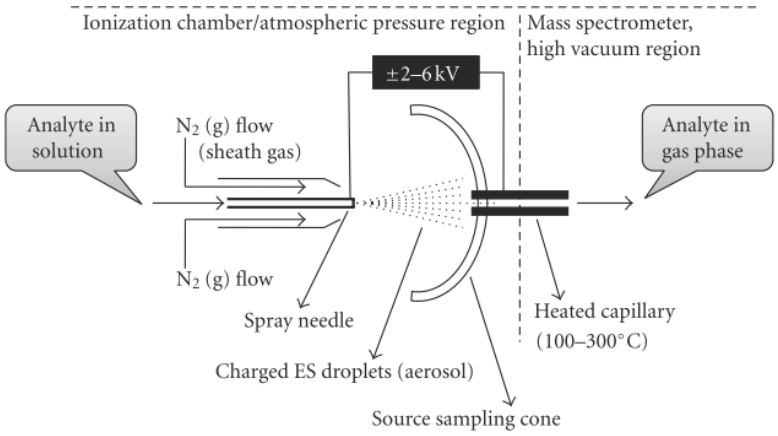


Figure 1. 21. A schematic representation of the ESI-ion source (Banerjee and Mazumdar, 2012).

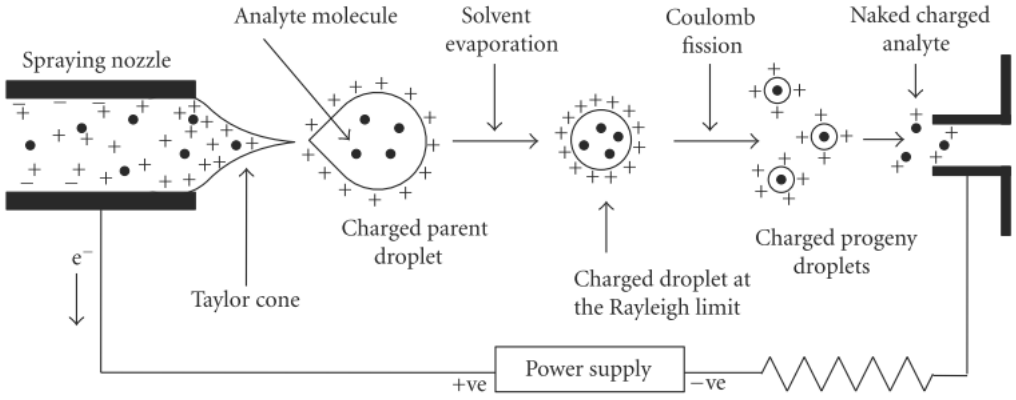


Figure 1. 22. Schematic representation of the electrospray ionization process (Banerjee and Mazumdar, 2012).

1.10.2. Mass Analyzers

The major function of mass spectrometers is to measure the m/z (mass-to-charge) ratio of proteins, peptides or peptide fragments. Mass analyzers are built upon three different separation principles: separation by time-of-flight (TOF-MS), quadrupole

electric fields generated by metal rods (quadrupole MS), or selective ejection of ions from a three-dimensional trapping field (ion trap or Fourier transform ion cyclotron MS). However, these three separate techniques are just for m/z (mass-to-charge) identification; in order to gain a more thorough information on the proteins, peptide sequences are to be found out. This is achieved by using a combination of the two of these three principles mentioned above. This is called tandem mass spectrometry (MS/MS). The fact that MALDI produces short pulses of ions make it convenient to be coupled with TOF analyzers; whereas the fact that ESI produces a continuous spray of droplets make it more suitable for coupling with quadrupole and ion-trapping analyzers (Mann et al., 2001). Depending on these three principles, there arise four general categories of mass analyzers except for their combinations (MS/MS): Quadrupole mass analyzers (Q), Time-of-flight (TOF), Quadrupole ion traps (QIT), Fourier transform ion cyclotron resonance (FT-ICR) (Wysocki et al., 2005).

1.10.2.1. Quadrupole (Q) Mass Analyzers

Quadrupole analyzers are composed of four parallel rods, through two of which direct current (DC) potential passes and from the rest of which an alternating radio-frequency (rf) potential passes. After the ions are formed in the ionizer component of the mass spectrometer, they are directed to a quadrupole by an electric field. Since each couple of rods have rf or DC potentials, the ions entering the quadrupole analyzer will follow an oscillating trajectory. Some of them depending on their mass/charge ratio will enter the detector following the oscillating trajectory, whereas the others with m/z ratios not appropriate for the applied potential will end up with colliding one of the rods. Therefore, by changing the applied potential across the rods, it is possible to select a certain m/z ratio for ions to pass through the analyzer and thus to be detected (El-Aneed et al., 2009). Though quadrupole analyzers occupy a small area, they are cheap and their maintenance is easy, they limited range of m/z to be detected from 0 to 4000 m/z (Figure 1.23) (El-Aneed et al., 2009; Wysocki et al., 2005).

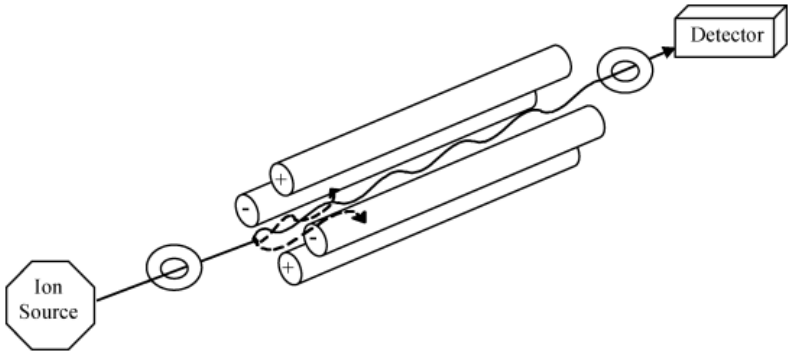


Figure 1. 23. Representation of a quadrupole mass analyzer. Four parallel electrical rods with varying direct current and alternating radio-frequency potentials. Only one m/z value will possess the "right" trajectory and survive the path to reach the detector (solid line). The rest will collide with rods and will be ejected (dotted line). (El-Aneed et al., 2009).

However, it was shown that if a quadrupole is coupled to two other quadrupoles, also known as triple quadrupole analyzers, it can be used as tandem mass spectrometry (MS/MS), first and third quadrupole acting as mass filters and second as a collision cell from which the structural information can be driven (Figure 1. 24) (Westermeier et al., 2008).

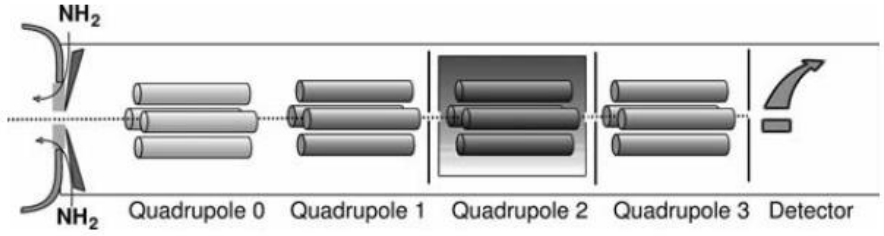


Figure 1. 24. Schematic of a triple quadrupole analyzer. (Westermeier et al., 2008).

1.10.2.2. Time-of-Flight (TOF) Mass Analyzers

The working mechanism behind the TOF analyzers depends on the free flight of ions coming from the ionization source in a 1-2 m-long tube before being detected by a detector. As an example, two ions formed at the same time and carrying the same charge but with different masses will travel through tube in the analyzer but the one with the lower mass will reach the detector earlier than the one with the higher mass. Hence, all the ions that are formed in the ionization chamber will be detected by TOF analyzers in contrast to the quadrupole analyzers which can detect ions with a limited m/z range (El-Aneed et al., 2009). Linear TOF analyzers has two main components: ion acceleration region and free-flight region also called a drift tube (Figure 1. 25) (Merchant and Weinberger, 2000). In the ion acceleration tube, there are also two elements: repeller element and a ground aperture. Some potential is given to the repeller plate while ground aperture has a ground level of potential. As a result, an electric field is generated along the ion acceleration region. Given a certain energy, ions entering the ion accelerator will be accelerated and enter the drift tube. They continue to move through the drift tube until they reach the detector (Merchant and Weinberger, 2000). Time of flight for each ion can be calculated according to the Equation 1.3 below (El-Aneed et al., 2009):

$$tf^2 = \left(\frac{m}{z}\right).\frac{(2s+x)^2}{2Es} \qquad \text{Equation 1. 3}$$

t_f= time of flight, s= the length of the ion acceleration region, x= the length of the free flight region, E= voltage applied

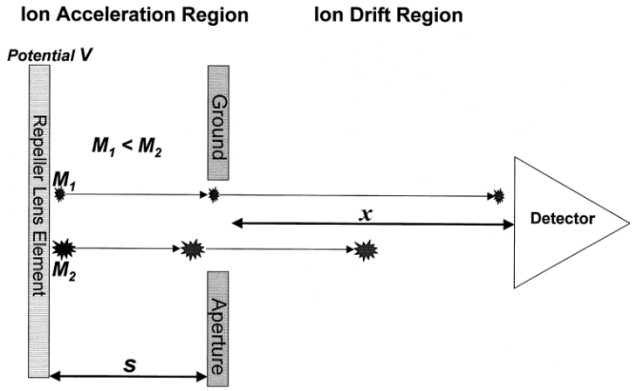


Figure 1. 25. A simple linear TOF-mass spectrometer. Two ions with masses of "M_1" and "M_2", where "M_2" is heavier than "M_1" are accelerated to constant energy. "V" is the acceleration potential, "s" is the acceleration distance, and x is the free flight distance (Merchant and Weinberger, 2000).

Considering that distance and voltage applied are constant values, t_f is directly related with the mass-to-charge ratio of the ions (Westermeier et al., 2008). However, initial kinetic energies of ions entering the TOF analyzer from ionization source can be different, thus as a disadvantage of linear TOF analyzers, -in contrast to what is expected-, ions with the same masses may not strike the detector at the same time and hence, linear TOF analyzers have a low resolving power as well as low mass accuracy. In order to eliminate this problem, reflectron (electrostatic ion mirror) technology has been incorporated into linear TOF analyzers. Reflectrons, as the name suggests, changes the trajectory that the ions in the drift tube are following (Figure 1. 26), it is a kind of an ion mirror to which the ions with higher kinetic energies penetrate deeper and thereby get slowed down; furthermore, by the change of trajectory, ions follow a longer path to reach the detector; both conditions increase the resolution and the mass accuracy of the TOF analyzers considerably (Cotter, 1999; Westermeier et al., 2008; El-Aneed et al., 2009).

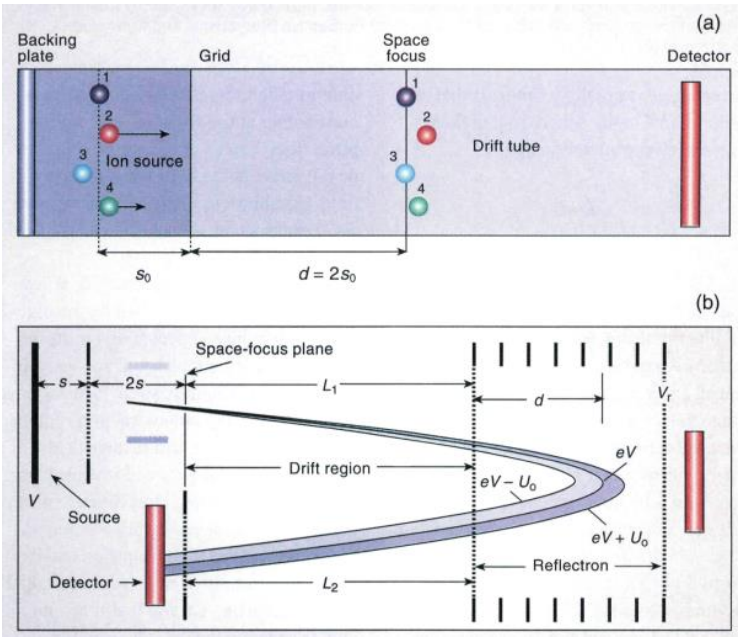


Figure 1. 26. TOF mass spectrometer designs. (a) Linear TOF mass spectrometer showing ions formed with different initial kinetic energies, ions formed at different locations in the ion source, and ions focused at the space-focus plane, (b) Single-stage reflectron mass spectrometer in which the ions are retarded, turned around, and focused by a constant electric field constructed from a series of lenses with linearly increasing voltages (Cotter, 1999).

1.11. The Aim of the Present Study

S. clavuligerus produces several antibiotics, the most important ones being Cephamycin C (CC) and Clavulanic acid (CA), as explained earlier. Using genetic/metabolic engineering approaches, two different CC overproducers, namely *S. clavuligerus* AK39 and TB3585 had been constructed in our laboratory. In the present study, these two strains as well as an industrial strain of *S. clavuligerus*, namely DEPA, which was developed through successive random mutagenesis and selection programs and used for commercial CA production for many years by DEPA Pharmaceuticals Co. (İzmit, Turkey) were subjected to proteomic analyses to elucidate their differential protein expression profiles when compared to that of the standard strain *S. clavuligerus* NRRL 3585. Two proteomics approaches, nongel-based LC-MS/MS and 2DE coupled

to MALDI-TOF/MS, were exploited for gaining a comprehensive understanding of global changes in protein profiles of these strains during the stationary phase at which morphological differentiation and secondary metabolism predominate. By this study, it was shown that targeted and random mutants of *S. clavuligerus* display a number of markedly different responses in many functional categories like carbohydrate and amino acid metabolism (primary metabolism), translation, secondary metabolism, global regulation, stress responses, etc. There were also up- and downregulations of many hypothetical and unknown proteins that remain to be characterized in order to enlighten the complex physiological activities and networks of the strains. With some potentially crucial changes in the levels of many proteins, the present work shed light on proteome-wide alterations that could contribute to increased production of CC and CA.

CHAPTER 2

MATERIALS AND METHODS

2.1. Bacterial Strains and Plasmids

The strains of *S. clavuligerus* which were used throughout the study are given in the Table 2.1.

Table 2. 1. Bacterial strains and their properties

S. clavuligerus **Strains**	**Description**	**Source or Reference**
NRRL 3585	Wild type, cephamycin C and clavulanic acid producer	Higgens and Kastner (1971)
AK39	Null mutant of *S. clavuligerus* NRRL 3585, *hom::kan*	Yılmaz et al. (2008)
TB3585	*S. clavuligerus* NRRL 3585 carrying pTB486, $Thio^R$	Özcengiz et al. (2010)
Industrial strain DEPA	Clavulanic acid overproducer	DEPA Pharmaceuticals Co., İzmit, Turkey

TB3585 contains multiple copies of the *ask* gene and AK39 is a *hom*-deleted mutant of NRRL 3585.

2.2. Media and Culture Conditions

S. clavuligerus strains were grown at 28°C at 220 rpm in baffled flasks in Tryptic Soya Broth (TSB) for stock preparation. For TB3585, the medium was supplemented with 50 µg/mL thiostrepton and for AK39, the medium was supplemented with 200 µg/mL kanamycin.

Seed cultures were prepared by adding 200-600 µL of *S. clavuligerus* stock cultures into 50 mL of TSB medium and incubated at 28 °C at 220 rpm for 24-48 h. Growth was determined by optical density measurement of the seed cultures, which was adapted from Malmberg et al. (1993). 3 mL of distilled water and 0.5 mL of 2.5 M HCl were mixed with 0.5 mL of sample from seed culture and then subjected to sonication (Ultrasonic Processor, Cole Parmer) at 50% amplititude for 3x30 seconds. When the optical density measured at 600 nm reached 0.7-0.8, 25 mL of seed culture was precipitated by centrifugation at 4000 g for 10 min at 4 °C, and thereafter pellet was washed with fresh medium (TSB or SA). Pellets were re-suspended in 100 ml TSB or SA medium. For the proteome analyses, AK39, TB3585 and NRRL 3585 were grown in TSB medium (Appendix A) while industrial strain (DEPA) and NRRL 3585 were grown in Starch-Asparagine (SA) medium (Appendix A) at 28°C at 220 rpm for 48h.

2.3. Proteomic Analyses

2.3.1. Protein Extraction

Protein extraction was adapted from Faurobert et al. (2007) with slight modifications. Cell cultures which were grown in TSB or SA medium were harvested at 48h by centrifugation at 6000 g for 15 min. Pellets were washed once with the fresh medium and then immmediately frozen in liquid nitrogen and stored at -80 °C until use. For the protein extraction, frozen cell cultures were ground in liquid nitrogen and suspended in 3 ml of protein extraction buffer (Appendix A) for each 1 g of sample in 15 mL Falcon tube. Thereafter, they were gently vortexed and left in incubation by 10 min-shaking on ice in order to prevent any probable protease activity. After adding an equal volume of Tris-buffered phenol (pH 8.0, Sigma), solution was mixed with gentle inversions and incubated for 10 min by shaking at RT. In order to separate the organic phase which contains proteins from the aqueous phase, solution was centrifuged at 3400g for 15-20 min. Phenolic phase was recovered carefully without touching the interphase and transferred into a new Falcon tube. The phenolic phase was back-titrated with protein extraction buffer (Appendix A) by mixing 3 ml of extraction buffer with the phenolic phase and shaking the solution for 3 min and then vortexing.

Mixture was centrifuged at 3400g for 15-20 min to separate the phases. Phenolic phase on the top was collected and mixed with 4 volume of cooled precipitation solution (Appendix A) by shaking through several inversions. Afterwards, sample was incubated for at least 4 h or o/n at -20°C to precipitate the proteins. After incubation, proteins were completely precipitated by centrifugation for 10 min at 3400g. They were washed with cooled precipitation solution three times and then with cooled acetone once. Each washing step consisted of mixing the solutions, incubating them at -20°C for 30 min and then centrifugation at 3400g for 10-15 min. After the washing steps, pellet was dried either overnight at -20°C or under Speed-Vac.

2.3.2. Determination of Protein Concentration

Protein pellets were dissolved in rehydration buffer (Appendix A) (Yin et al., 2008) and protein concentration was determined according to a modified Bradford assay described by Ramagli and Rodriguez (1985) (Appendix A) using bovine serum albumin as the standard. Bovine serum albumin was used for the preparation of the standards with final concentrations of 0, 2, 4, 6, 8, 10, 12, 14, 16, 20 µg/ml. Both standards and the samples contained 5 µL HCl (0.1 M) and 40 µL dH_2O and final volume was completed to 1800 µL with Bradford dye (1X) (Appendix A). Tubes were vortexed immediately and incubated for 5 min at dark. Absorbance of samples and standards were measured at 595 nm. Protein concentration of the samples was determined using the calibration curve constructed.

2.3.2. Preparations for MALDI-TOF Analyses

2.3.2.1. 2DE

After the protein extraction, samples were dissolved in rehydration buffer (Appendix A). Thereafter, 350 µg of protein sample in 400 µl of rehydration buffer was used for the passive rehydration of linear IPG strips (18 cm, pH 4-7, BioRad) on BioRad PROTEAN IEF Cell for 12-16 h. Upon the completion of passive rehydration, the first dimension of 2DE, namely isoelectric focusing (IEF), was performed with Immobilised pH Gradient (IPG) System “Ettan IPGphor3” (Amersham Biosciences, GE Healthcare). IEF of 2DE, during which a certain voltage is applied to the IPG strips

so as to mobilize proteins until they reach their corresponding pI regions, consisted of the following steps: 250V (Gradient) for 1.5h, 500V (Gradient) for 1.5h, 1000V (Gradient) for 3.0h, 5000V (Gradient) for 4.0h, 8000V (Gradient) for 12h, 500V (Steep) for 2.0h. At the end of the IEF, IPG strips were either kept at -20°C or directly equilibrated at RT. First, gels were equilibrated with equilibration buffer Solution I (Appendix A) for 15 min on a shaker and then with equilibration buffer Solution II (Appendix A) for 15 min on a shaker. Gel strip was then placed on 12% acrylamide/bis-acrylamide gel with a 4 % stacking gel on the top (Appendix A) using Bio-Rad Cell system (Bio-Rad, USA) to complete the second dimension of 2DE. 27 mA was applied for each gel during the run. Proteins were visualized by staining the gels with Coomassie Blue G-250 (Neuhoff et al., 1988) (Appendix A).

2.3.2.2. Image Analysis

Coomassie stained gels were digitized by using a scanner (Epson Perfection V750). Spot pattern analyses were accomplished by using the 2D image analysis software Delta2D version 3.4 (Decodon, Germany). Three biological replicates were used for WT, AK39 and TB3585 which were grown in TSB while two biological replicates were used for DEPA and WT and they were processed likewise during both the gel image analysis and identification. Relative quantities of the spot (%V) were used for the quantitation on Delta2D. Volume of each spot on the gel is calculated by assuming one gray unit on the spot as an absolute quantitiy 1. The total quantity of the spots, that is the whole gray units on the gel, are equalized to 100 %. Thereby, relative quantity of each spot (%V) is calculated with respect to total spot volume (100 %) by Delta2D and normalization is automatically performed by the software. By dividing the %V of each spot on the comparison groups, the spot ratio is calculated. Spots on the mutant gels having a ratio with 2.5-fold difference with respect to parental gel were selected for identification by MALDI-TOF/MS analysis.

2.3.2.3. Sample Preparation from 2DE Gels for MALDI-TOF MS Analyses

Protein spots selected for the identification were cut either manually or by Ettan Dalt Spot Cutter. Excised spots were transferred into 96 well microtiter plates. The

following procedure was performed automatically by Ettan Spot Handling Workstation (Amersham Biosciences, Uppsala, Sweden) for the analysis by MALDI-TOF/MS:

Gel pieces were first washed with 100 µl of 50% Acetonitrile (ACN) and 50% 50 mM NH_4HCO_3 solution for 30 min and this step repeated once more. Gel pieces were washed with 100 µl of 75% ACN for 10 min and were dried at 37°C for 17 min. Henceforth, 10 µl of trypsin solution (20 ng/µl trypsin) (Promega, Madison, WI, USA) was added and incubated at 37°C for 120 min. For the extraction of the proteins from the gel pieces, 60 µl of 0.1% trifluoroacetic acid (TFA) in 50% ACN was put onto the spots and they were incubated at 40°C for 30 min. Supernatants containing peptides were transferred into another microtiter plate and 60 µl of 0.1% trifluoroacetic acid (TFA) in 50% ACN was added onto the supernatants. Supernatants were dried completely at 40°C for 220 min and then dried samples were dissolved in 2 µl matrix solution consisting of α-cyano-4-hydroxycinnamic acid in 50% CH_3CN / 0.5% TFA. 0.7 µl of the dissolved peptides was used to spot on the MALDI target. The same volume of a saturated α-cyano-4-hydroxy cinnamic acid solution in 70% ACN was also added and both solutions were mixed by aspiration five times. The samples were dried on the target for 10 min or 15 min before MALDI-TOF measurements. The MALDI-TOF measurements for the peptide mass determination were carried out on 4800 MALDI-TOF/TOF Proteomics Analyser (Applied Biosystems, Foster City, CA, USA) which analyzes the data by using 4700 ExplorerTM Software. Spectra were recorded in a mass range of 900 to 3700 Da, focus mass was 2000 Da. When the signal-to-noise ratio (S/N) of autolytic fragments of trypsin with mono-isotopic $(M + H)^+$ m/z at 1045.556 and 2211.104 reached a level of 20, these peaks were used automatically for a two-point internal calibration. Standard mass deviation was smaller than 0.15. "Peak to mascot" script of the 4700 ExplorerTM Software was used to generate the peak lists with a S/N of 10–15. Peak search was filtered with a peak density of 15 peaks per 200 Da. For TOF/TOF measurements, two highest peaks in a spectrum were used, when possible. One-point internal calibration was automatically performed with the mono-isotopic arginine $(M + H)^+$ peak m/z at 175.119 or the lysine $(M + H)^+$ peak

m/z at 147.107 when peaks showed at least S/N ratio of 5. Peak lists were prepared with a S/N ratio of 10.

2.3.2.4. MALDI-TOF Protein Identification

S. clavuligerus protein data sequence was retrieved from UniProt database using http://www.uniprot.org/proteomes/UP000002357 web address in April, 2015.

Peptide Mass Fingerprint" and "MS/MS Ion Search" engines of MASCOT software (Matrix Science Inc., Boston, MA, USA) were used to analyze the peak lists of each protein spot (peptide mass fingerprint and MS/MS data) obtained through MALDI TOF/MS measurement with regard to the data retrieved from UniProt for *S. clavuligerus.* Results having a probability score higher than 53 were selected and used for protein identification.

2.3.3. Procedures for LC-MS/MS Analyses

2.3.3.1. Filter-Aided Sample Preparation (FASP)

In order to prepare extracted proteins for LC-MS/MS analyses, proteins were treated with FASP™ Protein Digestion Kit (Expedeon) with some modifications on the procedure (Wiśniewski et al., 2009). In the original procedure (Wiśniewski et al., 2009), only one enzyme (Trypsin) was used for digesting the proteins. However, in the modified procedure double digestion method was used as suggested by Wiśniewski and Mann (2012) who also named the multienzyme digestion procedure combined with FASP method as (MED) FASP.

30 µg of protein extract was completed to 200 µl with Urea Sample Solution (Appendix A) containing 10 mM DTT in 1.5 ml Eppendorf tubes. Samples were incubated at 25°C at 750 rpm for 1 h and then added into Spin Filters. Spin Filters were centrifuged 14,000 g for 15 min. Liquid collected in the collection tube was discarded. 200 µl of Urea Sample Solution (Appendix A) containing 10 mM DTT was added to each Spin Filter again and samples were centrifuged at 14,000 g for 15 min. Flow-through in the collection tube was discarded. 10 µl of 1X Iodoacetamide Solution (Appendix A) was added to each Spin Filter, samples were vortexed for 1

min and then incubated in dark for 20 min. Spin Filters were centrifuged at 14,000 g for 10 min. 100 µl of Urea Sample Solution (Appendix A) was added to the Spin Filters and samples were centrifuged at 14,000 g for 10 min. This step was repeated once more. Flow-through collected in the collection tube was discarded. 100 µl of 50 mM Ammonium Bicarbonate Solution provided by the kit was added to the Spin Filter and then samples were centrifuged at 14,000 g for 15 min. Collection tube was replaced with a new one and 65 µL of Digestion Solution I (Appendix A) was added to the Spin Filter, after a brief vortex followed by the spin down, top of the tubes was covered with the paraffin film to minimize the evaporation. Samples were incubated at 37°C for 16 h. After the incubation, 15 µl of Digestion Solution II (Appendix A) was added to each Spin Filter. Following a brief vortex and spin down, samples were incubated at 37°C for 6 h. Spin Filters were centrifuged at 14,000 g for 10 min and 30 µl of 0.5 M NaCl Solution provided by the kit was added into the Spin Filters. Samples were centrifuged at 14,000 g for 10 min. Filters were removed and the tubes containing the elution were evaporated under Speed-Vac till the volume of the samples is lowered to 60 µl so that final concentration of the proteins in the sample would be 0.5 µg/µl. Samples can be stored at -20°C until use.

FASP method was first described by Wiśniewski et al. (2009) and based on the retention of the extracted proteins (high-molecular weight molecules) in the filter and the removal of the salts, detergents and other impurities from the extracted proteins and finally elution of proteins as peptides from the filter with enzymatic treatment. This method was improved especially to isolate membrane proteins by removing SDS which decreases the efficiency of the enzymatic digestion and interferes with the mass spectra. FASP was shown to successfully remove SDS from the samples; furthermore, compared to other in-solution methods, higher coverage of membrane protein sequences were achieved with FASP method. Even the peptide identification results with FASP method were comparable to the ones obtained with SILAC experiments which yield high protein identification rates (Wiśniewski et al., 2009). Wiśniewski and Mann (2012) further improved FASP method to MED-FASP, MED standing for Multiple Enzymes for sample Digestion. With MED-FASP method, they compared

one-, two- and three-step digestions using FASP protocol and showed that among the two-step digestions performed with the combination of trypsin and other enzymes, Lys-C + trypsin couple gave the best results as for both the number of identified peptides and the coverage of their sequences. In three-step digestion, they showed there was just 1.5 - 2.5 µg more peptide material yielded when they used chymotrypsin with Lys-C + trypsin, compared to two-step digestion with just Lys-C + trypsin (Wiśniewski and Mann, 2012).

2.3.3.2. Zip-Tip Purification

In order to purify the peptide mixture from protein extract prepared by FASP procedure, Zip-tip® Millipore™ cleaning tips were used according to protocol described below:

All the solvents used for the procedure were MS-grade.

Tips were first wet with 100% Acetonitrile (ACN) twice. After discarding the 100% ACN, the tips were wet with Wetting Solution twice (Appendix A), subsequently equilibrated with Equilibration Solution twice (Appendix A). After the tips were removed of any impurities and equilibrated by these two steps, 10 µl of peptide mixture was pipetted 10 times to attach them onto ZipTip columns completely. Tips were washed with Washing Solution (Appendix A) twice for desalting of the peptides on the column. Finally, peptides were eluted with Elution Solution (Appendix A) by pipetting 10 µl of the solution 2-3 times in the tube and then transferring it to low protein binding tubes. 15 µl of MS-grade dH_2O was added to each tube. They were spun down and dried completely under Speed-Vac. They can be stored at -20°C until use, or suspended directly in 5 µl of dH_2O for LC-MS/MS analysis.

2.3.3.3. LC-MS/MS Analysis

AB SCIEX TripleTOF 5600+ instrument (AB SCIEX, Redwood City, CA, USA) coupled to Eksigent expert nano-LC 400 system (AB SCIEX) was used for tandem mass spectrometric analyses. LC system was comprised of a trap column (180µm x 20mm column, 300 ., nanoACQUITY UPLC® 2G-VM Trap 5µm Symetry® C18,

Waters, UK) and a separation column (75µm, x 150mm column, nanoACQUITY UPLC® 1.8µm 120 Å HSS T3, C18, Waters, UK). 2 µl of peptide solution was loaded onto the LC system having 10 µl loop volume in solvent A (0.1% formic acid (v/v) in H_2O) for 10 min at a constant flow rate of 1µl/min with trapping. Peptides were eluted in 4-96% solvent B (0.1% formic acid in 100% ACN) with a linear gradient and a low-rate of 250 nl/min for 310 min. The gradient program was adjusted as the following: The system was preconditioned with 96% solvent A (0.1% formic acid (v/v) in H_2O) and 4% solvent B (0.1% formic acid in 100% ACN) for 10 min. The ratio of solvent B (0.1% formic acid in100% ACN) was increased from 4% to 40% gradually in linear mode for a time period of 270 min. Thereafter, the ratio of solvent B was increased to 96% within 15 min and the concentration of the mixture was maintained as 96% solvent B and 4% solvent A for 10 min. Then solvent A to B ramping cleaning for the next run was established for 15 minutes. NanoSpray III source with a 10 µm ID nanospray emitter tip (New Objective, Woburn, MA) was used to inject the elute into the mass spectrometer. Voltage for electrospray ionization was kept at 2400 V. Analyst® TF v.1.6 (AB SCIEX) software was used for MS and MS/MS data acquisition. Survey scans at a resolution of >35K in the m/z range from 350 to 1250 were performed and then 20 most abundant precursor ions were selected for the fragmentation using high sensitivity on MS/MS mode with resolution in the m/z range of 100 to 1800. During the collision-induced dissociation (CID) fragmentation, a normalized collision energy of 10 and a declustering potential of 100 V were used. Henceforth, fragment ions were recorded. Mass tolerance was set to ±50 mDa. Mass spectrometer calibration was performed by using 25 fmol/µL β-galactosidase digest as the standard and the device was recalibrated at every third sample.

2.3.3.4. LC-MS/MS Data Analysis

Mass spectrometry data was analyzed by using ProteinPilot 4.5 Beta (AB SCIEX) for the peptide identification. After all the files were extracted, peak list generation as well as database search were performed by using *.wiff files. *S. clavuligerus* protein data was downloaded from UniProt database (http://www.uniprot.org/) in January, 2016. Parameters for the database searches included trypsin protease with the allowance of

one missed cleavage and oxidation of methionine. A precursor ion mass error window of 10 ppm and a fragment ion mass error window of 0.1 Da was accepted. 1% false discovery rates (FDRs) were calculated by searching the raw data against the decoy database. Proteins with at least two matching unique peptides were selected and considered as identified.

2.4. Databases for Protein Categorization, Function and Associations

Functional categories of the proteins were mostly determined by using COG database provided by NCBI (ftp://ftp.ncbi.nih.gov/pub/COG/COG2014/static/lists/listStrbin.html). PSORTb version 3.0.2 (http://www.psort.org/psortb/), Gpos-mPloc (http://www.csbio.sjtu.edu.cn/bioinf/Gpos-multi/#) and ngLoc (http://genome.unmc.edu/ngLOC/index.html) databases were used to determine the cellular localization of the proteins. Expasy pI/Mw tool was used to calculate the theoretical pI and Mw values of the proteins (Gasteiger et al. 2003). Sequence similarity search for comparison of two protein was performed through BLASTP program (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). Protein associations and pathways were predicted using STRING 10.0 (http://string-db.org/), which is a global source of genes from which protein associations can be inferred and predicted (von Mering et al., 2003), and KEGG mapper (http://www.kegg.jp/kegg/tool/annotate_sequence.html), respectively. EggNog version 4.5 was used to find out orthologous groups of proteins within the same genera and/or different taxonomic groups and their predicted functions (Huerta-Cepas et al., 2015). Theoretical map of *S. clavuligerus* total proteome was prepared by using JVirGel version 2.0, which uses theoretical Mw and calculated pI values of the proteins so as to predict the migration pattern of each protein (Hiller et al. 2003; Hiller et al., 2006)

2.5. Statistical Analyses for LC-MS/MS and MALDI-TOF Results

Statistical analyses for 2DE experiments were performed by using Delta2D 4.3 statistical tools. For comparison of DEPA and NRRL 3585 strains, paired t-test was used with all-permutations, Welch approximation. α value was equal to 0.05. For the

comparison of NRRL 3585, TB3585 and AK39, one-way ANOVA analysis was performed. Number of permutations was selected as 10,000 and α value was equal to 0.05.

Statistical analysis for LC-MS/MS experiments was adapted from Old et al. (2005). Spectral count quantitation method was performed. The equation shown below and described by Old et al. (2005) was used to find out the differential expression profiles of the proteins between the standard and mutant strains (Equation 2.1).

$$R_{SC} = \log_2[(n_2 + f)/(n_1 + f)] + \log_2[(t_1 - n_1 + f)/(t_2 - n_2 + f)] \quad \text{Equation 2. 1}$$

Rsc value refers to $\log_2$ ratio of the protein abundance between two samples, n_1 and n_2 represent the spectral counts of the samples that come from LC-MS/MS analyses, t_1 and t_2 are the total spectral counts in each sample. f value is the correction factor and has been set to 1.25 as stated by Old et al. (2005). By using the formula above, differential protein expression patterns between the mutant and parental strains were compared.

CHAPTER 3

RESULTS

3.1. Comparative Proteome Analysis of *S. clavuligerus* NRRL 3585, AK39, TB3585 and DEPA Strains

NRRL 3585, AK39 and TB3585 strains were grown in TSB medium and samples from each culture were collected at 48th h when they enter the stationary phase for the analysis of their protein profiles. Analysis at this phase is important because secondary metabolic activities and thus related protein profiles of the microorganisms show a dramatic change when compared to the active growth phase. Given that TB3585 and AK39 strains were genetically engineered in favor of cephamycin C overproduction, comparison of protein expression patterns at stationary phase might reveal important details about the metabolic activities of these microorganisms. NRRL 3585 and DEPA strains were also compared in this manner.

A theoretical map for the total proteome of *S. clavuligerus* NRRL 3585 was generated by using JVirGel software (Figure 3. 1), which was explained briefly in Section 2. 4. *S. clavuligerus* protein data was downloaded from UniProtKB and uploaded to JVirGel software in order to generate the virtual 2DE map within a pI range of 3 to10 and Mw range of 10-to-120 kDa. The region that lies within the pI 4 and 7 best represented the most densely populated area with proteins. Thus, we preferred to use a pI range of 4-7 for the same gel length in order to improve the migration pattern and protein focusing.

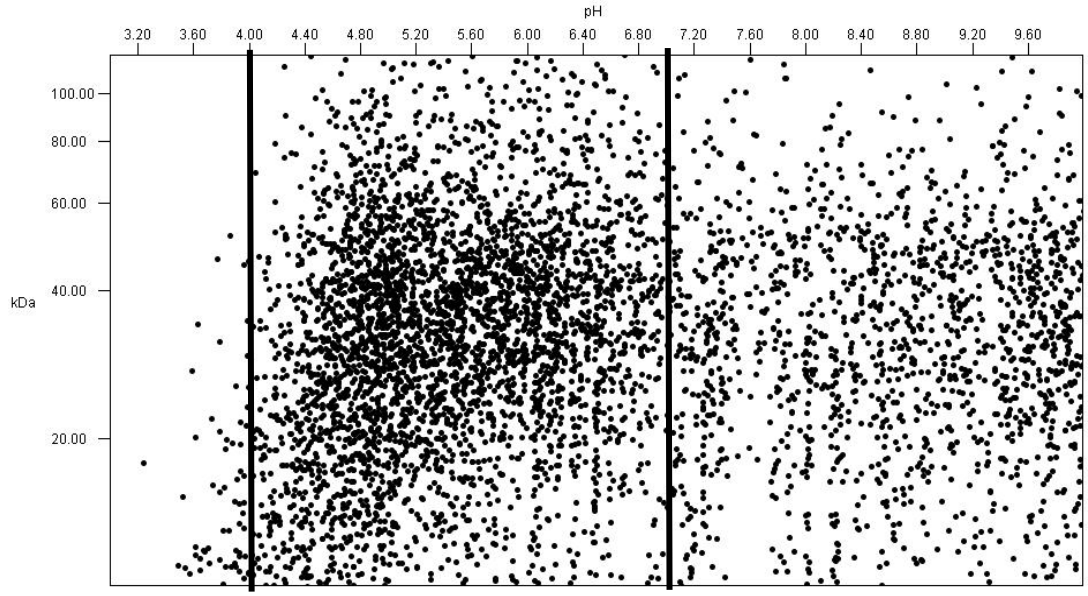


Figure 3. 1. Theoretical 2DE map of *S. clavuligerus* total proteome predicted by JVirGel Version 2.0 tool (pI range 3-10, Mw range 10-120 kDA).

3.2. Comparative Proteome Analysis of NRRL 3585 and AK39 Strains by 2DE MALDI-TOF/MS and LC-MS/MS

A fused image of cytosolic proteomes of NRRL 3585 and AK39 is shown in Figure 3. 2. Of the proteins that were identified as differentially expressed in AK39 strain, 13 proteins were significantly upregulated and 32 proteins were significantly downregulated (Table 3. 1 and Table 3. 3, respectively) in MALDI-TOF/MS analysis. On the other hand, 28 proteins were shown to be upregulated and 40 proteins were downregulated in AK39 strain through LC-MS/MS analysis (Table 3. 2 and Table 3. 4, respectively).

In Figure 3. 2, red spots represent the proteins expressed in AK39 strain whereas green spots represent the ones expressed in the standard strain. Proteins that are expressed in both strains overlap each other. Depending on the expression levels, the color of the overlapping spots are displayed in different scales of yellow.

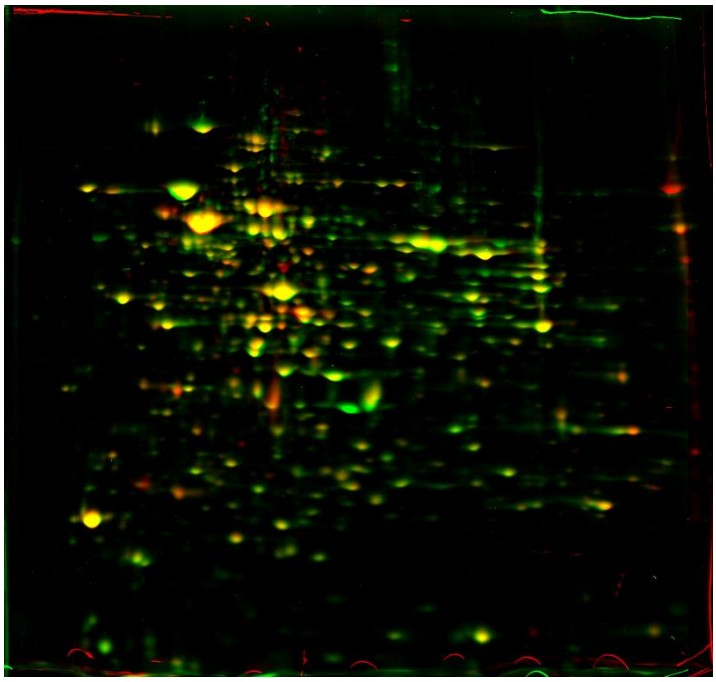


Figure 3. 2. Dual channel 2DE imaging of *S. clavuligerus* strains, NRRL 3585 (green) and AK39 (red) at 48h in TSB medium, pI 4-7. Overlapping spots are represented in yellow.

3.2.1. Upregulated Proteins in AK39

Figure 3. 3 shows protein spots on AK39 gel that showed more than 2.5-fold increase in expression when compared to their counterparts in NRRL 3585. A list of upregulated proteins identified through MALDI-TOF/MS and LC-MS/MS analyses is given in Table 3. 1 and Table 3. 2, respectively.

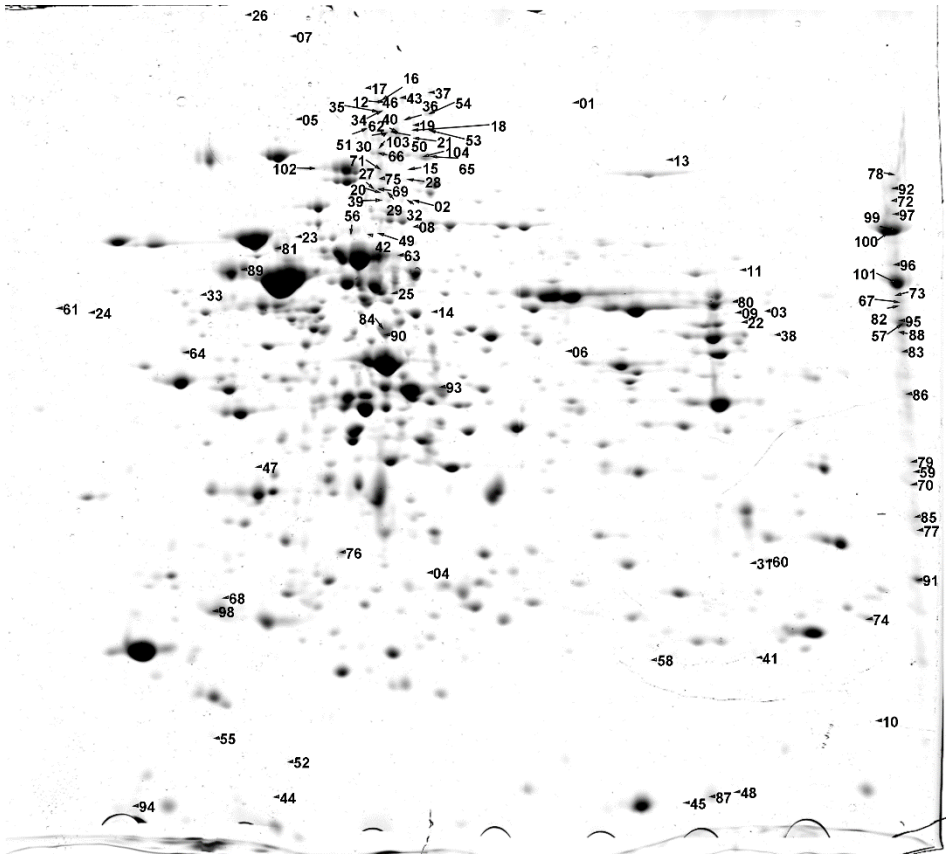


Figure 3. 3. Protein spots on AK39 2DE gel that showed more than 2.5-fold increase when compared to those on NRRL 3585 2DE gel.

Table 3. 1. List of upregulated proteins identified on AK39 2DE gel.

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spots	PTM *	Genome Location **	Spot #
Amino Acid Metabolism	B5GWI7	Xaa-Pro aminopeptidase I	*pepP*	Cytoplasm	5.8	4.67	54,35	–	–	Chr.	89
Lipid Metabolism	B5H0X6	Enoyl-[acyl-carrier-protein] reductase [NADH]	*fabI*	Cytoplasm	8.7	5.47	27,606	–	–	Chr.	60
	B5H1P4	Acetyl/propionyl CoA carboxylase	*accD*	Cytoplasm	3.0 (a); 2.8 (b)	6.91	57,553	2 (a, b)	M	Chr.	73, 101
DNA Replication, Recombination, Repair, Transcripton	E2QA28	Protein RecA	*recA*	Cytoplasm	2.6	6.58	44,081	–	–	Chr.	88
Stress-Related, Protein Turnover, Chaperones	E2PWA4	Proteasome subunit alpha	*prcA*	Cytoplasm	2.7	4.88	27,591	–	–	Chr.	47
Secondary Metabolism	E2Q5Q8	Carboxyethylarginine synthase	*ceaS2*	Cytoplasm	3.3	5.01	62,348	–	–	Chr.	42
	B5GLB5	Deacetoxycephalosporin C hydroxylase	*cefF*	Cytoplasm	2.7	4.79	34,599	–	–	Chr.	47
	B5GLB7	Positive regulator (ccaR)	*ccaR*	Cytoplasm	8.7	6.02	29,346	–	–	Chr.	60
Energy Production/Electron-Iron Transfer	B5GL53	ATP synthase gamma chain	*atpG*	Cytoplasm; Inner membrane	5546.4 (a); 299.5 (b)	8.60	32,81	2 (a, b)	M	Chr.	77, 85
Inorganic/Organic Molecule Transport	B5H413	Phosphoenolpyruvate-protein phosphotransferase	*SCLAV_0624*	Cytoplasm	5.8	4.74	57,216	–	–	Chr.	89
Hypothetical/Unknown Proteins	B5GRL8	DUF1906 domain-containing protein	*SCLAV_4982*	Unknown	3.4 (a); 6.1 (b)	6.92	81,717	2 (a, b)	M	Chr.	82, 100

Table 3.1. List of upregulated proteins identified on AK39 2DE gel. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spots	PTM *	Genome Location **	Spot #
Hypothetical/Unknown Proteins	B5H453	Predicted acyltransferase	*SCLAV_p1317*	Unknown	93.8	6.8	31,731	–	–	pSCL4	91
Others/General Function	B5GZB8	Aldehyde dehydrogenase	*aldH*	Cytoplasm	3.8	5.16	55,339	–	–	Chr.	14

*PTM refers to possible posttranslational modification. C denotes charge modification, M denotes mass modification, C-M shows both charge and mass modification. ** "Chr." refers to chromosome, "pSCL4" refers to plasmid pSCL4.

Table 3. 2. List of upregulated proteins in AK39 strain identified by LC-MS/MS.

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Amino Acid Metabolism	E2Q0I1	Histidine ammonia-lyase	*SCLAV_5457*	Cytoplasm	2.763	6.03	59,762	Chr.
Lipid Metabolism	D5SJB1	3-oxoacyl-acp synthase	*SCLAV_p0514*	Cytoplasm	2.763	4.90	40,640	pSCL4
DNA Replication, Recombination, Repair, Transcripton	E2Q4U1	ATP-dependent RNA helicase	*SCLAV_4026*	Cytoplasm	2.297	9.91	87,198	Chr.
	E2Q0M0	Two-component system sensor kinase	*SCLAV_1334*	Cytoplasm	2.297	5.30	24,917	Chr.
	B5H061	DNA-binding protein HU	*SCLAV_4504*	Cytoplasm	2.549	11.40	22,873	Chr.
Stress-Related, Protein Turnover, Chaperones	E2Q1R3	Chaperone	*SCLAV_5622*	Cytoplasmic	3.115	4.98	92,409	Chr.
	B5GYC7	Starvation-induced DNA protecting protein	*mrgA*	Cytoplasm	2.549	4.96	16,686	Chr.
Secondary Metabolism	B5GRF0	Polyketide synthase	*SCLAV_p0512*	Cytoplasmic membrane	3.634	4.96	64,145	pSCL4
	B5GMY7	Phosphotransferase	*aphD*	Cytoplasm	4.097	4.99	33,629	pSCL4
	B5GUT9	Alkaline d-peptidase	*SCLAV_4822*	Extracellular	2.049	9.19	42,034	Chr.
	E2PY60	NocE-like protein	*nocE*	Unknown	3.398	8.80	147,287	Chr.
Cell Processes (Shape/Division/Motility)	B5GMU0	Plasmid partitioning protein ParB	*parB*	Cytoplasm	2.297	4.94	51,212	pSCL4
Inorganic/Organic Molecule Transport	B5H413	Phosphoenolpyruvate-protein phosphotransferase	*SCLAV_0624*	Cytoplasm	2.376	4.74	57,216	Chr.
	B5H412	Phosphoenolpyruvate-dependent sugar phosphotransferase	*ptsA*	Cytoplasm	2.950	4.40	15,185	Chr.

Table 3. 2. List of upregulated proteins in AK39 strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Hypothetical/Unknown Proteins	D5SK44	Uncharacterized protein	*SCLAV_p0800*	Cytoplasm	2.763	4.38	19,528	pSCL4
	E2Q0Y5	Putative alpha-glucosidase	*SCLAV_1450*	Cytoplasm	3.837	4.78	60,086	Chr.
	B5GMR8	Uncharacterized protein	*SCLAV_p1086*	Unknown	2.549	6.18	21,169	pSCL4
	B5H1P6	Putative uncharacterized protein	*SCLAV_1941*	Cytoplasm	2.549	5.05	28,161	Chr.
	B5GP37	Putative uncharacterized protein	*SCLAV_3208*	Unknown	2.763	9.04	21,165	Chr.
	B5GNB3	Uncharacterized protein	*SCLAV_p0717*	Unknown	2.950	4.25	15,552	pSCL4
	B5H3M9	Putative 2-isopropylmalate synthase	*SCLAV_4478*	Cytoplasmic	3.115	5.42	58,271	Chr.
	E2Q5V3	Phage tail sheath protein, putative	*SCLAV_0037*	Cytoplasmic membrane	3.521	5.17	80,627	Chr.
	B5H459	Uncharacterized protein	*SCLAV_p1312*	Unknown	2.297	7.77	17,316	pSCL4
	D5SKY6	Methyl-accepting chemotaxis sensory transducer	*SCLAV_p1093*	Unknown	2.549	9.89	24,315	pSCL4
	D5SJB0	KR domain family protein	*SCLAV_p0513*	Cytoplasmic	3.115	5.59	39,149	pSCL4
Others/General Function	E2PWH9	ATPase	*SCLAV_0935*	Cytoplasm	2.019	4.91	41,412	Chr.
	E2PU67	PRC-barrel domain protein	*SCLAV_0610*	Unknown	2.297	4.46	14,31	Chr.
	D5SKX8	Peptidase S1 and S6	*SCLAV_p1085*	Extracellular	2.549	5.10	72,908	pSCL4

* "Chr." refers to chromosome, "pSCL4" refers to plasmid pSCL4.

3.2.1.1. Amino Acid Metabolism

Xaa-Pro aminopeptidase coded by *pepP* I is involved in protein degradation. It is a proline-specific peptidase that can be differentiated from broad specificity aminopeptidases by being able to degrade imide bonds that precede the proline residue or amide bonds that come after the proline. Though broad specificity aminopeptidases can release most of the amino acid residues from the N-terminus of peptides, they cannot act on the bonds where proline is present due to its unique cyclic structure (Nandan et al., 2011; Nandan and Nampoothiri, 2014). With 5.8 fold increased levels in AK39 as compared to the parental NRRL 3585 in 2DE MALDI-TOF/MS analysis, it must have a specific function that must be demonstrated by future studies.

Histidine ammonia-lyase enzyme was shown to be upregulated by LC-MS/MS analysis in this category and catalyzes the conversion of histidine to urocanic acid by the removal of ammonia (Givot et al., 1969). This is the first step that initiates a cascade of reactions to produce glutamate from histidine (Schwede et al., 1999). Glutamate can be used as both carbon and nitrogen sources in *Streptomyces* and also used in the production of several important intermediates that are shunted towards primary and secondary biosynthetic pathways (Rigali et al., 2008). For example, valine and arginine precursors that are necessary for CC and CA production, respectively, can be derived from glutamate.

3.2.1.2. Lipid Metabolism

Three proteins in this category, enoyl-[acyl-carrier-protein] reductase (FabI), 3-oxoacyl-acp synthase and acetyl/propionyl CoA carboxylase [AccD]), were shown to be upregulated in AK39 strain.

FabI protein is an enoyl-[acyl-carrier-protein] reductase (enoyl-ACP reductase) and plays an important role in fatty acid biosynthesis in bacteria. Bacterial fatty acid biosynthesis is carried out by type II fatty acid synthase (FAS) system which is responsible for the production of essential components of the cell such as cell envelope, phospholipids, lipoproteins, lipopolysaccharides, and mycolic acids. The actors of the FAS system in bacteria are usually small, soluble, monofunctional

enzymes that act sequentially to carry out the biosynthetic process (Heath et al., 2000, Heath et al., 2001; Kitagawa et al., 2007). *E.coli* FAS II system begins with the formation of malonyl-ACP catalyzed by malonyl-CoA:ACP transacylase (FabD). Type II fatty acid biosynthesis cycle begins with the β- ketoacyl-ACP synthase III (FabH) enzyme catalyzing the condensation of malonyl-ACP with acetyl-CoA. The β-group on the substrate is then reduced by the NADPH-dependent β-ketoacyl-ACP reductase (FabG) and the resultant product is dehydrated by the β-hydroxyacyl-ACP dehydratase (FabA or FabZ) to an enoyl-ACP. NAD(P)H-dependent enoyl-ACP reductase (FabI, FabK, or FabL) reduced the enoyl chain to produce an acyl-ACP. β-ketoacyl-ACP synthase (FabB or FabF) enzymes catalyze the additional elongation of steps by adding two carbons to form a β-ketoacyl-ACP. When the fatty acid chains reach a length of 14-18 carbons, the reactions are terminated to incorporate the end-products into the membrane phospholipids (Zhang et al., 2004).

3-oxoacyl-acp synthase protein carries the 3-oxoacyl-[acyl-carrier-protein (ACP)] synthase III domain according to UniProtKB database. As explained above, this protein is involved in type 2 fatty acid biosynthesis system observed in bacteria and initiate the fatty acid biosynthesis cycle by condensing the malonyl-ACP with acetyl-CoA (Taguchi et al., 2006; Kitagawa et al., 2007). However, this protein was not identified as FabH but as SCLAV_p0514. Therefore, it is likely to be an alternative to *fabH* from a paralogous gene.

Another protein shown to be upregulated in fatty acid metabolism is acetyl/propionyl CoA carboxylase enzyme. Acetyl-CoA carboxylase (ACC) catalyzes the first committed step for the fatty acid biosynthesis. It was found in *E. coli* that acetyl-CoA carboxylases are composed of four different subunits encoded by *accA, accB, accC* and *accD* genes (Dunn et al., 2004). The main function of the enzyme is to carboxylate acetyl-CoA to produce malonyl-CoA (Gago et al., 2011) which is not only the primer for the fatty acid biosynthesis (Rodríguez et al., 2001) but also the most common precursor used as extender units for the biosynthesis of many polyketide antibiotics (Rodríguez and Gramajo, 1999). Several ACC ezymes isolated from actinomycetes were shown to carboxylate other substrates than malonyl-CoA, such as propionyl- or

butryryl-CoA (Rodríguez et al., 2001) which are incorporated into several other polyketide antibiotics. Thus, ACC enzymes are more appropriately called acyl-CoA carboxylases to show the differential substrate specificity of the enzymes and they provide a linkage between the primary and secondary metabolism producing the key components necessary for both the fatty acid and polyketide synthesis in the organism (Dunn et al., 2004; Rodríguez and Gramajo, 1999). Hence, upregulation of this protein might suggest a probable increase in polyketide biosynthesis in AK39 strain. Moreover, acetyl/propionyl CoA was represented by two spots on the AK39 2DE gel within the same pI range, suggesting that it might have undergone a posttranslational modification that altered the mass of the protein.

3.2.1.3. DNA Replication, Recombination, Repair, Transcription

In this category, there were 4 proteins: RecA, two-component system sensor kinase, DNA-binding protein HU and ATP-dependent RNA helicase.

RecA is a multifunctional protein involved in homologous recombination, DNA strand exchange, and recombinational DNA repair mechanisms. In case of DNA damage and the resultant single stranded DNAs, RecA enzyme is activated and induces a series of activation and inactivation of genes related with SOS response (Vierling et al., 2001). Moreover, Volff and Altenbuchner (1998) suggested that RecA is involved in amplification mechanism and control of genetic instability in *Streptomyces*. Long linear chromosomes of *Streptomyces* undergo large deletions as well as DNA rearrangements which are the source of genetic instability observed in *Streptomyces*. Genetic instability affects many properties of the organism varying from morphological differentiation and production of secondary metabolites to secretion of extracellular enzymes (Volff and Altenbuchner, 1998). The effect of RecA on genetic instability was shown on a mutant strain of *S. lividans* which expressed a truncated RecA protein. The mutants, when compared to parental strain, showed 70 times more genetic instability (Vierling et al., 2001). Another property observed in *Streptomyces* is the presence of high-copy number DNA amplifications. A *recA* mutant of *S. lividans* complemented with *recA* gene was shown to restore the amplification of a certain element (*AUD1*) in the genome (Volff and Altenbuchner, 1998). The upregulation of

this protein might suggest an increased need for this protein as a response to an increased genetic instability and/or mutations in AK39.

As explained earlier in Chapter I, sensor kinase is a component of the two-component system, which controls both the secondary metabolism and morphological differentiation in *Streptomyces*, by acting together with cognate response regulator. Sensor kinases detect the environmental changes and activate the cognate response regulators to regulate the gene expression (Martín and Liras, 2010; Rodríguez et al, 2013).

DNA-binding protein HU is a bacterial histone-like protein, the main function of which is to change the topology of the DNA. Therefore, these proteins are involved in all DNA-dependent reactions such as gene regulation, DNA repair, recombination as well as replication (Grove, 2011).

UniProtKB database suggests that ATP-dependent RNA helicase (SCLAV_4026) protein contains a DEAD box helicase domain. These proteins are involved in RNA metabolism, act on short RNA duplex structures and can even recognize foreign RNA duplexes. In addition, RNA helicases with DEAD box domain are proposed to be involved in translation initiation, RNA decay and unwinding, ribosome biogenesis. Although these proteins are called as helicases, they can have much more diverse functions in the cell (Cordin et al., 2006; Iost and Dreyfus, 2006). Moreover, overexpression of ATP-dependent RNA helicase can be exploited by the organism as a fine-tuning mechanism for the translation of certain proteins through ribosomes.

3.2.1.4. Stress-Related, Protein Turnover, Chaperones

There were three proteins identified under this category: Proteasome subunit alpha (PcrA), chaperone and starvation-induced DNA protecting protein.

Proteasome subunit alpha (PrcA) is one of the two subunits of 20S proteasome, the other subunit being proteasome subunit beta (Nagy et al, 1998). Prokaryotic 20S proteosomes are found in archaebacteria and actinomycetes (De Mot et al., 2007). They are self-assembling proteases which consist of inner and outer rings. Inner rings are composed of α subunits and the outer rings are of β subunits. Catalytic site is

confined to the interior part of the proteasome (Nagy et al, 1998). There are not many reports on 20S proteasome in *Actinomycetes* except for one reporting on 20S proteasome found in *M. tuberculosis* that provides the organism protection against oxidative and nitrosative stress (Darwin et al., 2003) and the other about the 20S proteasome in *S. lividans* which was shown to display no effects under either normal or stress conditions, but only involved in possible degradation of certain secreted heterologous proteins (Hong et al., 2005). Although overrepresentation of these two proteins are reminiscent of some kinds of stress over AK39, the function of this protein remains to be further characterized.

Chaperone proteins or widely known as molecular chaperones prevent misfolded or unfolded polypeptides to aggregate and also help proteins fold appropriately especially under stress conditions (Bucca et al., 2003).

Starvation-induced DNA protecting protein coded by *mrgA* belongs to DPS family of proteins. MrgA protein is induced by H_2O_2 and and also by metal-ion starvation in *Bacillus subtilis*. It forms multimeric complexes with DNA especially during the stationary phase, thereby protecting the cells against oxidative stress (Chen and Helmann, 1995).

Upregulation of these proteins might indicate the presence of an oxidative stress over AK39, probably as a result of metabolic alteration used to construct it.

3.2.1.5. Secondary Metabolism

Carboxyethylarginine synthase, deacetoxycephalosporin C hydroxylase, positive regulator (CcaR) identified in this category were directly related with CC production. Considering the metabolic engineering applied to this strain (primary metabolism was re-directed towards CC precursor molecules), approximately 3 fold increase in the levels of two key biosynthetic enzymes of CC is quite expected. The need for this increase seems to be met by as high as 8.7 fold overrepresented levels of CC- and CA-pathway specific transcriptional activator CcaR by yet unknown global regulatory mechanism(s).

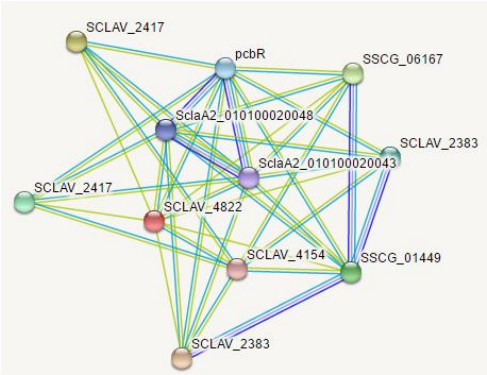
Polyketide synthase (SCLAV_p0512) level was also shown to be upregulated in AK39 strain. Polyketide synthases (PKSs) are involved in the biosynthesis of polyketide (PK) antibiotics. Polyketides are produced from small carboxylic acids (2- and 3-carbon acyl groups derived from coenzyme A thioesters through sequential condensation reactions) (Katz and Donadio, 1993; Pfeifer and Khosla, 2001). They are assembled using several common biosynthetic precursors, by the carboxylation of acetyl-CoA and butyryl-CoA/crotonyl-CoA, respectively, while numerous pathways give rise to methylmalonyl-CoA (Cronan and Thomas, 2009). Polyketide synthesis evolutionarily descends from fatty acid biosynthesis, except that polyketide synthetases display highly programmed mechanisms while FAS system is restricted with several options (Cronan and Thomas, 2009; Hopwood and Sherman, 1990). PKs are complex natural products that show great biological diversity. For example, antibiotics (erythromycin, tetracycline), anticancer agents such as doxorubicin, enediynes, antiparasitic agents (avermectin, nemadectin), antifungals (arnphotericin, grisenfulvin), cardiovascular agents (lovastatin, compactin), and veterinary products (monensin, tylosin) are of polyketide origin (Carreras and Santi, 1998) or well known FK506 is a 23-membered polyketide macrolide with immunosuppressant activity (Mo et al., 2009). Furthermore, more than half of the biologically active, natural polyketides are produced by actinomycetes (Pfeifer and Khosla, 2001). In *S. clavuligerus*, there exist six PKS gene clusters located in the 1.8 Mb megaplasmid pSCL4 (Medema et al., 2010). Generally, PKSs are divided into three categories in bacteria: Type I PKSs, type II PKSs, type III PKSs (Shen, 2003). PKSs are multifunctional enzymes. Type I PKSs are comprised of polypeptides, each of which contains multicatalytic sites while type II PKSs are comprised of polypeptides, each of which contains only one active site. Both type I and type II PKSs contain one or more modular structures that usually contain three domains: AT domain, ketosynthase (KS), ACP domain. There can be other domains on the modules, as well. Type III PKSs, on the other hand, are comprised of monofunctional enzymes (Liou and Khosla, 2003). The questions of whether or not (i) the fatty acid and polyketide synthetic pathways share protein components, and (ii) play an indirect role in providing building blocks for synthesis of the polyketides have been tried to be addressed almost exclusively in *Streptomyces* spp. due to the

remarkable diversity of polyketides in these bacteria. The studies has mainly focused on type II polyketide synthases which are composed of discrete monofunctional proteins (analogous to Fas II) rather than the large, highly modular type I polyketide synthases (Cronan and Thomas, 2009).

Phosphotransferase (AphD) is a member of an aminoglycoside phosphotransferases (APHs) which belong to a protein kinase family. Aph enzymes confer bacteria resistance against aminoglycoside antibiotics (such as gentamicin, amikacin, tobramycin, kanamycin through phosphorylation. Aminoglycosides target bacterial ribosomes interfering with the translation and their effect is bactericidal rather than bacteriostatic (Wright and Thompson, 1999). APH enzymes not only provide protection against foreign aminoglycosides but also against the ones that are produced by the bacteria *per s*e (Wright and Thompson, 1999; Anderson et al., 2002). Aminoglycoside phosphotransferases (APHs), just like other aminoglycoside modifying enzymes (aminoglycoside-*N*-acetyltransferase [AAC], aminoglycoside-*O*-adenyltransferase [AAD]) can be either plasmid- or chromosome-borne and usually encoded on transposable elements (Rouch et al., 1987). The expression of this gene can indicate that there might be aminoglycoside(s) produced by the microorganism although draft genome sequence could not reveal such a cluster in *S. clavuligerus*. However, despite the fact that the same genome sequence could not locate the gene cluster for holomycin, this antibiotic was shown to be produced by this organism (Li and Walsh, 2010). Alternatively, the organism might have acquired the resistance gene by horizontal gene transfer which might explain the propable absence of the gene cluster. Furthermore, gene coding for AphD is located on the plasmid pSCL4 in *S. clavuligerus*, further strengthening the assumption of the horizontal gene transfer. We have also detected a 22.7 fold downregulation of a related protein, namely a putative aminoglycoside 2-N-acetyltransferase Aac2, in CA overproducer DEPA (Section 3.4.2.8).

Alkaline d-peptidase is a strictly d-stereospecific endopeptidase and is active against oligopeptides containing D-phenylalanine residues such as D-$(Phe)_3$ and D-$(Phe)_4$. It was shown that this enzyme could also catalyze the oligomerization of D-

phenylalanine with D-phenylalanine methyl ester by using the latter as substrate (Komeda and Asano, 2003). Komeda and Asano (2000) suggested that these proteins are penicillin-binding proteins and Asano et al. (1996) showed that alkaline d-peptidase that they isolated from *B. cereus* showed β-lactamase activity acting on ampicillin and penicillin G (Asano et al., 1996). Furthermore, when the protein sequence of alkaline d-peptidase (SCLAV_4822) was searched through STRING database (Figure 3. 4), some of the proteins that it associates with were predicted as putative penicillin-binding exported protein, putative D-alanyl-D-alanine carboxypeptidase, penicillin binding protein, D-alanyl-D-alanine carboxypeptidase, which suggests that this protein might contribute to the antibiotic resistance in the microorganism. The sequence was also searched through EggNog database to find out its orthologues and their predicted functions in the kingdom of bacteria and the results showed that the protein resembles β-lactamases.



Your Input:

		Neighborhood	Gene Fusion	Cooccurence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
SCLAV_4822	*Alkaline d-peptidase (395 aa)*									

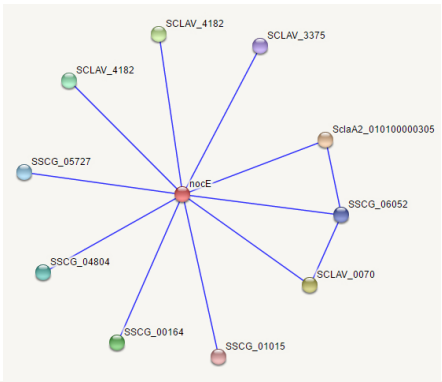
Predicted Functional Partners:

SCLAV_2383	*Putative penicillin-binding exported protein (549 aa)*									0.866
SCLAV_2417	*Putative D-alanyl-D-alanine carboxypeptidase (470 aa)*									0.866
SSCG_06167	*Penicillin binding protein (581 aa)*									0.866
SSCG_01449	*D-alanyl-D-alanine carboxypeptidase (455 aa)*									0.866
SCLAV_2417	*no annotation available (470 aa)*									0.866
SCLAV_2383	*no annotation available (549 aa)*									0.866
pcbR	*PcbR (551 aa)*									0.634
SclaA2_010100020048	*Orf18 (494 aa)*									0.634
SclaA2_010100020043	*Orf19 (717 aa)*									0.634
SCLAV_4154	*Putative penicillin-binding protein (490 aa)*									0.634

Figure 3. 4. STRING data showing the associaton of alkaline d-peptidase with other proteins.

nocE gene is found in nocardicin A biosynthetic gene cluster in *Nocardia uniformis* subsp. *tsuyamanensis* (Davidsen and Townsend, 2012). Nocardicin A is a monocyclic β-lactam antibiotic and shows a broad spectrum of activity against Gram-negative bacteria. This antibiotic was shown to be relatively resistant to β-lactamases most probably due to their stable monocyclic β-lactam ring and their low affinity against the β-lactamases (Gunsior et al., 2004). Kojo et al. (1988) compared the β-lactamase resistance of this antibiotic with other nine β-lactam antibiotics including cefoxitin and cefuroxime. Of all the antibiotics tested, nocardicin was shown to be the most resistant to several β-lactamases produced by different species except for the ones of *Klebsiella oxytoca* and *Proteus vulgaris*. Nocardicin A biosynthetic gene cluster was shown to contain 14 open reading frames in *Nocardia uniformis* and of these, *nocE* which is the downstream gene of the cluster was shown to code for a large protein of 1414 amino

acid residues in *Nocardia uniformis* with unkown function (Gunsior et al., 2004; Davidsen and Townsend, 2012). Davidsen and Townsend (2009) showed that *nocE* gene was not essential for the nocardicin biosynthesis and Gunsior et al. (2004) suggested that NocE protein might just mark the pathway boundaries. Since there is not much definitive information about the NocE-like protein, the sequence of the protein was queried in STRING database (Figure 3. 5). Although some of the associated proteins are unknown; it is predicted to be associated with a transcriptional regulator (XRE family) and three NsdA proteins. NsdA protein is a regulatory protein which is widely common and conserved in *Streptomyces* and was shown to negatively regulate both the sporulation and antibiotic processes in *S. coelicolor* and *Streptomyces bingchengensis* (Li et al., 2006; Wang et al., 2009).



Your Input:

		Neighborhood	Gene Fusion	Cooccurence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
nocE	*NocE (1381 aa)*									

Predicted Functional Partners:

SclaA2_010100000305	*no annotation available (975 aa)*			•						0.736
SCLAV_0070	*Transcriptional regulator, XRE family (994 aa)*			•						0.734
SCLAV_4182	*Orf16 (401 aa)*			•						0.711
SSCG_00164	*no annotation available (401 aa)*			•						0.711
SCLAV_4182	*no annotation available (401 aa)*			•						0.711
SSCG_04804	*NsdA (502 aa)*			•						0.704
SSCG_05727	*NsdA (498 aa)*			•						0.687
SSCG_06052	*Putative uncharacterized protein (706 aa)*			•						0.685
SCLAV_3375	*NsdA (479 aa)*			•						0.684
SSCG_01015	*no annotation available (479 aa)*			•						0.684

Figure 3. 5. STRING data showing the associaton of NocE-like protein with other proteins.

Interestingly, upregulated proteins in the“Secondary Metabolism” of AK39 strain, which were detected through LC-MS/MS analysis, showed quite diverse functions and no proteins functional in CC biosynthetic pathway could be identified. That might be related with (i) the different sensitivities of LC-MS/MS and MALDI-TOF/MS analyses, because of the difference in relative abundance of peptide fragments in mixtures, (ii) the pI and Mw restrictions imposed by 2DE method and (iii) For example, pI values of alkaline d-peptidase and NocE-like protein are higher than 7.0, which constituted the farthest edge in the alkaline pI range that was used in 2DE method in this study. Probably, the abundance of these proteins was much higher than that of the ones related with CC biosynthesis so that the peptides coming from CC biosynthetic proteins were masked in LC-MS/MS analysis. These results demonstrate that there is not an ultimate method to study proteomics; on the contrary, different techniques used in this area are complementary to one another and provide a much better comprehension for proteomic profiles of the organisms.

3.2.1.6. Energy Production/Electron-Iron Transfer

ATP synthase gamma chain identified under this category is a subunit of F_1 hydrophilic part of the ATP synthase (F_0F_1) (Feniouk et al., 2007). ATP synthase catalyzes the ATP synthesis from adenosine diphosphate (ADP) and inorganic phosphate accompanied with proton translocation through transmembrane (F_0) part, a process powered by proton motive force (Stock et al., 1999; Feniouk et al., 2007). F_0 is the hydrophobic part embedded in the membrane layer and is responsible for the proton translocation, while F_1, which is comprised of 5 subunits ($\alpha_3\beta_3\gamma_1\delta_1\epsilon_1$), is the hydrophilic part that protrudes from the membrane layer and catalyzes the ATP hydrolysis/synthesis (Stock et al., 1999; Gaballo et al., 2006; Feniouk et al., 2007). Although ATP synthase gamma chain has a pI of 8.6, the protein was observed at the pH 7.0 region of of the gel and it was represented by two spots within the same pI range. It is thus possible that this protein could have undergone a mass modification as well as charge modification. Moreover, the presence of this protein with a probable PTM could be indicative of extensive energy requirement of AK39 strain.

3.2.1.7. Cell Processes (Shape/Division/Motility)

With its gene located on pSCL4, plasmid partitioning protein ParB was the sole overrepresented protein.

Plasmid partitioning protein ParB along with ParA protein regulates DNA segregation (Ebersbach and Gerdes, 2004). However, in bacteria, there are also homologs of ParAB proteins encoded by the genes on the chromosome which are crucial for the chromosome segregation, as well (Jensen and Gerdes, 1999).

3.2.1.8. Inorganic/Organic Molecule Transport

There were two proteins in this category which are phosphoenolpyruvate(PEP)-protein phosphotransferase and phosphoenolpyruvate-dependent sugar phosphotransferase. Phosphoenolpyruvate(PEP)-protein phosphotransferase was the common protein identified by both MALDI-TOF/MS and LC-MS/MS analyses.

Phosphoenolpyruvate-dependent sugar phosphotransferase coded by *ptsA* and phosphoenolpyruvate-protein phosphotransferase (PTS), which are the parts of the PTS system, comprise a transport mechanism for sugar moieties such as monosaccharides, disaccharides, amino sugars, polyols, and other sugar derivatives in bacteria (Dörschug et al., 1984; Deutscher et al., 2006). The upregulation of these proteins might indicate an increased energy requirement in AK39.

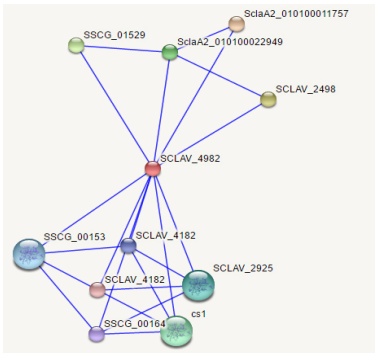
3.2.1.9. Hypothetical/Unknown Proteins

There were a total of 13 unknown/hypothetical proteins shown to be upregulated in this category. Of these, 5 of them were uncharacterized proteins.

DUFs refer to "domains of unknown function" and include a large set of families in the Pfam database with proteins of no known functions. Furthermore, although they are called DUFs, it is not certain whether they contain one or more domains. (Bateman et al., 2010). Moreover, by using structural genomics programmes, Jaroszewski et al. (2009) showed that about two thirds of them are likely to be the divergent members of existing protein superfamilies. As Goodacre et al. (2014) showed in 16 model bacterial

species, approximately 67% of DUFs are biologically essential and hence called as "essential DUFs" or eDUFs.

DUF1906 domain-containing protein was represented on two spots on the gel within the same pI range but distinctly located, suggesting a possible PTM on the protein that changed the mass of the protein and perhaps it suggests a possible cleavage. According to STRING database, DUF1906 is associated with clavaminic acid synthase and clavaminate synthase (Figure 3. 6). Furthermore, EggNog database was searched for the orthologous proteins of DUF1906 and for their predicted function based on sequence similiarity. It was predicted to be a peptidoglycan binding domain protein, having 51 orthologues in 41 species. Although DUF1906 domain-containing protein is predicted to have a relationship clavam synthesis or peptidoglacan binding proteins, it cannot go beyond a hypothesis until it is proven experimentally.



Your Input:

		Neighborhood	Gene Fusion	Cooccurence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
SCLAV_4982	*DUF1906 domain-containing protein (750 aa)*									

Predicted Functional Partners:

SclaA2_010100011757	*no annotation available (576 aa)*			•						0.697
SCLAV_2498	*Putative membrane protein (595 aa)*			•						0.694
SSCG_01529	*Putative uncharacterized protein (575 aa)*			•						0.688
SclaA2_010100022949	*Putative membrane protein (904 aa)*			•						0.610
cs1	*Clavaminic acid synthase 1 (324 aa)*			•						0.590
SCLAV_2925	*no annotation available (324 aa)*			•						0.590
SSCG_00153	*Clavaminate synthase (325 aa)*			•						0.578
SCLAV_4182	*Orf16 (401 aa)*			•						0.539
SSCG_00164	*no annotation available (401 aa)*			•						0.539
SCLAV_4182	*no annotation available (401 aa)*			•						0.539

Figure 3. 6. STRING data showing several associations of DUF1906 domain-containing protein with other proteins.

As for the predicted acetyltransferase (SCLAV_p1317), EggNog database shows that the orthologues of this protein represent a GCN5-related N-acetyltransferase. GCN5-related N-acetyltransferase (GNAT) superfamily is a very large family housing enzymes with very different substrate preferences. However, the main function of all the enzymes in this superfamily is to transfer the acetyl group of acetyl-CoA to an acceptor (primary amine) (Dyda et al., 2000).

For the uncharacterized protein (SCLAV_p0800), STRING database could not retrieve any protein information. When searched through EggNog database, there were only two protein entries showing resemblance to the protein queried. Tellurite resistance protein showed the similarity with a higher probability. Furthermore this protein coded by *SCLAV_p0800* is located on plasmid pSCL4; therefore, it is highly probable that this protein is involved in the secondary metabolism than the primary metabolism.

As for uncharacterized protein (SCLAV_p1086) and uncharacterized protein (SCLAV_p1312), neither STRING database nor EggNog could retrieve any results.

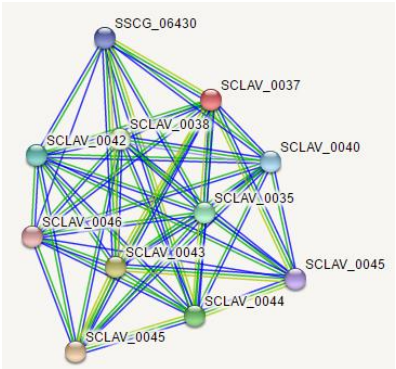
When putative uncharacterized protein (SCLAV_1941) was searched through EggNog database, it returned with 36 orthologous proteins from 34 species but all were unknown ones. STRING database predicted that it can be associated with 10 proteins, only three of which were annotated while the others were unknown (data not shown).

There were found to be 7 orthologous proteins from 5 species in EggNog database for the putative uncharacterized protein (SCLAV_3208) and the function of them could not be predicted. STRING database found that this protein is associated with 4 proteins two of which are unknown (data not shown).

When KR domain family protein was searched through EggNog database, there were 16 orthologous proteins depicted as luciferase-alpha subunit in 15 species of the bacterial kingdom.

Methyl-accepting chemotaxis sensory transducer was shown to have DUF312 domain with no known function according to UniProtKB database.

Putative phage tail sheath protein (SCLAV_0037) is another protein with unknown function. Association network predicted by STRING database (Figure 3. 7) showed that this protein might be interacting or associated with two Rhs element Vgr proteins and two phage tail region proteins. The fact that it has functional partners like phage region and Vgr proteins suggests a relation of this protein with contact-dependent Type VI secretion system which was identified in *Vibrio cholerae* (Pukatzki et al., 2006) and *Pseudomonas aeruginosa* to export Hcp (Haemolysin-Coregulated Protein) and a class of proteins named Vgr (Val-Gly Repeats) (Mougous et al., 2006). T6SS is mostly confined to pathogenic bacteria but also found in non-pathogenic ones. Their functions range from stress sensing and bacterial interactions to targeting cells (Schwarz et al., 2010; Costa et al., 2015). Moreover, apart from their involvement in pathogenesis, Vgr proteins were suggested to play roles in self-recognition of the bacteria to set boundaries with non-self bacteria in a niche (Gibbs et al., 2008). These finding has a prominent importance in that they constitute the first evidence suggesting T6SS in *S. clavuligerus*. In the following Sections (Section 3.4.1.11 and 3.4.1.12), two proteins that were upregulated in DEPA were also shown to have a network of interaction with this protein (SCLAV_0037).



Your Input:

SCLAV_0037 *Phage tail sheath protein, putative (754 aa)*							

Predicted Functional Partners:

	Neighborhood	Gene Fusion	Cooccurence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
SCLAV_0045 *Putative GPW/gp25 family protein (134 aa)*	•	•							0.983
SCLAV_0043 *Rhs element Vgr protein (393 aa)*	•	•							0.982
SCLAV_0038 *Phage tail region protein (151 aa)*	•	•							0.981
SCLAV_0044 *Rhs element Vgr protein (234 aa)*	•	•							0.973
SCLAV_0035 *Putative uncharacterized protein (192 aa)*	•	•							0.962
SCLAV_0042 *Peptidoglycan-binding LysM (223 aa)*	•	•							0.960
SCLAV_0040 *Phage tail region protein (172 aa)*	•	•							0.954
SSCG_06430 *no annotation available (134 aa)*	•	•							0.942
SCLAV_0045 *no annotation available (134 aa)*	•	•							0.942
SCLAV_0046 *Putative uncharacterized protein (648 aa)*	•	•							0.939

Figure 3. 7. STRING data showing the associaton of phage tail sheath protein with other proteins.

Another protein categorized in this group was putative alpha glucosidase. Alpha glucosidases are responsible for the release of D-glucose moieties from the carbohydrate substrates (Chiba, 1997).

3.2.1.10. Others/General Function

There were 4 proteins identified in this category: Aldehyde dehydrogenase (ALDH), ATPase, PRC-barrel domain protein, peptidase S1 and S6.

Aldehyde dehydrogenase (ALDH) comprises a large superfamily that contain enzymes catalyzing the oxidation of a variety of aromatic and aliphatic aldehydes. Depending on the specificities of the enzymes, they can be categorized as (i) semialdehyde dehydrogenases, (ii) non-specific ALDHs (iii) betaine dehydrogenases (iv) non-phosphorylating glyceraldehyde 3-phosphate dehydrogenases, (v) phenylacetaldehyde dehydrogenases, (vi) lactaldehyde

dehydrogenases, (vii) ALDH-like proteins. (Lindahl, 1992; Sophos and Vasilio, 2003). At the level of our proteomics study, it is quite hard to attribute any functions to any of these enzymes and suggesting any mechanism that could be related to CC overproduction.

PRC-barrel domain proteins show differential functions depending on species and the protein which they are a part of. For example, in purple bacteria, they are involved in photosynthesis, and in *E. coli*, Rim protein containing PRC barrel domain is involved in translation and 16S RNA processing (Anantharaman and Aravind, 2002).

Peptidase S1 and S6 are serine peptidases which are also known as S1 (chymotrypsin)/family S6 (Hap) serine peptidases (Kahel-Raifer et al., 2010). Members of S1 (chymotrypsin) family are endopeptidases with a range of biological functions (Rawlings and Barrett, 1993).

3.2.2. Downregulated Proteins in AK39

The spots with more than 2.5-fold decreased intensity as compared to their counterparts in NRRL 3585 are shown in Figure 3. 8 that indicates the proteins that are downregulated in both AK39 and TB3585 strains. For this, NRLL3585 gel was used for a much better visualization of the differentially expressed proteins. Downregulated proteins in AK39 according to MALDI-TOF/MS and LC-MS/MS analyses are listed in Table 3. 3 and Table 3. 4, respectively. 32 proteins were shown to be downreguated by 2DE MALDI-TOF/MS method while 40 proteins were downregulated in LC-MS/MS.

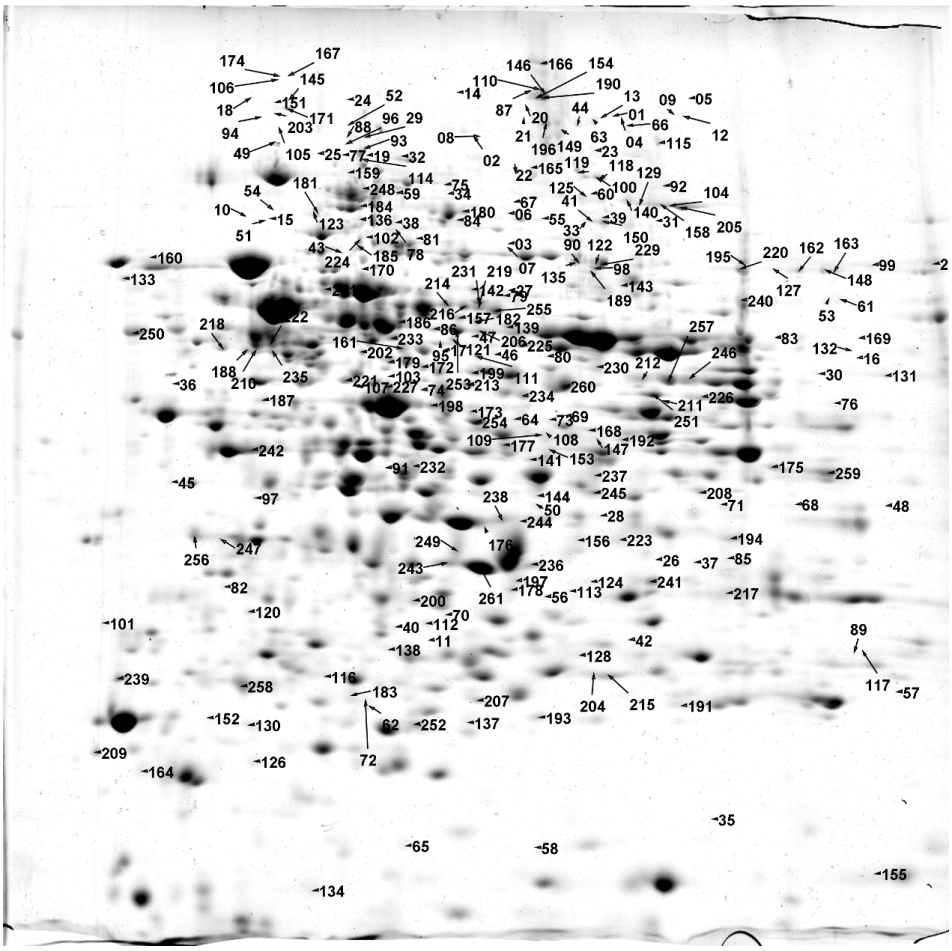


Figure 3. 8. Cytosolic protein spots on NRRL 3585 2DE gel that showed more than 2.5-fold higher intensity when compared to their counterparts on AK39 and TB3585 gels.

Table 3. 3. List of downregulated proteins identified on AK39 2DE gel.

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spots	PTM **	Genome Location ***	Spot #
Nucleotide Metabolism (DNA/RNA Processing)	B5GUB0	Uracil phosphoribosyltransferase	*upp*	Cytoplasm	-3.5	4.97	22,843	–	–	Chr.	207
Carbohydrate Metabolism	B5GL42	UDP-glucose 4-epimerase	*SCLAV_4282*	Cytoplasm	-3.8	5.57	35,148	–	–	Chr.	223
	E2PVD8	Putative isocitrate dehydrogenase	*icdA*	Cytoplasm	-3.2	4.97	79,048	–	–	Chr.	181
	E2Q387	Succinyl-CoA ligase [ADP-forming] subunit beta	*sucC*	Cytoplasm	-2.9	4.74	40,845	–	–	Chr.	242
Amino Acid Metabolism	B5GVF3	Serine hydroxymethyltransferase	*glyA*	Cytoplasm	-3.5	5.94	45,125	–	–	Chr.	226
	B5GSL5	4-hydroxyphenylpyruvate dioxygenase	*SCLAV_2046*	Cytoplasm	-47.6	5.11	41,873	–	–	Chr.	232
	B5GUR5	5-carboxymethyl-2-hydroxymuconate delta-isomerase	*SCLAV_4800*	Cytoplasm	-2.5	5.34	29,825	–	–	Chr.	236
	D5SLL7	Methionine synthase II	*SCLAV_p1324*	Cytoplasm	-2.9	4.86	38,157	–	–	pSCL4	242
	B5GWG1	Fumarylacetoacetase	*SCLAV_3578*	Unknown	-3.4	5.46	43,994	–	–	Chr.	254
	B5H1V1	D-aminoacylase	*SCLAV_3902*	Cytoplasm	nd*	5.32	57,695	–	–	Chr.	182
	B5GWB7	Glutamate-1-semialdehyde 2,1-aminomutase	*hemL*	Cytoplasm	-3.4	5.52	47,229	–	–	Chr.	254
	B5GZV9	E1-alpha branched-chain alpha keto acid dehydrogenase	*bkdA1*	Cytoplasm	-6.6	5.68	40,843	–	–	Chr.	192

Table 3. 3. List of downregulated proteins identified on AK39 2DE gel. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spots	PTM **	Genome Location ***	Spot #
Lipid Metabolism	E2Q061	Enoyl-CoA hydratase	*echA2*	Cytoplasm	nd*	5.16	27,211	–	–	Chr.	112
	B5GQ29	Acyl CoA isomerase	*SCLAV_4002*	Cytoplasm	-2.7	5.52	27,714	–	–	Chr.	215
	B5H1P2	Hydroxymethylglutaryl-CoA lyase	*SCLAV_1945*	Cytoplasm	-13.9	5.49	33,866	–	–	Chr.	176
	B5GMB5	Acyl-CoA dehydrogenase domain-containing protein	*SCLAV_p0938*	Cytoplasm	-3.8	5.46	40,818	–	–	.pSCL*4*	237
	B5GZV9	E1-alpha branched-chain alpha keto acid dehydrogenase	*bkdA1*	Cytoplasm	-6.6	5.68	40,843	–	–	Chr.	192
DNA Replication, Recombination, Repair, Transcripton	D5SII7	Transcriptional regulator TraR	*traR*	Cytoplasm	-2.8	5.15	32,335	–	–	pSCL4	112
	E2Q4E3	Exodeoxyribonuclease V	*recD2*	Cytoplasm	-3.2	6.77	79,569	–	–	Chr.	181
Translation, Ribosomal Structures	B5GX56	Elongation factor G	*fusA2*	Cytoplasm	-3.2	4.99	77,548	–	–	Chr.	181
Stress-Related, Protein Turnover, Chaperones	E2PYD4	Ectoine hydroxylase	*ectD*	Unknown	-50.0	5.64	32,419	–	–	Chr.	124
	E2Q3P1	Trigger factor	*tig*	Cytoplasm	-5.0	4.50	55,724	–	–	Chr.	133
Secondary Metabolism	E2PYD4	Ectoine hydroxylase	*ectD*	Unknown	-50.0	5.64	32,419	–	–	Chr.	124
Energy Production/Electron-Iron Transfer	B5GL54	ATP synthase subunit alpha	*atpA1*	Cytoplasm	nd*	5.05	57,652	–	–	Chr.	182

Table 3. 3. List of downregulated proteins identified on AK39 2DE gel. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spots	PTM **	Genome Location ***	Spot #
Energy Production/Electron-Iron Transfer	B5GL52	ATP synthase subunit beta	*atpB*	Cytoplasmic membrane	-2.8	4.81	52,409	–	–	Chr.	235
Cell Processes (Shape/Division/Motility)	E2PZS0	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase	*murE*	Cytoplasm	-6.3	5.39	58,036	–	–	Chr.	219
Inorganic/Organic Molecule Transport	B5GTX3	BldKB	*SCLAV_3976*	Cell membrane	-3.4	8.34	65,261	–	–	Chr.	220
Hypothetical/Unknown Proteins	B5H0U3	DUF574 domain-containing protein	*SCLAV_0131*	Cytoplasm	-5.1	4.82	30,332	–	–	Chr.	120
	E2PZM3	Putative uncharacterized protein	*SCLAV_1253*	Cytoplasm	-2.6	6.46	23,056	–	–	Chr.	57
Others/General Function	E2Q811	Amidase	*SCLAV_4472*	Cytoplasm	nd*	5.28	50,210	–	–	Chr.	111
	E2Q3Q1	Acetyltransferase	*SCLAV_1797*	Cytoplasm	-4.3	5.44	44,545	–	–	Chr.	177
	E2PXK3	Oxidoreductase	*SCLAV_1047*	Cytoplasm	-3.5	5.35	25,729	–	–	Chr.	207
	E2Q3I4	Flavoprotein disulfide reductase	*SCLAV_3827*	Cytoplasm	-6.3	5.48	55,168	–	–	Chr.	219
	E2Q6Q8	LmbE family protein	*SCLAV_4283*	Cytoplasm	-2.6	6.30	25,212	–	–	Chr.	57

*nd: not determined. ** PTM refers to possible postranslational modifications. ***"Chr." refers to chromosome and "pSCL4" refers to plasmid pSCL4.

Table 3. 4. List of downregulated proteins in AK39 strain identified by LC-MS/MS.

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Amino Acid Metabolism	B5GL71	Threonine synthase	*SCLAV_4251*	Cytoplasm	-2,945	5.58	37,34	Chr.
	E2PYZ0	Threonine synthase	*SCLAV_3250*	Cytoplasm	-2,310	5.52	47,294	Chr.
	E2Q6M5	Homoserine dehydrogenase	*SCLAV_4250*	Cytoplasm	-2,738	5.46	44,983	Chr.
Lipid Metabolism	E2Q773	Methylmalonyl-CoA mutase small subunit	*mutA*	Cytoplasm	-2,582	5,16	65,042	Chr.
	E2Q711	Biotin carboxylase	*SCLAV_0182*	Cytoplasm	-2,407	5.14	46,512	Chr.
DNA Replication, Recombination, Repair, Transcripton	B5GPA7	TetR-family transcriptional regulator	*SCLAV_3146*	Cytoplasm	-2,407	5.16	28,388	Chr.
Stress-Related, Protein Turnover, Chaperones	E2PV72	Peroxidase	*SCLAV_0743*	Cytoplasm	-2,945	5.88	55,001	Chr.
	E2PYD1	L-2,4-diaminobutyric acid acetyltransferase	*ectA1*	Cytoplasm	-2,207	6.43	21,913	Chr.
Secondary Metabolism	D5SIN8	Non-ribosomal peptide synthetase	*SCLAV_p0290*	Cytoplasm	-3,897	5.46	154,508	pSCL4
	E2Q805	Hybrid NRPS / PKS	*SCLAV_4466*	Cytoplasm	-3,665	5.51	525,974	Chr.
	B5GS02	Non-ribosomal peptide synthetase	*SCLAV_p0303*	Cytoplasm	-3,125	5.34	139,482	pSCL4
	D5SIN9	Non-ribosomal peptide synthetase	*SCLAV_p0291*	Cytoplasm	-2,497	5.34	72,96	pSCL4
	E2PWJ7	Clavaminate synthase 1	*cas1*	Cytoplasm	-2,497	5.31	35,37	Chr.

Table 3. 4. List of downregulated proteins in AK39 strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Secondary Metabolism	D5SIQ2	Hybrid PKS / NRPS	*SCLAV_p0304*	Cytoplasmic membrane	-2,310	5.07	135,795	pSCL4
	E2PWJ8	Aldo/keto reductase	*cvm1*	Cytoplasm	-2,096	6.41	36,65	Chr.
	E2PWJ4	Putative pyridoxal phosphate-dependent aminotransferase	*cvm6*	Cytoplasm	-2,207	5.36	48,065	Chr.
	Q9X5G6	Putative ribulose-5-phosphate epimerase	*cvm2*	Cytoplasm	-2,407	5.27	16,877	Chr.
	E2PYD1	L-2,4-diaminobutyric acid acetyltransferase	*ectA1*	Cytoplasm	-2,207	6.43	21,913	Chr.
	Q9X5G9	Flavin-dependent oxidoreductase, F420-dependent methylene-tetrahydromethanopterin reductase	*cvm5*	Cytoplasm	-3,384	5.85	44,863	Chr.
Energy Production/Electron-Iron Transfer	E2Q1D4	NADH-quinone oxidoreductase subunit D	*nuoD1*	Cytoplasm	-2,497	5.29	50,204	Chr.
	B5GWE3	Electron transfer oxidoreductase	*SCLAV_3561*	Cytoplasm	-2,310	7.71	46,765	Chr.
Hypothetical/Unknown Proteins	E2PWW5	Putative M28-family peptidase	*SCLAV_3043*	Extracellular	-2,879	6.85	52,975	Chr.
	B5H0U3	DUF574 domain-containing protein	*SCLAV_0131*	Cytoplasm	-2,096	4.82	30,332	Chr.
	E2Q809	Putative secreted lipase	*SCLAV_4470*	Extracellular	-4,141	5.36	33,675	Chr.
	E2Q9C1	Putative oxidoreductase	*SCLAV_0465*	Unknown	-3,068	6.22	37,219	Chr.
	E2Q829	Putative sensor-like histidine kinase	*SCLAV_4490*	Cytoplasm	-2,662	5.72	122,486	Chr.

Table 3. 4. List of downregulated proteins in AK39 strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Hypothetical/Unknown Proteins	B5GZ79	Putative zinc-binding dehydrogenase	*SCLAV_1520*	Cytoplasm	-2,207	5.27	34,906	Chr.
	E2Q9K7	Putative uncharacterized protein	*SCLAV_0551*	Cytoplasm	-2,310	5.13	61,991	Chr.
	E2Q3B3	Putative ornithine aminotransferase	*SCLAV_3756*	Cytoplasm	-2,207	5.82	43,699	Chr.
	E2PW58	Secreted protein	*SCLAV_0814*	Cell wall, Extracellular	-2,407	8.45	47,613	Chr.
	D5SIN7	Methylase involved in ubiquinone/menaquinone biosynthesis-like protein	*SCLAV_p0289*	Cytoplasm	-2,879	5.22	37,995	pSCL4
Others/General Function	E2Q811	Amidase	*SCLAV_4472*	Cytoplasm	-2,945	5.28	50,21	Chr.
	E2Q6I6	Cellulose-binding protein	*SCLAV_2237*	Cytoplasm	-2,407	4.70	32,156	Chr.
	E2Q6R1	Radical SAM domain protein	*SCLAV_4286*	Cytoplasm	-3,235	5.89	38,475	Chr.
	E2Q6R2	NAD-dependent epimerase/dehydratase	*SCLAV_4287*	Cytoplasm	-2,207	5.93	30,116	Chr.
	E2Q6R0	GCN5-related N-acetyltransferase	*SCLAV_4285*	Cytoplasm	-2,407	7.09	34,671	Chr.
	E2Q0X0	Integral membrane protein	*SCLAV_1434*	Cytoplasmic membrane	-2,662	7.92	53,605	Chr.
	E2Q1E6	Phenylacetate-coenzyme A ligase	*SCLAV_5505*	Cytoplasm	-2,945	6.20	48,951	Chr.
	B5GV03	Nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	*cobT*	Cytoplasm	-2,207	4.60	40,158	Chr.

Table 3. 4. List of downregulated proteins in AK39 strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Others/General Function	E2PW77	Phenylacetaldehyde dehydrogenase	*feaB*	Cytoplasm	-2,310	4.76	52,612	Chr.
	E2Q806	Short-chain dehydrogenase/reductase SDR	*SCLAV_4467*	Cytoplasm	-2,096	9.59	55,878	Chr.

* "Chr." refers to chromosome, "pSCL4" refers to plasmid pSCL4.

3.2.2.1. Nucleotide Metabolism (DNA/RNA Processing)

There was one protein identified in this category. Uracil phosphoribosyltransferase (UPRT) encoded by *upp* gene is found in the pyrimidine salvage pathway and converts the uracil and 5-phosphoribosyl-α-1-pyrophosphate (PRPP) to UMP and pyrophosphate (Arsène-Ploetze et al., 2006). All pyrimidine nucleotides are synthesized using UMP as precursor. By the pyrimidine salvage pathway, exogenous or endogenous uracil which is the product of degradation of nucleic acids in the cells can be recycled to produce pyrimidine nucleotides (Rasmussen et al., 1986; Andersen et al., 1992). In *Streptomyces*, stationary phase marks the morphological differentiaton from vegetative mycelium to aerial hyphae formation. Pyrimidine salvage pathway is one of the pathways that the organism resorts to for re-cycling of the metabolites to conserve energy (Hughes et al., 2005).

3.2.2.2. Carbohydrate Metabolism

There were shown to be three proteins downregulated in this category: UDP-glucose 4-epimerase, putative isocitrate dehydrogenase, and succinyl-CoA ligase [ADP-forming] subunit beta.

UDP-glucose 4-epimerase is one of the proteins involved in carbohydrate metabolism of *Streptomyces. S. clavuligerus* differs from other *Streptomyces* strains in glucose utilization. It cannot grow in a medium containing glucose as the sole carbon source, neither can it utilize hexoses (e.g., mannose, galactose or fructose), pentoses (e.g., xylose), sugar alcohols (e.g., mannitol or sorbitol), or disaccharides (e.g., lactose or sucrose) (Pérez-Redondo et al., 2010). Therefore, at the first glance, it was not expected to observe the expression of UDP-glucose 4-epimerase in *S. clavuligerus*, which is responsible for the reversible conversion of galactose to glucose 1-phosphate (Thoden et al., 1996). However, *S. clavuligerus* contains genes for both the transport and utilization of glucose (Pérez-Redondo et al., 2010). In order to reveal which system is defective in glucose utilization, Pérez-Redondo et al. (2010) conducted a search on glucose permease (*glcP*) and glucose kinase (*glkA*) genes and showed that once glucose is transported into the cell, it can be utilized successfully, proving that *glkA* genes are functional. They also show that *glcP* (glucose permease) gene

displayed a very weak expression which was not enough for the growth of the organism on glucose. Therefore, they expressed *glcP* gene of *S. clavuligerus* by using *S. coelicolor glcP1* promoter and showed that *S. clavuligerus* could grow on glucose, which confirmed that it was the weak promoter of *glcP* gene that conferred *S. clavuligerus* unable to grow on glucose (Pérez-Redondo et al., 2010). Although *glcP* is a weak promoter, it shows some residual activity, in other words, gene is expressed but with very low efficiency not sufficient for the organism to grow on glucose (Garcia-Dominguez et al., 1989; Pérez-Redondo et al., 2010). Garcia-Dominguez et al., (1989) had also confirmed that it was the transport system that prevented the organism from utilizing glucose. They could isolate *gut-1* mutant of *S. clavuligerus* which restored the glucose transport into the cell and this restored system permitted the uptake of not only glucose but also galactose molecules (Garcia-Dominguez et al., 1989). As a result, the fact that downregulation of UDP-glucose 4-epimerase is not unexpected since wild type most probably takes up both glucose and galactose at residual levels and that the pathways utilizing glucose and galactose are intact in *S. clavuligerus*. Therefore, it might be possible that uptake of these carbon sources might have decreased significantly in AK39 strain compared to the parental one. Recently, it has been found that TunF protein which was isolated from *S. chartreusis* is involved in tunicamycin biosynthesis and is a UDP-GlcNAc 4-epimerase (Wyszynski et al., 2012) Hence, the amino acid sequences of TunF (also known as Tun6) isolated from *S. chartreusis* and UDP-glucose-4-epimerase (SCLAV_4282) identified in this study were compared. Blastp results showed 100% sequence coverage and 76% identity between these two proteins. However, there has been no other report on the tunicamycin biosynthesis in *S. clavuligerus* except for the finding of a tunicamycin-like complex by Kenig and Reading (1979). Medema et al. (2010) who reported the first draft genome sequence of *S. clavuligerus* along with Song et al. (2010a) did not demonstrate the tunicamycin-related antibiotic gene cluster in the genome. It then becomes necessary to conduct further studies to find out the main function of UDP-glucose-4-epimerase (SCLAV_4282) and its involvement in the secondary metabolism of this organism.

Isocitrate dehydrogenase is common to 3 kingdoms of life as a result of its importance in carbon metabolism (Zhang et al., 2013b). It is involved in citric acid cycle and converts the isocitrate to α-ketoglutarate which is also an important precursor for the biosynthesis of several metabolites in cells (Hurley et al., 1996; Takahashi-Iñiguez et al., 2014). One of the metabolites that is produced from α-ketoglutarate is glutamate which is involved in the biosynthesis of several amino acids such as arginine, aspartate, valine and serine.

Succinyl-CoA ligase, also known as succinyl-CoA synthetase, catalyzes the reversible reaction from succinate to succinyl-CoA in the citric acid cycle as shown below. N denotes adenosine or guanosine (Fraser et al., 1999).

Succinyl-CoA + NDP + Pi ↔ Succinate + CoA + NTP

Downregulated proteins in the "Carbohydrate Metabolism" of AK39 mostly comprised proteins that lead to production of intermediates for the primary metabolism of the cell, which suggests that central metabolism in this targeted mutant has been significantly altered.

3.2.2.3. Amino Acid Metabolism

There were 11 proteins which we showed as downregulated in amino acid metabolism of AK39 strain.

Serine hydroxymethyltransferase is involved in one-carbon metabolism and catalyzes the reversible conversion of serine to glycine and tetrahydrofolate to 5,10-methylene tetrahydrofolate, which is important for the biosynthesis of purine, thymidine, choline, and methionine (Trivedi et al., 2002). The generation of one-carbon units in the cell mostly depends on this reaction catalyzed by SHMT and commonly found in three kingdoms (Barra et al., 1983). Morover, serine is also used as a precursor for cysteine biosynthesis. L-cysteine is one of the three precursors necessary for the biosynthesis of cephamycin C in *Streptomyces*. Moreover, in clavam biosynthesis gene clusters and paralogue gene clusters of them, there were shown to be genes coding for serine hydroxymethyltransferases (Zelyas et al., 2008; Álvarez-Álvarez et al., 2014). For example, a recently identified paralog gene cluster for alanylclavam biosynthesis in *S.*

clavuligerus was shown to contain *orfA* that encodes a putative serine hydroxymethyltransferase, the absence of which inhibits the production alanylclavam but not other other clavams including CA (Zelyas et al., 2008). Hence, this protein might not only be involved in the generation of intermediates for basic cellular processes but also in the the secondary metabolism of *Streptomyces*. Downregulation of this protein is not unexpected given its importance in the production of CA and/or clavams. Furthermore, genomic data shows that there are four genes encoding serine hydroxymethyltransferase on the genome of *S. clavuligerus*.

4-hydroxyphenylpyruvate dioxygenase is involved in tyrosine catabolism, the products of which can be used in a variety of biological processes (Brownlee et al., 2004). L-tyrosine is degraded to 4-hydroxyphenylpyruvatic acid and homogentisic acid intermediates to produce acetoacetic acid and fumaric acid end products through which the conversion of 4-hydroxyphenylpyruvatic acid to homogentisic acid is catalyzed by 4-hydroxyphenylpyruvate dioxygenase. Homogentisic acid, on the other hand, is known as a precursor for the synthesis of secondary metabolites like tocopherols and tocopherylquinones produced by photosynthetic microbes and plants (Luckner, 2013). Through the tyrosine catabolism, fumarylacetoacetase, catalyzes the last step leading to the production of fumarate and acetoacetate from 4-fumaryl-acetoacetate in the tyrosine catobolism (Kvittingen, 1985). Fumarate and acetoacete products are required for the growth of the organism and are involved in Krebs cycle (Yang et al., 2007).

5-carboxymethyl-2-hydroxymuconate delta-isomerase is involved in the catabolism of aromatic compounds which can be used as energy and carbon sources (Yan et al., 2009) Downregulation of this protein along with 4-hydroxyphenylpyruvate dioxygenase and fumarylacetoacetase might suggest that aromatic amino acid catabolism is somehow downregulated in AK39.

Cobalamin-independent methionine synthase (Methionine synthase II) identified in this study is responsible for the production of tetrahydropteroyltriglutamate and methionine by tranferring a methyl group from 5-methyltetrahydropteroyltriglutamate

to homocysteine in the presence of magnesium and phosphate ions (Whitfield et al., 1970) and it was extensively explained in Section 3.4.2.3.

D-aminoacyclases convert L-amino acids to D-amino acids. Moriguchi and Ideta (1988) and Tsai et al. (1988) reported the presence of D-aminoacyclases, in *Alcaligenes*, which are also called "abnormal amino acids", as D-amino acids are used by some microorganisms to synthesize antibiotic peptides (Sugie and Suzuki, 1978; Muniz-Lozano et al., 1998). Furthermore, D-amino acids such as D-valine and D-phenylglycine are used as intermediates for the preparation of various pesticides, bioactive peptides, and antibiotics (Tsai et al., 1988). These amino acids seem to have important roles for the biological activity of peptides (Sugie and Suzuki, 1978). Morover, D-aminoacyclases have been lost from AK39 as revealed by 2DE analysis, suggesting that the primary and secondary metabolite production by AK39 might have been altered more than we predicted in our metabolic engineering program.

Glutamate-1-semi-aldehyde 2,1-aminomutase is one of the proteins identified in amino acid metabolism shown to be significantly downregulated. This protein is involved in C5 pathway which gives rise to δ -aminolevulinic acid (ALA). ALA is formed through three steps in C5 pathway. First, tRNA is charged with glutamate (Glu) which is then reduced to semialdehyde and finally glutamate-1-semi-aldehyde 2,1-aminomutase catalyzes the transamination of the semialdehyde to produce ALA (Petříček et al., 2006). δ-aminolevulinic acid (ALA) is important in that the biosynthesis of all tetrapyrroles starts with the formation of ALA (Heinemann et al., 2008). Tetrapyrrole pathway produces hemes, chlorophylls, billins, and corrinoids using ALA as the precursor. (Petříček et al., 2006). The C5 pathway is observed in plants and commonly in some Gram-positive and Gram-negative bacteria (Hashimoto et al., 1997) Alternative pathway to C5 to produce ALA is Shemin pathway in which ALA is formed by the condensation of succinyl-CoA and glycine. Both C5 and Shemin pathways were discovered in *S. nodosus* subsp. *asukaensis* (Heinemann et al., 2008). However, only ALA produced throughout Shemin pathway was shown to give rise to the production of antibiotic production such as asukamycin, reductiomycin, and moenomycin in *Streptomyces* (Petříček et al., 2006). Different types of tetrapyrroles

can have diverse functions in cells and can be used as cofactors of a large variety of enzymes involved in diverse processes such as primary metabolism, transcriptional regulation, energy metabolism as well as detoxification mechanisms (de Orué Lucana et al., 2004; Zappa et al., 2010; Takano et al., 2015). For example, a knock-out mutant of *S. coelicolor* lacking the synthesis of a certain tetrapyyrole was shown to be defective in both the primary metabolism and the developmental growth (Takano et al., 2015). Although δ -aminolevulinic acid production might contribute to diverse cellular processes, Given that δ -aminolevulinic acid production might contribute to diverse cellular processes and the proteins leading to the production of them is involved in C5 pathway, they most probably contribute to the primary metabolism of the organism; yet the pathway leading to its production and its role in *S. clavuligerus* should be searched further.

E1-alpha branched-chain alpha keto acid dehydrogenase is a part of the multienzyme complex (BCDH complex) that is composed of four components: a BCDH and decarboxylase (E1 α and E1 β), a dihydrolipoamide acyltransferase (E2), and a dihydrolipoamide dehydrogenase (E3). Oxidative decarboxylations of α-keto-isovalerate, α-keto-β-methylvalerate, and α-ketoisocaproate, which are the deamination products of the branched-chain amino acids valine, isoleucine, and leucine, respectively, are carried out by BCDH complex and acyl-CoA analogs and NADH are produced as a result (Denoya et al., 1995). BCDH in *B. subtilis* was shown to have a crucial role in fatty acid synthesis (Skinner et al, 1995). Denoya et al. (1995) further showed that for the polyketide antibiotic biosynthesis in *Streptomyces avermitilis*, activity of BCDH complex is essential. Branched-chain amino acid catabolism is another aspect of the enzyme, showing its relationship with the amino acid metabolism (Harris et al., 1990). Given the versatile function of E1-alpha branched-chain alpha keto acid dehydrogenase, it might be contributing to the production of polyketides in *S. clavuligerus* which is a rich source of secondary metabolites. Hence, the decreased expression of this protein in AK39 strain might suggest the decreased production of polyketides in our targeted mutant to increase the CC production in the cell.

Homoserine dehydrogenase catalyzes the conversion of aspartic semialdehyde (branch point intermediate) to homoserine in aspartate pathway and directs the branch that starts with homoserine formation and ends up with the production of threonine, methionine and isoleucine (Figure 1.9). The other branch, also known as lysine branch, leads to the CC biosynthesis via formation of α-aminoadipic acid (Mendelovitz and Aharonowitz, 1982; Yılmaz et al., 2008). Actually, since the gene coding for homoserine dehydrogenase was genetically disrupted in AK39, the level of this protein should have been zero. As a matter of fact, protein expression of the protein was completely absent from AK39 strain according to the raw data obtained from LC-MS/MS analysis.

Threonine synthase acts on the homoserine intermediate produced by homoserine dehydrogenase through aspartate pathway and catalyzes the conversion of L-homoserine to L-threonine (Yılmaz et al., 2008). There were two threonine synthases identified as downregulated, actually not synthesized at all in this study. In the absence of the homoserine dehydrogenase, threonine synthase activity is no longer needed. The required threonine and methionine is externally supplied to the auxotrophic cells in synthetic medium which is even not needed in a complex medium like TBS.

3.2.2.4. Lipid Metabolism

There were 7 proteins downregulated in lipid metabolism. E1-alpha branched-chain alpha keto acid dehydrogenase was explained in Section 3.2.2.3.

Enoyl-CoA hydratase is involved in fatty acid β-oxidation pathway (Hofstein et al, 1999).

According to UniProtKB database, acyl CoA isomerase (SCLAV_4002) has also enoyl-CoA hydratase/isomerase conserved site and hence is involved in fatty acid metabolism, but its main function needs to be elucidated.

Hydroxymethylglutaryl-CoA lyase catalyzes the last step of ketogenesis and leucine catabolism and may play a role in fatty acid biosynthesis (Ashmarina et al., 1994; Li et al., 2005a).

Acyl-CoA dehydrogenase domain-containing protein is another protein in this category. Acyl-CoA dehydrogenases are flavoproteins and have a large variety of members (Ghisla and Thorpe, 2004). Although their substrates vary, all are responsible for the α, β dehydrogenation in fatty acid metabolism (Tiffany et al., 1997).

Methylmalonyl-CoA mutase catalzes the reversible conversion of succinyl-CoA to methylmalonyl-CoA during fatty acid degradation; moreover, methylmalonyl-CoAs produced can be used for the polyketide biosynthesis, as well (Jost et al., 2015). Methylmalonyl-CoA as well as malonyl-CoA are common precursors exploited as extender units to be incorporated into the polyketides (Walton et al., 2006).

Biotin carboxylase is one of the four components of acetyl-CoA carboxylases (Salie and Thelen, 2016). As explained in Section 3.2.1.2 in detail, acetyl-CoA carboxylases are involved in production of malonyl-CoA which forms the precursors for both fatty acid and polyketide biosynthesis (Rodríguez and Gramajo, 1999; Rodríguez et al., 2001).

3.2.2.5. DNA Replication, Recombination, Repair, Transcripton

There were three proteins in this category: Transcriptional regulator TraR, exodeoxyribonuclease V and TetR-family transcriptional regulator.

Transcriptional activator TraR protein is an autoinducer-dependent trancriptional regulator, (Vannini et al., 2002) and a quorum-sensing protein (Zhu and Winans, 1998). TraR protein has been widely studied in *Agrobacterium tumefaciens* which is a plant pathogen causing crown gall tumors. The mechanism of infection by *Agrobacterium tumefaciens* is via Ti plasmids that carry *vir* genes required for the tranfer of the oncogenic DNA. Once the DNA is transferred to plant cells, they begin to produce several nutrients for the pathogen such as opines and phytohormones (Vannini et al., 2002). Some of the released opines induce *tra/trb* regulon which, in turn, induces interbacterial DNA transfer (Zhu and Winans, 1998). TraR-TraI regulates the expression of this regulon through an opine-activated protein, which activates the transcription of *traR* gene (Zhu et al., 2000). TraR complexed with an autoinducer (AAI), which is synthesized by TraI, binds to *tra* boxes, activating the

transcription of related genes (e.g., *traB*) for the conjugal transfer (Vannini et al., 2002). *Streptomyces* are notorious for their large and linear plasmids and these plasmids have the necessary ammunition for the conjugal tranfer of not only the plasmid genes but also the chromosomal genes. As in *Agrobacterium*, the transfer of DNA requires plasmid-encoded *tra* gene on plJ101 plasmid in *S. lividans* (Pettis and Cohen, 1996). It was shown that *Streptomyces* cannot transfer DNA conjugally in liquid cultures (Hopwood et al., 1985); however, Pettis and Cohen (1996) showed that the concentration of Tra proteins is higher in liquid-grown cultures than those found in solid-grown cultures and that there was little or no expression of Tra protein during secondary metabolite production, suggesting that Tra protein is important during vegatative growth before the initiation of morphological differentiation in the microorganism (Pettis and Cohen, 1996). It can be hypothesized that the downregulation of TraR protein in AK39 is possibly linked to CC overproduction, yet this issue is to be elucidated through further studies.

Exodeoxyribonuclease V is responsible for binding ds-DNA and producing ss-DNA by unwinding the strands and can also hydrolyze ss-DNAs. It has crucial roles in several cellular activities such as homologous genetic recombination, recovery from DNA damage, maintenance of cell viability, and the destruction of damaged and foreign DNA (Amundsen et al., 1986). It might be proposed that DNA recombination or repair activities have been decreased in AK39 strain compared to the standard strain to reduce the energy cost for certain metabolic processes. Upregulation of RecA protein in this strain (Section 3.2.1.3) might also support this proposal.

TetR family of proteins are transcriptional regulators that mostly function as repressors but sometimes as activators. Although they are widely distributed in bacterial genomes, function of only 85 members of this family have been elucidated (Ramos et al., 2005). Furthermore, some *Streptomyces* species contain more than 100 TetR family protein coding genes, which must be a result of the complex intertwined differentiation and secondary metabolic processes observed in streptomycetes (Guo et al., 2013). The regulation of diverse processes by this family members range from

multidrug resistance, catabolic pathways, osmotic stress and antibiotic biosynthesis to pathogenicity (Ramos et al., 2005; Wei et al., 2014).

3.2.2.6. Translation, Ribosomal Structures

Elongation factor G was the sole protein shown to be downregulated in AK39 strain in this category, and as it is well known, is responsible for the translocation of transfer RNAs (tRNAs) and messenger RNA (mRNA) during translation by the ribosome (Gao et al., 2009).

3.2.2.7. Stress-Related, Protein Turnover, Chaperones

There were 4 proteins shown to be downregulated in this category: Ectoine hydroxylase, L-2,4-diaminobutyric acid acetyltransferase, trigger factor and peroxidase.

L-2,4-diaminobutyric acid acetyltransferase functions in 3-step ectoine biosynthesis pathway (Bursy et al., 2008). In the first step of the pathway, L-aspartate-β-semialdehyde is converted to L-2,4 diaminobutyric acid by L-2,4-diaminobutyrate transaminase (EctB). L-2,4-diaminobutyrate is then converted to. N^{γ}-acetyl-L-2,4-diaminobutyric acid by L-2,4-diaminobutyrat acetyltransferase (EctA) enzyme. At the last step, ectoine is synthesized by ectoine synthase (EctC). Ectoin hydroxylase (EctD) acts on the final product, namely ectoine, and converts it to 5-hydroxyectoine (Bursy et al., 2008). Only some microorganisms can convert the ectoin to 5-hydroxyectoine by EctD. Ectoine was first discovered in *Ectothiorhodospira halochloris*, an extremely halophilic phototrophic purple sulfur bacterium while 5-hydroxyectoine, which is the hydroxylated derivative of ectoin, was discovered in *Streptomyces parvulus* (Galinski et al., 1985; Inbar and Lapidot, 1985; Bursy et al., 2007). These molecules are compatible solutes produced during stress in the cell; for example, salt stress was shown to induce the production of 5-hydroxyectoine in *Streptomyces clavuligerus*, *S. griseus*, *S. parvulus*, *Streptomyces peucetius*, and *Streptomyces antibioticus* (Malin and Lapidot, 1996). Moreover, both ectoin and 5-hydroxyectoine have stabilizing properties on proteins, and they were shown to affect the DNA melting temperature (Tm) (ectoine increased Tm while 5-hydroxyectoine decreased Tm) (Schnoor et al.,

2004). Bursy et al. (2008) showed that *Streptomyces coelicolor* A3(2) accumulated ectoin and 5-hydroxyectoine under heat stress within the cells and suggested that the accumulation of these compatible solutes in the cell could change the melting temperature of DNA, especially at promoter regions, thus affecting the transcription through those promoter regions and the interaction of proteins with the target DNA sequences. In this study, ectoin hydroxylase is downregulated in AK39 strain. Response to stress conditions probably differ in parental and AK39 strains as a result of activation of the differential regulatory mechanism in both strains. AK39 might also be decreasing the metabolic burden like synthesizing certain molcules in the cell to cope with increased antibiotic production. Furthermore, the fact that both the ectoin production and the CC production requires the same precursor, L-aspartate- β-semialdehyde, and that AK39 strain has an increased level of CC production indicate that the precursor is channelled towards antibiotic production in AK39 strain rather than compatible solutes to protect the cell.

Trigger factor is a chaperone that binds the polypeptide chains while they are being translated on the ribosome thus preventing them from misfolding or aggregating (Wong and Houry, 2004). Downregulation of this protein is difficult to interpret, yet it is in accord with downregulation of elongation factor G (Section 3.2.2.6).

Peroxidases which include a broad group of enzymes can be divided in two categories: Haem peroxidases and non-haem peroxidases (le Roes-Hill et al., 2011). They use H_2O_2 to oxidize several compounds; for example, intracellular molecules such as steroids, fatty acids, prostaglandins, and leukotrienes and extracellular molecules such as drugs, insecticides, solvents, and hydrocarbons (Youn et al., 1995; Ricoux et al., 2008). They are widely distributed in plants, animals and microorganisms. In prokaryotes, there were also found peroxidases with both peroxidase and catalase activity (Mliki and Zimmermann, 1992).

The low expression of peroxidase, trigger factor and proteins involved in ectoine and hydroxyectoine biosynthesis might suggest that AK39 strain prefers to utilize other proteins like proteasome, chaperone and starvation-induced DNA protecting protein, the levels of which were found increased.

3.2.2.8. Secondary Metabolism

A total of 12 important proteins were underrepresented in this group. L-2,4-diaminobutyric acid acetyltransferase along with ectoine hydroxylase which was explained in Section 3.2.2.7 is related with ectoine biosynthetic pathway. As explained there, the ectoine biosynthesis pathway uses L-aspartate-β-semialdehyde as the precursor, which is also used for the production of CC. Hence, as a result of competition between both systems for the same precursor, the ectoine biosynthesis must be the one on the unfavored side.

Nonribosomal peptide synthetases (NRPSs) found in bacteria and fungi are multifunctional enzymes that catalyze the synthesis of highly diverse secondary metabolites (Challis, 2000; Tanovic et al., 2008; Bloudoff et al., 2016). The nonribosomal products include antifungals (bacillomycin), antibacterials (daptomycin), antivirals (luzopeptin), antitumors (actinomycin D), siderophores (enterobactin), and immunosuppressants (cyclosporin) since NRPSs can exploit more than 500 possible amino acid substrates (Felnagle et al., 2008; Bloudoff et al., 2016). They can even introduce modifications on them during peptide synthesis (Bloudoff et al., 2016). Peptides are not only in linear forms but can also be synthesized in cyclic and branched cyclic structures (Du and Lou, 2010). Each NRPS is composed of modules and each module is composed of at least three domains: adenylation (A) domain, thiolation (peptide carrier protein [PCP]) domain and a condensation (C) domain (Tanovic et al., 2008; Du and Lou, 2010). “A” domains are substrate-specific and activate a certain amino acid so that it can be transferred and covalently linked to 4’-phosphopantetheinyl cofactor of the PCP domain, thereby a thioester, aminoacyl-S-PCP is formed. Transportation and elongation of the peptide chains take place in PCP domains (Strieker et al., 2010; Du and Lou, 2010). A dipeptide is produced by the catalysis of amide bond by the C domain; for this, the amino acid residue (in the form of aminoacyl-S-PCP) is donated from an upstream PCP domain to the downstream PCP domain on the module and the condensation reaction is catalyzed through the activity of C-domain and elongation of the peptide chains continues on the PCP domain until it is terminated by the C-terminal thioesterase domain (TE domain)

by the transfer of the peptide chain from terminal PCP domain to TE domain and NRP is released from the enzyme (Kohli et al., 2001; Sieber and Marahiel, 2005; Du and Lou, 2010). Additional domains can be located on NRPS modules so that monomers used as substrates can be modified; for instance, an epimerization domain just between the PCP and C domains can catalyze the epimerization of L-amino acids to D-amino acids before the condensation at the C domain (Du and Lou, 2010; Calcott et al., 2014). In this study, three different NRPSs as well as a hybrid NPRS/PK were found as significantly downregulated. As suggested before, in order to overproduce CC, suppression of other secondary metabolite biosynthetic pathways is quite logical in this strain.

Polyketides are natural products just like nonribosomal peptides (NRPs) with high diversity and activities and produced by polyketide synthases (PKSs) which were explained in Section 3.2.1.5 in great detail. PKSs are very similar to NRPSs in that they are also multifunctional proteins with modular structures that catalyze the formation of the products in a stepwise fashion. The only difference is that while NRPSs use amino acids to elongate the peptide chains, PKSs use carboxylic acids to produce the PKs. The modules function similarly; the precursors for the PKs are activated by acyl transferase domains (AT domains). During the biosynthesis of PKs, they remain covalently attached to acyl-carrier proteins (ACPs), and condensation reactions between monomers are carried out by ketosynthase (KS) domains (Weinig et al., 2003). Finally, termination of the chain elongation is catalyzed by a thioesterase (TE) domain just like in NRPSs (Du and Lou, 2010). Although PKS and NRPS systems can function individually, there are reports on the combination of both systems functioning together, so called PKS-NRPS hybrids that generate products containing amino acids and short carboxylic acids in their backbones (Weinig et al., 2003; Williams, 2013). They can be found in either PKS/NRPS or NRPS/PKS order. For example, BaeJ responsible for the production of bacillaene antibiotic is a typical example of NRPS/PKS system in which an NRPS module is followed by PKS module I and PKS module II (Fisch, 2013). In secondary metabolism of downregulated AK39 proteins, there seems to be two PKS-NRPS hybrids, one in PKS/NRPS and the other in NRPS/PKS order. Since NRPSs and PKs can be intimately related and also function

in a quite similar fashion, the under-expression of either PKSs, NRPSs or PKS/NRPS hybrid systems is not unexpected. However, contrary to these results, a PKS (SCLAV_p0512) was upregulated in AK39 (Section 3.2.1.5). Given that *S. clavuligerus* can have quite diverse PKS-coding genes on its genome, their differential expression possibly under the regulation of different signals is understandable.

Clavaminate synthase (CAS) enzyme catalyzes the conversion of proclavaminic acid to clavaminic acid and thus is involved in the early steps of both CA and clavam biosynthesis (Hamed et al., 2013). Moreover, CAS enzymes has two paralog genes in *S. clavuligerus*: *cas1* and *cas2*. *cas1* is located in clavam gene cluster whereas *cas2* is located in CA gene cluster (Tahlan et al., 2007; Jensen, 2012). *cas1* and *cas2* genes are differentially regulated (Tahlan et al., 2004a). Moreover, it was shown that elimination of *cas1* gene resulted in loss of clavam biosynthesis, but not CA (Tahlan et al., 2007). The protein identified in this category is the one encoded by *cas1* gene, suggesting the decreased production of 5S clavams in AK39 strain in favor of CC overproduction.

Aldo/keto reductase coded by *cvm1* is involved in the biosynthesis of 5S clavams such as clavam-2-carboxylate, 2-formyloxymethylclavam, 2-hydroxymethylclavam, and alanylclavam in *Streptomyces clavuligerus* and is located in the clavam gene cluster (Zelyas et al., 2008); however, the function of it has not been elucidated clearly yet (Song et al., 2010b). *cvm1* mutants of *S. clavuligerus* were shown to lack clavam-2-carboxylate, 2-hydroxymethylclavam, and alanylclavam production (Mosher et al., 1999).

Flavin-dependent oxidoreductase, F_{420}-dependent methylene-tetrahydromethanopterin reductase coded by *cvm5* is also located in clavam gene cluster like *cvm1* and is involved in 5S clavam biosynthesis (Zelyas et al., 2008). Tahlan et al. (2007) showed that *cvm5* mutants in *S. clavuligerus* could not produce any of the known 5S clavams, instead, a new metabolite 2-carboxymethylideneclavam was accumulated (Tahlan et al., 2007; Song et al., 2010b). It is suggested that the protein coded by *cvm5* might be catalyzing the conversion of 2-

carboxymethylideneclavam to 2-formyl-oxymethylclavam through Baeyer–Villiger oxidation (Jensen, 2012).

Putative ribulose-5-phosphate epimerase is coded by *cvm2* which is also located in clavam gene cluster. It encodes a protein with a limited similarity to isomerases and the absence of the gene in *cvm2* mutants severely hinders the production of 5S clavams (Jensen, 2012). Within the clavam gene cluster, the presence of *cvm1*, *cvm2* and *cvm5* genes are absolutely necessary for the production of all 5S clavams (Tahlan et al., 2007).

Putative pyridoxal phosphate-dependent aminotransferase is coded by *cvm6*. *cvm6* is located downstream of *cvm5* in clavam gene cluster and the product of it shows similarity to the aminotransferases (Jensen, 2012). However, their function in clavam biosynthesis could not be revealed yet since the mutants lacking *cvm6* showed no phenotypic properties (Tahlan et al., 2007).

To sum up these findings, the, proteins of secondary metabolism directly related with 5S clavam biosynthesis, polyketide biosynthesis, ectoin synthesis and the formation of NRPS's are severely impaired in AK39.

3.2.2.9. Energy Production/Electron-Iron Transfer

In this category, there were 4 proteins: ATP synthase subunit alpha, ATP synthase subunit beta, NADH-quinone oxidoreductase subunit D and electron transfer oxidoreductase.

ATP synthase α and β subunits of the ATP synthase complex were shown to be downregulated. ATP synthase produces ATP depending on transmembrane proton translocation as powered by proton motive force and explained in Section 3.2.1.6 in detail. ATP synthase γ chain of the ATP synthase complex, on the other hand, was shown to be upregulated in AK39 strain. Interestingly, it was represented by two protein spots showing a mass shift and a localization at an unexpected pI region. Therefore, it can be possible that this subunit might have undergone a posttranslational modification conspicously in AK39 strain, the reason of which is open to further research.

NADH-quinone oxidoreductase is involved in the respiratory chain of the bacteria. Two electrons from NADH is removed and transferred through several redox centers to ubiquinone by this enzyme (Walker, 1992). For two electrons transferred to ubiquinone, 4 protons are translocated across the membrane thereby generating an electrochemical gradient (Yano and Ohnishi, 2001).

According to UniProtKB database, electron transfer oxidoreductase belongs to geranylgeranyl reductase family and is involved in oxidation-reduction reactions.

3.2.2.10. Cell Processes (Shape/Division/Motility)

UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase protein was the sole protein downregulated in this category.

UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase (MurE) protein is involved in peptidoglycan layer biosynthesis in bacteria. The formation of peptidoglycan layer occurs in two steps. First one takes place in the cytoplasm and is mediated by MurA, MurB, MurC, MurD, MurE and MurF enzymes. MurA catalyzes the condensation of phospho(enol)pyruvate (PEP) and UDP-N-acetylglucosamine. MurB reduces the pyruvate moiety to D-lactate ending up with UDP-N-acetylmuramate. Thereafter, MurC, MurD, MurE and MurF catalyze the addition of pentapeptide side chain on the newly reduced D-lactyl group sequentially, which is a process involving the non-ribosomal peptide bond formation followed by the hydrolysis of ATP (El Zoeiby et al., 2001; Ramos et al., 2004). Second stage begins with the transport of the precursors across the membrane and is finalized with the addition of these precursors to the growing cell layer by penicillin-binding proteins (PBPs) (El Zoeiby et al., 2001). Six fold underrepresentation of MurE is thought to account for somewhat slower growth of AK39 that was recorded earlier (Çaydaşı, 2006).

3.2.2.11. Inorganic/Organic Molecule Transport

BldKB belongs to the oligopeptide permease component A (OppA) family and is one of the most commonly found proteins in membrane. Opp complexes are ATP-binding cassette (ABC) transporters and responsible for transporting extracellular signalling

molecules, peptides and nutrients (Li et al., 2014). In *S. coelicolor* A3(2), *bldKA* and *bldKC* genes encode permeases and *bldKD* and *bldKE* code for ATPases. Together with BldKB which is a solute binding protein, they form BldK ABC transporter system (Akanuma et al., 2011). BldK ABC transporter system is responsible for the import of some extracellular signaling molecule, probably an oligopeptide to initiate the signaling cascades for morphological differentiation in *Streptomyces* given that *bldKB* mutants cannot form aerial mycelia (Nodwell et al., 1996; Sugawara et al, 2002). For instance, without the transport of morphogenic oligopeptide Bld261 by BldKB in *Streptomyces coelicolor*, mycelium differentiation could not be completed (Nodwell et al., 1996; Chávez et al., 2011). Secondary metabolism and morphological differentiation processes are nearly inseparable in *Streptomyces*. The signal that initiates one also directly or indirectly induces the other. *bld* genes are pleiotropic effectors in *Streptomyces* affecting both the morphological differentiation and the secondary metabolism. For example, in rich media, *bld261, bldA, bldB, bldC, bldF, bldG, bldH* and *bldI* genes were shown to be important for the aerial mycelium formation (Nodwell et al., 1996). *bldA* gene encodes a rare $tRNA^{leu}$ which mostly has a direct effect on the morphological development as well as secondary metabolism of *Streptomyces* (Chater and Chandra, 2008). According to extracellular complementation studies, presence of *bldH/bldA* genes are enough to initiate the morphological differentiation even in the absence of *bldK* gene that codes for a transporter (Nodwell et al., 1996). Hence, downregulation of this protein might not directly indicate a halt or a deficiency in the differentiation process of AK39 strain.

3.2.2.12. Hypothetical/Unknown Proteins

There were 11 proteins categorized in this group. DUF574 domain-containing protein was identified in both MALDI-MS/MS and LC-MS/MS analyses.

As explained before (Section 3.2.1.9), DUF proteins are unknown proteins. In order to predict functional partners of DUF574, the sequence of the protein has been submitted to STRING database (data not shown) and all the predicted proteins were found to be DNA-binding proteins.

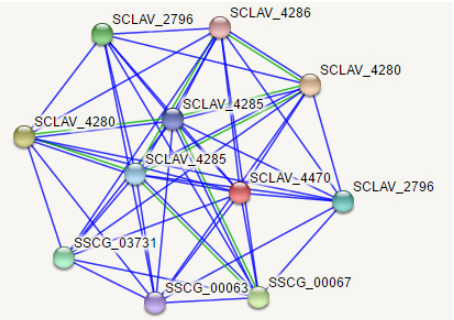
There is a putative uncharacterized protein (SCLAV_0551), the sequence of which was searched through STRING database in order to find out the possible associations of this protein with other proteins (data not shown). According to the prediction, most of the proteins that putative uncharacterized protein is associated with are unknown. The known proteins that were shown to be possibly interacting with these proteins are GCN5-related N-acetyltransferase, a member of GNAT superfamily, which has diverse functions in both the anabolic and catabolic reactions (Dyda et al., 2000), ATP-dependent Clp protease adapter protein, ClpS, which is related with protein degradation (Román-Hernández et al., 2011) and, cytochrome c oxidase subunit XV assembly proteinS. Cytochrome c oxidase is critical for the cellular respiration (Capaldi, 1990).

M28 family proteases are a group of metallopeptidases. They are unique in both having aminopeptidase activity and carboxypeptidase activity at the same time and also in requiring two metal ions in their active site for their catalytic activity (Nagase, 2001). M28 family proteins contain two zinc ions in their active site (Fundoiano-Hershcovitz et al., 2004). All the metallopeptidases that have been identified so far and that have possessed two metal ions in their active sites are exopeptidases (Nagase, 2001) Moreover, proteins in M28 family show diverse structural differences (Rawlings and Barret, 1997). For example, *Streptomyces griseus* aminopeptidase (SGAP) belongs to M28 family proteins (Arima et al., 2006) Aminopeptidases remove amino acid residues from N-termini of the peptides (Arima et al., 2006; Yoo et al., 2010). Arima et al. (2006) found that the activity of bacterial aminopeptidases were affected by penultimate residue and length of the peptide (Arima et al., 2006). The function of putative M28 family protease found in this study and the importance of its secretion into the environment needs to be further characterized.

The function of the secreted protein (SCLAV_0814) is unknown. When searched through EggNog database, there were orthologues of this protein in 33 species of *Actinobacteria* phylum; however, the function of them have not yet been revealed. When searched in all bacterial species, it showed the highest similarity to YngK protein which is involved in cell wall/membrane/envelope biogenesis processes. In

STRING database, on the other hand, it was shown to be associated with 3-hydroxybutyryl-CoA dehydrogenase, phage protein, and an unknown protein (data not shown).

Putative secreted lipase protein was predicted to have alpha/beta hydrolase fold according to UniProtKB database. Proteins with alpha/beta hydrolase folds are found in diverse hydrolytic enzymes which can have distinct functions (Ollis et al., 1992). STRING data (Figure 3. 9) suggested that this protein is associated or interacts with GCN5-related N-acetyltransferase (SCLAV_4285) and radical SAM domain protein (SCLAV_4286), both of which were categorized as downregulated proteins in AK39 strain in Section 3.2.2.13 in the following section (Others/General Function). These proteins might represent a regulon or signaling pathway that was suppressed in AK39 strain.



Your Input:

		Neighborhood	Gene Fusion	Cooccurence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
SCLAV_4470	*Putative secreted lipase (303 aa)*									

Predicted Functional Partners:

SCLAV_4280	*Type I phosphodiesterase/nucleotide pyrophosphatase (518 aa)*			•						0.774
SCLAV_4280	*no annotation available (518 aa)*			•						0.774
SSCG_00067	*Putative uncharacterized protein (606 aa)*			•						0.773
SCLAV_2796	*TlmR2 (401 aa)*			•						0.768
SSCG_03731	*no annotation available (401 aa)*			•						0.768
SCLAV_2796	*no annotation available (401 aa)*			•						0.768
SCLAV_4285	*GCN5-related N-acetyltransferase (322 aa)*			•						0.739
SCLAV_4285	*no annotation available (322 aa)*			•						0.739
SSCG_00063	*Putative uncharacterized protein (359 aa)*			•						0.738
SCLAV_4286	*Radical SAM domain protein (338 aa)*			•						0.727

Figure 3. 9. STRING data showing the associaton of putative secreted lipase with other proteins.

Ornithine aminotransferases are involved in the biosynthesis of L-glutamate (Yasuda et al., 1979), which is a versatile amino acid that can be converted to several other amino acids like proline or can be used as intermediate for diverse processes in the cell.

3.2.2.13. Others/General Function

There were identified 14 proteins in this category.

Amidase proteins generally catalyze the hydrolysis of the amide bonds; however, they are composed of a large group of enzymes, the substrate specificities of which show considerable variety (Valiña et al., 2004).

Acetyltransferase (SCLAV_1797) protein was predicted to have GNAT domain by UniProtKB database. GCN5-related N-acetyltransferase (GNAT) superfamily is comprised of a large variety of members and show different substrate preferences as well as diverse functions (Dyda et al., 2000).

Flavoprotein disulfide reductase family consists of members showing functional and structural homology. They usually form homodimers and contain nonflavin redox center and noncovalently but tightly bound flavin adenine dinucleotide (FAD) molecule within their structure.They catalyze the transfer of electrons from reduced pyridine nucleotide to their substrates using FAD and redox-active disulfide located on each monomer. Lipoamide dehydrogenase (LipDH), glutathione reductase (GR), trypanothione reductase, mycothione reductase, thioredoxin reductase are several examples belonging to flavoprotein disulfide reductase family (Argyrou et al., 2004; Argyrou and Blanchard, 2004).

LmbE protein was first thought to take part in lincomycin biosynthesis in *Streptomyces lincolnensis* (Tanaka et al., 2004). However, lmbE-like enzymes have now been shown to form a superfamily that hydrolyzes substrates with N-acetylglucosamine or N-acylglucosamine core. Although all their substrates contain either of these two core structures, they show high functional diversity due to different moieties coupled with these core structures. For example, N-acetylglucosamine core can be ornamented with various carbohydrate, lipid, and amino acid moieties, which contribute to the functional differences that LmbE family proteins display. Functions can range from the biosynthesis of protective reducing agents, antibiotics to cell membrane components. MshB, BshB, GlcNAc-PI de-N-acetylases, Orf2 and BtrD are all LmbE family proteins. MshB and BshB are responsible for the production of primary reducing agents which are mycothiols in *Mycobacterial* sp. and bacillithiol in *Bacillus* sp., respectively (Viars et al., 2014), GlcNAc-PI de-N-acetylases are involved in glycosylphosphatidylinositol biosynthesis in eukaryotes, Orf2 and BtrD are related with synthesis of lipoglycopeptide antibiotics (Deli et al., 2010). LmbE family protein identified in this study has not been specified, thus, no speculations about its downregulation in AK39 could be made

Phenylacetate-coenzyme A ligase catalyzes the conversion of phenylacetate into phenylacetate-CoA which is an important precursor for the biosynthesis of penicillin G in *Penicillium chrysogenum* (Koetsier et al., 2009). However, penicillin G cannot

be produced by *S. clavuligerus*; hence, the enzyme should be involved in another pathway that is to be elucidated.

Cobalamin is a tetrapyrrole that is used in the biosynthesis of coenzymic form known as coenzyme B12 or adenosylcobalamin (AdoCbl). CobT enzyme is one of the proteins involved in the biosynthesis of adenosylcobalamin (Claas et al., 2010). Adenosylcobalamin acts as coenzyme in more than 10 enzymatic reactions (Ishida et al., 1993). It is intriguing that cobalamin is synthesized by only some bacteria but have no observed importance for fungi, plants or some other bacteria (Roth et al., 1996).

Short-chain dehydrogenase/reductase SDR proteins are NAD(P)(H)-dependent oxidoreductases are involved in a large variety of biological processes (Kavanagh et al., 2008). GAF domains confer proteins diverse functional properties such as binding to small molecules or protein-protein interactions, etc. Nonetheless, most of the GAF domains await for the functional characterization (Heikaus et al., 2009).

Many NAD-dependent epimerase/dehydratases found in *Streptomyces* are usually located in secondary metabolite gene clusters. For example, validamycin A antibiotic produced by *Streptomyces hygroscopicus* var. *limoneus* is used as an agent against *Rhizoctonia solani* that causes plant diseases. Validamycin antibiotic gene cluster consists of *vldABC* and divergently located *vldDEFGH*. *vldD* gene codes for VldD which was predicted to belong to a NAD-dependent epimerase/dehydratese family and it showed approximately 25-30 % identities to dTDP-4-keto-L-rhamnose reductase of *S. griseus*, dTDP-glucose 4,6-dehydratase of *S. avermitilis* or UDP-glucose 4-epimerase from *S. coelicolor* (Singh et al., 2006). Moreover, SoxR regulon found in *S. coelicolor* contained several genes, the expression of which are developmentally regulated. Among them is a predicted NAD-dependent epimerase/dehydratase (SCO1178) protein. SoxR regulon in *Streptomyces* was suggested to be related with the antibiotic production (Dela Cruz et al., 2010). As another example, fostriecin produced by *Streptomyces pulveraceus* is a phosphate monoester antibiotic that acts as a protein phosphatase. Within the fostriecin gene cluster resides *fosM* gene that codes for a protein belonging to NAD-dependent epimerase/dehydratase family (Kong et al., 2013). Putative epimerase/dehydratase protein found in this study showed 100

% sequence coverage and 91 % identity with NAD dependent epimerase/dehydratase family protein from *Streptomyces ipomoeae* and 100 % sequence coverage and 91 % identity with NAD-dependent dehydratase from *Streptomyces tsukubensis*, NAD-dependent dehydratase from *Streptomyces* sp. NRRL WC-3773, epimerase from *Streptomyces viridochromogenes*. This protein needs to be further characterized as for its function and relation with the the secondary metabolism of *S. clavuligerus*.

Radical SAM proteins belong to a new superfamily of proteins that consist of unusual Fe-S centers and catalyze the production of free radicals through SAM cleavage. They are involved in a large variety of reactions from unusual methylation, isomerization, and sulfur insertion to ring formation and take place in different cellular processes such as vitamin, cofactor or antibiotic biosynthesis (Sofia et al., 2001).

GCN5-related N-acetyltransferase (GNAT) superfamily is a very large family housing enzymes with very different substrates preference and hence diverse functions (Dyda et al., 2000). However, the main function of all the enzymes in this superfamily is to transfer the acetyl group of acetyl-CoA to an acceptor (primary amine) (Dyda et al., 2000). They can be involved in the acetylation such substrates as antibiotics, hormones, tRNA, histones, metabolic enzymes and transcription factors (Tucker and Escalante-Semerena, 2013). The first identified members of this family were aminoglycoside acetyltransferases that modify aminoglycosides so that bacteria gain resistance against these antibiotics (Davies and Wright, 1997; Vetting et al., 2005). They now constitute one of the largest superfamilies (Tucker and Escalante-Semerena, 2013).

3.2.3. Summary for AK39 Upregulated and Downregulated Proteins.

Proteomic approaches that we used for the comparative proteomic analysis of mutants compared to parental strain showed that MALDI-TOF-MS/MS and LC-MS/MS results were complementary to each other. A total of 40 proteins were upregulated whereas 50 proteins were shown to downregulated by both techniques. 2DE coupled to MALDI-TOF/MS results indicated that proteins related directly with CC production were upregulated in AK39 strain while LC-MS/MS results also indicated the presence

of several other proteins involved in the biosynthesis of unrelated secondary metabolites. According to LC-MS/MS results, a considerable number of secondary metabolites related with PKS, NRPS, clavam biosynthesis were downregulated, probably for favoring CC production. The distribution of the total number of proteins which were identified and shown to be upregulated and downregulated to functional categories are shown in Figure 3. 10 and Figure 3. 11, respectively. Among the upregulated protein categories, "Hypothetical/Unknown Proteins" category harbored most of the proteins and was followed by "Secondary Metabolism". Third most populated category was the "DNA Replication, Recombination, Repair, Transcription" category.

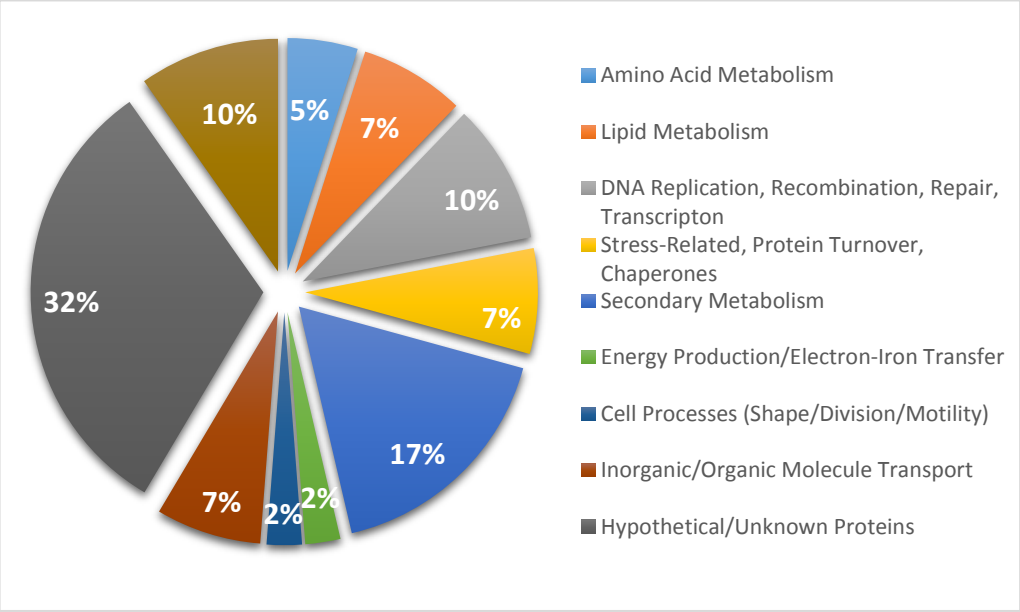


Figure 3. 10. Relative quantitative distribution of upregulated proteins identified in AK39 to the functional categories. Percentage of each category was calculated by dividing the number of proteins in each category to the total number of proteins in all categories.

Upregulation of TCS sensor kinase, DNA binding protein HU and RecA proteins included in the third category might be important in that they can affect the

transcription of genes related with CC production. For example, SCLAV_1334 identified as TCS sensor kinase might be a part of a positive regulation system for the production of CC. RecA protein, apart from being involved in transcription, was also suggested to be responsible for the preservation of the genetic stability in the mutant strain. Although proteins identified in all categories except for the ones in"Hypothetical/Unknown Proteins" could be explained in terms of their probable roles in the secondary as well as primary metabolism of the mutant, "Hypothetical/Unknown Proteins" might require much of the attention. For example, SCLAV_0037 protein was proposed to be involved in T6SS, the presence of which has not been shown in *S. clavuligerus* before this study.

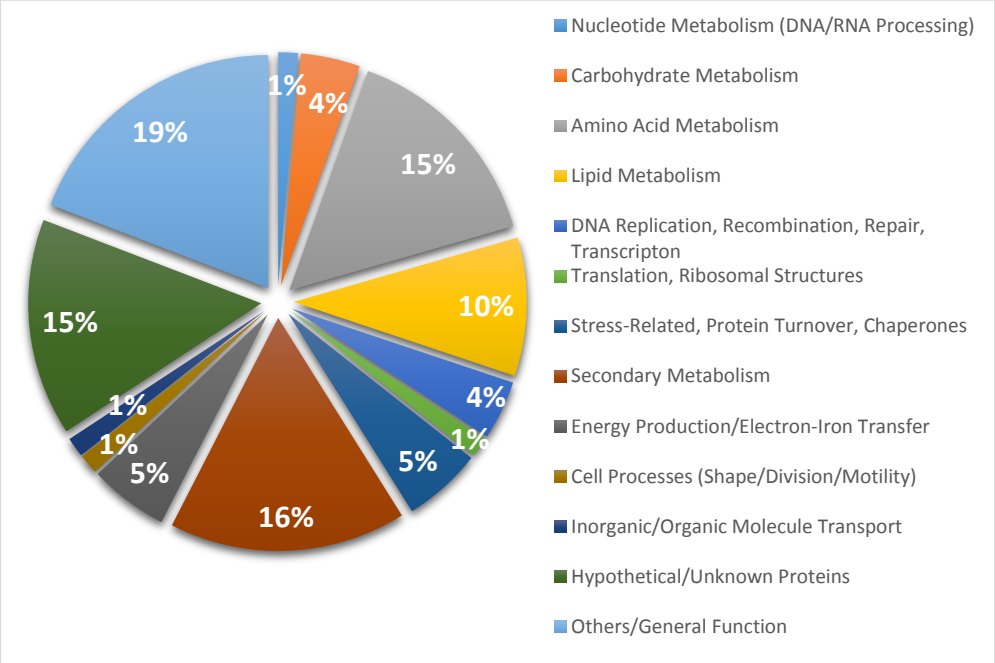


Figure 3. 11. Relative quantitative distribution of downregulated proteins identified in AK39 to the functional categories. Percentage of each category was calculated by dividing the number of proteins in each category to the total number of proteins in all categories.

Aside from upregulated proteins, downregulated proteins in an organism also gives important clues about its physiological state. Among the total downregulated protein categories, “Others/General Proteins” showed the highest ratio compared to the others Second most populated category was “Secondary Metabolism”, which was plausible, since biosynthesis of many other secondary metabolites must be suppressed in order to overproduce CC. Moreover, proteins residing in the rest of the categories mostly represented proteins that take part in the central metabolism of the organism like cell growth, suggesting that such metabolic pathways are minimized in favor of CC production.

3.3. Comparative Proteome Analysis of NRRL 3585 and TB3585 Strains by 2DE MALDI-TOF/MS and LC-MS/MS

A fused image of total proteomes of NRRL 3585 and TB3585 is shown in Figure 3. 12. Of the proteins that were identified as differentially expressed in TB3585 strain, 23 proteins were significantly upregulated and 28 proteins were significantly downregulated (Table 3. 5 and Table 3. 7, respectively) in MALDI-TOF/MS analysis. On the other hand, 23 proteins were shown to be upregulated and 22 proteins were downregulated in TB3585 strain through LC-MS/MS analysis (Table 3. 6 and Table 3. 8, respectively).

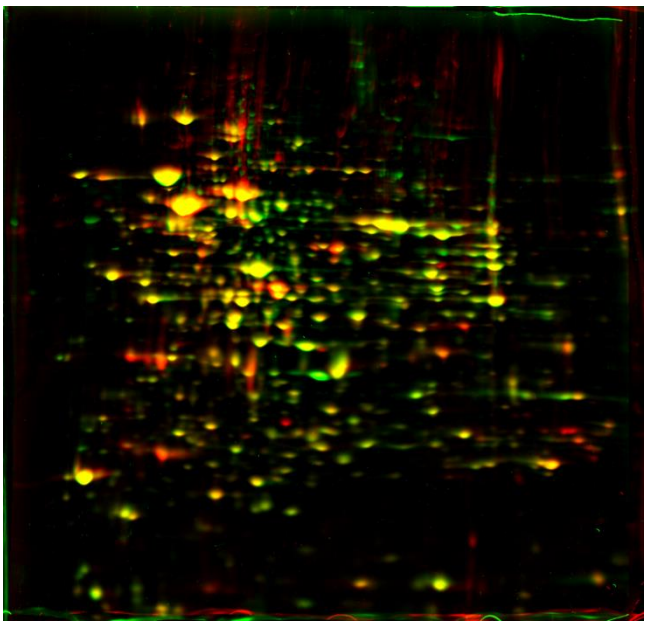


Figure 3. 12. Dual channel 2DE imaging of *S. clavuligerus* NRRL 3585 strain (green) and TB3585 strain (red) when the strains are grown for 48 h in TSB medium, pI 4-7. Overlapping spots are represented in yellow.

3.3.1. Upregulated Proteins in TB3585

Figure 3. 13 shows protein spots on TB3585 gel that showed more than 2.5-fold increase in expression when compared to their counterparts on NRRL 3585 gel. Upregulated proteins of TB3585 strain were tabulated in Table 3. 5 and Table 3. 6. 23 proteins were identified as upregulated in both MALDI-TOF/MS and LC-MS/MS results.

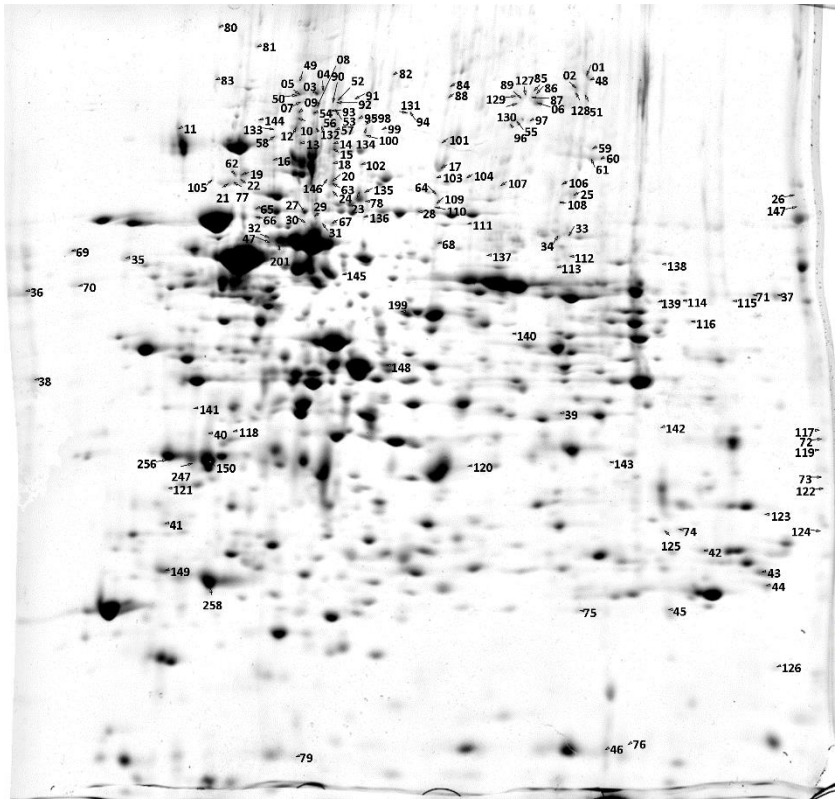


Figure 3. 13. Protein spots on TB3585 gel that showed more than 2.5 fold increase compared to the parental strain NRRL 3585.

Table 3. 5. List of upregulated proteins detected on TB3585 2DE gel.

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spot	PTM *	Genome Location **	Spot #
Carbohydrate Metabolism	E2PV06	Aconitate hydratase	*acnA*	Cytoplasm	21	4.84	98,214	–	–	Chr.	22
Amino Acid Metabolism	B5H1S3	Dihydrodipicolinate synthase	*dapA*	Cytoplasm	5.5	5.98	31,484	–	–	Chr.	74
	B5GLC1	L-lysine-epsilon aminotransferase	*lat*	Cytoplasm	3.0	5.36	49,889	–	–	Chr.	199
Lipid Metabolism	B5H1P3	Acetyl/propionyl CoA carboxylase alpha subunit	*SCLAV_1944*	Cytoplasm	4.1	5.44	67,249	–	–	Chr.	111
DNA Replication, Recombination, Repair, Transcripton	E2Q279	Two-component system response regulator	*SCLAV_3595*	Cytoplasm	3.0	5.73	24,298	–	–	Chr.	75
Stress-Related, Protein Turnover, Chaperones	B5GTB1	Chaperone protein dnaK	*dnaK*	Cytoplasm	3.0	4.79	66,533	–	–	Chr.	65
	E2PWA4	Proteasome subunit alpha	*prcA (SCLAV_0860)*	Cytoplasm	2.9	4.88	27,591	–	–	Chr.	247
	E2Q3Z8	Mycothiol S-conjugate amidase	*mca*	Cytoplasm	3.1	4.63	32,257	–	–	Chr.	256
Secondary Metabolism	B5GLB5	Deacetoxycephalosporin C hydroxylase	*cefF*	Cytoplasm	5.2 (a); 2.9 (b)	4.79	34,599	2 (a, b)	C-M	Chr.	40, 247
	B5GLB7	Positive regulator (CcaR)	*ccaR*	Cytoplasm	4.3 (a); 5.5 (b)	6.02	29,346	2 (a, b)	C-M	Chr.	42, 74
	E2Q5Q5	Clavaminate synthase 2	*cas2*	Cytoplasm	2.5	5.18	35,796	–	–	Chr.	148
	E2Q5Q8	Carboxyethylarginine synthase	*ceaS2*	Cytoplasm	2.9	5.01	62,348	–	–	Chr.	201

Table 3. 5. List of upregulated proteins detected on TB3585 2DE gel. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spot	PTM *	Genome Location **	Spot #
Secondary Metabolism	B5GLB3	Cephalosporin hydroxylase CmcI	*cmcI*	Cytoplasm	3.2	4.80	27,584	–	–	Chr.	258
	E2Q5Q3	ABC-type dipeptide transport system, periplasmic component	*oppA1*	Cytoplasm (Periplasmic component)	2.9	5.00	62,301	–	–	Chr.	201
	B5GLC1	L-lysine-epsilon aminotransferase	*lat*	Cytoplasm	3.0	5.36	49,889	–	–	Chr.	199
Cell Processes (Shape/Division/Motility)	B5GPR6	Bifunctional protein GlmU	*glmU*	Cytoplasm	3.3	5.94	49,548	–	–	Chr.	116
	B5H1S3	Dihydrodipicolinate synthase	*dapA*	Cytoplasm	5.5	5.98	31,484	–	–	Chr.	74
Inorganic/Organic Molecule Transport	E2Q5Q3	ABC-type dipeptide transport system, periplasmic component	*oppA1*	Cytoplasm (Periplasmic component)	2.9	5.00	62,301	–	–	Chr.	201
Hypothetical/Unknown Proteins	E2PXW4	Putative MerR-family transcriptional regulator	*SCLAV_3133*	Cytoplasm	4.3	6.88	31,757	–	–	Chr.	42
	D5SLL2	Putative transcriptional regulator AraC family	*SCLAV_p1319*	Unknown	13.6	8.55	25,424	–	–	pSCL4	45
	B5H453	Predicted acyltransferase	*SCLAV_p1317*	Unknown	10.9	6.80	31,731	–	–	pSCL4	124
Others/General Function	E2PU37	Oxidoreductase	*SCLAV_0580*	Cytoplasm	21	6.50	33,141	–	–	Chr	72
	B5GQF0	ATP-binding protein	*SCLAV_4119*	Cytoplasmic membrane	3.7	6.19	28,072	–	–	Chr.	43
	E2Q9Q1	Acetyl-coenzyme A synthetase	*acsA*	Cytoplasm	73.6	5.09	72,963	–	–	Chr.	135

Table 3. 5. List of upregulated proteins detected on TB3585 2DE gel. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spot	PTM *	Genome Location **	Spot #
Others/General Function	E2Q6Q8	LmbE family protein	*SCLAV_4283*	Cytoplasm	6.4	6.30	25,212	–	–	Chr.	44
	B5GP52	Sulfurtransferase	*SCLAV_3193*	Cytoplasm	3.1	4.68	31,672	–	–	Chr.	256

* PTM refers to possible postranslational modifications. **"Chr." refers to chromosome and "pSCL4" refers to plasmid pSCL4.

Table 3. 6. List of upregulated proteins in TB3585 strain identified by LC-MS/MS.

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
DNA Replication, Recombination, Repair, Transcripton	E2Q1X6	MerR-family transcriptional regulator	*SCLAV_1527*	Cytoplasm	3.486	4.76	32,554	Chr.
Stress-Related, Protein Turnover, Chaperones	E2Q5J7	Lon protease	*lonA*	Cytoplasm	2.246	5.14	86,139	Chr.
Secondary Metabolism	B5GLB7	Positive regulator	*ccaR*	Cytoplasm	2.556	6.02	29,346	Chr.
	E2PYR6	Alkaline d-peptidase	*adp2*	Extracellular	2.125	9.63	46,079	Chr.
	E2PY60	NocE-like protein	*nocE*	Unknown	3.712	8.80	147,287	Chr.
Cell Processes (Shape/Division/Motility)	B5GMU0	Plasmid partitioning protein parB	*parB*	Cytoplasm	2.125	4.94	51,212	pSCL4
Inorganic/Organic Molecule Transport	B5GUY9	Cation/multidrug efflux protein	*SCLAV_1380*	Cytoplasmic Membrane	2.125	5.13	108,947	Chr.
	E2Q6Q4	ABC transporter related protein	*SCLAV_4279*	Cytoplasmic Membrane	2.556	6.77	33,757	Chr.
Hypothetical/Unknown Proteins	E2PXW4	Putative MerR-family transcriptional regulator	*SCLAV_3133*	Cytoplasm	5.199	6.88	31,757	Chr.
	E2Q633	Putative uncharacterized protein	*SCLAV_0117*	Unknown	2.125	5.08	29,941	Chr.
	B5GP37	Putative uncharacterized protein	*SCLAV_3208*	Unknown	2.125	9.04	21,165	Chr.

Table 3. 6. List of upregulated proteins in TB3585 strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Hypothetical/Unknown Proteins	B5H122	Uncharacterized protein	*SCLAV_p1179*	Cytoplasmic Membrane	2.125	5.25	19,646	pSCL4
	E2PUX9	Putative plasmid partitioning protein, Parb2	*SCLAV_4803*	Cytoplasm	2.357	10.70	35,473	Chr.
	E2Q564	Putative nicotinamidase	*SCLAV_2032*	Cytoplasm	2.460	5.11	22,736	Chr.
	B5GPX3	Putative transmembrane transport protein	*SCLAV_2301*	Cytoplasmic Membrane	2.647	5.42	77,276	Chr.
	B5GU15	Putative AfsR-like transcriptional regulator	*SCLAV_0412*	Cytoplasm	2.812	6.77	83,477	Chr.
	E2Q1T4	Putative ABC transporter ATP-binding protein	*SCLAV_1483*	Cytoplasmic Membrane	2.960	5.35	35,99	Chr.
	B5GVX3	GPP34 domain-containing protein	*SCLAV_p0135*	Cytoplasm; Cell membrane	2.246	9.48	24,509	pSCL4
Others/General Function	B5H3A5	Methyltransferase type 12	*SCLAV_2606*	Cytoplasm	2.125	4.98	31,423	Chr.
	B5GNM7	Peptidase S1 and S6	*SCLAV_3506*	Extracellular	2.246	5.76	43,479	Chr.
	E2Q172	Peptidase S1 and S6, chymotrypsin/Hap	*SCLAV_3505*	Extracellular	2.357	5.64	45,579	Chr.
	E2Q1T1	NH3-dependent NAD+ synthetase	*nadE*	Cytoplasm	2.795	4.98	63,031	Chr.
	B5GSN1	Nicotinate phosphoribosyltransferase	*pncB*	Cytoplasm	3.388	5.76	46,605	Chr.

* "Chr." refers to chromosome, "pSCL4" refers to plasmid pSCL4.

3.3.1.1. Carbohydrate Metabolism

There was found to be one protein in this category: Aconitate hydratase, which is also known as aconitase, is involved in TCA cycle and catalyzes the reversible isomerization of citrate to isocitrate (Viollier et al., 2001). Being an intermediate in the TCA cycle, citrate is converted to oxaloacetate which can be exploited either as an intermediate in the TCA cycle or as a precursor for the production of pyruvate and L-aspartate. L-aspartate initiates the aspartate pathway that leads to the production of CC production. Hence, the upregulation of aconitase supplies the cells with higher amounts of metabolic intermediates to overproduce CC.

3.3.1.2. Amino Acid Metabolism

There were two proteins shown to be upregulated in this category: Dihydrodipicolinate synthase and L-lysine-ε aminotransferase.

Lysine is an important intermediate in aspartic acid pathway leading to the CC biosynthesis (Figure 1. 9). L-lysine-ε aminotransferase (LAT) catalyzes the conversion of lysine to α-aminoadipic acid semialdehyde which is spontaneously cyclized to piperideine-6-carboxylic acid (P6C) (Coque et al. 1991; Liras, 1999; Martín, 1998). Piperideine-6-carboxylate dehydrogenase (P6C-DH) converts P6C to α-aminoadipic acid which is one of the three precursors necessary for cephamycin C production (Martín et al., 2010; Özcengiz and Demain, 2013).

Dihydrodipicolinate synthase is also involved in aspartate pathway and responsible for the conversion of L-aspartate 4-semialdehyde to 4-hydroxy-2,3,4,5-tetrahydrodipicolinate, after which diaminopimelic acid (DAP) is produced through sequential reactions. Lysine production is dependent on diaminopimelic acid (DAP) intermediate produced in the pathway, which is also used as a cell wall precursor (Mendelovitz and Aharonowitz, 1982; Azevedo et al., 2006).
Upregulation of both proteins were expected since the increase in the levels of these proteins indicate higher levels of metabolites leading to the overproduction of CC.

3.3.1.3. Lipid Metabolism

Acetyl/propionyl CoA carboxylase alpha subunit was the sole protein in this category. As explained for AK39 (Section 3.2.1.2), acetyl-CoA carboxylase (ACC) proteins are not only important in providing precursors for fatty acid biosynthesis but also for the polyketide biosynthesis (Rodríguez and Gramajo, 1999).

3.3.1.4. DNA Replication, Recombination, Repair, Transcripton

Two-component system response regulator and MerR-family transcriptional regulator proteins were identified in this category.

Two component systems (TCSs), as explained in detail in Chapter 1, control both the secondary metabolism and differentiation in *Streptomyces* and are triggered by changing environmental conditions such as nutritional changes or other stress factors (Martín and Liras, 2010; Rodríguez et al, 2013). TCSs are mainly composed of two components: Histidine kinases responsible for detecting environmental stimuli and a response regulator that will activate the cellular response by regulating the transcription of target genes (Rodríguez et al., 2013). The contribution of this protein to secondary metabolism and/or the morphological differentiation of *S. clavuligerus* needs further research.

MerR family members usually act as transcriptional activators and are responsive to several effectors such as heavy metals, oxygen radicals, or cytotoxic compounds (Hayashi et al., 2013). For instance, a MerR family protein in *E. coli* was reported to function as an activator in the presence of mercury but as a repressor in the absence of it. Although MerR family proteins are widely distributed in prokaryotes, most of them were reported in Gram-negative bacteria and only a few of them could be described in *Streptomyces*, one of them being TipA protein. TipA protein is induced by thiostrepton activating certain genes (Romero-Rodríguez et al., 2015). The protein identified in this study is only known to be one of the MerR family proteins and increased levels of this protein indicate the presence of certain effector molecules in the environment. Thiostrepton antibiotic that is added into the culture medium of TB3585 strain might

be one of the inducers of the protein. Nevertheless, it needs to be further characterized to find out the effectors and the function of the protein within the cell.

3.3.1.5. Stress-Related, Protein Turnover, Chaperones

There were four proteins categorized under this group: Chaperone protein dnaK, proteasome subunit alpha, mycothiol S-conjugate amidase and Lon protease.

DnaK protein is a heat shock protein (Hsp) and helps newly formed peptides fold in the right configuration (Bucca et al., 1997). Just like Hsp60/GroE chaperones, Hsp70/DnaK chaperones are produced under both normal and stressful environmental conditions (Bucca et al., 2000).

Proteasome subunit alpha (PrcA) is also shown to be upregulated in AK39 strain (Section 3.2.1.4). It is a part of prokaryotic 20S proteasome (Nagy et al, 1998), which is found in archeabacteria and actinobacteria (De Mot et al., 2007) and was shown to protect *M. tuberculosis* against oxidative and nitrosative stress (Darwin et al., 2003).

Mycothiol S-conjugate amidase (Mca) is an important enzyme in the mycothiol (MSH)-dependent detoxification mechanism in cells (Steffek et al., 2003). Mycothiol (MSH) is a low molecular weight thiol composed of large numbers of cysteine derivatives, 2-(N-acetylcysteinyl)amido-2-deoxy-α-D-glucopyranosyl-(1→1)-myo-inositol (Newton et al., 1996). These molecules were shown to be produced only in *Actinomycetes*, especially in mycobacteria and act as reducing agent and a storage form of cysteines (Ung and Av-Gay, 2006; Nilewar and Kathiravan, 2014). Mycothiol is maintained at reduced state by mycothiol disulfide reductase and is responsible for protecting cell against oxidative stress, electrophilic toxins, etc. It is shown to play roles in detoxification of formaldehyde and nitric oxide, neutralizing the alkylating agents and metabolism of certain antibiotics (Steffek et al., 2003). The antioxidant property of MSH comes from the sulfur atom on the cysteine residues of the molecule and the reduced –SH group is what makes MSH functional. By forming MSH S-conjugates with the xenobiotics, toxins are detoxified. That is when mycothiol S-conjugate amidase (Mca) comes into play. After the formation of MSH S-conjugates, Mca cleaves these into GlcN-Ins which is recycled to MSH and modified toxin

AcCysR (R representing the toxin) which is then excreted out of the cell. (Ung and Av-Gay, 2006).

Lon protease is an ATP-dependent serine protease and found ubiquitously in both prokaryotes and eukaryotes; however, their substrates as well as biological roles differ depending on the organism (de Crécy-Lagard et al., 1999). Lon proteases contain the ATPase and protease domains on the same polypeptide and are mainly composed of three domains: N-terminal domain (LON), a central ATPase domain (AAA+ module) and a C-terminal proteolytic domain (Lon_C). They are divided into two categories as LonA and LonB proteases which are named according to the conserved amino acid residues around the catalytic serine residue. They show a strict sequence specific degradation pattern and not only take part in the degradation of the misfolded proteins but can also regulate cell differentiation, sporulation, pathogenicity and stress response in bacteria (Tripathi and Sowdhamini, 2008).

Altered metabolic activitis of TB3585 strain might have caused an increase in reactive oxygen species or toxic oxidants that lead to the oxidative stress in the cell, possibly accounting for increased expression of DnaK and PrcA, Lon protease as well as mycothiol S-conjugate amidase.

3.3.1.6. Secondary Metabolism

There were 9 proteins upregulated in this category. Of these proteins, positive regulator CcaR was shown to be upregulated in both MALDI-TOF/MS and LC-MS/MS results.

Lysine-ϵ-aminotransferase (LAT) and its importance for CC synthesis are already explained in Section 3.3.1.2. It forms the branch point leading to the cephamycin C biosynthesis in aspartate pathway.

Deacetoxycephalosporin C hydroxylase also known as deacetylcephalosporin C synthase (DACS) encoded by *cefF* was also shown to be 2.7 fold upregulated in AK39 strain and is involved in CC biosynthetic pathway (Liras and Martín, 2006). In TB3585, this protein was represented by two spots (with 5.2 and 2.9 fold increases)

which showed a slight shift within both the pI and mass range, suggesting a PTM with both charge and mass modification on the protein, which was not the case for AK39.

Cephalosporin hydroxylase (CmcI) is another enzyme in CC biosynthetic pathway and carries out the C-7 methoxylation of deacetylcephalosporin C intermediate with CmcJ protein to produce CC (Figure 1. 10) (Enguita et al., 1996; Liras and Martín, 2006).

Being also 8.7 fold upregulated in AK39, CcaR is a SARP regulating both the CC and CA biosynthetic gene. It can bind the promoters of several genes in CC gene cluster such as *lat, cefF, cefD–cmcI* to activate transcription (Santamarta et al., 2011). Hence, increase in the expression of LAT, deacetoxycephalosporin C hydroxylase coded by *cefF* and CmcI are not surprising. Moreover, there were two spots (with a 4.3 and 5.5 fold increases) representing CcaR on the gel, the locations of which were changed with respect to both mass and pI, suggesting that CcaR might have undergone PTM in TB3585.

An increase of 2.5 fold was detected in clavaminate synthase (CAS2) that catalyzes the conversion of proclavaminic acid to clavaminic acid which is the common precursor for both 5S clavams and CA production. There are two paralog genes of clavaminate synthase on *S. clavuligerus* genome: *cas1* located in clavam gene cluster and *cas2* located in CA gene cluster (Tahlan et al., 2007; Jensen, 2012). Although they catalyze the same reaction, they are regulated differently; for example, while *cas2* is expressed in both complex soy medium and SA medium, transcription of *cas1* was observed in SA medium but not in soy medium. (Tahlan et al., 2004b). Moreover, it was shown that *cas1* and *cas2* genes were downregulated in *ccaR* mutants of *S. clavuligerus* (Álvarez-Álvarez et al., 2014). Hence, expression of CAS2 encoded by *cas2* is controlled by CcaR and the increased levels of CAS should be associated with the upregulation of CcaR in TB3585.

Carboxyethylarginine synthase (CeaS2) catalyzing the first reaction in the CA and clavam biosynthesis (Jensen, 2012) was also shown to be upregulated in AK39 strain (Section 3.2.1.5).

oppA1 is located in clavulanic acid gene cluster (Álvarez-Álvarez et al., 2013) and its product (ABC-type dipeptide transport system [OppA1]) is involved in the transport of oligopeptide signaling molecules required for CA production (Lorenzana et al., 2004). OppA1 also showed approximately 36% similarity to BldKB protein (Mackenzie et al., 2010) which is responsible for the transport of a certain oligopeptide to initiate the aerial mycelium formation in *Streptomyces* (Nodwell et al., 1996).

An alkaline d-peptidase protein which was also shown to be upregulated in AK39 (Section 3.2.1.5) consistently overrepresented in TB3585. The same applies to Noc-E like protein (Section 3.2.1.5).

Proteins under the category of "Secondary Metabolism" in TB3585 are mostly related with CC and CA and/or clavam biosynthesis. *cas2* and *ceaS2* are known as the early biosynthetic genes for both the CA and clavam biosynthesis. Hence, we have not yet determined if CC overproduction is accompanied with CA or clavam overproduction. However, as will be explained in the following section, several proteins directly related with clavam biosynthesis were downregulated in TB3585, suggesting CA overproduction rather than that of clavam is probable. Apart from the positive effect of multi-copy aspartokinase (*ask*) gene expression, *ccaR* overexpression in a yet unknown mechanism seems to account for CC overproduction by TB3585 as was the case for AK39 which was constructed by deleting *hom* gene. On the other hand, upregulation of acetyl/propionyl CoA carboxylase (Sections 3.3.1.3 and 3.2.1.2) might be an indicator of the biosynthesis of polyketides. Moreover, two-component system response regulator (SCLAV_3595) which was 3 fold overrepresented in TB385 (Section 3.3.1.4) might also contribute to CC overproduction.

3.3.1.7. Cell Processes (Shape/Division/Motility)

There were three proteins in this category: Dihydrodipicolinate synthase, bifunctional protein GlmU and plasmid partitioning protein ParB.

UDP-GlcNAc is the cytoplasmic precursor of bacterial peptidoglycan cell wall. A 4-step process is required for the synthesis of UDP-GlcNAc. The first reaction involving the conversion of fructose-6P to GlcN-6P is catalyzed by glucosamine-6P synthase

(GlmS). The second step is the conversion of GlcN-6P to GlcN-1P, which is carried out by phosphoglucosamine mutase (GlmM). Final reaction is catalyzed by the bifunctional enzyme GlmU (N-acetylglucosamine-1-phosphate uridyltransferase) with acetyltransferase and uridylyltransferase activities. GlcN-1P is first converted to GlcNAc-1P and then to UDP-GlcNAc by GlmU (Mengin-Lecreulx and van Heijenoort, 1996; Świątek et al., 2012).

Dihydrodipicolinate synthase was the other upregulated (5.5 fold) protein in TB3585 and has already been mentioned (Section 3.3.1.2). As explained before, it leads to the production of diaminopimelic acid which is used as a precursor for cell wall synthesis and lysine production, which is required for the CC biosynthesis.

Plasmid partitioning protein ParB was also shown to be upregulated in AK39 (Section 3.2.1.7).

3.3.1.8. Inorganic/Organic Molecule Transport

There were two proteins categorized under this group: Cation/multidrug efflux protein and ABC transporter related protein.

According to UniProtKB database, cation/multidrug efflux protein has multidrug efflux transporter AcrB TolC docking domain. ArcB protein together with ArcA and TolC proteins form the AcrAB-TolC transport system in *E. coli* which transports drugs out of the cell by using proton motive force across the membrane (Wang et al., 2011). Therefore, they are important in conferring microorganism resistance against a variety of toxic compounds (Paulsen et al., 1996). The inclusion of thiostrepton into the culture medium of TB3585 for recombinant plasmid maintenance could explain the abundance of this cation/multidrug efflux protein.

ABC transporters are one of the largest family of transporters functioning as exporters to excrete toxins, drugs and lipids across membranes or as importers to transport diverse molecules, primarily nutrients (Rees et al., 2009).

3.3.1.9. Hypothetical/Unknown Proteins

There were 12 proteins in this category. Putative MerR-family transcriptional regulator was the common protein shown to be upregulated in both MALDI-TOF/MS and LC-MS/MS analyses.

MerR family members are mostly transcriptional activators responsive to such effectors as heavy metals, oxygen radicals, or cytotoxic compounds (Hayashi et al., 2013). This protein was explained in detail in Section 3.3.1.4.

AraC family proteins, just like two component systems, represent a signal transduction system in prokaryotes. However, in contrast to TCSs, they comprise one-component system and thus have both the sensory and output functions on the same polypeptide. Well-known examples of these proteins are related with the sugar metabolism, e.g., arabinose catabolism. On the other hand, some can recognize molecules other than sugars and can take part in virulence, morphological development and antibiotic production of the organism (Cuthbertson and Nodwell, 2013). Its overrepresentation of 13.6 fold points to its relation with CC overproduction with a yet unkown mechanism.

As for the putative uncharacterized protein (SCLAV_0117), half of the proteins that were shown to be associated with it are unknown, and one of the proteins is a regulatory protein according to STRING data (data not shown). EggNog, on the other hand, showed that the orthologues of this protein is involved in secondary metabolism and might be related to non-ribosomal peptide synthetase.

EggNog showed that all the orthologues of putative uncharacterized protein (SCLAV_3208) were unknown. STRING data (data not shown) showed that this protein is associated with 4 proteins, 2 of which are unknown and one is a transcriptional regulator, suggesting that this protein might be involved in transcriptional regulatory processes in the microorganism.

Uncharacterized protein (SCLAV_p1179) was described as unknown by EggNog regarding the orthologues of this protein. According to STRING database (Figure 3. 14), protein is associated with a transcriptional regulator and β-lactamase domain protein.

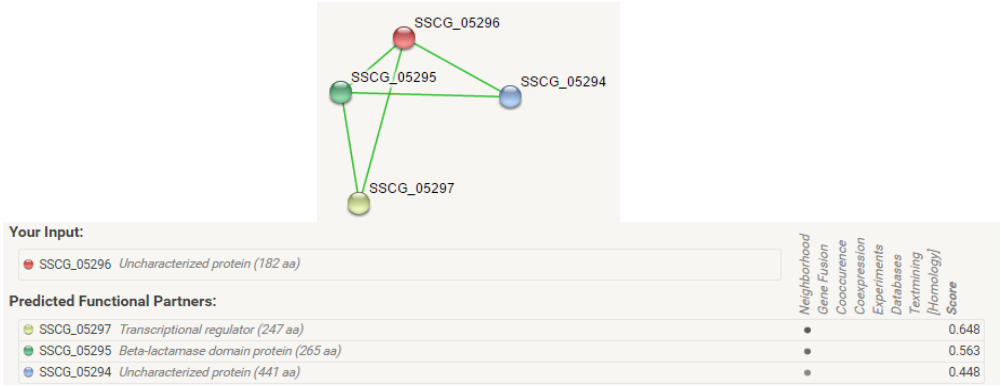
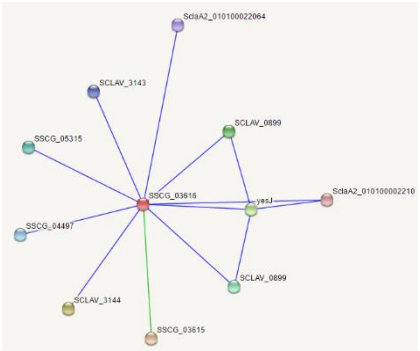


Figure 3. 14. STRING data showing the association of uncharacterized protein (SCLAV_p1179) with other proteins.

AfsR protein is a global regulator that affects not only the secondary metabolism but also the morphological differentiation in *Streptomyces*. AfsK/AfsR/AfsS system was first described in *S.coelicolor* and generally AfsK acts as a kinase of AfsR which, upon phosphorylation, binds the promoter region of *afsS*, activating the transcription of it. Thereafter, AfsS protein stimulates the transcription of pathway-specific regulatory genes, which in turn regulate the antibiotic production (actinorhodin and undecylprodigiosin) in *S. coelicolor*. However, AfsK/AfsR/AfsS system does not only include the interaction of these three proteins but also other players come into the stage to control the activity of AfsR and the expression of AfsS. Furthermore, this system controls phosphate, nitrogen, and sulfate assimilation processes other than secondary metabolism. The overexpression of AfsR protein was also shown to positively affect the antibiotic production in other microorganisms such as in *S. lividans*, *S. clavuligerus*, *S. griseus*, *S. peucetius*, and *S. venezuelae* (Santos-Beneit et al., 2011b).

According to the STRING database, this putative AfsR-like transcriptional regulator interacts with three regulatory proteins and several unknown proteins (data not shown).

As for GPP34 domain-containing protein, EggNog found 47 orthologues in 41 species, and none of them had a predicted function. Bell et al. (2001) first described the presence of Golgi-associated protein of 34 kDa (GPP34) predicted to be a peripheral membrane protein in Golgi and it is conserved from yeast to humans and suggested to be involved in Golgi trafficking (Bell et al., 2001). According to the STRING database (Figure 3.15), one of the proteins that GPP34 domain-containing protein is associated with is σ^{24} subunit belonging to the so called extracytoplasmic function (ECF) family, the name of which was coined by Lonetto et al. (1994). Most members of ECF σ factors are responsive to extracellular signals (Paget and Helmann, 2003). Moreover, most of them are cotranscribed with its cognate anti-σ factor that are transmembrane proteins with an extracytoplasmic domain that that senses the environmental stimuli and intracellular domain that binds to the σ factor to inhibit their activity. Upon stimulation of anti- σ factor by certain environmental signals such as small molecules or unfolded proteins, σ factor is released to bind to RNA polymerase so as to initiate the transcription through certain genes. This system is usually involved in transport, secretion, extracytoplasmic stress (Helmann, 2002). For example, in *S. coelicolor*, there were predicted to be 65 σ factors, fifty of which were estimated to be ECF σ factors. Nevertheless, only three of them, σ^{E}, σ^{R} and σ^{BldN}, could be described in detail with respect to their function (Paget et al., 2002; Helmann, 2002). Hence it is probable that GPP34 domain-containing protein identified in this study can function as an anti-sigma factor involved in the transport of yet-to-be characterized molecules and the function and importance of this so-called "anti-σ-σ" factor system for TB3585 strain needs to be characterized.



Your Input:

		Neighborhood	Gene Fusion	Cooccurence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
SSCG_03616	*GPP34 domain-containing protein (227 aa)*									

Predicted Functional Partners:

SSCG_03615	*Uncharacterized protein (215 aa)*	•								0.653
SCLAV_3144	*RNA polymerase, sigma-24 subunit, ECF subfamily (507 aa)*			•						0.572
yesJ	*Acetyltransferase (186 aa)*			•						0.546
SCLAV_0899	*Putative uncharacterized protein (62 aa)*			•						0.538
SCLAV_0899	*no annotation available (62 aa)*			•						0.538
SSCG_05315	*Putative uncharacterized protein (229 aa)*			•						0.523
SSCG_04497	*no annotation available (119 aa)*			•						0.516
SCLAV_3143	*Putative regulator (651 aa)*			•						0.507
SclaA2_010100022064	*no annotation available (139 aa)*			•						0.494
SclaA2_010100002210	*no annotation available (356 aa)*			•						0.491

Figure 3. 15. STRING data showing the association of GPP34 domain-containing protein with other proteins.

3.3.1.10. Others/General Function

In this category, 10 proteins were upregulated.

Acetyl-CoA synthetase is found in all kingdoms of life and catalyzes the reversible conversion of acetate to Ac-CoA, which is used in both the catabolic and anabolic reactions in the cell. Thus, it plays a crucial role in the central metabolism of the organism (Gulick et al., 2003; Gardner et al., 2006)

According to UniProtKB database, sulfurtransferase, which is identified in this study and coded by *SCLAV_3193*, contains rhodonase-like domain. Although the function of this protein has not yet been clearly elucidated, there is only one study in the literature on this protein (Nárdiz et al., 2011). The authors identified a protein overrepresented in the proteome of *oppA2* mutants of *S. clavuligerus*, which was then revealed to be the product of *SCLAV_3193*. They named the gene as *rhlA* and prepared *rhlA* mutants in order to understand the function of the protein. Though it was known

that *rhlA* codes for a rhodonase-like protein, it did not show ordinary rhodanase-like properties. Rhodonases are involved in thiosulfate utilization by reducing one of the sulfur atoms to sulfide and oxidizing the other to sulfite. However, *ΔrhlA* and control strains showed a similar growth pattern on thiosulfate containing medium. More importantly, *rhlA* mutants showed decreased levels of CC, holomycin and CA. In contrast to CC and holomycin, CA does not have sulfur atoms in its structure. Therefore, for the reduced production of CA, Nárdiz et al. (2011) suggested that a probable metabolic stress imposed on the microorganism by excess thiosulfate in the medium induces the production of sulfur containing antibiotics such as CC and holomycin, which, in turn, indirectly affects the production of CA and other antibiotics. Although excess thiosulfate-containing medium was not the case in our study, 3 fold upregulation of sulfurtransferase points to some kind of relation of this enzyme with CC biosynthesis. Yet, upregulation of mycothiol S-conjugate amidase (Section 3.3.1.5) might be indicative of the accumulation of mycothiols that are composed of large numbers of cysteine derivatives and act as reducing agents and storage forms of cysteines in *Actinomycetes*, as reported earlier. (Newton et al., 1996; Ung and Av-Gay, 2006; Nilewar and Kathiravan, 2014).

Oxidoreductase identified in this study was predicted to belong to GFO/IDH/MOCA family in UniProtKB/Swiss-Prot database (http://www.ebi.ac.uk/interpro/entry/IPR000683). However, there is not much information on this protein and its function in *Streptomyces*. Hence, it requires further characterization.

LmbE like enzymes form a large family, in which each member can perform diverse functions ranging from the biosynthesis of protective reducing agents and antibiotics to cell membrane components. For example, MshB responsible for the production of mycothiols in *Mycobacteria* are LmbE like enzymes (Viars et al., 2014). This protein was also shown to be downregulated in AK39 (Section 3.2.2.13).

According to UniProtKB database, methyltransferase type 12 contains a domain found in S-adenosyl-L-methionine-dependent methyltransferases. These proteins can be a

member of a methyltransferase superfamily and the class I-like SAM-binding methyltransferase superfamily.

Peptidase S1 and S6 were also shown to be upregulated in AK39 and explained in Section 3.2.1.10.

NH_3-dependent NAD^+ synthetase coded by *nadE* catalyzes the conversion of nicotinic acid adenine dinucleotide (NaAD) to NAD^+, being the final enzyme in the NAD^+ biosynthesis pathway (Devedjiev et al., 2001; Bieganowski et al., 2003). NAD^+ is crucial for many biological processes such as numerous oxidation-reduction reactions as well as DNA repair, DNA recombination and protein ADP ribosylations. There are two routes leading to the biosynthesis of NAD^+: a *de novo* pathway or pyridine nucleotide salvage pathway (Rizzi et al., 1996). For example, in *M. tuberculosis*, 5 genes (*nadA*, *nadB*, *nadD*, *nadE* and *nadF*) involved in *de novo* biosynthetic pathway of NAD^+ was shown to be present on the genome (Bellinzoni et al., 2005).

Nicotinate phosphoribosyltransferase coded by *pncB* is the first enzyme in NAD^+ salvage pathway (Chow, 2015) and catalyzes the formation of nicotinate mononucleotide (NAMN) from nicotinic acid (NA) and α-d-5-phosphoribosyl-1-pyrophosphate (PRPP) through ATP hydrolysis (Liang et al., 2012). In *M. tuberculosis*, *pncB* genes were shown to be upregulated under nutrient limitation and hypoxic conditions (Chow, 2015). Both the salvage pathway and *de novo* pathway merge after NaMN production.

3.3.2. Downregulated Proteins in TB3585

The spots with more than 2.5-fold decreased intensity as compared to their counterparts in NRRL 3585 are shown in Figure 3. 8 that indicates the proteins that are downregulated in both AK39 and TB3585 strains. For this, NRLL3585 gel was used for a much better visualization of differentially expressed proteins. Of the identified proteins, 28 proteins were found to be significantly downregulated in MALDI-TOF/MS analysis while 22 proteins were downregulated in LC-MS/MS analysis (Table 3. 7 and Table 3. 8, respectively).

Table 3. 7. List of downregulated proteins detected by 2DE in TB3585 strain.

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-Spots	PTM *	Genome Location **	Spot #
Carbohydrate Metabolism	E2PYJ1	Glucose-6-phosphate 1-dehydrogenase	*zwf*	Cytoplasm	-3.0	5.08	56,973	–	–	Chr.	186
	E2Q599	Malate oxidoreductase	*SCLAV_2067*	Cytoplasm	-3.4	5.03	48,951	–	–	Chr.	221
	B5GV08	Dihydrolipoyl dehydrogenase	*SCLAV_1398*	Cytoplasm	-3.3	5.94	49,402	–	–	Chr.	220
Amino Acid Metabolism	E2Q7X7	Acetolactate synthase	*ilvB1*	Cytoplasm	-5.6	5.65	65,779	–	–	Chr.	143
	E2PXE0	Alanine dehydrogenase	*SCLAV_0984*	Cytoplasm	-4.7	5.76	41,935	–	–	Chr.	175
	B5H1V1	D-aminoacylase	*SCLAV_3902*	Cytoplasm	-16.7	5.32	57,695	–	–	Chr.	182
	B5GZV9	E1-alpha branched-chain alpha keto acid dehydrogenase	*bkdA1*	Cytoplasm	-2.6	5.68	40,843	–	–	Chr.	192
	D5SIP1	Putative ornithine cyclodeaminase	*arcB*	Cytoplasm	-3.0	6.06	38,557	–	–	pSCL4	259
	B5GSL5	4-hydroxyphenylpyruvate dioxygenase	*SCLAV_2046*	Cytoplasm	-3.4	5.11	41,873	–	–	Chr.	232
	B5GUR5	5-carboxymethyl-2-hydroxymuconate delta-isomerase	*SCLAV_4800*	Cytoplasm	-7.5	5.34	29,825	–	–	Chr.	261
	E2PWS1	Branched-chain alpha-keto acid dehydrogenase subunit E2	*SCLAV_2999*	Cytoplasm	-5.6	5.85	51,496	–	–	Chr.	143
Lipid Metabolism	E2PWS1	Branched-chain alpha-keto acid dehydrogenase subunit E2	*SCLAV_2999*	Cytoplasm	-5.6	5.85	51,496	–	–	Chr.	143

Table 3. 7. List of downregulated proteins detected by 2DE in TB3585 strain. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-Spots	PTM *	Genome Location **	Spot #
Lipid Metabolism	B5GZV9	E1-alpha branched-chain alpha keto acid dehydrogenase	*bkdA1*	Cytoplasm	-2.6	5.68	40,843	–	–	Chr.	192
	B5GMB4	3-hydroxyisobutyrate dehydrogenase	*SCLAV_p0937*	Cytoplasm	-9.8	5.66	45,134	–	–	pSCL4	237
	B5GMB5	Acyl-CoA dehydrogenase domain-containing protein	*SCLAV_p0938*	Cytoplasm	-9.8	5.46	40,818	–	–	pSCL4	237
	B5H0X6	Enoyl-[acyl-carrier-protein] reductase [NADH]	*fabI*	Cytoplasm	-4.1	5.47	27,606	–	–	Chr.	42
Translation, Ribosomal Structures	E2PV67	Histidine--tRNA ligase	*hisS*	Cytoplasm	-2.5	5.63	46,521	–	–	Chr.	168
Stress-Related, Protein Turnover, Chaperones	E2PYD4	Ectoine hydroxylase	*ectD*	Unknown	-10.0	5.64	32,419	–	–	Chr.	124
	B5H3I2	ATP-dependent Clp protease	*SCLAV_2385*	Cytoplasm	-6.25	5.77	93	–	–	Chr.	129
	E2Q3P1	Trigger factor	*tig*	Cytoplasm	-53	4.50	55,724	–	–	Chr.	133
Secondary Metabolism	E2PYD4	Ectoine hydroxylase	*ectD*	Unknown	-10.0	5.64	32,419	–	–	Chr.	124
Cell Processes (Shape/Division/Motility)	E2PZS0	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase	*murE*	Cytoplasm	-4.1	5.39	58,036	–	–	Chr.	219
	E2PVV1	UDP-N-acetylmuramate--L-alanine ligase	*murC*	Cytoplasm	-3.4	5.63	54,761	–	–	Chr.	221
	B5GX14	Phosphoglucosamine mutase	*glmM*	Cytoplasm	-3.4	5.03	46,621	–	–	Chr.	221

Table 3. 7. List of downregulated proteins detected by 2DE in TB3585 strain. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-Spots	PTM *	Genome Location **	Spot #
Inorganic/Organic Molecule Transport	B5GTX5	ABC transporter intracellular ATPase subunit	*SCLAV_3978*	Cytoplasmic membrane	-2.6	5.56	37,838	–	–	Chr.	192
Hypothetical/Unknown Proteins	E2PUV4	Putative two-component system response regulator	*SCLAV_4778*	Cytoplasm	-4.1	6.01	27,422	–	–	Chr.	42
	B5H0U3	DUF574 domain-containing protein	*SCLAV_0131*	Cytoplasm	-3.6	4.82	30,332	–	–	Chr..	120
Others/General Function	E2Q3Q1	Acetyltransferase	*SCLAV_1797*	Cytoplasmic	-2.7	5.44	44,545	–	–	Chr.	177
	B5GZV5	Dehydrogenase	*paaZ*	Cytoplasm	-3.0	5.11	59,495	–	–	Chr.	186
	E2Q3I4	Flavoprotein disulfide reductase	*SCLAV_3827*	Cytoplasm	-4.1	5.48	55,168	–	–	Chr.	219
	B5GZB8	Aldehyde dehydrogenase	*aldH*	Cytoplasm	-9.2	4.53	55,339	–	–	Chr.	161

*PTM refers to possible posttranslational modification. C denotes charge modification, M denotes mass modification, C-M shows both charge and mass modification. ** "Chr." refers to chromosome, "pSCL4" refers to plasmid pSCL4.

Table 3. 8. List of downregulated proteins in TB3585 strain identified by LC-MS/MS.

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Amino Acid Metabolism	D5SLL7	Methionine synthase II	*SCLAV_p1324*	Cytoplasm	-2.148	4.86	38,157	pSCL4
	B5GLV1	Cysteine synthase	*SCLAV_p1477*	Cytoplasm	-2.003	5.24	31,830	pSCL4
Lipid Metabolism	E2Q711	Biotin carboxylase	*SCLAV_0182*	Cytoplasm	-2.710	5.14	46,512	Chr.
Translation, Ribosomal Structures	B5GXN9	Ribosomal RNA small subunit methyltransferase H	*rsmH*	Cytoplasm	-2.148	6.48	34,541	Chr.
Stress-Related, Protein Turnover, Chaperones	B5H0Q8	Alkaline serine protease	*SCLAV_0093*	Extracellular	-2.003	9.98	42,961	Chr.
	E2Q479	Catalase	*katA*	Cytoplasm	-2.279	5.66	55,348	Chr.
Secondary Metabolism	D5SIN9	Non-ribosomal peptide synthetase	*SCLAV_p0291*	Cytoplasm	-2.800	5.34	72,960	pSCL4
	E2PWJ7	Clavaminate synthase 1	*cas1*	Cytoplasm	-2.800	5.31	35,370	Chr.
	Q9X5G6	Putative ribulose-5-phosphate epimerase	*cvm2*	Cytoplasm	-2.710	5.27	16,877	Chr.
	E2PWJ4	Putative pyridoxal phosphate-dependent aminotransferase	*cvm6*	Cytoplasm	-2,510	5.36	48,065	Chr.
	E2PWJ8	Aldo/keto reductase	*cvm1*	Cytoplasm	-2,399	6.41	36,650	Chr.
	Q9X5G9	Flavin-dependent oxidoreductase, F420-dependent methylene-tetrahydromethanopterin reductase	*cvm5*	Cytoplasm	-3,687	5.85	44,863	Chr.
	E2PZM4	Cytochrome P450 monooxygenase	*SCLAV_1254*	Cytoplasmic Membrane	-2,260	6.26	49,791	Chr.

Table 3. 8. List of downregulated proteins in TB3585 strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Inorganic/Organic Molecule Transport	B5GXJ3	Branched-chain amino acid transport ATP-binding protein	*SCLAV_1201*	Cytoplasmic Membrane	-2,148	6.93	25,373	Chr.
Hypothetical/Unknown Proteins	E2Q9C1	Putative oxidoreductase	*SCLAV_0465*	Cytoplasm	-3,371	6.22	37,219	Chr.
	B5GZ79	Putative zinc-binding dehydrogenase	*SCLAV_1520*	Cytoplasm	-2,510	5.27	34,906	Chr.
	B5GST5	Putative uncharacterized protein	*SCLAV_3868*	Unknown	-2,003	9.03	14,662	Chr.
	E2Q2N4	Putative GMC-family oxidoreductase	*SCLAV_5678*	Cytoplasmic Membrane	-2,003	8.38	60,996	Chr.
	E2PW58	Secreted protein	*SCLAV_0814*	Cell wall; Extracellular	-2.710	8.45	47,613	Chr.
	B5GQE3	Extracellular small neutral protease	*SCLAV_4112*	Extracellular	-2.148	9.03	22,984	Chr.
Others/General Function	B5GMR3	Short-chain dehydrogenase/reductase SDR	*SCLAV_p1083*	Cytoplasm	-2,003	5.37	25,558	pSCL4
	E2Q158	N5,N10-methylenetetrahydromethanopterin reductase	*SCLAV_3491*	Cytoplasm	-2.279	5.72	33,007	Chr.

* "Chr." refers to chromosome, "pSCL4" refers to plasmid pSCL4.

3.3.2.1. Carbohydrate Metabolism

There were three proteins shown to be downregulated in this category: Glucose-6-phosphate 1-dehydrogenase, malate oxidoreductase and dihydrolipoyl dehydrogenase.

Glucose-6-phosphate dehydrogenase is one of the key enzymes in pentose phosphate pathway (Agarwal and Auchus, 2005). Pentose phosphate pathway is important for the generation of NADPH, the breakdown of carbon sources, and the recruitment of essential metabolites for nucleic acids, amino acids and vitamins. This pathway can be divided into two branches as oxidative and non-oxidative pentose phosphate pathway (Sprenger, 1995). Glucose-6-phosphate dehydrogenase is the major player of the oxidative pentose phosphate pathway (Sprenger, 1995) which produces NADPH from $NADP^+$ needed for biosynthetic processes (Kruger and von Schaewen, 2003).

Malate oxidoreductase catalyzes the conversion of malate to pyruvate, an intermediate used in a large variety of metabolic processes within the cell (Loeber et al., 1991). As mentioned earlier, pyruvate is not only used as an intermediate to supply cells with energy but can also be converted to several other precursors including amino acids to be used in diverse metabolic processes.

Dihydrolipoyl dehydrogenase enzyme is one of the components of the pyruvate dehydrogenase multienzyme complex which is composed of three enzymes: pyruvate dehydrogenase, dihydrolipoyl transacetylase, and dihydrolipoyl dehydrogenase (Moe and Hammes, 1974). They catalyze the oxidative decarboxylation of pyruvate and their sequential enzymatic activity is shown below (Figure 3. 24) (Moe et al., 1974).

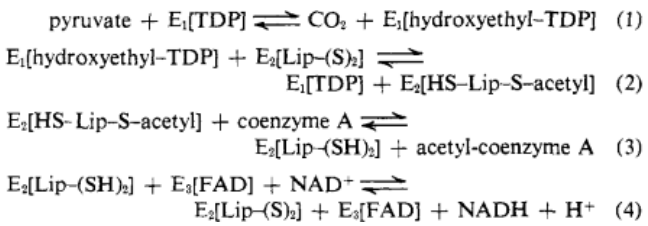


Figure 3. 16. Oxidative decarboxylation of pyruvate by pyruvate dehydrogenase multienzyme complex. E_1= pyruvate dehydrogenase, E_2= dihydrolipoyl transacetylase, E_3= dihydrolipoyl dehydrogenase, TDP= thiamine diphosphate, Lip-$(S)_2$ = oxidized form of lipoic acid, Lip-$(SH)_2$ = reduced form of lipoic acid (Moe et al., 1974).

Proteins under this category are related with the production of basic metabolites to be used in diverse cellular processes like energy production or biosynthetic reactions. The decreased expression of these proteins might reflect a kind of trade-off for the production of secondary metabolites.

3.3.2.2. Amino Acid Metabolism

There were 10 proteins shown to be downregulated in this category. Of these, E1-alpha branched-chain alpha keto acid dehydrogenase, 4-hydroxyphenylpyruvate dioxygenase, D-aminoacylase, 5-carboxymethyl-2-hydroxymuconate delta-isomerase and methionine synthase II were also shown to be downregulated in AK39 (Section 3.2.2.3).

Branched-chain alpha-keto acid dehydrogenase subunit E2 is one of the four components of BCDH complex, just like E1-alpha branched-chain alpha keto acid dehydrogenase (Denoya et al., 1995), and is involved in fatty acid and polyketide biosynthesis as well as branched chain amino acid catabolism.

Acetohydroxy acid synthase (AHAS), also known as acetolactate synthase, is important for the biosynthesis of branched chain amino acids. It first catalyzes the decarboylation of pyruvate and then the condensation of it with another pyruvate to form acetolactate which is the precursor for both valine and leucine (Elisáková et al., 2005).

L-Alanine dehydrogenase catalyzes the reversible deaminating reaction of L-alanine to pyruvate as shown below (Kasarenini and Demain, 1994).

L-alanine + NAD^+ + H_2O <--> pyruvate + NH_3+ NADH + H^+

L-alanine dehydrogenase activity was shown to be stimulated by higher concentrations of alanine or NH_4^+ in *S. clavuligerus* and suggested to have considerable contribution to the ammonium assimilation process in the organism only when the ammonium concentration is high in the environment (Braña et al., 1986). Moreover, Kasarenini and Demain (1994) showed that as the activity of alanine dehydrogenase increased, alanine was accumulated in the cell, which had inhibitory effect on CC production by negatively affecting the activity of ACV synthetase, cyclase and expandase enzymes in *S. clavuligerus*. Given that this enzyme produces alanine that has inhibitory effect

on cephamycin C production in *S. clavuligerus*, downregulation of it in a CC overproducer is not much intriguing.

Putative ornithine cyclodeaminase encoded by *arcB* is responsible for the irreversible conversion of L-ornithine to L-proline by the removal of the ammonia (Goodman et al., 2004). Ornithine is the precursor of arginine that is necessary for the biosynthesis of CA.

Cysteine synthase catalyzes the formation of cysteine from O-acetyl-O-serine and hydrogen sulfide (Saito et al., 1994). Cysteine, apart from being building block amino acid of proteins, is one of the three precursors required for the synthesis of CC biosynthesis (Liras and Martín, 2006). Morover, it is also a component of mycothiols which are produced by *Actinomycetes* and are low molecular weight molecules used as reducing agents as well as storage form of cysteines in the cells (Lee et al., 2005; Ung and Av-Gay, 2006; Nilewar and Kathiravan, 2014). Cysteines are used for the biosynthesis of sulfur-containing amino acids (Chang and Vining, 2002). In other words, cysteine is a versatile amino acid that can be used in several biological pathways. *Streptomyces clavuligerus* genome contains 9 cysteine synthase-coding genes including putative genes (Medema et al., 2010).

Given that the the conversion of cysteine to cystathionine marks the initiation of the trans-sulfuration pathway leading to methionine biosynthesis (Chang and Vining, 2002), concomitant downregulation of cysteine synthase with methionine synthase in TB3585 provides some clue for decreased synthesis of methionine in this CC overproducer. Methionine can be used as an important precursor for diverse cellular processes in the central metabolism. Hence, the downregulation of these two proteins along with the others in this category might suggest that certain central metabolic pathways might be suppressed to shunt more precursors towards CC biosynthetic pathway.

3.3.2.3. Lipid Metabolism

There were 6 proteins shown to be downregulated in TB3585 under this category: Branched-chain alpha-keto acid dehydrogenase subunit E2 (Sections 3.2.2.3 and 3.3.2.2), E1-alpha branched-chain alpha keto acid dehydrogenase (Sections 3.2.2.3

and 3.3.2.2), 3-hydroxyisobutyrate dehydrogenase, enoyl-[acyl-carrier-protein] reductase [NADH] (FabI), acyl-CoA dehydrogenase domain-containing protein and biotin carboxylase.

FabI protein which is an enoyl-[acyl-carrier-protein] reductase (enoyl-ACP reductase) that takes role in type II fatty acid biosynthesis system in bacteria and important for the biosynthesis of cell envelope, phospholipids, lipoproteins, lipopolysaccharides, etc. (Heath et al., 2000, Heath et al., 2001; Kitagawa et al., 2007) was shown to be upregulated in AK39 strain (Section 3.2.1.2).

3-Hydroxyisobutyrate dehydrogenase catalyzes the reversible oxidation of L-3-hydroxyisobutyrate to methylmalonate semialdehyde. This enzyme plays an important role in valine catabolism (Chowdhury et al., 1996). Methylmalonate semialdehyde is converted to propionyl-CoA which is turned to methylmalonyl-CoA by propionyl-CoA carboxylase. Polyketides produced by *Streptomyces* are assembled through several extender units such as malonyl-CoA, methylmalonyl-CoA, and occasionally, ethylmalonyl-CoA (Liu and Reynolds, 2001). Thus 3-hydroxyisobutyrate dehydrogenase plays a pivotal role in the production of methylmalonyl-CoA which can be used as precursor for polyketide biosynthesis as well as the valine metabolism.

Biotin carboxylase is one of the four components of acetyl-CoA carboxylases which is involved in the production of malonyl-CoA production that can be used as precursors for both the fatty acid as well as polyketide biosynthesis (Rodríguez and Gramajo, 1999; Rodríguez et al., 2001). This protein was also shown to be downregulated in AK39 (Section 3.2.2.4). As suggested there, polyketide biosynthesis might be suppressed in our overproducers in order to to increase CC production, thereby decreasingother metabolic burden.

Acyl-CoA dehydrogenase domain-containing protein was also shown to be downregulated in AK39 strain (Section 3.2.2.4)

3.3.2.4. Translation, Ribosomal Structures

Histidine--tRNA ligase is also known as histidyl-tRNA synthetase and functions to produce histidyl-transfer RNAs which are used during protein synthesis to incorporate histidines into peptides (Freist et al., 1999).

Ribosomal RNA small subunit methyltransferase H (RsmH) is an AdoMet-dependent MTase and responsible for the N^4-methylation of C1402 in *E. coli* 16S rRNA (Wei et al., 2012). Modification of rRNA nucleotides is a well-known and common mechanism that affects ribosome activities such as RNA-RNA interactions or antibiotic resistance (Kimura and Suzuki, 2010).

3.3.2.5. Stress-Related, Protein Turnover, Chaperones

There were 5 proteins categorized in this group: Ectoine hydroxylase, ATP-dependent Clp protease, trigger factor, alkaline serine protease and catalase.

Ectoine hydroxylase and the trigger factor are also shown to be downregulated in AK39 strain and explained in detail in Section 3.2.2.7. It is worth noting that downregulation of these two proteins is drastical in both AK39 and TB3585, being 10 fold and 53 fold, respectively, in TB3585 strain according to MALDI-TOF/MS results.

Clp proteins are ATP-dependent proteases. *E. coli* Clp proteases are composed of two parts: ATPase/chaperone (ClpA, ClpX, ClpY) and proteolytic subunits (ClpP, ClpQ) (Schelin et al., 2002). These proteases can also be divided into two categories depending on the proteolytic core. For example, if the proteolytic core is ClpP, chaperone/ATPase subunit complexed with it will be ClpA or ClpX. If the proteolytic core is ClpQ, the subunit complexed with the core will be ClpY. ClpP is a serine type protease and commonly found in bacteria but also in plants and mammals, whereas ClpQ has a threonine active site and is more like 26S proteasome family (Porankiewicz et al., 1999). Substrates to be degraded are recognized by the chaperone units which unfold proteins and direct them to the proteolytic chambers, in which they are eventually degraded into smaller peptide fragments (Schelin et al., 2002). Moreover, it was also shown that Clp proteases usually target key regulatory proteins (de Crécy-Lagard et al., 1999). WhiB1, which is a transcriptional repressor, is also a target of Clp protease in *Mycobacterium tuberculosis*. If WhiB protein degradation by Clp proteases is blocked, stabilized WhiB proteins in the cell turn out to be toxic for the organism (Raju et al., 2014).

Catalase enzyme encoded by *katA* is responsible for converting H_2O_2 to H_2O and O_2 (Zou et al., 1999), thereby preventing the production of highly reactive hydroxyl

radical ($HO^{\cdot}$) and contributing to the survival and adaptation processes of the microorganism (Lee et al., 2005). KatA is a monofunctional catalase enzyme that shows only catalase activity in contrast to bifunctional catalases which have both catalase and peroxidase functions (Lee et al., 2005).

Actinomycetes are known to produce a large variety of proteases such as neutral alkaline, and thermostable proteases (Ramesh et al., 2009). Alkaline proteases, which are extracellular proteins, have a high pH optimum of 8.5 to 10 and can be produced by bacteria, molds, yeasts and also observed in mammalian tissues (Yum et al., 1994). The nitrogen and carbon sources in the environment as well as the growth rate are important factors that affect the secretion of alkaline proteases in *Actinomycetes* (Mehta et al., 2006).

3.3.2.6. Secondary Metabolism

There were 8 downregulated proteins in this category, seven of which were common with the downregulated proteins of the same category in AK39; namely one NRPS (SCLAV_p0291; plasmid-encoded), clavaminate synthase 1, putative ribulose-5-phosphate epimerase, putative pyridoxal phosphate-dependent aminotransferase, aldo/keto reductase, flavin-dependent oxidoreductase and ectoine hydroxylase. They were were extensively discussed in Section 3.2.2.8.

Cytochrome P450 monooxygenase is found in all kingdoms (Wong, 1998, Hannemann et al., 2007) and involved in a plethora of processes (De Mot and Parret, 2002) such as hydroxylation, N-, O- and S-dealkylation, sulphoxidation, epoxidation, deamination, desulphuration, dehalogenation, peroxidation, and N-oxide reduction (Bernhardt, 2006). However, they are especially known for the metabolism of xenobiotics (e.g. petroleum products, pesticides) as well as antibiotic biosynthesis (O'Keefe and Harder, 1991; Rodríguez et al., 1995; De Mot and Parret, 2002). *Actinomycetes*, and more specifically *Streptomyces*, which are versatile producers of secondary metabolites were found to be rich sources of xenobiotic metabolizing enzymes, cytochrome P450s. Apart from xenobiotic metabolizing properties in *Actinomycetes*, cytochrome P450 monooxygenases can have roles in secondary metabolism like macrolide biosynthesis. For example, it was shown that a P450 encoded by the *eryF* gene catalyzes an intermediate step during the production of

erythromycin A, which is a macrolide antibiotic produced by *Saccharopolyspora erythraea* (O'Keefe and Harder, 1991). Since the type of cytochrome P450 that was identified in this study could not be specified in LC-MS/MS analyses, the association network and the predicted function of its orthologues were scrutinized through both the STRING (data not shown) and EggNog databases, respectively. According to STRING data, of the 10 proteins shown to be associated with cytochromes P450, only three of them are known. One of the predicted proteins is squalene-hopene cyclase which is involved in hopanoid biosynthesis (Ghimire et al., 2015). Hopaniods in bacteria are equivalents of sterols in the eukaryotes and stabilize the bacterial membrane by integration into the cytoplasmic membranes (Siedenburg and Jendrossek, 2011). However, the presence of the hopanoids is hard to detect, mainly due to two reasons: (i) hopanoids are rigid structures which make them difficult to extract by analytic techniques, (ii) synthesis of the hopanoids is suggested to be strictly regulated (Kannenberg and Poralla, 1999). Evidence for the latter comes from the fact that even though *S. peucetius* was shown to contain a cluster genes of hopanoid biosynthesis, there was no hopanoid detection after the isolation of lipids (Ghimire et al., 2015). Another protein shown to be associated is prenyltransferase/squalene oxidase which is suggested to be related with secondary metabolism (Zhou et al., 2013). Furthermore, in EggNog database, the orthologues of the protein were categorized under secondary metabolism. This protein requires further characterization in order to find out its function in *S. clavuligerus*.

3.3.2.7. Cell Processes (Shape/Division/Motility)

There were 3 proteins shown to be downregulated in this category: UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase (MurE), UDP-N-acetylmuramate--L-alanine ligase (MurC), phosphoglucosamine mutase (GlmM).

UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase (MurE) protein was also shown to be downregulated in AK39 strain (Section 3.2.2.10). MurE with UDP-N-acetylmuramate--L-alanine ligase (MurC) protein is involved in peptidoglycan layer biosynthesis in bacteria.

GlmM protein, as explained in Section 3.3.1.7, is also involved in peptidoglycan cell wall synthesis.

3.3.2.8. Inorganic/Organic Molecule Transport

There were two proteins in this category: ABC transporter intracellular ATPase subunit and branched-chain amino acid transport ATP-binding protein.

(ABC) transporter superfamily is one of the largest transporter classes. Their action mechanism depends on the binding and hydrolysis of ATP to provide sufficient energy to transport diverse substances varying from ions to macromolecules across the membrane. They can be importers or transporters and found in all types of organisms. Importers are exclusively found in the prokaryotes while exporters are present in both prokaryotes and eukaryotes. ABC transportes are composed of 4 structural domains: two transmembrane domains (TMDs) and two nucleotide binding domains (NBDs) (Rees et al., 2009). Although NBDs are highly conserved within the family, TMDs show a great deal of variety as a result of the chemically diverse molecules to be transported across the membrane (Hung et al., 1998; Rees et al., 2009).

According to UniProtKB database, branched-chain amino acid transport ATP-binding protein contains ABC transporter-like domain.

3.3.2.9. Hypothetical/Unknown Proteins

Eight proteins were identified under this category.

DUF574 domain-containing protein was also shown to be downregulated in AK39 (Section 3.2.1.12) and as was stated before, it is possible for this protein to be a transcriptional regulator or is a DNA binding protein *per se* since STRING data predicted its association with some DNA binding proteins.

Also explained in detail earlier, two component systems (TCSs) are important regulatory mechanisms that control both the secondary metabolism and morphological differentiation in *Streptomyces*. (Martín and Liras, 2010; Rodríguez et al, 2013). A two component system (SCLAV_3595) was shown to be upregulated in TB3585 (Section 3.3.1.4). Another two-component system (SCLAV_4778), a putative one, appeared as downregulated in the same strain.

Putative zinc-binding dehydrogenase was predicted to belong to zinc-type alcohol dehydrogenase superfamily by UniProtKB database. ADHs catalyze a large variety of reactions. Their substrates also differ considerably ranging from normal and branched chain alcohols to aldehydes and ketones. Furthermore, organisms usually contain more than one ADHs (Reid and Fewson, 1994).

In order to predict the association network and the orthologues of the putative uncharacterized protein (SCLAV_3868), STRING (Figure 3. 17) and EggNog databases were searched. In STRING database, association of four proteins with unknown functions was predicted. Furthermore, putative uncharacterized protein was predicted to be VapC toxin by sequence similarity. VapBC (virulence-associated protein) family members comprise toxin-antitoxin (TA) modules found in *Mycobacterium tuberculosis* genome and VapC toxin acts as a ribonuclease, but it binds specific RNA sequences, inhibiting the translation of mRNAs (Sharp et al., 2012). EggNog database retrieved data with orthologues of unknown functions.

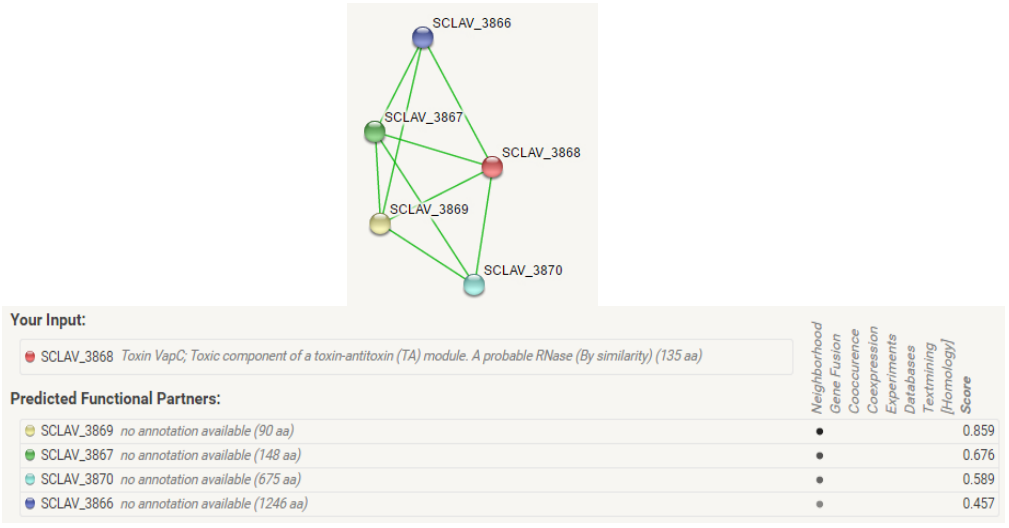
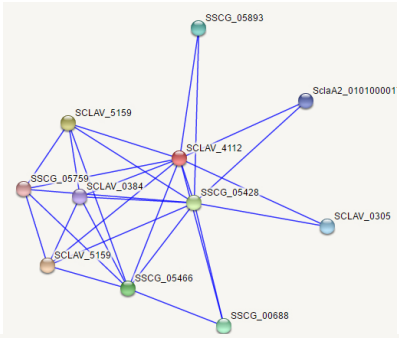


Figure 3. 17. STRING data showing the associaton of putative uncharacterized protein (SCLAV_3868) with other proteins.

GMC (glucose–methanol–choline) oxidoreductases contain a wide variety of enzymes, catalyze diverse reactions and are found in both prokaryotes and eukaryotes. Though they show diverse biological functions, novel sequences were also detected to contain GMC patterns with no known function and thus awaiting characterization. (Zámocký et al., 2004).

The same secreted protein (SCLAV_0814) was also shown to be downregulated in AK39 strain and explained in Section 3.2.2.12.

When orthologues of extracellular small neutral protease (SCLAV_4112) was searched through EggNog database, there was no assigned function for them. According to the STRING database (Figure 3. 18), SCLAV_4112 is associated with some membrane and secreted proteins as well as a glycosyltransferase and IgA peptidase. Glycosyltransferases are responsible for transferring a sugar moiety to several substrates such as a saccharide, protein, lipid, DNA. Therefore, they can be involved in diverse processes in the cell. For example, MurG is involved in peptidoglycan biosynthesis and is responsible for transferring N-acetyl-D-glucosamine (GlcNAc) from UDP-GlcNAc to N-acetylmuramoyl pentapeptide. Bacteriophage T4 β -glucosyltransferase (BGT), on the other hand, transfers glucose moieties to 5-hydroxymethylcytosine residues found in its own DNA, thereby evading from the attacks of host nucleases (Ünligil and Rini, 2000). However, glycosyltransferases are also prominent in antibiotic producing microorganisms. Mgt, from *Streptomyces lividans* and OleD from *Streptomyces antibioticus* are involved in the inactivation of oleandomycin by transferring glucose moiety. Moreover, several glycosyltransferases were shown to be involved in the biosynthesis of antibiotics such as baumycin and daunorubicin in *Streptomyces peucetius*, or tylosin biosynthesis in *Streptomyces fradiae* and mithramycin biosynthesis in *Streptomyces argillaceus* (Fernández et al., 1998). IgA peptidase is another protein that small neutral protease was shown to be associated with. It is a secreted protein especially found in bacteria that use humans as hosts such as pathogenic bacteria (Lenart et al., 2013).



Your Input:

		Neighborhood	Gene Fusion	Cooccurence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
SCLAV_4112	*Extracellular small neutral protease (217 aa)*									

Predicted Functional Partners:

SCLAV_5159	*IgA Peptidase M64 (638 aa)*			•						0.777
SCLAV_5159	*no annotation available (602 aa)*			•						0.777
SSCG_05428	*Glycosyl transferase (948 aa)*			•						0.754
SSCG_05466	*LamG domain protein (1461 aa)*			•						0.751
SSCG_00688	*Putative transmembrane transport protein (325 aa)*			•						0.729
SSCG_05893	*Secreted protein (502 aa)*			•						0.727
SCLAV_0305	*Putative secreted protein (528 aa)*			•						0.724
SclaA2_010100001705	*no annotation available (505 aa)*			•						0.723
SCLAV_0384	*Putative membrane protein (495 aa)*			•						0.716
SSCG_05759	*no annotation available (495 aa)*			•						0.716

Figure 3. 18. STRING data showing the associaton of putative uncharacterized protein (SCLAV_4112) with other proteins.

3.3.2.10. Others/General Function

In this category, there were 6 proteins downregulated in TB3585: Acetyltransferase, dehydrogenase, flavoprotein disulfide reductase, aldehyde dehydrogenase, short-chain dehydrogenase/reductase SDR and N5,N10-methylenetetrahydromethanopterin reductase.

Aldehyde dehydrogenase family is composed of members with different substrate specificities and can be subdivided into 7 groups. This protein was shown to be upregulated in AK39 (Section 3.2.1.9).

paa genes in *E. coli* are involved in phenylacetic acid degradation. There are 14 *paa* genes that are transcribed as 3 transcription units (Ferrández et al., 1998). Of these PaaZ, which was identified in this study as dehydrogenase enzyme, catalyzes the conversion of oxepin-CoA to 3-oxo-5,6-dehydrosuberyl-CoA in phenylacetic acid degradation pathway (Teufel et al., 2010).

Flavoprotein disulfide reductase which is coded by *SCLAV_3827* and acetyltransferase which is coded by *SCLAV_1797* were also shown to be downregulated in AK39 strain (Section 3.2.1.13). Downregulation of both proteins in two different CC overproducers might suggest the presence of similar mechanisms to be exploited for breaking down the controls that exist to limit the secondary metabolism.

Short-chain dehydrogenase/reductase SDR (SCLAV_p1083) is one of the largest family of proteins and is found in all kingdoms. They are NAD(P)(H)-dependent oxidoreductases with similar mechanisms. They were shown to be involved in a large variety of biological processes such as lipid, amino acid, carbohydrate, cofactor, hormone, xenobiotic metabolism and redox sensor mechanisms (Kavanagh et al., 2008; Oppermann et al., 2003). Another short-chain dehydrogenase/reductase SDR protein (SCLAV_4467) was also shown to be downregulated in AK39 in Section 3.2.2.13.

Methylenetetrahydromethanopterin reductase (Mer) is responsible for the reversible reduction of N^5, N^{10}-methylenetetrahydromethanopterin to methyltetrahydromethanopterin during which $F_{420}H_2$ is also oxidized to F_{420} (Aufhammer et al., 2005). Mer is commonly found in methanogenic and sulfate reducing archaea (Thauer and Kunow, 1995; Shima et al., 2000). Moreover, the sequence comparison of Mer isolated from *Methanobacterium thermoautotrophicum* showed 27% similarity to the F_{420}-dependent reductase (LmbY) protein known to be part of lincomycin biosynthesis in *Streptomyces lincolnensis* (Vaupel and Thauer, 1995).

3.3.3. Summary for TB3585 Upregulated and Downregulated Proteins

According to MALDI-TOF/MS and LC-MS/MS results, a total of 45 proteins were upregulated while 50 proteins were downregulated in TB3585. Furthermore, distribution of upregulated and downregulated proteins identified by both techniques to functional classes are shown in Figure 3. 19 and Figure 3. 20, respectively.

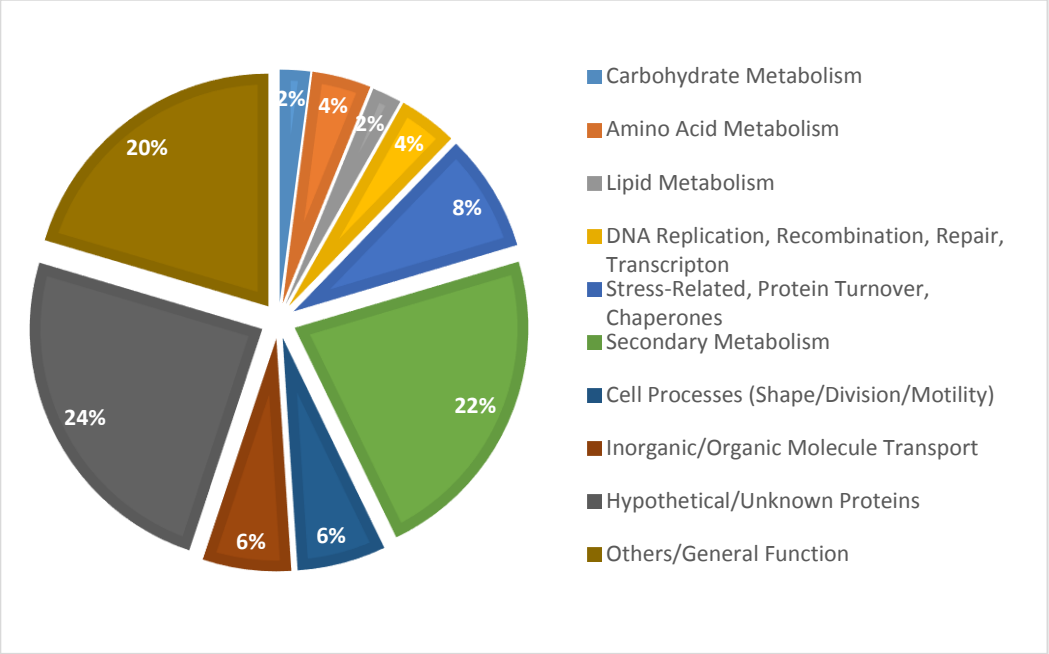


Figure 3. 19. Relative quantitative distribution of upregulated proteins identified in TB3585 to the functional categories. Percentage of each category was calculated by dividing the number of proteins in each category to the total number of proteins in all categories

Upregulated proteins in "Hypothetical/Unknown Proteins" and "Secondary Metabolism" categories constituted approximately 50% of the proteins found in all categories. Moreover, overrepresented proteins in "Secondary Metabolism" identified by the MALDI-TOF/MS analysis showed direct relation with the CC biosynthesis whereas LC-MS/MS analysis could also identify unrelated proteins such as NocE-like and alkaline d peptidase involved in other secondary metabolic pathways. These two proteins were also shown to be upregulated in AK39, suggesting that two CC overproducers might resort to similar regulatory mechanisms in order to increase the antibiotic production. A two component system regulator (SCLAV_3595) or putative transcriptional regulators identified in the upregulated proteins of TB3585 might have roles in the regulation of CC production.

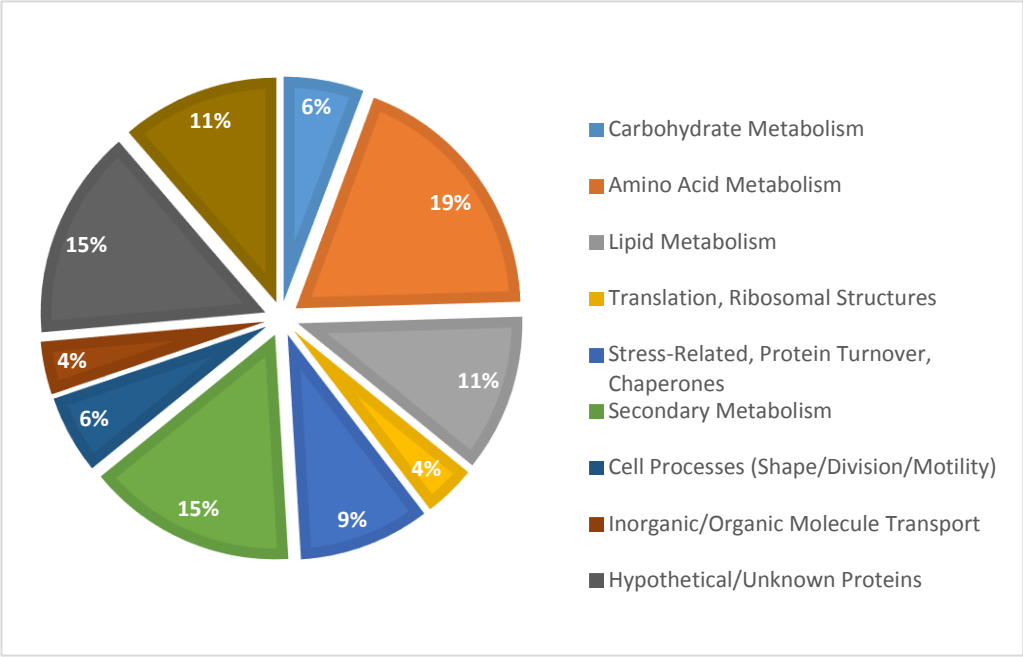


Figure 3. 20. Relative quantitative distribution of downregulated proteins identified in TB3585 strain to the functional categories. Percentage of each category was calculated by dividing the number of proteins in each category to the total number of proteins in all categories.

Among the downregulated protein categories, "Amino Acid Metabolism" had the second highest number of proteins, which was followed by "Secondary Metabolism" and "Hypothetical/Unknown Proteins". Moreover, downregulated proteins in "Secondary Metabolism" in LC-MS/MS results were mostly comprised of the ones involved in clavam biosynthesis, suggesting that clavam biosynthesis might be negatively affected by the overproduction of CC.

3.4. Comparative Proteome Analysis of NRRL 3585 and DEPA Strains by 2DE MALDI-TOF/MS and LC-MS/MS

The iterative cycles of random mutagenesis introduced to the industrial strain DEPA has introduced several important genetic modifications (random damage to the DNA through strand breakage, rearrangements, addition, deletion, or substitution of bases) leading to an industrial CA producer. Although classical strain improvement methodology is slow and laborious, its long history of success still fascinates industrial

fermentation researchers, especially with the availability of high throughput screening and analytical technologies today in the post-"omics" era (Patnaik, 2008). In order to provide insight into the modifications that DEPA strain has undergone, the wild type strain NRRL 3585 and DEPA were analyzed by comparative proteomics based via MALDI-TOF/MS and LC-MS/MS technique.

A fused image of the total soluble proteome of NRRL 3585 and DEPA strains are shown in Figure 3. 21. Of the proteins that were identified in DEPA strain, 33 proteins were shown to be significantly up-regulated and 60 proteins were shown to be significantly downregulated compared to the parental strain in MALDI-TOF/MS (Table 3. 9 and Table 3. 11, respectively). On the other hand, 50 proteins were upregulated while 120 proteins were shown to be downregulated in DEPA in LC-MS/MS analysis (Table 3.10 and Table 3.12, respectively).

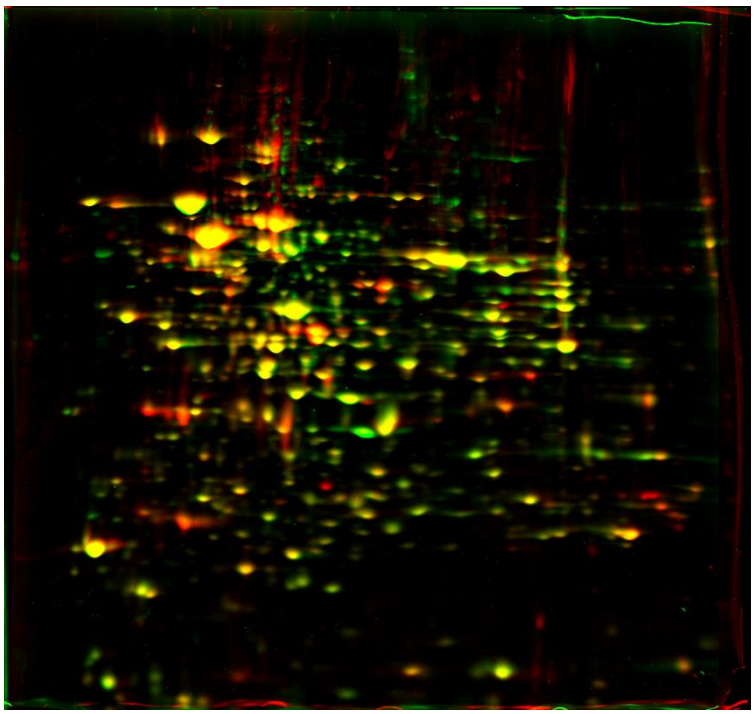


Figure 3. 21. Dual channel 2DE imaging of *S. clavuligerus* strains, NRRL 3585 (green) and DEPA (red), when the strains were grown for 48 h in SA medium, pI 4-7.

3.4.1. Upregulated Proteins in DEPA Strain

Figure 3. 22 shows protein spots on DEPA gel that showed more than 2.5-fold increase in expression when compared to their counterparts on NRRL 3585 gel. Upregulated proteins of DEPA strain are tabulated in Table 3. 9 and Table 3. 10. 33 proteins were identified as upregulated by MALDI-TOF/MS analysis whereas 50 proteins were shown to be upregulated by LC-MS/MS analysis.

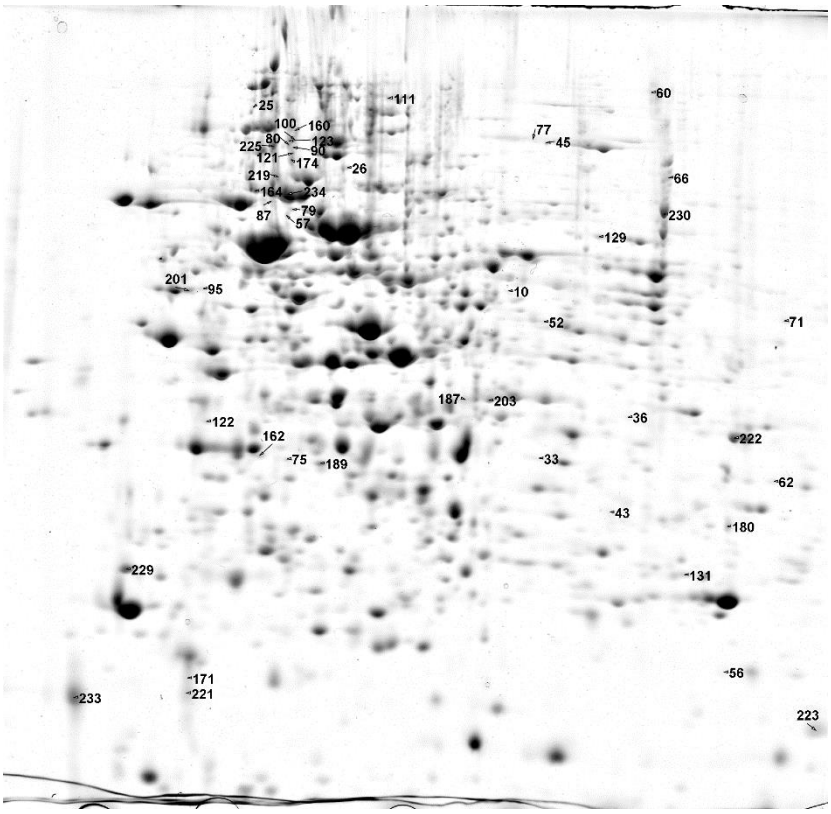


Figure 3. 22. Protein spots on DEPA 2DE gel that showed more than 2.5 fold increase compared to those on NRRL 3585 gel.

Table 3. 9. List of upregulated proteins identified on DEPA 2DE gel.

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spot	PTM *	Genome Location **	Spot #
Nucleotide Metabolism (DNA/RNA Processing)	B5H1T0	Polynucleotide phosphorylase/polyadenylase	*pnp*	Cytoplasm	11.0	5.00	79,508	–	–	Chr.	26
Carbohydrate Metabolism	E2PV06	Aconitate hydratase	*acnA*	Cytoplasm	4.6	4.84	98,214	–	–	Chr.	25
	E2PVD8	Putative isocitrate dehydrogenase	*icdA*	Cytoplasm	22.8 (a); 6.1 (b)	4.97	79,048	2 (a, b)	M	Chr.	100, 160
	B5GV08	Dihydrolipoyl dehydrogenase	*SCLAV_1398*	Cytoplasm	364.0	5.94	49,402	–	–	Chr.	60
Amino Acid Metabolism	B5H2V9	Glycine hydroxymethyltransferase	*glyA*	Cytoplasm	3.3	5.78	45,508	–	–	Chr.	52
	E2Q7M7	Pyrroline-5-carboxylate reductase	*proC*	Cytoplasm	3.5	5.85	29,718	–	–	Chr.	131
	E2Q396	Bifunctional protein FolD	*folD*	Cytoplasm	2.6	6.26	30,143	–	–	Chr.	222
DNA Replication, Recombination, Repair, Transcripton	E2QA66	ATP-dependent DNA helicase	*SCLAV_4693*	Cytoplasm	1344	5.96	78,933	–	–	Chr.	66
	B5GPP5	Anti-sigma factor	*SCLAV_2541*	Cytoplasm	7.3 (a); 3.3 (b)	4.55	14,964	2 (a, b)	M	Chr.	171, 221
Secondary Metabolism	E2Q5Q8	Carboxyethylarginine synthase	*ceaS2*	Cytoplasm	11	5.01	62,348	–	–	Chr.	26
	B5GLD3	Clavaldehyde dehydrogenase	*car*	Cytoplasm	8.2	6.17	26,456	–	–	Chr.	180
	B5GLC7	Carboxyethyl-arginine beta-lactam synthase	*bls2*	Cytoplasm	20.6	5.92	54,53	–	–	Chr	230

Table 3. 9. List of upregulated proteins identified on DEPA 2DE gel. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spot	PTM *	Genome Location **	Spot #
Secondary Metabolism	E2PZ87	Thioredoxin reductase	*SCLAV_5275*	Cytoplasm	3.3	6.52	37,503	–	–	Chr.	203
Energy Production/Electron-Iron Transfer	B5GWF3	NADH-quinone oxidoreductase subunit	*nuoG1*	Cytoplasm; Cell membrane	5.4 (a); 2.7 (b)	5.78	87,947	2 (a, b)	C	Chr.	45, 77
	B5H1L3	Ferredoxin-nitrite reductase	*SCLAV_4899*	Cytoplasm	3.2	5.77	62,969	–	–	Chr.	129
Inorganic/Organic Molecule Transport	E2Q5C6	Protein translocase subunit SecA	*secA*	Cytoplasm	24.1	5.10	103,997	–	–	Chr.	111
Hypothetical/ Unknown Proteins	B5H0R6	Putative phosphoesterase	*SCLAV_0101*	Cytoplasm	3.5	5.87	35,762	–	–	Chr.	36
	E2Q5V1	Putative uncharacterized protein	*SCLAV_0035*	Cytoplasm	3.1	6.12	20,971	–	–	Chr.	56
	E2Q783	Putative acetyltransferase	*SCLAV_0254*	Cytoplasm	3.0 (a); 2.8 (b)	5.01	30,907	2 (a, b)	C	Chr.	75, 189
	E2PY35	Putative uncharacterized protein	*SCLAV_5138*	Unknown	5.2 (a); 2.9 (b)	4.66	47,919	2 (a, b)	C	Chr.	95, 201
	D5SK44	Uncharacterized protein	*SCLAV_p0800*	Cytoplasm	12.8	4.38	19,528	–	–	pSCL4	233
	B5GMI9	Beta-lactamase domain protein	*SCLAV_p1007*	Cytoplasm	(a-l) ; 3.6, 14.2,23325, 5.9, 19,23, 184.4, 10.6, 28.7, 2.9, 6.2, 4.6	4.90	70,62	12 (a-l)	C-M	pSCL4	57, 79, 80, 87,90, 121, 123, 164, 174, 219, 225, 234

Table 3. 9. List of upregulated proteins identified on DEPA 2DE gel. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spot	PTM *	Genome Location **	Spot #
Others/General Function	B5H4A8	Rhs element Vgr protein	*SCLAV_0043*	Cytoplasm	3.8	5.68	43,514	–	–	Chr.	10
	B5H160	Siderophore-interacting protein	*SCLAV_0843*	Cytoplasm	4.7	5.70	31,519	–	–	Chr.	33
	B5H0U2	DNA-binding protein	*SCLAV_0130*	Cytoplasm	28.8	6.20	33,689	–	–	Chr.	43
	B5GN32	Aminotransferase class-III	*SCLAV_p0794*	Cytoplasm	3.3	5.67	50,462	–	–	pSCL4	52
	E2PU72	FHA domain protein	*SCLAV_0615*	Cytoplasm	9.1	6.97	34,128	–	–	Chr.	62
	B5GZH0	Amino-transferase	*SCLAV_4765*	Cytoplasm	2.7	6.34	47,299	–	–	Chr.	71
	E2Q6I6	Cellulose-binding protein	*SCLAV_2237*	Cytoplasm	3.5	4.70	32,156	–	–	Chr.	122
	B5GZG7	Methyltransferase	*SCLAV_4762*	Cytoplasm	3.5	4.85	30,471	–	–	Chr.	162
	D5SK16	Methylenetetrahydromethanopterin reductase	*SCLAV_p0772*	Cytoplasm	2.8	5.33	38,276	–	–	pSCL4	182
	B5H3N0	YceI family protein	*SCLAV_4479*	Unknown	2.9	4.60	21,705	–	–	Chr.	229
	B5GX76	Nucleotide-binding protein	*SCLAV_3612*	Cytoplasm	2.9	6.37	17,984	–	–	Chr.	223

*PTM refers to possible posttranslational modification. C denotes charge modification, M denotes mass modification, C-M shows both charge and mass modification. ** "Chr." refers to chromosome, "pSCL4" refers to plasmid pSCL4.

Table 3. 10. List of upregulated proteins in DEPA strain identified by LC-MS/MS.

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Carbohydrate Metabolism	E2PVV7	Beta-glucosidase	*SCLAV_4874*	Cytoplasm	2.065	4.89	50,206	Chr.
	E2Q0G5	Pyruvate phosphate dikinase	*SCLAV_5441*	Cytoplasm	3.255	7.28	61,735	Chr.
Amino Acid Metabolism	B5H3S7	Aspartate aminotransferase	*SCLAV_1074*	Cytoplasm	2.065	5.23	37,815	Chr.
	E2Q7L2	Glutamyl-tRNA reductase	*hemA*	Cytoplasm	2.065	5.23	57,494	Chr.
	B5GZD0	Glutamate-1-semialdehyde 2,1-aminomutase	*SCLAV_5663*	Cytoplasm	2.279	6.96	47,307	Chr.
DNA Replication, Recombination, Repair, Transcripton	B5GYV1	Two component transcriptional regulator, LuxR family	*bldM*	Cytoplasm	2.466	10.09	21,955	Chr.
	E2Q5K2	Histidine kinase	*SCLAV_4141*	Cytoplasm	2.466	5.56	101,818	Chr.
	E2PZH1	Two-component system response regulator	*SCLAV_1202*	Cytoplasm	2.466	4.82	24,352	Chr.
	E2PUN2	CRISPR-associated protein Cas4	*SCLAV_2739*	Cytoplasm	2.279	5.86	41,732	Chr.
	E2PXV5	ATP-dependent RNA helicase	*SCLAV_3124*	Cytoplasm	2.065	5.87	82,053	Chr.
	B5GQK1	Helix-turn-helix domain-containing protein	*SCLAV_1642*	Cytoplasm	2.279	7.91	31,548	Chr.
	E2QA55	LexA repressor	*lexA*	Cytoplasm	2.779	6.99	28,412	Chr.
	B5GSX2	DeoR family transcriptional regulator	*SCLAV_3828*	Unknown	2.466	9.76	33,405	Chr.
Translation, Ribosomal Structures	E2PW68	50S ribosomal protein L35	*rpmI*	Cytoplasm	2.065	11.49	7,119	Chr.

Table 3. 10. List of upregulated proteins in DEPA strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Stress-Related, Protein Turnover, Chaperones	B5GWY1	Peptidyl-prolyl cis-trans isomerase	*SCLAV_0740*	Unknown	2.065	9.55	27,605	Chr.
	B5H3S5	Diaminobutyrate--2-oxoglutarate transaminase	*ectB1*	Cytoplasm	2.466	5.48	46,083	Chr.
Secondary Metabolism	B5GLB7	Positive regulator	*ccaR*	Cytoplasm	2.065	6.02	29,346	Chr.
	B5H3S5	Diaminobutyrate--2-oxoglutarate transaminase	*ectB1*	Cytoplasm	2.466	5.48	46,083	Chr.
	E2PUT1	Polyprenyl synthetase	*SCLAV_4756*	Cytoplasm	2.279	7.23	45,996	Chr.
	E2PZ87	Thioredoxin reductase	*SCLAV_5275*	Cytoplasm	3.899	6.52	37,503	Chr.
	E2Q954	Phosphotransferase	*SCLAV_4596*	Cytoplasm	2.779	8.26	31,905	Chr.
	B5GLB4	7-alpha-cephem-methoxylase P8 chain	*cmcJ*	Cytoplasm	2.914	5.53	34,043	Chr.
Cell Processes (Shape/Division/Motility)	E2Q316	CHAP domain-containing protein	*SCLAV_1694*	Extracellular	2.779	6.08	33,650	Chr.
	E2PZF8	Exopolysaccharide phosphotransferase	*SCLAV_5352*	Cytoplasm	2.631	8.93	69,266	Chr.
Inorganic/Organic Molecule Transport	B5H183	Secreted protein	SCLAV_2084	Unknown	2,631	4.94	48,486	Chr.
Hypothetical/Unknown Proteins	B5GZG2	Putative uncharacterized protein	*SCLAV_4757*	Cytoplasm	2.027	4.97	44,959	Chr.
	E2Q0G6	Putative transcriptional regulator	*SCLAV_5442*	Unknown	2.065	5.13	22,796	Chr.
	E2Q753	Putative uncharacterized protein	*SCLAV_0224*	Unknown	2.065	5.16	22,930	Chr.
	D5SL16	Putative methyltransferase	*SCLAV_p1123*	Unknown	2.065	5.60	31,869	pSCL4

Table 3. 10. List of upregulated proteins in DEPA strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Hypothetical/Unknown Proteins	E2PU21	Putative ATP/GTP-binding protein	*SCLAV_0564*	Cytoplasm	2.065	6.59	22,782	Chr.
	B5GZG0	Putative uncharacterized protein	*SCLAV_4755*	Unknown	2.065	8.90	12,384	Chr.
	B5GNM2	Putative uncharacterized protein	*SCLAV_3510*	Unknown	2.065	9.16	33,917	Chr.
	E2PYV7	Putative uncharacterized protein	*SCLAV_3217*	Unknown	2.065	9.43	39,512	Chr.
	E2Q041	Putative P450-like hydroxylase	*SCLAV_3389*	Cytoplasmic membrane	2.279	4.82	47,384	Chr.
	E2Q752	Putative uncharacterized protein	*SCLAV_0223*	Cytoplasm	2.279	5.59	37,747	Chr.
	B5GZ98	Putative MaoC-like dehydratase	*SCLAV_1540*	Unknown	2.279	8.80	37,282	Chr.
	E2Q4K4	Putative mutase	*SCLAV_1969*	Cytoplasmic membrane	2.779	5.80	22,354	Chr.
	E2PVT5	Putative protoporphyrinogen oxidase	*SCLAV_4852*	Cytoplasm	2.779	6.16	51,382	Chr.
	E2Q4B3	Putative uncharacterized protein	*SCLAV_1878*	Cytoplasm	3.037	5.48	28,555	Chr.
	B5H484	Putative uncharacterized protein	*SCLAV_2624*	Unknown	3.255	5.67	44,014	Chr.
	B5GL09	DUF124 domain-containing protein	*SCLAV_4315*	Cytoplasm	2.065	5.57	28,076	Chr.
	E2PW10	Sporozoite_P67 domain-containing protein	*SCLAV_4927*	Cytoplasm	2.065	5.71	65,732	Chr.
	D5SL37	Peptidoglycan-binding domain 1 protein	*SCLAV_p1144*	Extracellular	2.466	8.72	34,781	pSCL4

Table 3. 10. List of upregulated proteins in DEPA strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Hypothetical/Unknown Proteins	E2Q5G9	Putative polyprenyl diphosphate synthase	*SCLAV_4108*	Cytoplasm	2.520	5.21	42,006	Chr.
	E2Q1P7	Cell surface protein	*SCLAV_5606*	Cytoplasm	2.779	4.57	53,513	Chr.
	E2Q428	DUF946 domain-containing protein	*SCLAV_3892*	Cytoplasmic membrane	3.833	6.83	49,275	Chr.
Others/General Function	B5GZH0	Amino-transferase	*SCLAV_4765*	Cytoplasm	2.631	6.34	47,299	Chr.
	B5H0U2	DNA-binding protein	*SCLAV_0130*	Cytoplasm	2.779	6.20	33,689	Chr.
	B5GSK5	Dehydrogenase	*SCLAV_2055*	Cytoplasm	2.065	5.19	81,016	Chr.
	B5GQ70	NDP-hexose 4-ketoreductase	*SCLAV_4039*	Unknown	2.466	8.63	33,242	Chr.
	B5GV80	Oxidoreductase-like protein	*SCLAV_1870*	Cytoplasm	2.466	6.16	40,168	Chr.

* "Chr." refers to chromosome, "pSCL4" refers to plasmid pSCL4.

3.4.1.1. Nucleotide Metabolism (DNA/RNA Processing)

There was just one protein in this category: Polynucleotide phosphorylase/polyadenylase (PNPase). This protein is coded by *pnp* gene and is mainly responsible for RNA degradation being a 3'-5' exoribonuclease (Wu et al., 2009; Sohlberg et al., 2003). It is also a component of RNA degradosome, made up of PNPase, enolase, endoribonuclease, RNase E, ATP-dependent RNA helicase and RhlB (Bralley and Jones, 2004; Xiao et al., 2011). During the degradation of RNA molecules, PNPase can either function alone or for the processing of highly structured RNA molecules, it can be complexed with RhlB or whole degradosome. PNPase also has poly(A) polymerase activity and it was shown in *E. coli* that it can degrade RNAs or in the poly(A) polymerase enzyme lacking mutants of *E. coli*, add poly(A) tails to RNAs (Mohanty and Kushner, 2000; Bralley and Jones, 2004). Moreover, PNPase seems to be a critical component of the cell under oxidative stress conditions. When *E. coli pnp* mutants were subjected to H_2O_2, the cells accumulated 8-oxoG in cellular RNA; however, when the cells were complemented with *pnp* gene, the amount of oxidized RNAs were minimized (Wu et al., 2009). This can be a regulatory mechanism for the translation of certain proteins in the cells or it can contribute to the stress response in DEPA strain against emerging nutritional and oxidative stress.

3.4.1.2. Carbohydrate Metabolism

There were 5 proteins in this category: Aconitate hydratase, dihydrolipoyl dehydrogenase, putative isocitrate dehydrogenase, β-glucosidase and pyruvate phosphate dikinase.

In this study, aconitate hydratase was shown to be upregulated by 4.6 fold in DEPA strain. Aconitate hydratase (also known as aconitase) is responsible for the reversible conversion of citrate to isocitrate (Viollier et al., 2001). However, a microarray analysis of *S. clavuligerus* industrial strain which overproduces CA was shown to express citrate synthase and aconitase in TCA cycle lesser levels (Medema et al., 2011). These contradictory results must stem from the different mutagenesis processes that both strains had undergone. Hence, metabolic activities of the strains as well as the precursor pools that they use to produce CA must differ from each other.

Dihydrolipoyl dehydrogenase (E3) is one of the three components of the bacterial pyruvate dehydrogenase multienzyme complex and catalyzes the oxidative decarboxylation of pyruvate to form acetyl-CoA (Koike et al., 1960; Moe et al., 1974). It is known that this enzyme is inhibited when one or more of the three following ratios are increased: ATP/ADP, NADH/NAD^+ and acetyl-CoA/CoA. While it was found as a significantly upregulated protein (over 3 hundred folds) in this study, transcriptomics work reported by Medema et al. (2011) reported it as an over two-fold downregulated enzyme of carbohydrate metabolism.

Isocitrate dehydrogenase encoded by *icdA* is common in 3 kingdoms of life as a result of its importance in carbon metabolism (Zhang et al., 2013b). It is involved in citric acid cycle and converts the isocitrate to α-ketoglutarate which is also an important precursor for the biosynthesis of several metabolites in the cell (Hurley et al., 1996; Takahashi-Iñiguez et al., 2014). This protein was represented by two spots increased by 23- and 6-fold with the same pI but differing in their mass.

β-glucosidase is responsible for converting cellobiose to glucose units. Although this enzyme is mostly intracellular in some species such as *Streptomyces* and *Micromonospora*, extracellular forms of the enzyme is also present, for example, in *Cellulomonas fimi* (Spiridonov and Wilson, 2001). Moreover, *Streptomyces antibioticus* and *S. venezuelae* can use extracellular β-glucosidase to remove the glucose moiety on the macrolides that they produce in order to render them active after they have been excreted into the environment (Zmudka et al., 2013).

Pyruvate phosphate dikinase is important for the glycolytic pathway and is involved in the reversible conversion of AMP, PP_i, and phosphoenolpyruvate (PEP) to ATP, P_i, and pyruvate (Eisaki et al., 1999). Phosphoenolpyruvate and pyruvate are involved and used as precursors in diverse biological processes as well as energy production. Moreover, feeding studies showed that pyruvate together with ornithine or arginine increases the production of CA yield in cultures (Saudagar et al., 2008).

Proteins in this category are related to the production of metabolites that are exploited in a variety of biosynthetic processes, including α-ketoglutarate family of amino acids with arginine as a precursor of CA.

3.4.1.3. Amino Acid Metabolism

There were 6 proteins upregulated in "Amino Acid Metabolism": Glycine hydroxymethyltransferase, pyrroline-5-carboxylate reductase, bifunctional protein FolD, aspartate aminotransferase, glutamyl-tRNA reductase, glutamate-1-semialdehyde 2,1-aminomutase.

Glycine hydroxymethyltransferase, also known as serine hydroxymethyltransferase (SHMT) was shown to be downregulated in AK39 and explained there in detail (Section 3.2.2.3)

Pyrroline-5-carboxylate reductase (P5C reductase) is the last enzyme involved in proline synthesis from glutamate. (Deutch et al., 1982). Proline biosynthesis is a three-enzyme-catalyzed process in which glutamate kinase and γ-glutamyl phosphate reductase are the first two enzymes that catalyze the conversion of glutamate to glutamate-4-semialdehyde which spontaneously cyclizes to pyrroline-5-carboxylate (P5C). Third enzyme, P5C reductase, converts P5C to proline (Smith et al., 1995).

Bifunctional protein FolD which is also known as 5,10-methylene-tetrahydrofolate dehydrogenase/5,10-methylenetetrahydrofolate cyclohydrolase (Kim et al., 2015b) catalyzes a two-step reaction. First, 5,10-methylenetetrahydrofolate is oxidized to 5,10-methenyltetrahydrofolate and then 5,10-methenyltetrahydrofolate is hydrolyzed to 10-formyltetrahydrofolate (Xia et al., 2015). Tetrahydrofolate (THF) and derivatives coming from THF are important cofactors involved in one-carbon metabolism (Schirch, 1998; Kim et al., 2015b). Folate-activated 1-carbon atoms are used for the biosynthesis of purines and thymidylate and also for the remethylation of homocysteine to methionine. Methionine is an important molecule involved in a range of processes varying from protein synthesis to production of S-adenosylmethionine which is expoited in a variety of reactions such as methylation of proteins, cytosine bases on DNA, phospholipids, and numerous small molecules (Stover, 2009).

Aspartate aminotransferase catalyzes the reversible transamination reaction shown below (Rhee et al., 1997):

L-Aspartate + α-ketoglutarate $\rightleftharpoons$ L-glutamate + oxaloacetate

Reaction products catalyzed by this enzyme are important for the carbon and the nitrogen metabolism of the organism since they can be used as intermediates in diverse biological processes (Torre et al., 2006). One of the pathways in which aspartate is used is the aspartate pathway that leads to the production of CC biosynthesis. Moreover, glutamate is used in diverse metabolic processes and gives rise to several amino acids. It also leads to the production of ornithine which is the precursor of arginine in clavam and CA biosynthesis.

Glutamyl-tRNA reductase and glutamate-1-semialdehyde 2,1-aminomutase are involved in 5-aminolevulinic acid (ALA) biosynthesis which is the precursor of such tetrapyrroles as hemes, chlorophylls, billins, and corrinoids through C5 pathway. Different types of tetrapyrroles are used as cofactors of enzymes involved in diverse metabolic processes such as central metabolism, transcriptional regulation, energy metabolism and detoxification mechanisms (de Orué Lucana et al., 2004; Zappa et al., 2010; Takano et al., 2015).

3.4.1.4. DNA Replication, Recombination, Repair, Transcripton

In this category, there were 10 proteins shown to be upregulated in DEPA strain.

According to UniProtKB data, ATP-dependent DNA helicase identified in this study belongs to RecQ type ATP-dependent DNA helicase family. RecQ type helicases were shown to be responsible for genome maintenance and their homologs showed unwinding of the paired DNA or acted on Holliday junctions or G-quartet DNA, suggesting that they might be involved in DNA recombination processes (Cobb et al., 2002).

Sigma (σ) factors complexing with RNA polymerases form holoenzyme structures and are crucial for RNA polymerases to recognize the promoter regions to initiate the trancription. There are two types of sigma factors: primary σ factors and alternative σ factors. Primary σ factors control the expression of housekeeping genes, while

alternative σ factors are activated depending on environmental or stress conditions. Anti-σ factors regulate the function of alternative σ factors. Anti-σ factors bind their cognate sigma factors so that the transcription of related genes are prevented until the conditions are met for their release (Helmann, 1999). Anti-sigma factor identified in the present study was represented by two spots suggesting that it might have undergone PTM. Moreover, further characterization of the protein is needed in order to explain its main function and if there is any, its probable involvement in the secondary metabolism of the organism.

CRISPR (clustered regularly interspaced short palindromic repeat) is a non-coding DNA sequence widely distributed in bacterial genome and gives rise to non-coding RNAs (Deveau et al., 2010). CRISPR sequences are separated by spacer sequences which are derived from the invading DNA sequences like virus or plasmid DNA (Zhang et al., 2012). RNAs arising from CRISPR arrays together with CRISPR-associated (Cas) proteins target invading nucleic acids, thereby constituting the bacterial immune system. They are first transcribed as a large primary transcript and then processed by Cas proteins into short RNA sequences also known as CRISPR RNAs (crRNA) that contain invader-targeting sequences (Hale et al., 2009). Cas4 protein together with Cas1 and Cas2 proteins are hypothesized to be involved in the spacer acquisition from the invader sequences (Zhang et al., 2012). The upregulation of Cas4 protein might suggest increased spacer acquisition from the environment, i.e. activation of CRISPR system for a yet unknown reason.

UniProtKB database suggests that ATP-dependent RNA helicase protein (SCLAV_3124) identified in DEPA strain contains a DEAD box helicase domain. The same protein, but with a different gene entry (SCLAV_4026), was also among the upregulated proteins of AK39 (Section 3.2.1.3). We have already discussed the possible roles of this protein in relevant sections.

Two component transcriptional regulator (BldM) is a putative response regulator (Molle and Buttner, 2000), which is involved in the aerial mycelium formation (Chater, 2001). *Streptomyces* morphogenesis is initiated with the aerial hypae production and the mutants lacking aerial hyphae are called bald (*bld*) mutants

(Sprusansky et al., 2003). *bld* gene products are usually pleiotropic regulators that affect, for example, carbon catabolite repression, signalling processes, morphogenesis as well as antibiotic production (Molle and Buttner, 2000). *bldM* expression seems to be dependent on *bldN* gene which is, in turn, dependent on *bldG* and *bldH* transcription (Bibb et al., 2000), the products of which affect both the antibiotic production and morhological differentiation (Bignell et al., 2005). Increased level of this protein in DEPA is thus as expected from an industrial overproducer. Two-component regulators identified as overrepresented (encoded by *bldM* and *SCLAV 1202*) in DEPA has not been shown in AK39 and TB3585.

Helix-turn-helix (HTH) domain is commonly found in transcriptional factors as the main part that interacts with the DNA in all three kingdoms of life as descended from a common ancestor (Aravind et al., 2005). However, HTH domain containing protein identified in this study needs to be further characterized with respect to the promoters that it interacts with in DEPA strain.

The global regulator LexA repressor negatively regulates the transcription of genes belonging to SOS regulon that comprises the functions of error-prone DNA polymerases, DNA repair enzymes and cell division inhibitors that act until the damage is rectified. The winged HTH motif of LexA is a variant form of HTH DNA binding motif. Under normal physiological conditions, LexA protein represses the transcription of most of the genes in this regulon; however, when there is a stressful DNA damage, RecA protein which senses DNA damage at stalled replication forks is activated and by forming a two-component regulatory system with LexA, it causes LexA to self-cleave itself (autoproteolysis), relieving the repression on the genes of SOS regulon (Domain et al., 2004). The overexpression of LexA protein in DEPA clearly shows that SOS regulon is suppressed, which is probably a kind of adaptation by the organism to iterative rounds of mutagenesis that this organism has been subjected to. It can be speculated that overexpression of LexA is a means to prevent error-prone DNA repair as well as cell division delay.

DeoR family of proteins are transcriptional regulators mostly functioning as repressors of carbohydrate metabolism. (Elgrably-Weiss et al., 2006). For instance, in

Corynebacterium glutamicum, DeoR-type regulator SugR induced by fructose-6-phosphate was shown to inhibit the transcription of *ptsG* gene which codes for a protein involved in phosphoenolpyruvate-dependent sugar phosphotransferase system (PTS) (Engels and Wendisch, 2007).

3.4.1.5. Translation, Ribosomal Structures

50S ribosomal protein L35was the sole protein identified in this category.

50S ribosomal protein L35 is one of the proteins found in the large subunit of the ribosomes (Kashiwagi and Igarashi, 1987). The abundant expression of this protein might suggest the increased fidelity of the translation of RNAs from the ribosomes.

3.4.1.6. Stress-Related, Protein Turnover, Chaperones

Two proteins were identified in this category: Peptidyl-prolyl cis-trans isomerase, diaminobutyrate--2-oxoglutarate transaminase.

Peptidyl-prolyl cis-trans isomerase (PPIase) is involved in protein folding (Stoller et al., 1995). Protein folding generally occurs spontaneously and this information is encoded on the primary sequence of the peptide. According to the refolding studies, folding of some globular polypeptides takes place in a matter of seconds or miliseconds. However, one conformational change was shown to occur much more slowly, which is the *cis-trans* isomerization of peptidyl-prolyl bonds (Göthel and Marahiel, 1999). PPIase catalyzes the *cis-trans* isomerization of X-Pro peptide bonds, which is the rate limiting step during protein folding. It was shown to accelerate the refolding of denatured proteins in vitro and is suggested to be essential for the folding of newly synthesized proteins by accelerating the rate limiting proline isomerization step (Kofron et al., 1991).

Diaminobutyrate--2-oxoglutarate transaminase is encoded by *ectB1*. Ectoine is synthesized in three successive enzymatic reactions starting from aspartic β-semialdehyde (Bursy et al., 2008). The genes involved in the biosynthesis are called *ectA*, *ectB* and *ectC* which encode the enzymes L-2,4-diaminobutyric acid acetyltransferase, L-2,4-diaminobutyric acid transaminase, and L-ectoine synthase,

respectively. In our study, L-2,4, diaminobytyric acetyltransferase (*ectA1*) was found downregulated in AK39, as revealed by LC-MS/MS studies (Section 3.2.2.7). Another enzyme used by a subgroup of the ectoine producers to convert ectoine into 5-hydroxyectoine is ectoine hydroxylase (EctD), a member of the nonheme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily. This enzyme was demonstrated as downregulated in 2DE MALDI TOF/MS analysis of both AK39 (Section 3.2.2.7) and TB3585 (Section 3.3.2). In the present part of our study, we found the second enzyme of ectoin biosynthesis, diaminobutyrate--2-oxoglutarate transaminase encoded by *ectB1* as an upregulated enzyme in DEPA strain. Since L-aspartate-β-semialdehyde is a precursor for both ectoine and CC synthesis, we suggested earlier that AK39 and TB3585 produce much more CC than NRRL 3585 by using L-aspartate-β-semialdehyde so that ectoine biosynthesis is drastically downregulated in favor of CC overproduction. However, since the strain DEPA is a CA overproducer, it is not influenced from ectoin biosynthesis.

3.4.1.7. Secondary Metabolism

There were 9 proteins found to be upregulated in this category. Thioredoxin reductase is the common protein that was shown to be upregulated in both MALDI-TOF/MS and LC-MS/MS results. Diaminobutyrate--2-oxoglutarate transaminase was already explained in Section 3.4.1.6.

Positive regulator (CcaR), as explained before, is a transcriptional regulator that positively affects the CC and CA production in *S. clavuligerus* by binding the promoter regions of certain genes in both CC and CA gene cluster. For example, CcaR can bind *claR* gene which is a transcriptional regulator that regulates transcription of late genes in CA gene cluster (Santamarta et al., 2011; Ferguson et al., 2016). Overexpression of CcaR is definitely an indication of increased transcription through *claR*, thereby accounting for increased levels of CA in DEPA.

7-alpha-cephem-methoxylase P8 chain (CmcJ) is involved in the final steps of the CC biosynthesis. It is suggested that CmcJ complexed with CmcI is responsible for the the C-7 methoxylation of DAC, which is one of the two last steps in CC producton (Öster

et al., 2006). The increased levels of CmcJ can be a result of the upregulation of CcaR since CcaR can bind the promoter regions of several CC gene clusters such as *lat, cefF, cefD–cmcI* to activate the transcription through them (Santamarta et al., 2011).

Carboxyethylarginine synthase (CeaS2) catalyzes the condensation of L-arginine with glyceraldehyde-3-phosphate, which is the first step in the CA and clavam production (Jensen, 2012).

Clavaldehyde dehydrogenase (Car) catalyzes the last step in CA production (Nicholson et al., 1994; Pérez-Redondo et al., 1998).

Carboxyethyl-arginine beta-lactam-synthase (Bls2) catalyzes the second step in CA biosynthesis and introduces monocyclic β-lactam ring into N^2-(2-carboxyethyl)arginine and as such forms the deoxyguanidinoproclavaminate (DGPC) (Bachmann et al. 1998).

Both Ceas2 and Bls2 proteins are grouped under early biosynthetic genes that are required for both the CA and clavam biosynthesis. On the other hand, Car which catalyzes the last step of CA production is not involved in clavam biosynthesis. Since DEPA is an industrial strain that has been subjected to numerous screening and unknown mutagenesis processes and is a prolific producer of CA, upregulation of these proteins was as expected.

Thioredoxin reductase is a member of the family of dimeric flavoenzymes, the function of which is to transfer electrons from pyridine nucleotides to disulfide/dithiol compounds. They are functional as homodimers and each monomer has FAD as prosthetic group, NADPH-binding site and redox-active disulfide site. (Williams et al., 2000; Štefanková et al., 2006). Thioredoxin reductase (SCLAV_5275) protein identified in this study was shown to be located in holomycin biosynthesis gene cluster after the draft genome sequence of *S. clavuligerus* was revealed in 2010, in which the location of holomycin gene cluster could not be predicted accurately on the genome (Li and Walsh, 2010; Huang et al., 2011; Medema et al., 2010). Holomycin belongs to dithiolopyrrolone class of antibiotics which display a broad spectrum of activity and is responsible for interfering with RNA biosynthesis just like other dithiolopyrrolone

antibiotics (Li and Walsh, 2011; Liras, 2014). The presence of holomycin in *S. clavuligerus* culture was first described by Kenig and Reading (1979). Thereafter, de la Fuente (2002) showed that the mutants that are defective in the late steps of CA production such as ΔORF12 and ΔORF15, but not the early steps, produced considerably high levels of holomycin. Afterwards, mutants of *SCLAV 5275* gene, which was later described as either *homI* or *hlmI*, were shown not to be able to produce holomycin (Huang et al., 2011). Moreover, Li and Walsh (2011) also showed that when wt and ΔORF15 strains of *S. clavuligerus* were disrupted in *hlmI* gene, they produced substantially reduced levels of holomycin. The function of HlmI was suggested to be the disulfide bond formation in the reduced holothin structure during the holomycin biosynthesis (Li and Walsh, 2010; Huang et al., 2011). Furthermore, transcriptomic analyses of *ccaR*- and *claR*-deleted mutants suggested the presence of cross-regulation between holomycin and CA biosynthetic pathways (Álvarez-Álvarez et al., 2014; Martínez-Burgo et al., 2015). Given that holomycin and CA were shown to have an inverse relationship in terms of biosynthesis, it is probable that thioredoxin reductase, SCLAV_5275, identified in this study can be also involved in some other pathways aside from holomycin biosynthesis.

Polyprenyl synthetases are involved in isoprenoid biosynthesis pathways. Isoprenoids are natural products showing a high variety in both the structure and biological function (Barkovich and Liao, 2001; Dairi, 2005). For example, steroids in eukaryotes are isoprenoids that are responsible for maintaining membrane fluidity as well as having roles as hormones while carotenoids in photosynthetic organisms are antioxidants; ubiquinone, menaquinone and plastoquinone which are also isoprenoids, on the other hand, are involved in electron transport (Barkovich and Liao, 2001). Some industrially important compounds such as flavors, antibiotics and plant hormones are also comprised of isoprenoids. All isoprenoids are built from five carbon precursor molecules called isopentenyl diphosphates (IPP) (Dairi, 2005). Polyprenyl synthetase enzyme found in DEPA is responsible for catalyzing the condensation reactions between isoprene units (Ashby and Edwards, 1990). Isoprenoid molecules can further undergo cyclizations accompanied with modifications by certain enzymes, producing biologically and functionally diverse compounds. *Actinomycetes* seem to produce

numerous isoprenoid compounds such as as 2-methylisoborneol, geosmin, squalene-hopene (Dairi, 2005).

Phosphotransferase (SCLAV_4596) protein identified in our study belongs to "aminoglycoside/hydroxyurea antibiotic resistance kinase" family according to InterPro entry (http://www.ebi.ac.uk/interpro/entry/IPR006748). Aminoglycoside phosphotransferases (APHs) provide resistance against aminoglycoside antibiotics (Wright and Thompson, 1999). An overexpressed aminoglycoside phosphotransferase (AphD) was also shown by LC-MS/MS analysis of AK39 (Section 3.2.1.5) as well as an underexpressed putative aminoglycoside 2-N-acetyltransferase in 2DE MALDI TOF/MS analysis of DEPA (Section 3.4.2.8) for which we could not bring about any explanation since no clusters in the genome of *S. clavuligerus* encoding enzymes related to aminoglycoside biosynthesis were reported (Medema et al., 2010).

3.4.1.8. Energy Production/Electron-Iron Transfer

Two proteins identified under this category were NADH-quinone oxidoreductase subunit and ferrodoxin-nitrite reductase.

Proton-translocating NADH-quinone oxidoreductase protein is a very complex energy transferring complex consisting of at least 6 iron-sulfur clusters and is found in both prokaryotes and eukaryotes (Nakamaru-Ogiso et al., 2005). There were two protein spots located on the gel representing this protein. Thus, it is possible that this protein might have undergone a PTM with a charge modification.

Ferrodoxin-nitrite reductase catalyzes the reduction of nitrite to ammonia (Ramirez et al., 1966).

3.7.1.9. Cell Processes (Shape/Division/Motility)

There were two proteins categorized in this category: CHAP domain-containing protein and exopolysaccharide phosphotransferase.

CHAP domain containing proteins were independently discovered by Bateman and Rawlings (2003) and Rigden et al. (2003). Proteins possesing CHAP domains also contain amidase domains which are associated with peptidoglycan hydrolytic

properties. For example, autolysin LytA from *Staphylococcus aureus* and endolysin from *S. aureus* bacteriophage Twort contain CHAP domains and were shown to have hydrolytic activities towards peptidoglycan layers (Bateman and Rawlings, 2003; Rigden et al., 2003).

Exopolysaccharide (EPS) phosphotransferase is involved in exopolysaccharide biosynthesis. EPSs produced by bacteria are long-chain polysaccharides which are comprised of sugar or sugar derivatives showing a repeating pattern (Sun et al., 2007) and excreted out of the cells (Wang et al., 2003). The functions of them vary considerably depending on the species, including the protection of the microorganism against immune system recognition or increasing the antibacterial tolerance (Wang et al., 2003; Vuong et al., 2004). EPSs were also suggested to have some clinical importance due to showing anti-tumorigenic, immunostimulatory or blood cholesterol-lowering properties. Although EPSs are a major component of the biofilms produced by biofilm-producing bacteria and were investigated widely in Gram-negative bacteria, the function of EPSs or the biosynthesis and the modification of them in Gram-positive bacteria are largely unknown (Vuong et al., 2004). In the UniProtKB database, exopolysaccharide phosphotransferase (SCLAV_5352) is suggested to be a stealth protein involved in the biosynthesis of polysaccharides and imparts the microorganism evasion ability from host immune system (Sperisen et al., 2005).

3.4.1.10. Inorganic/Organic Molecule Transport

Protein translocase subunit SecA and secreted protein (SCLAV_2084) were the two proteins categorized in this group.

Sec pathway is responsible for the translocation of the unfolded proteins across the membrane or the insertion of them into the cell membrane layer. Sec translocase system is composed of four proteins (SecY, SecE, SecG, SecA), Of these 4 subunits, SecY, SecE and SecG form the protein conducting channel (PCC) while SecA is a peripheral ATPase protein that generates the required energy to translocate the unfolded proteins through the channel (Natale et al., 2008). The secretion of certain

proteins might have increased in industrial strain or there might be proteins that are incorporated into the membrane layer of the organism through this pathway. The possible importance of Sec pathway for DEPA strain remains to be further searched in order to understand what kind of advantage it might provide the industrial strain.

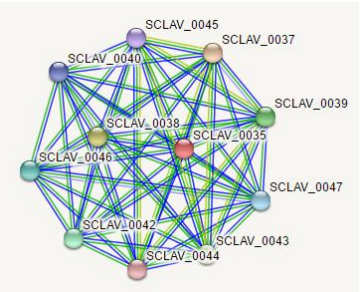
Secreted protein (SCLAV_2084) was denoted to belong to "solute-binding family 1" by UniProtKB database. Solute-binding proteins are part of the ABC transporters (Singh and Röhm, 2008) which are the largest family of transporters and found in all kingdoms of life. They can be either exporters of toxins, drugs and lipids across membranes, etc. or importers of a large variety of molecules, especially nutrients (Rees et al., 2009). ABC importers require a specific solute-binding protein in their constitution so as to selectively bind the solutes to be transported across the membrane. The solute binding proteins are found as membrane-associated lipoproteins in both archaea and Gram-positive bacteria while they are found as dissolved proteins in the periplasm of Gram-negative bacteria (Singh and Röhm, 2008). STRING database also shows its association with ABC transporters (data not shown). However, the specificity of this protein regarding its substrate remains to be elucidiated.

3.4.1.11. Hypothetical/Unknown Proteins

There were 27 proteins upregulated in this category: Of these, putative acetyltransferase (SCLAV_0254) and putative uncharacterized protein (SCLAV_5138) were represented by two spots showing a charge modification while β-lactamase domain protein (SCLAV_p1007) was represented by as many as 12 spots indicating both charge and mass modifications. Thus, it is very likely that the latter protein is subjected to extensive PTMs. Moreover, high number of proteins with unknown functions in DEPA could point to a kind of unusual gene expression in response to its modified cellular activities with many mysteries.

When the associations of putative uncharacterized protein (SCLAV_0035) were scrutinized through STRING, it revealed a very complex network with a diverse variety of proteins (Figure 3.23). The fact that it has functional partners like phage region and Vgr proteins suggests a relation of this protein with contact-dependent Type

VI secretion system (T6SS) which is known to function in resisting predation, pathogenesis, stress sensing, regulating bacteria-bacteria interactions, targeting other bacterial cells, therefore help in competition towards a specific niche in different bacteria (Schwarz et al., 2010; Costa et al., 2015). In Figure 3.32, this protein is shown to be interacting/associated with the putative phage tail sheath protein (SCLAV_0037) which was upregulated in AK39 and suggested to be involved in T6SS (Section 3.2.1.9) and Rhs element Vgr protein (SCLAV_0043) which will be explained in the following section (Section 3.4.1.12).



Your Input:

SCLAV_0035 *Putative uncharacterized protein (192 aa)*							

Predicted Functional Partners:

	Neighborhood	Gene Fusion	Cooccurence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
SCLAV_0037 *Phage tail sheath protein, putative (754 aa)*	•		•						0.962
SCLAV_0038 *Phage tail region protein (151 aa)*	•		•						0.940
SCLAV_0043 *Rhs element Vgr protein (393 aa)*	•		•						0.933
SCLAV_0039 *Putative phage tail sheath protein (500 aa)*	•		•						0.932
SCLAV_0042 *Peptidoglycan-binding LysM (223 aa)*	•		•						0.932
SCLAV_0046 *Putative uncharacterized protein (648 aa)*	•		•						0.931
SCLAV_0047 *Putative uncharacterized protein (684 aa)*	•		•						0.930
SCLAV_0040 *Phage tail region protein (172 aa)*	•		•						0.927
SCLAV_0045 *Putative GPW/gp25 family protein (134 aa)*	•		•						0.918
SCLAV_0044 *Rhs element Vgr protein (234 aa)*	•		•						0.916

Figure 3. 23. STRING data showing the associaton of putative uncharacterized protein (SCLAV_0035) with other proteins.

When the associations of putative uncharacterized protein (SCLAV_5138) were searched through STRING, it gave a simple scheme of network (data not shown).

β-lactamase domain protein was represented by many spots in a clustered pattern on 2DE gel, which suggested that it has undergone an extensive posttranslational

modification of both “mass” and “charge-type”. As mentioned in Chapter 1, genome sequencing of *S. clavuligerus* ATCC 27064 revealed a giant linear plasmid of 1.84 Mb in length named pSCL4 (Medema et al., 2010). While it contains no protein-coding genes essential for primary metabolism, the plasmid harbors an exceptionally large number of gene clusters for secondary metabolite production, including a large number of putative antibiotics, such as staurosporine, moenomycin, β-lactams, and enediynes. Putative non-ribosomal peptide synthetase (NRPS) modules of secondary metabolism fused to β-lactamase domain (PFAM number PF00144) were also remarkable. No significant resemblance of this domain to any experimentally studied protein could be detected; therefore, the authors could only speculate about its function as a transpeptidase or a protein that binds to, but not hydrolyzes, a β-lactam compound that is then attached to the peptide synthesized by the NRPS. Generally, these domains are highly important in β-lactam producing organisms since they can provide resistance to its own antibiotics of *S. clavuligerus.* A total of 22 proteins with a predicted β-lactamase domain were detected on the chromosome or megaplasmid of the organism. Although *S. clavuligerus* β-lactamases have not yet been experimentally characterized, posttranslational modifications like lysine carboxylation in Class D ones (Li et al., 2005b; Che et al., 2014) as well as proteolytic modification of zinc-based or metallo β-lactamases belonging to class B which are not inhibited by clavulanic acid were reported (Koh et al., 2008). UniProtKB database suggests that this protein has an alkyl sulfatase domain and a metallo-β-lactamase domain. Alkylsulfatases is a sub-group of metallo-β-lactamases and also grouped under Type III sulfatases. They catalyze the degradation of sulfate esters. To date, there are only three type III sulfatases reported. One is SdsA1 from *Pseudomonas aeruginosa*, the other Pisa1 from *Pseudomonas* sp. Another one recently identified is YjcS from *E. coli*. SdsA1 and YjcS were shown to degrade SDS which is toxic to the cells (Hagelueken et al. 2006; Liang et al., 2014), while Pisa hydrolyzes the secondary alkyl sulfates (Knaus et al., 2012). Furthermore, according to protein blast results in NCBI database, β-lactamase identified in this study showed 98% sequence coverage and 42% identity to the YjcS protein of *E. coli*, while it showed 87% sequence coverage and 44% identity to the SdsA1 protein of *P. aeruginosa.* There arise two presumptions given the importance of β-lactamases in

Streptomyces. One is that this protein is a metallo-β-lactamase and its predicted interaction with thioredoxin might suggest the involvement of this alkylsulfatase protein in stress related conditions. The second is that this metallo β-lactamase protein with some isoforms hundreds-fold overexpressed favors the CA production in the industrial strain in a mechanism as yet to be revelaed.

The association network of putative phosphoesterase (SCLAV_0101) was searched through STRING database (data not shown), most of the proteins predicted to be associated or interacting with putative phosphoesterase are unknown. One of the proteins that could be predicted is Nudix hydrolase. These proteins are commonly found in organisms and in *Actinomycetes*, some *Streptomyces* sp. genes encoding them can vary in number from 20 to 90. They are mainly responsible for the hydrolysis of such substrates as nucleoside di- and triphosphates, dinucleoside and diphosphoinositol polyphosphates, nucleotide sugars (McLennan, 2006). However, the function of this protein and its interacting partners need to be searhed further.

Another protein in this category is putative acetyltransferase (SCLAV_0254) which was represented by two spots indicating a charge modification on the protein. According to STRING data, some of the enzymes that this protein was predicted to be associated or interacts with are responsible for the biosynthesis of N-acetylglucosamine which is one of the precursors of the peptidoglycan cell wall in bacteria (data not shown). For example, L-glutamine: D-fructose-6-phosphate amidotransferase also known as glucosamine-6-phosphate synthetase catalyzes conversion of fructose-6-phosphate to glucosamine-6-phosphate. Thereafter, glucosamine-6-phosphate is converted to UDP-N-acetylglucosamine through sequential reactions catalyzed by several enzymes such as phosphoglucosamine mutase and GlcN-1P acetyltransferase (Chmara et al., 1984; Durand et al., 2008). Moreover, another protein predicted to be associated with putative acetyltransferase is phosphomannomutase that is involved in the synthesis of GDP-mannose, the biosynthesis of which also requires fructose-6-phosphate as precursor. GDP-mannose can be used as an intermediate to derive sugar residues from it so that they are added to several compounds (Lochlainn and Caffrey, 2009) such as bleomycin which is an

anticancer agent (Du et al., 2000), neocarzinosatin (Liu et al., 2005) and hygromycin A antibiotic (Palaniappan et al., 2006). These sugar derivatives coming from GDP-mannose can be important in modifying the activity of a molecule within the cell. Moreover, it was suggested that GDP-mannose might also be important in the biosynthesis of glycoconjugate that is found in spore walls of *S. nodosus* (Lochlainn and Caffrey, 2009). Given the association network prediction, we can hypothesize that this protein might be related with the biosynthesis of peptidoglycan cell layer in *S. clavuligerus*. However, the function of this protein and the importance of probable PTM that it undergoes need to be further characterized.

Putative uncharacterized protein (SCLAV_4757) can be a regulatory protein or might be associated with them given the association network of it with other proteins in STRING database (data not shown). There are 3 regulatory proteins and a geranylgeranyl pyrophosphate synthase protein that it is associated with. Geranylgeranyl pyrophosphate synthase is involved in isoprenoid biosynthesis.

Putative uncharacterized protein (SCLAV_4755) was also searched through STRING database for association prediction (data not shown). This protein was predicted to be associated with 2 other proteins one of which is a geranylgeranyl pyrophosphate synthase.

Putative uncharacterized protein (SCLAV_3217) is also searhed through STRING database (data not shown). It was associated with an ent-copalyl diphosphate synthase protein that is responsible for the cyclization of several isoprenoid products, a process contributing to the diversity of the isoprenoids (Ikeda et al., 2007). Another protein is prenyltransferase and squalene oxidase repeat protein. Prenyltransferase/squalene oxidases were suggested to be involved in secondary metabolism (Zhou et al., 2013).

Another protein under this category is putative MaoC-like dehydratase. MaoC-like dehydratases show enoyl-CoA hydratase activity and are responsible for channeling the PHA (polyhydroxyalkanoates) precursors from the β-oxidation pathway to PHA biosynthetic pathway; in other words, they are important in connecting β-oxidation and PHA pathways (Park and Lee, 2003). PHAs are used as carbon- and energy-

storage material in various bacteria and composed of polyesters of hydroxycarboxylic acids (Wang et al., 2009)

There were two proteins contaning DUF in this category: One of them is DUF124 domain-containing protein. According to STRING data, this protein forms a complex network of association with other proteins; however, none of the proteins in this network is known (data not shown).

Second proten containing DUF domain is DUF946 domain-containing protein.This protein was searched through EggNog database in order to find the orthologues and their predicted function. It was shown to be an unknown protein and described as plant protein of unknown function and found in 4 proteins in 3 species of bacterial kingdom. Furthermore, in UniProtKB database, it was described as vacuolar protein sorting-associated protein 62 responsible for transporting proteins to vacuoles (Bonangelino et al., 2002). Since this protein was also predicted to be located in cytoplasmic membrane, it might be involved in the transport of the proteins through the membrane.

STRING data shows that Sporozoite_P67 domain-containing protein is associated with just one protein of unknown function (data not shown) EggNog database was searched to find the orthologues of this protein in bacterial kingdom and the query returned 24 proteins from 24 species with unknown functions.

Cell surface protein (SCLAV_5606) was found to contain WD40/YVTN repeat-like-containing domain in InterPro database directed from UniProtKB. The sequence of the protein was also searched through EggNog database to find out the orthologues of the protein and their predicted functions. In bacterial kindom, the orthologues were shown to contain 40-residue yvtn family beta-propeller repeat protein and more specifically, the orthologues in *Actinobacteria* were shown to contain WD-40 repeat-containing protein, indulged in transcription. In eukaryotes, proteins with WD-40 motifs are usually involved in protein-protein interactions and take part in several cellular processes like signaling, cell cycle control, and cytoskeleton assembly. These motifs are rarely observed in prokaryotes and mostly found in actinobacteria (especially actinomycetes), cyanobacteria, and proteobacteria (Charusanti et al., 2012). YVTN-

type repeat domain, on the other hand, are found in surface layer proteins of archaeobacteria to protect the cell against extreme environmental conditions (Jing et al., 2002).

Another protein in this category was putative polyprenyl diphosphate synthase. According to UniProtKB database, this protein belongs to polyprenyl synthetase family. As explained in Section 3.7.1.2, polyprenyl synthetases are involved in the biosynthesis of isoprenoids which show great deal of structural and functional diversity (Barkovich and Liao, 2001; Dairi, 2005).

Peptidoglycan-binding domain 1 protein upregulated is encoded by SCLAV_p1144 which is pSCL4-borne. Peptidogycan binding domain proteins are usually found in enzymes that have cell wall degrading properties. For example, phage endolysins usually consist of a C-terminal cell wall binding domain and an N-terminal catalytic domain (Loessner, 2005; Briers et al., 2007). KZ144 and EL188 proteins are lytic peptidoglycan hydrolases found in *Pseudomonas aeruginosa* bacteriophage with their N-terminal containing a substrate binding domain (Briers et al., 2007). Moreover, many lytic cell wall enzymes found in Gram-positive bacteria contain ligand binding domains at or near their C- or N-terminal domains. Muramoyl pentapeptide carboxypeptidase from *S. albus*; for instance, hydrolyze the interpeptide bridges in peptidoglycan cell wall structure and has an N-terminal binding domain (Ghuysen et al., 1994). On the other hand, *Aeromonas hydrophila* has a ExeAB inner membrane complex to transport protein toxins out of the cells. ExeA protein has a peptidoglycan binding domain at its C-terminal which comprise the periplasmic domain of the protein (Li and Howard, 2010).

3.4.1.12. Others/General Function

There were 14 distinct proteins categorized under this group. Aminotransferase and DNA binding proteins were shown to be upregulated by both MALDI-TOF/MS and LC-MS/MS analyses.

UniProtKB database suggests that DNA binding domain protein (SCLAV_0130) has a lambda repressor like domain. Its association network was queried through SRING

database (data not shown). It was shown to be associated with several unknown proteins as well as a regulator protein. Furthermore, 28 fold increase of this protein in DEPA according to MALDI-TOF/MS results suggests that this protein might have a crucial role in the CA biosynthesis.

One of the proteins identified in this category is Rhs element Vgr protein (SCLAV_0043). Rhs elements are accessory repetitious sequences found in *E. coli* that show high structural and distributional variety (Wang et al., 1998). Their GC contents are quite different than that of *E. coli* (Zhao and Hill, 1995). Rhs elements are the major source of the chromosomal rearrangements in laboratory cultures. Eight distinct Rhs elements have been described in *E. coli*, all containing open reading frames with unclear functions. Rhs elements are composed of two basic components: Rhs core and core-extension which together form the Rhs-core-ORF. Core regions are approximately 3.7 kb while core extensions are about 0.4-0.6 kb. Core regions show homology with core regions of other Rhs elements but the core extensions show a great deal of variety. There is also dsORF (downstream ORF) which either overlaps or follows closely the core extensions. dsORFs are unique sequences and usually N-terminal of them are predicted to contain a signal peptide for transport across the membrane. There is another component called *vgr* (named for Val-Gly dipeptide repetition) that is found only in *RhsE* and *RhsG* elements in *E. coli*. *vgr* genes in *E. coli* were found to be located upstream of the core regions (Hill, 1999) (Figure 3.24).

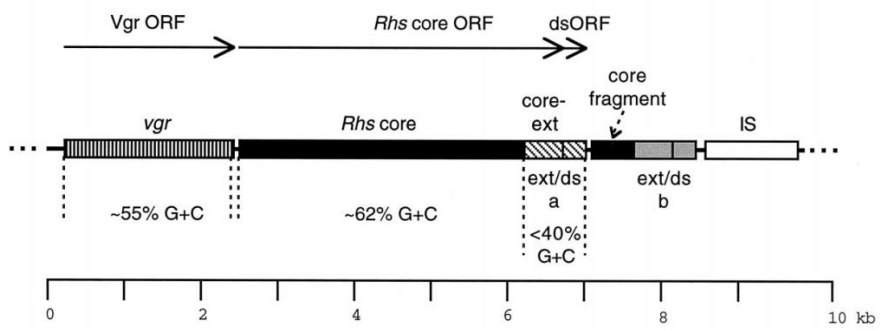


Figure 3. 24. A prototypical Rhs element showing its components and organization. The most conserved Rhs component is the Rhs core; the Rhs core, when joined to a core-extension, comprises the Rhs core-ORF. In some cases, an additional core-extension linked to a 3'-core fragment is found elsewhere in the element. A segment encoding a dsORF is linked to each core-extension; DNA segments coding both a core-extension and the adjacent dsORF are designated ext/ds. The vgr component is associated with RhsE and RhsG elements, but is absent from the others. Rhs elements generally contain one or more insertion sequences, although the ISs may be defective. The G+C contents of the major components are shown (Hill, 1999).

Although the function of Vgr proteins and Rhs elements have not been described, they are predicted to code for large, hydrophilic proteins that may serve as ligand-binding proteins (Hill et al., 1994). Wilderman et al. (2001) further suggests that *vgr* genes might be involved in horizontal gene transfer or more precisely tranfer of a foreign DNA into the cell given that *vgr* genes described so far are usually located close to the genes that might have been horizontally transferred from other species. In order to find out what kind of associations Vgr protein that was described in our results can have with other proteins, amino acid sequence of the protein was searched through STRING database (Figure 3.25).

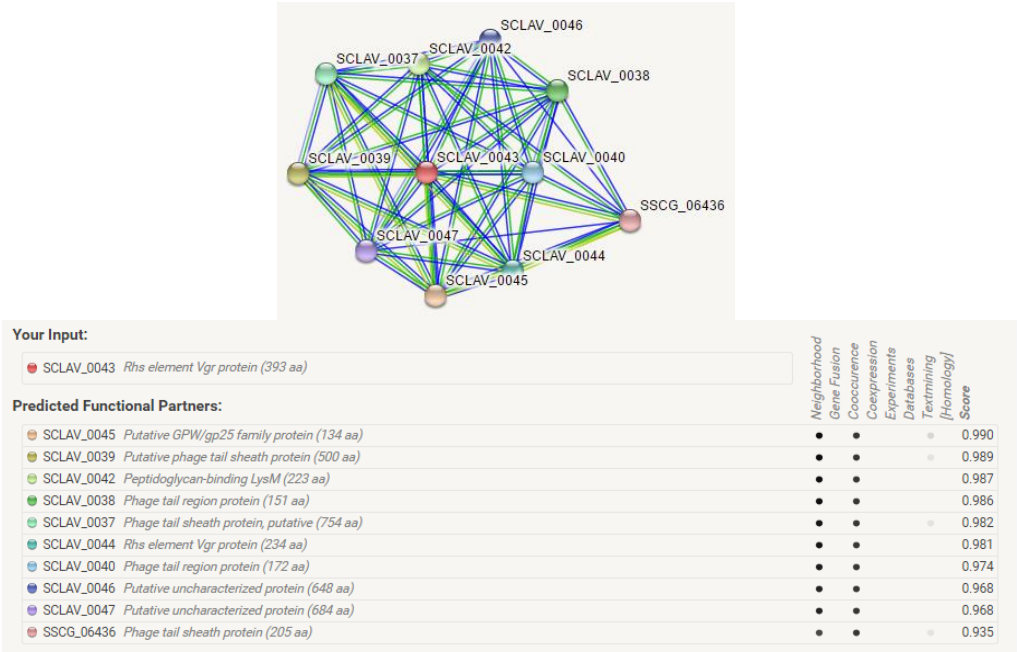


Figure 3. 25. STRING data showing the associaton of Rhs element Vgr protein with other proteins.

STRING data suggest that Vgr protein is related with the phage tail proteins as well as another Vgr protein. In Section, 3.4.1.11, putative uncharacterized protein (SCLAV_0035) and in Section 3.2.1.9, putative phage tail sheath protein (SCLAV_0037) upregulated in AK39 were also shown to associate with this protein (SCLAV_0043) (Figure 3. 25), strengthening the assumption of the presence of T6SS in this organism. Indeed, *S. clavuligerus* has been known to possess mainly three different types of general protein secretion systems: Sec pathway, Esx secretion system (T7SS) and the Tat pathway (Chater et al., 2010). In this respect, an evidence for T6SS in *S. clavuligerus* has been reported for the first time in the present study. Considering the functions of this secretion system other than pathogenesis, overexpression of this protein might confer advantages like better stress sensing, bacteria-bacteria interactions and competitiveness to the industrial mutant. However the function of both Rhs element Vgr protein (SCLAV_0043) and putative

uncharacterized protein (SCLAV_0035) require an elaborate study to find out the main functions and importance of these proteins in *S. clavuligerus* DEPA strain.

Siderophores are iron chelators produced by the microorganisms. They are exported into the environment in which they bind to ferric iron strongly forming ferric-siderophore complexes and then are transported back into the cell by active transport mechanisms (Patel et al., 2010). Though siderophore-mediated-iron transfer is not the sole mechanism to transport iron into the cells, it is the most commonly used one that microorganisms has adapted to supply iron from the environment (Tu et al., 2014). The ability to acquire iron under low-iron conditions is related to the virulence of a variety of bacterial pathogens. SIP (siderophore-interacting protein) or FSR (ferric siderophore reductase) superfamily proteins are the major families that are observed in bacteria that have siderophore gene clusters and both families are responsible to reduce the iron-siderophore complexes to release iron (Li et al., 2015). ViuB in *V. cholerae* was the first SIP protein described and is important for iron utilization. Expression of it depends on *fur* gene. Another SIP protein described was YqjH from *E. coli* and responsible for catalyzing the release of iron from several iron-chelators. On the other hand, SIP protein from *R. anatipestifer* strain CH3 is important for iron acquisition as well as the virulence of the organism (Tu et al., 2014). Besides the classical functions of iron acquisition and virulence, siderophores do also have non-classical biological functions like non-iron metal transport, toxic metal sequestration, protection from oxidative stress and molecular signaling (Johnstone and Nolan, 2015).

As a putative periplasmic protein, the crystal structure of YceI family of proteins suggests that they play an important role in isoprenoid quinone metabolism and/or transport and/or storage. Though the functions of YceI family proteins vary considerably depending on the microorganism (respiratory electron transport, controlling oxidative stress and gene regulation), these proteins are known to be involved in transport and/or storage of amphipathic compounds under osmotic stress and fatty acid or amide sequestration related to the acid stress response (Handa et al.,2005; Zammit et al., 2012). YceI proteins are overexpressed under acidic stress conditions, together with other essential virulence factors. For instance, *Helicobacter pylori* acidic stress response factor HP1286 is a YceI homolog (Sisinni et al, 2010).

Fork-head associated (FHA) domain term was first coined by Hofmann and Bucher in 1995 (Hofmann and Bucher, 1995). FHA domains contain conserved sequences comprised of 65-100 amino acid residues. They are mostly found in eukaryotic proteins but were also shown to prevail in those of prokaryotes (Durocher et al., 1999). FHA modules recognize the phosphoprotein peptides, usually phosphothreonine peptides (Pallen et al., 2002). Most of the FHA domain containing proteins in eukaryotes are found in the nucleus and are associated with transcription, DNA repair, cell cycle regulation and chromosome segregation; furthermore, they can also be involved in processes such as protein degradation, signal transduction and vesicular transport (Durocher et al., 1999; Pallen et al., 2002). In *Mycobacterium tuberculosis*, two proteins EmbR and Rv1747 were found to contain FHA domain and interact with eukaryotic-like serine/threonine protein kinases (STPKs) which are are important players for the signaling cascades through phosphorylation reactions in the cell. Moreover, the fact that several STPKs were shown to be interacting with several other FHA domain containing proteins in *M. tuberculosis* and that FHA-domain containing proteins might be direct targets of the kinases suggested the the importance of the interplay between these proteins and the complex nature of the signaling pathways in the organism (Molle et al., 2003; Molle et al., 2004;. Grundner et al., 2005). Protein identified in this study was searched through STRING database (data not shown) in order to find out possible interactions or associations of it with other proteins. It was predicted to interact/associate with two serine/threonine protein kinases and a regulatory protein, suggesting its probable relation with transcriptional regulation and/or certain signaling pathways. However, its mechanism of function is to be searched further.

3.4.2. Downregulated Proteins in DEPA Strain

Of the identified proteins, 60 proteins were found to be significantly downregulated in MALDI-TOF/MS analysis while 120 proteins were downregulated in LC-MS/MS analysis (Table 3. 11 and Table 3. 12, respectively). The spots with more than 2.5-fold decreased intensity as compared to their counterparts in NRRL 3585 are shown in

Figure 3. 26. For this, NRLL3585 gel was used for a much better visualization of differentially expressed proteins.

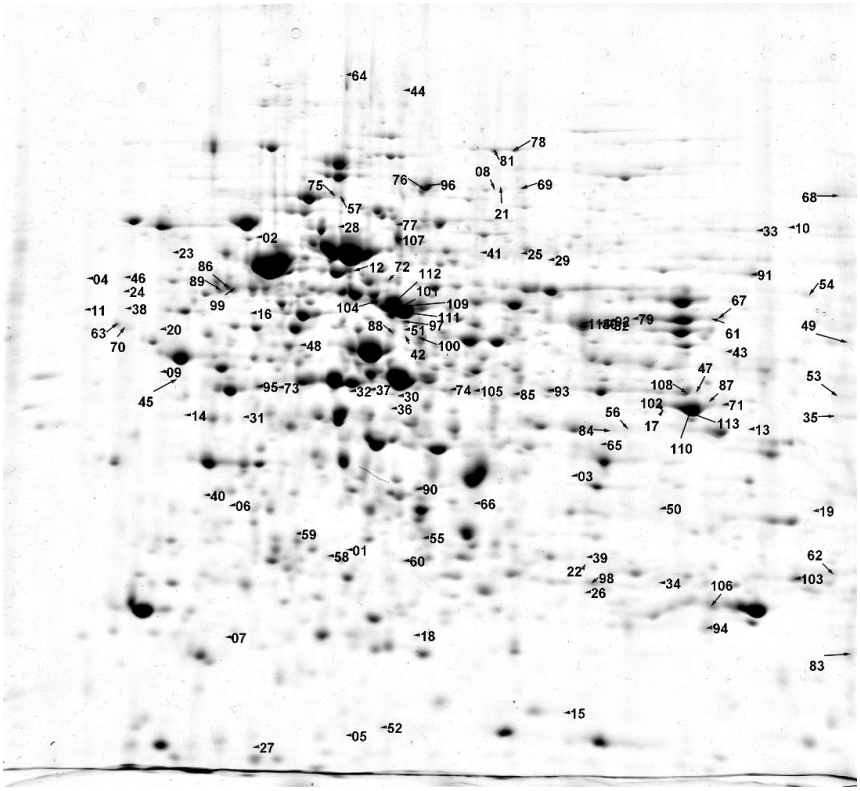


Figure 3. 26.Total protein spots on NRRL 3585 2DE gel that showed more than 2.5-fold more intensity compared to their counterparts on DEPA gel.

Table 3. 11. List of downregulated proteins identified on DEPA 2DE gel.

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spot	PTM *	Genome Location **	Spot #
Nucleotide Metabolism (DNA/RNA Processing)	E2Q7J2	Purine nucleoside phosphorylase	*SCLAV_2325*	Cytoplasm	-14.1	5.25	30,470	–	–	Chr.	90
	E2PYR8	5'-nucleotidase	*SCLAV_3178*	Cell membrane; Cytoplasm	-3.5	7.00	64,328	–	–	Chr.	91
	E2Q359	Inositol-5-monophosphate dehydrogenase	*SCLAV_3702*	Cytoplasm	-2.9	5.58	40,012	–	–	Chr.	93
Carbohydrate Metabolism	E2PYV9	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	*gpmA*	Cytoplasm	-4.7	8.42	32,180	–	–	Chr.	22
	B5GYJ1	Phosophoglycerate kinase	*pgk*	Cytoplasm	-4.9	5.06	41,926	–	–	Chr.	32
	E2PVD8	Putative isocitrate dehydrogenase	*icdA*	Cytoplasm	-3.2	4.97	79,048	–	–	Chr.	57
	B5GYK0	Oxppcycle protein	*opcA*	Cytoplasmic membrane	-3.9	5.39	37,739	–	–	Chr.	105
Amino Acid Metabolism	B5GZC5	Cystathionine gamma-synthase	*SCLAV_5668*	Cytoplasm	-83.3 (a); -3.5 (b); -4.3 (c); -8.3 (d); -16.9 (e)	6.04	41,497	5 (a,b,c,d,e)	C-M	Chr.	17,71 ,102, 110, 113
	B5GX10	S-adenosylmethionine synthetase	*metK*	Cytoplasm	-2.6	4.92	43,523	–	–	Chr.	48
	D5SKH4	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	*metE*	Cytoplasm	-2.9 (a); -2.6 (b)	5.46	88,844	2 (a, b)	M	pSCL4	76, 96

Table 3. 11. List of downregulated proteins identified on DEPA 2DE gel. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spot	PTM *	Genome Location **	Spot #
Amino Acid Metabolism	D5SLL7	Methionine synthase II	*SCLAV_p1324*	Cytoplasm	-4.2	4.86	38,157	–	–	pSCL4	95
	E2Q1K0	O-acetylhomoserine aminocarboxypropyltransferase	*SCLAV_5559*	Cytoplasm	-6.5	5.81	53,426	–	–	Chr.	20
	E2PZN1	Histidinol-phosphate aminotransferase	*hisC*	Cytoplasm	-5.7	5.23	41,009	–	–	Chr.	30
	E2Q7M7	Pyrroline-5-carboxylate reductase	*proC*	Cytoplasm	-4.1	5.85	29,718	–	–	Chr.	34
	B5GWZ1	3-dehydroquinate synthase	*aroB*	Cytoplasm	-2.9	5.38	38,227	–	–	Chr.	74
	B5GQA8	3-phosphoshikimate 1-carboxyvinyltransferase	*aroA*	Cytoplasm	-4.4	5.25	46,003	–	–	Chr.	51
	E2PXE0	Alanine dehydrogenase	*SCLAV_0984*	Cytoplasm	-2.6 (a); -3.4 (b)	5.76	41,935	2 (a, b)	C	Chr.	47, 108
	B5GSL5	4-hydroxyphenylpyruvate dioxygenase	*SCLAV_2046*	Cytoplasm	-35.7 (a); -4.9 (b)	5.11	41,873	2 (a, b)	C	Chr.	37, 32
	E2PZM5	Imidazole glycerol phosphate synthase subunit HisF	*hisF*	Cytoplasm	-32.3	5.21	26,809	–	–	Chr.	60
Lipid Metabolism	B5H0X6	Enoyl-[acyl-carrier-protein] reductase [NADH]	*fabI*	Cytoplasm	-4.5	5.47	27,606	–	–	Chr.	39
	B5H0X7	3-oxacyl-(ACP) reductase	*SCLAV_1028*	Cytoplasm	-3.2	6.17	24,657	–	–	Chr.	62
	E2PXB5	Acyl CoA dehydrogenase	*SCLAV_0960*	Cytoplasm	-3.5	6.15	42,936	–	–	Chr.	71
	E2PVS8	Acetyl-CoA acetyltransferase	*SCLAV_4845*	Cytoplasm	-2.9	5.30	43,27	–	–	Chr.	74

Table 3. 11. List of downregulated proteins identified on DEPA 2DE gel. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spot	PTM *	Genome Location **	Spot #
DNA Replication, Recombination, Repair, Transcripton	E2Q5D4	Two-component system response regulator	*SCLAV_2102*	Cytoplasm	-4.5	7.81	28,723	–	–	Chr.	98
	B5GPA7	TetR-family transcriptional regulator	*SCLAV_3146*	Cytoplasm	-2.8	5.39	28,388	–	–	Chr.	55
Translation, Ribosomal Structures	B5GXK8	30S ribosomal protein S1	*rpsA*	Cytoplasm	-4.8	4.53	54,731	–	–	Chr.	23
	B5GX54	Elongation factor Tu	*tuf2*	Cytoplasm	-2.6	5.07	43,624	–	–	Chr.	48
Stress-Related, Protein Turnover, Chaperones	B5GTB1	Chaperone protein DnaK	*dnaK*	Cytoplasm	-4.4	4.79	66,533	–	–	Chr.	2
	E2Q3P1	Trigger factor	*tig*	Cytoplasm	-37	4.50	55,724	–	–	Chr.	46
Secondary Metabolism	B5GLB5	Deacetoxycephalosporin C hydroxylase	*cefF*	Cytoplasm	-7.6	4.79	34,599	–	–	Chr.	31
	B5GLC8	Proclavaminate amidinohydrolase	*pah2*	Cytoplasm	-4.8	5.69	33,401	–	–	Chr.	65
	D5SLG8	Moenomycin biosynthesis protein MoeA5	*moeA5*	Cytoplasm	-4.7	6.05	43,072	–	–	pSCL4	87
	E2Q9C7	Gamma-butyrolactone biosynthesis protein	*avaA2*	Cytoplasm	-2.6	6.05	38,935	–	–	Chr.	47
	E2Q7Z9	Putative aminoglycoside 2-N-acetyltransferase	*aac2*	Cytoplasm	-22.7	5.20	19,65	–	–	Chr.	52
Cell processes (Shape/Division/ Motility)	B5GXQ0	Cell division protein FtsZ	*ftsZ*	Cytoplasmic membrane	-2.7	4.48	41,218	–	–	Chr.	9

Table 3. 11. List of downregulated proteins identified on DEPA 2DE gel. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spot	PTM *	Genome Location **	Spot #
Cell processes (Shape/Division/ Motility)	D5SI70	Teichoic acid biosynthesis protein	*tagC*	Unknown	-83.3	5.98	39,328	–	–	pSCL4	17
	B5GXP3	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase	*SCLAV_1299*	Cytoplasm; Cell membrane	-3.2	5.23	50,07	–	–	Chr.	97
	B5GXQ2	Alanine racemase domain protein	*SCLAV_1290*	Unknown	-4.7	5.68	26,067	–	–	Chr.	22
Hypothetical/ Unknown Proteins	E2Q9V4	Putative peptide hydrolase	*SCLAV_2609*	Cytoplasm	-4.4	4.80	65,285	–	–	Chr.	2
	E2PWQ8	DUF2587 domain-containing protein	*SCLAV_2986*	Cytoplasm	-2.8	5.11	23,192	–	–	Chr.	7
	E2Q2L3	Putative epimerase/dehydratase	*SCLAV_5657*	Cytoplasm	-3.4	5.83	34,785	–	–	Chr.	13
	B5H3Q2	Putative sugar hydrolase	*SCLAV_1830*	Cytoplasm	-3	5.21	44,795	–	–	Chr.	42
	E2PUJ1	Putative amino acid decarboxylase	*SCLAV_2698*	Cytoplasm	-7.4	6.08	48,833	–	–	Chr.	43
	B5GLV6	DUF1254 multi-domain protein	*SCLAV_p1482*	Cytoplasmic membrane	nd*	4.40	48,605	–	–	pSCL4	63
	E2PVS7	Putative fatty acid oxidation complex alpha-subunit	*SCLAV_4844*	Cytoplasm	-71.4	5.13	74,916	–	–	Chr.	77
	D5SLL2	Putative transcriptional regulator AraC family	*SCLAV_p1319*	Unknown	-2.7 (a); -13 (b)	8.55	25,424	2 (a, b)	M	pSCL4	106, 94
	E2PWW5	Putative M28-family peptidase	*SCLAV_3043*	Extracellular	-2.7 (a); -6.1 (b)	6.85	52,975	2 (a, b)	C	Chr.	61,67
	E2Q0Y5	Putative alpha-glucosidase	*SCLAV_1450*	Cytoplasm	-3.1 (a); -5.2 (b); -3.1 (c)	4.78	60,086	3 (a, b, c)	C	Chr.	89, 99, 86

Table 3. 11. List of downregulated proteins identified on DEPA 2DE gel. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spot	PTM *	Genome Location **	Spot #
Hypothetical/ Unknown Proteins	E2Q926	Putative aminotransferase	*SCLAV_4568*	Cytoplasm	-3	5.19	48,212	–	–	Chr.	42
Others/General Function	B5H0V5	Phospholipid-binding protein	*SCLAV_1005*	Cell membrane; Extracellular	-22.7	5.19	18,687	–	–	Chr.	52
	E2PYK8	UPF0042 nucleotide-binding protein	*SCLAV_1152*	Cytoplasm	-3.1	5.09	34,157	–	–	Chr.	66
	B5GZE3	Bi-domain oxidoreductase	*SCLAV_5650*	Cytoplasm	-13.2	6.86	76,903	–	–	Chr.	68
	E2Q928	Aldehyde dehydrogenase	*SCLAV_4570*	Cytoplasm	-3.3	5.10	51,754	–	–	Chr.	72
	B5GPD0	Aldehyde dehydrogenase	*SCLAV_2425*	Cytoplasm	-2.7	5.07	52,052	–	–	Chr.	104
	B5GZB8	Aldehyde dehydrogenase	*aldH*	Cytoplasm	-4.4	5.16	55,339	–	–	Chr.	111
	B5GLD8	Acetyltransferase GNAT family protein	*SCLAV_4185*	Cytoplasm	-14.1 (a); -2.8 (b)	5.70	36,293	2 (a, b)	C	Chr.	56,84
	D5SJT7	Bacterial luciferase domain-containing protein	*SCLAV_p0693*	Unknown	-14.1	5.22	31,646	–	–	pSCL4	90
	B5GXB2	Single-stranded DNA-binding protein	*SCLAV_2881*	Cytoplasm	-4.2	5.29	20,039	–	–	Chr.	18

Table 3. 11. List of downregulated proteins identified on DEPA 2DE gel. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Expression (Fold)	pI	Mw (Da)	Number of Multi-spot	PTM *	Genome Location **	Spot #
Others/General Function	B5GZC6	Amidohydrolase	*SCLAV_5667*	Cytoplasm	(a-f); -11.4, -7.2, -5.2, -5.9, -4.5, -5.3	5.64	45,975	6 (a-f)	C	Chr.	16, 38, 80, 82, 92, 114
	B5GTG6	Oxygenase	*SCLAV_2696*	Cytoplasmic membrane	-5.3	5.18	35,198	−	−	Chr.	36

*PTM refers to possible posttranslational modification. C denotes charge modification, M denotes mass modification, C-M shows both charge and mass modification. ** "Chr." refers to chromosome, "pSCL4" refers to plasmid pSCL4.

Table 3. 12. List of downregulated proteins in DEPA strain identified by LC-MS/MS.

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Nucleotide Metabolism (DNA/RNA Processing)	E2Q7N6	Exopolyphosphatase	*SCLAV_2369*	Cytoplasm	-2.023	5.84	34,982	Chr
Carbohydrate Metabolism	B5GN11	Phosphoenolpyruvate synthase	*SCLAV_p0814*	Cytoplasm	-2.981	7.39	104,516	pSCL4
	B5GRN8	Malate synthase	*masY*	Cytoplasm	-2.184	5.94	60,012	Chr.
	E2PXE4	Glycosyl hydrolase	*SCLAV_0988*	Cytoplasm	-2.023	5.06	68,307	Chr.
Amino Acid Metabolism	D5SLL7	Methionine synthase II	*SCLAV_p1324*	Cytoplasm	-3.719	4.86	38,157	pSCL4
	D5SKH4	5-methyltetrahydropteroyltriglutamate --homocysteine methyltransferase	*metE*	Cytoplasm	-2.173	5.46	88,844	pSCL4
	B5GZV9	E1-alpha branched-chain alpha keto acid dehydrogenase	*bkdA1*	Cytoplasm	-2.023	5.68	40,843	Chr.
	E2PWQ0	Branched-chain alpha keto acid dehydrogenase E1 beta subunit	*bkdB1*	Cytoplasm	-2.580	5.24	36,531	Chr.
	E2PWQ1	Putative dihydrolipoamide acyltransferase	*bkdC*	Cytoplasm	-2.328	5.65	51,335	Chr.
	E2Q9H8	Urease subunit alpha	*ureC*	Cytoplasm	-2.981	5.20	60,195	Chr.
	E2Q9H6	Urease accessory protein UreG	*ureG*	Cytoplasm	-2.981	5.24	24,489	Chr.
	E2PXH6	3-phosphoserine phosphatase	*serB*	Cytoplasm	-2.691	5.07	44,136	Chr.
	E2Q2P6	L-threonine 3-dehydrogenase	*tdh*	Cytoplasm	-2.184	5.75	39,187	Chr.

Table 3. 12. List of downregulated proteins in DEPA strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Amino Acid Metabolism	D5SIP1	Putative ornithine cyclodeaminase	*arcB*	Cytoplasm	-3.429	6.06	38,557	pSCL4
Lipid Metabolism	B5H1P3	Acetyl/propionyl CoA carboxylase alpha subunit	*SCLAV_1944*	Cytoplasm	-3.222	5.44	67,249	Chr.
	B5H1P4	Acetyl/propionyl CoA carboxylase	*accD*	Cytoplasm	-2.981	6.91	57,553	Chr.
	E2PXK9	Enoyl-CoA hydratase/isomerase	*SCLAV_1053*	Cytoplasm	-2.023	5.19	28,445	Chr.
	B5GZV9	E1-alpha branched-chain alpha keto acid dehydrogenase	*bkdA1*	Cytoplasm	-2.023	5.68	40,843	Chr.
	E2PWQ0	Branched-chain alpha keto acid dehydrogenase E1 beta subunit	*bkdB1*	Cytoplasm	-2.580	5.24	36,531	Chr.
	E2PWQ1	Putative dihydrolipoamide acyltransferase	*bkdC*	Cytoplasm	-2.328	5.65	51,335	Chr.
	B5GL03	Isobutyryl-CoA mutase	*SCLAV_4321*	Cytoplasm	-2.328	5.08	62,23	Chr.
	E2Q9B0	3-oxoacyl-acyl-carrier-protein synthase III	*SCLAV_0454*	Cytoplasm	-3.609	4.63	36,937	Chr.
	B5H3R1	3-oxoacyl-(Acyl carrier protein) synthase III	*SCLAV_0484*	Unknown	-3.552	4.96	35,908	Chr.
	B5H2L5	Acyl-CoA thioesterase	*SCLAV_0496*	Cytoplasm	-2.328	6.16	33,519	Chr.
	B5GQM0	FabD	*fabD*	Cytoplasm	-2.184	4.89	31,025	Chr.

Table 3. 12. List of downregulated proteins in DEPA strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Lipid Metabolism	E2PYA5	Acetyl-CoA acetyltransferase	*SCLAV_5209*	Cytoplasm	-2.184	5.76	43,749	Chr.
	E2Q9A7	Acyl-CoA dehydrogenase	*SCLAV_0451*	Cytoplasmic membrane	-4.313	5.62	66,407	Chr.
	E2Q4I1	Acyl-CoA dehydrogenase	SCLAV_1946	Cytoplasm	-2.580	5.35	41,851	Chr.
	E2PXC8	Transferase	*plsC2*	Cytoplasmic membrane	-2.023	11.58	23,999	Chr.
DNA Replication, Recombination, Repair, Transcripton	E2Q4N1	MutT-like protein	*mutT*	Unknown	-2,023	5.32	21,331	Chr.
	B5H322	LuxR family two-component response regulator	*SCLAV_0205*	Cytoplasm	-2.460	5.54	24,400	Chr.
	E2PZ58	TetR-family transcriptional regulator	*SCLAV_5246*	Cytoplasm	-2.328	9.29	22,446	Chr.
Translation, Ribosomal Structures	B5GXN9	Ribosomal RNA small subunit methyltransferase H	*rsmH*	Cytoplasm	-2.184	6.48	34,541	Chr.
	E2Q1J5	Methionyl-tRNA formyltransferase	*SCLAV_5554*	Cytoplasm	-2.023	6.18	63,497	Chr.
	D5SJT1	Isoform 2 of Ribosomal protein S6 modification-like protein B	*rimKLB*	Cytoplasm	-2.023	5.84	35,535	pSCL4
Stress-Related, Protein Turnover, Chaperones	B5H0Q8	Alkaline serine protease	*SCLAV_0093*	Extracellular	-2.580	9.98	42,961	Chr.
	D5SLB4	Secreted trypsin-like serine protease	*SCLAV_p1221*	Unknown	-2.184	8.77	26,216	pSCL4
Secondary Metabolism	E2Q9C7	Gamma-butyrolactone biosynthesis protein	*avaA2*	Cytoplasm	-3.363	6.05	38,935	Chr.

Table 3. 12. List of downregulated proteins in DEPA strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Secondary Metabolism	E2Q9A9	Polyketide synthase	*SCLAV_0453*	Cytoplasmic	-4.810	5.80	105,757	Chr.
	E2Q518	Modular polyketide synthase	*SCLAV_0012*	Unknown	-2.328	5.53	372,808	Chr.
	B5H3Q8	Daunorubicin-doxorubicin polyketide synthase	*SCLAV_0481*	Unknown	-3.609	5.15	37,167	Chr.
	B5H3R3	Cytochrome P450-like enzyme	*SCLAV_0486*	Cytoplasmic membrane	-4.539	5.17	44,768	Chr.
	E2Q524	Cytochrome P450	*SCLAV_0018*	Cytoplasmic membrane	-2.794	5.89	45,929	Chr.
	E2PWJ8	Aldo/keto reductase	*cvm1*	Cytoplasm	-3.820	6.41	36,650	Chr.
	Q9X5G6	Putative ribulose-5-phosphate epimerase	*cvm2*	Cytoplasm	-2.794	5.27	16,877	Chr.
	E2PWJ6	Homoserine O-acetyltransferase	*cvm4*	Cytoplasm	-2.691	5.63	34,641	Chr.
	Q9X5G9	Flavin-dependent oxidoreductase, F420-dependent methylene-tetrahydromethanopterin reductase	*cvm5*	Cytoplasm	-4.046	5.85	44,863	Chr.
	E2PWJ4	Putative pyridoxal phosphate-dependent aminotransferase	*cvm6*	Cytoplasm	-4.625	5.36	48,065	Chr.

Table 3. 12. List of downregulated proteins in DEPA strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Secondary Metabolism	B5GLB6	3'-hydroxymethylcephem-O-carbamoyltransferase	*cmcH*	Unknown	-2.194	5.52	57,370	Chr.
	E2Q9D0	Goadsporin biosynthetic protein	*SCLAV_0474*	Cytoplasmic membrane	-4.625	5.77	60,143	Chr.
	E2Q9A3	FkbH like protein (Fragment)	*SCLAV_0447*	Cytoplasm	-4.568	5.76	62,815	Chr.
	E2Q1F4	Puromycin N-acetyltransferase	*SCLAV_5513*	Unknown	-2.023	5.69	22,089	Chr.
Energy Production/Electron-Iron Transfer	B5GZU6	Ferredoxin	*SCLAV_2965*	Cytoplasm	-2.023	6.47	37,892	Chr.
Inorganic/Organic Molecule Transport	D5SJE7	ABC Fe transporter	*SCLAV_p0551*	Cytoplasmic membrane	-3.665	6.22	36,146	pSCL4
	B5GPA9	Amino acid ABC transporter amino acid-binding protein	*SCLAV_2403*	Cytoplasmic membrane	-2.580	5.88	33,418	Chr.
	B5GSJ9	Sugar transporter sugar-binding protein	*SCLAV_2061*	Extracellular	-2.023	7.61	45,399	Chr.
	B5GX67	Preprotein translocase subunit SecE	*SCLAV_3623*	Cytoplasmic membrane	-2.023	9.39	10,206	Chr.
	E2Q9E5	MMPL domain protein	*SCLAV_0489*	Cytoplasmic membrane	-2.581	5.33	80,74	Chr.

Table 3. 12. List of downregulated proteins in DEPA strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Inorganic/Organic Molecule Transport	B5GSU7	ABC transporter ATP-binding protein	*SCLAV_3857*	Cytoplasmic membrane	-2.328	5.22	36,363	Chr.
Hypothetical/Unknown Proteins	D5SLL2	Putative transcriptional regulator AraC family	*SCLAV_p1319*	Unknown	-4.088	8.55	25,424	pSCL4
	E2Q1E7	Putative uncharacterized protein	*SCLAV_5506*	Extracellular	-2.580	6.30	102,062	Chr.
	E2PWZ8	Putative secreted FAD-linked oxidase	*SCLAV_5002*	Unknown	-3.820	8.61	62,020	Chr.
	D5SLV4	Putative hydrolase	*SCLAV_p1415*	Unknown	-3.770	9.78	55,779	pSCL4
	E2Q9C1	Putative oxidoreductase	*SCLAV_0465*	Unknown	-3.609	6.22	37,219	Chr.
	E2Q9B6	Putative hydrolase	*SCLAV_0460*	Cytoplasmic	-2.794	11.24	36,542	Chr.
	E2Q9C4	Putative membrane protein	*SCLAV_0468*	Cytoplasmic membrane	-2.691	8.66	82,944	Chr.
	E2PWX6	Putative uncharacterized protein	*SCLAV_4980*	Unknown	-2.328	11.05	27,520	Chr.
	E2Q2P0	Putative SNF2/RAD54 family helicase	*SCLAV_5684*	Cytoplasm	-2.328	5.74	109,199	Chr.
	E2Q7G3	Putative esterase	*SCLAV_2296*	Cytoplasmic membrane	-2.184	9.78	43,052	Chr.
	E2Q0E3	Putative uncharacterized protein	*SCLAV_5419*	Cytoplasm	-2.184	4.30	34,615	Chr.
	E2Q9C3	SMP-30/gluconolaconase/LRE domain-containing protein	*SCLAV_0467*	Cytoplasm	-3.363	10.45	35,346	Chr.

Table 3. 12. List of downregulated proteins in DEPA strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Hypothetical/Unknown Proteins	E2Q7E7	Probable FAD-dependent monooxygenase	*SCLAV_2280*	Cytoplasm	-2.023	6.82	43,098	Chr.
	E2Q3B3	Putative ornithine aminotransferase	*SCLAV_3756*	Cytoplasm	-2.023	5.82	43,699	Chr.
	E2Q2R1	Putative diaminopimelate decarboxylase	*SCLAV_5705*	Cytoplasm	-2.023	7.87	45,540	Chr.
	B5GPM3	Putative uncharacterized protein	*SCLAV_2519*	Cytoplasm	-2.023	5.60	41,200	Chr.
	E2Q132	Putative TetR-family transcriptional regulator	*SCLAV_3466*	Cytoplasm	-2.023	9.57	23,363	Chr.
	E2Q1H2	Putative 3-hydroxybutyrate dehydrogenase	*SCLAV_5531*	Cytoplasm	-2.023	7.13	26,881	Chr.
	E2PXL3	Putative uncharacterized protein	*SCLAV_1057*	Cytoplasm	-2.023	6.20	53,066	Chr.
	E2Q0H4	Putative uncharacterized protein	*SCLAV_5450*	Unknown	-2.023	5.25	28,491	Chr.
	E2Q2S2	Putative uncharacterized protein	*SCLAV_5716*	Unknown	-2.023	6.00	25,200	Chr.
	E2Q8I3	Putative two-component system response regulator	*SCLAV_2398*	Cytoplasm	-2.023	5.82	24,311	Chr.
	E2Q1N8	Putative uncharacterized protein	*SCLAV_5597*	Cytoplasm	-2.023	5.64	17,402	Chr.
	B5H453	Predicted acyltransferase	*SCLAV_p1317*	Unknown	-2.891	6.80	31,731	pSCL4
	D5SJD5	Uncharacterized protein	*SCLAV_p0539*	Unknown	-4.205	5.79	248,698	pSCL4
	B5GMU8	Uncharacterized protein	*SCLAV_p0876*	Unknown	-3.552	5.91	107,431	pSCL4
	E2QAB7	Lipoprotein	*SCLAV_4744*	Unknown	-2.328	4.97	24,811	Chr.

Table 3. 12. List of downregulated proteins in DEPA strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Hypothetical/Unknown Proteins	D5SJG5	Uncharacterized protein	*SCLAV_p0569*	Cytoplasm	-2.460	5.19	63,034	pSCL4
	B5GRI3	Uncharacterized protein	*SCLAV_p0538*	Unknown	-2.460	9.10	15,000	pSCL4
	B5GMY6	Uncharacterized protein	*SCLAV_p0839*	Unknown	-2.328	10.18	26,714	pSCL4
	E2Q9B2	Hydrolase superfamily dihydrolipoamide acyltransferase-like protein	*SCLAV_0456*	Cytoplasm	-2.580	11.07	30,811	Chr.
	B5H017	CHAD domain-containing protein	*SCLAV_1772*	Cytoplasm	-2.580	9.75	56,671	Chr.
	E2PUG6	DUF89 domain-containing protein	*SCLAV_2673*	Cytoplasm	-2.328	5.14	44,904	Chr.
	B5GMY4	DUF1906 domain-containing protein	*SCLAV_p0841*	Unknown	-2.275	7.25	80,708	pSCL4
	B5GZ43	DUF574 domain-containing protein	*SCLAV_4571*	Cytoplasm	-2.184	5.99	29,884	Chr.
	E2Q2I4	DUF574 domain-containing protein	*SCLAV_5628*	Cytoplasm	-2.023	5.48	29,618	Chr.
	D5SKY9	DUF1906 multi-domain protein	*SCLAV_p1096*	Extracellular	-2.023	6.42	80,346	pSCL4
	E2Q1E2	WGR domain protein	*SCLAV_5501*	Cytoplasm	-2.023	6.27	51,083	Chr.
	E2Q5J1	ATPase, AAA family, putative	*SCLAV_4130*	Cytoplasm	-2.023	5.36	39,182	Chr.
Others/General Function	D5SJT7	Bacterial luciferase domain-containing protein	*SCLAV_p0693*	Unknown	-3.719	5.22	31,646	pSCL4
	E2Q0K7	2-hydroxyacid-family dehydrogenase	*SCLAV_5483*	Cytoplasm	-2.328	5.94	36,239	Chr.
	D5SJT4	Acetyltransferase	*SCLAV_p0690*	Cytoplasm	-2.184	6.12	31,009	pSCL4

Table 3. 12. List of downregulated proteins in DEPA strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Others/General Function	B5GZC6	Amidohydrolase	*SCLAV_5667*	Cytoplasm	-2.023	5.64	45,975	Chr.
	B5GLP1	Short-chain dehydrogenase/reductase SDR	*SCLAV_p1420*	Cytoplasm	-2.891	5.07	29,396	pSCL4
	B5GMR3	Short-chain dehydrogenase/reductase SDR	*SCLAV_p1083*	Cytoplasm	-2.460	5.37	25,558	pSCL4
	E2PVL1	Short chain dehydrogenase	*SCLAV_2852*	Cytoplasm	-2.328	10.57	75,106	Chr.
	B5GLZ2	NAD-dependent epimerase/dehydratase	*SCLAV_p1515*	Cytoplasm	-2.184	5.15	22,332	pSCL4
	E2Q9D8	AMP-dependent synthetase and ligase	*SCLAV_0482*	Cytoplasm	-4.625	4.97	55,891	Chr.
	E2Q9A5	AMP-binding domain-containing protein	*SCLAV_0449*	Cytoplasm	-4.348	5.47	65,18	Chr.
	B5GRS4	Aminotransferase	*SCLAV_5040*	Cytoplasm	-2.580	5.76	45,046	Chr.
	E2Q9D9	Decarboxylase	*SCLAV_0483*	Cytoplasm	-4.278	5.06	56,656	Chr.
	E2Q3H5	Aldehyde dehydrogenase	SCLAV_3818	Cytoplasm	-2.981	5.62	51,881	Chr.
	E2Q9K6	GCN5-related N-acetyltransferase	*SCLAV_0550*	Cytoplasm	-3.222	7.34	52,041	Chr.

Table 3. 12. List of downregulated proteins in DEPA strain identified by LC-MS/MS. (Cont'd)

Functional Categories	UniProt Accession Number	Protein Name	Gene	Cellular Localization	Rsc	pI	Mw (Da)	Genome Location*
Others/General Function	E2Q1Z6	Tryptase	*SCLAV_1547*	Unknown	-3.222	8.81	31,806	Chr.
	E2Q8Y7	Acetyltransferase	*SCLAV_4530*	Cytoplasm	-3.146	4.95	45,083	Chr.
	E2QAB8	Integral membrane protein	*SCLAV_4745*	Cytoplasmic membrane	-2.981	7.84	79,468	Chr.
	E2Q1H4	GAF domain protein	*SCLAV_5533*	Cytoplasm	-2.794	6.31	70,067	Chr.
	B5GVX9	Methyltransferase	*SCLAV_p0129*	Cytoplasm	-2.580	6.45	29,982	pSCL4
	E2Q1A2	Monooxygenase FAD-binding protein	*SCLAV_3535*	Cytoplasmic membrane	-2.580	8.07	52,028	Chr.
	D5SLF7	AMP-dependent synthetase and ligase	*SCLAV_p1264*	Cytoplasm	-2.460	7.43	59,693	pSCL4
	E2Q400	Regulatory protein	*SCLAV_3866*	Cytoplasm	-2.460	6.39	130,618	Chr.
	B5H2E5	Hydrogenase	*SCLAV_0281*	Cytoplasm	-2.050	9.79	45,022	Chr.
	B5GPW6	Oxidoreductase	*SCLAV_2295*	Cytoplasm	-2.023	4.90	27,527	Chr.

* "Chr." refers to chromosome, "pSCL4" refers to plasmid pSCL4.

3.4.2.1. Nucleotide Metabolism (DNA/RNA Processing)

There were 4 downregulated proteins categorized under this group: Purine nucleoside phosphorylase, 5'-nucleotidases, inositol-5-monophosphate dehydrogenase and exopolyphosphatase.

Purine nucleoside phosphorylase generates purine base and ribose(deoxyribose)-1-phosphate using Pi (inorganic orthophosphate) by cleaving the glycosidic bond of ribo- and deoxyribonucleosides. This protein is important in the salvage pathway of the purine metabolism, in which purine bases from metabolized purine ribo- and deoxyribonucleosides are recycled to produce purine nucleotides (Bzowska et al., 2000).

5'-nucleotidases hydrolyze 5'-nucleotides generating Pi (Ammerman and Azam, 1985). 5'-nucleotidases show differences in both function and the substrates that they utilize depending on the organism and even the tissue type. The localization of them can also show variability; they can be cytoplasmic or membrane-bound. Substrates that they hydrolyze range from 5'-mononucleotides, 5'-dinucleotides, 5'-trinucleotides to complex nucleotides like UDP-glucose or FAD (Zimmermann, 1992).

Inositol-5-monophosphate dehydrogenase (IMDH) converts inositol-5′-monophosphate to xanthosine 5′-monophosphate with the reduction of NAD to NADH. IMDH catalyzes the first step for guanosine 5′-triphosphate synthesis, in other words, it is related with guanine nucleoside metabolism. Since GTP is the precursor of ppGpp synthesis, the function of the enzyme can be important in regulating cell proliferation, differentiation and secondary metabolism (Le Maréchal et al., 2013; Chiu et al., 2011).

Exopolyphosphatase is responsible for the hydrolysis of inorganic polyphosphate (PolyP) that is accumulated in the cell. Inorganic polyphosphates are composed of hundreds of orthophosphate residues (Pi) coming together and used for diverse cellular processes. They can be sources of phosphorus and/or energy or they can be exploited in kinase reactions or stress conditions, metal chelation, regulatory interactions (Gray et al., 2014; Lindner et al., 2009).

3.4.2.2. Carbohydrate Metabolism

There were found to be 7 proteins downregulated in this category.

2,3-bisphosphoglycerate-dependent phosphoglycerate mutase catalyzes the reversible conversion of 3-phosphoglycerate and 2-phosphoglycerate and this reaction is required in both the glycolysis and gluconeogenesis pathways (van der Oost et al., 2002).

On the other hand, phosphoglycerate kinase enzyme is responsible for the conversion of 1,3-diphospho-D-glycerate to 3-phospho-D-glycerate accompanied by ADP to ATP production in glycolytic pathway. (Banks et al., 1979; Yon et al., 1990). In the transcriptome study of Medema et al (2011), the glycolytic enzymes in between glyceraldehyde-3-phosphate and phosphoenolpyruvate were found unchanged in their industrial strain. However, as found in this proteomics study, both of the enzymes were downregulated by 4.7 and 4.9 folds, respectively.

Isocitrate dehydrogenase (IDH) catalyzes the conversion of isocitrate to α-ketoglutarate in the citric acid cycle, which provides organism with energy and precursors for several biological processes (Hurley et al., 1996; Zhang et al., 2009). More importantly, the same protein was also shown to be among the overexpressed proteins of DEPA strain with approximately 5 fold increased levels (Section 3.4.1.2). However, as shown in Figure 3. 27, an additional protein spot possibly representing a different form of the same protein was 3 fold downregulated in DEPA. Notably, this protein was reported by Medema et al (2011) as as one of the central enzymes with unchanged levels in the transcriptome of the industrial strain of *S. clavuligerus*.

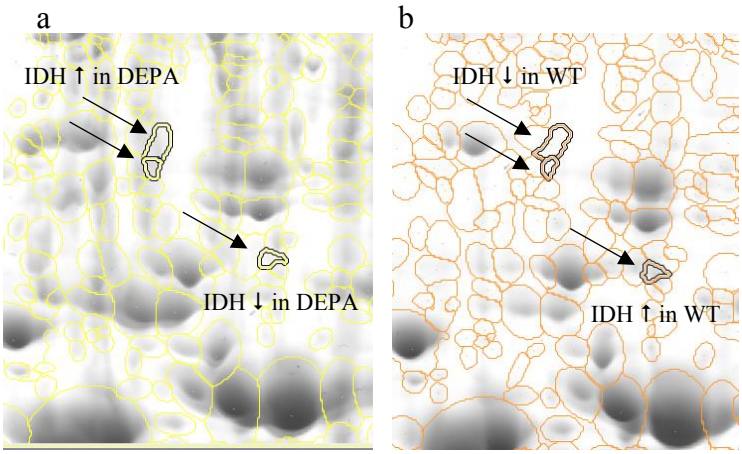


Figure 3. 27. Localized representation of protein spots detected on DEPA gel (a) and NRRL 3585 gel (b) by Delta2D software. Highlighted spots pointed with arrows show up- and downregulated isocitrate dehydrogenase (IDH) spots on DEPA (a) and NRRL 3585 (b) gel.

Oxppcycle protein (OpcA) encoded by *opcA* is suggested to be a positive allosteric effector of glucose-6-phosphate dehydrogenase (G6PD) which is one of the key enzymes in pentose phosphate pathway (Hagen and Meeks, 2001; Agarwal and Auchus, 2005). It was suggested that G6PD oligomer assembly, which is the activated form of G6PD, is dependent on the presence of OpcA proteins in *Synechococcus* 7942 (Sundaram et al., 1998). On the other hand, Hagen and Meeks (2001) showed that the assembly of G6PD into its native tetrameric form occurs in *N. punctiforme* even in the absence of OpcA or when the stoichiometry between OpcA and G6PD is greatly altered. However, although *N. punctiforme* G6PD is assembled correctly in the absence of OpcA, it is essentially inactive, because it has a very low affinity for G6P substrate (Hagen and Meeks, 2001).

Phosphoenolpyruvate synthase is responsible for the reversible conversion of pyruvate to phosphoenolpyruvate (PEP), which can be used for the biosynthesis of diverse components (Cooper and Kornberg, 1967).

Malate synthase is important in glyoxylate pathway and catalyzes the condensation of malate from acetate and glyoxylate. Malate synthesized in this pathway is returned to

TCA cycle to produce key intermediates for cellular processes (Chan and Sim, 1998; Loke et al., 2002)

Glycosyl hydrolases are responsible for the breakdown of glycosidic bonds in carbohydrates in order to yield smaller carbohydrate molecules. Lysozymes, viral neuraminidase and bacterial sialidase are examples for glycosyl hydrolases (Campbell et al., 1997).

Although we do not know yet very well the existing controls appeared in DEPA and the logic behind, proteins that showed lowerexpression in the carbohydrate metabolism of DEPA strain are mainly used for supporting the cell growth and central metabolism.

3.4.2.3. Amino Acid Metabolism

There were 21 proteins in this category. Of these, methionine synthase II and 5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase were the common proteins identified by both MALDI-TOF/MS and LC-MS/MS analyses. Moreover, enzymes involved in methionine biosynthesis in this category constituted a considerable fraction.

In *E. coli*, homoserine is activated by an *O*-succinyl group whereas in others its activation depends on an *O*-acetyl group from acetyl-CoA, which is followed by the generation of homocysteine. Cystathionine γ-synthase (*metB*) and cystathionine β-lyase (*metC*) are responsible for the synthesis of homocysteine from O-succinylhomoserine in a two-step reaction in *E. coli* (Figure 3. 28 and Figure 3. 29) (Ferla and Patrick 2014). Cystathionine γ-synthase enzyme is important in methionine biosynthesis by catalyzing the γ-replacement reaction during the synthesis (Johnston et al., 1979; Kreft et al., 1994). Till the study of Chang and Vining (2002) who demonstrated roles for cystathionine *γ*-synthase and transsulfuration, there were only limited information about interconversion of sulfur-containing amino acids in streptomycetes. Cystathionine γ-synthase was represented by 5 different spots on the gel forming a cluster-like pattern around a central protein spot, so it must have been subject to an extensive posttranslational modification that affected both the charge and

the mass of the protein. These spots were at 3.5 to 83 fold decreased and along with other downregulated enzymes of methionine biosynthesis (explained below), it most probably served for lowering the metabolic flux to methionine.

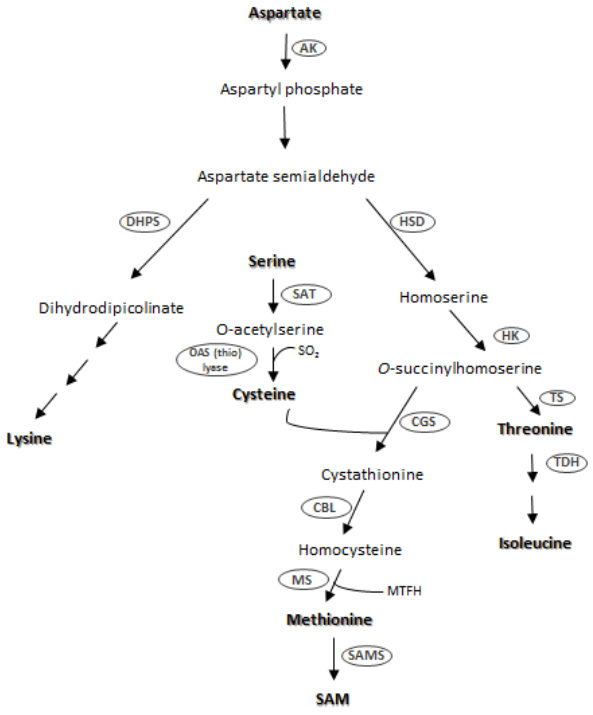


Figure 3. 28. Aspartatate family biosynthetic pathway.

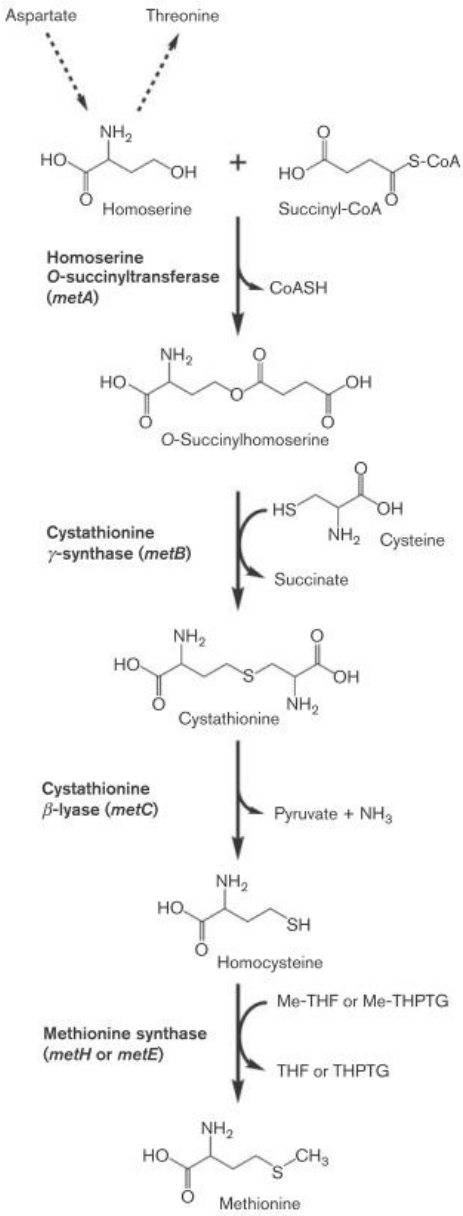


Figure 3. 29. Methionine biosynthesis in *E. coli*. Me-THF, 5-methyl-tetrahydrofolate; Me-THPTG, 5-methyltetrahydropteroyl tri-L-glutamate; THF, tetrahydrofolate; THPTG, tetrahydropteroyl tri-L-glutamate. (Ferla and Patrick, 2014).

O-acetylhomoserine aminocarboxypropyltransferase is involved in methionine biosynthesis and catalyzes the reaction shown below (Schatschneider et al., 2011),

which does not require homocysteine as the intermediate but catalyzes the synthesis of methionine directly from O-acetyl-L-homoserine (Ferla and Patrick, 2014).

O-acetyl-L-homoserine + methanethiol → L-methionine + acetate (http://www.brenda-enzymes.org/enzyme.php?ecno=2.5.1.49)

This enzyme, with a 6.5-fold downregulated level, provides evidence that streptomycetes utilize this single step conversion in addition to two-step trans-sulfuration in methionine biosynthesis.

The S-methylation of homocysteine is the final step in methionine biosynthesis. As stated earlier, there are two types of non-homologous methionine synthases in *E.coli*. One is cobalamin-dependent enzyme encoded by *metH* and the other is cobalamin-independent enzyme encoded by *metE* which is also called methionine synthase II (Banerjee and Matthews, 1990). Methionine synthase II, also known as 5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase, is responsible for the production of tetrahydropteroyltriglutamate and methionine by transferring a methyl group from 5-methyltetrahydropteroyltriglutamate to homocysteine in the presence of magnesium and phosphate ions (Whitfield et al., 1970). On the other hand, cobalamin-dependent enzyme, which is also known as 5-methyltetrahydrofolate-homocysteine methyltransferase, transfers the methyl group from 5-methyltetrahydrofolate to homocysteine using its cobalamin for the transfer of methyl group, producing tetrahydrofolate and methionine (Figure 3. 29 and Figure 3. 30) (Banerjee and Matthews, 1990). Since *E. coli* has lost cobalamin biosynthetic pathway, it uses MetH only in the presence of exogenous cobalamin, moreover it represses MetE in such conditions (Ferla and Patrick, 2014). MetE protein identified in this study was represented by two different spots with up to 3-fold downregulated levels, suggesting the presence of PTM. Another protein with the same function but named as SCLAV_p1324 was shown to be upregulated approximately 4 fold.

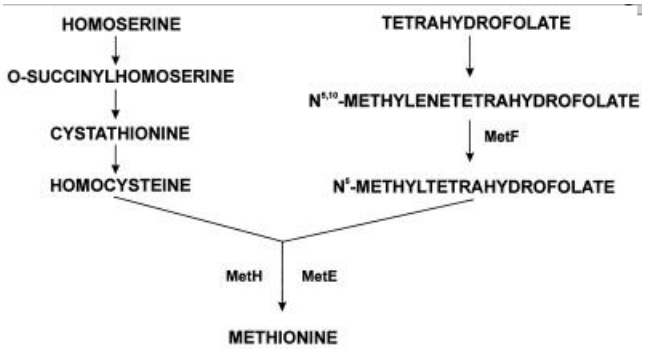


Figure 3. 30. Biosynthetic pathway of methionine showing the formation of the homocysteine moiety from homoserine (left branch) and the origin of the methyl group from the folate branch (right branch) (Blanco et al., 1998).

S-adenosylmethionine synthetase is encoded by *metK* and catalyzes the synthesis of S-adenosylmethionine (SAM) from ATP and methionine. The primary function of SAM is to donate methyl groups to diverse metabolites coming from primary or secondary metabolism as well as proteins, nucleic acids and polysaccharides (Yoon et al., 2006). However, SAMs have also been shown to affect the secondary metabolism and morphological differentiation in *Streptomyces* (Shin et al., 2006). It was shown that multicopy *metK* genes and exogenous addition of SAM to *S. coelicolor* cells resulted in the increase actinorhodin production (Okamoto et al., 2003). Antibiotics such as bicozamycin from *Streptomyces griseoflavus*, pristinamycin from *Streptomyces pristinaespiralis* and granaticin from *Streptomyces violaceoruber* were also shown to be produced in increased levels by the addition of SAM (Yoon et al., 2006). In *B. subtilis*, SAM has some roles in regulation of transcription by binding to several mRNA molecules (Shin et al., 2006). SAM *per se* was also shown to activate a transcriptional activator, namely actII-ORF4, thereby transcription through the actinorhodin biosynthesis gene clusters are enhanced resulting in increased levels of actinorhodin antibiotic in *Streptomyces lividans* TK23 (Kim et al., 2003). Hence, 2.6 fold reduction of this protein along with strongly downregulated levels of methionine biosynthesis might have a significant contribution to higher CA levels in DEPA strain.

Histidinol-phosphate aminotransferase is a member of aminotransferases. The aminotransferases catalyze the reversible transfer of amino groups from amino acids to oxo acids, one of the numerous transformations of amino acids that are performed by vitamin-B_6-dependent enzymes (Mehta et al., 1993).

Pyrroline-5-carboxylate reductase (P5C reductase) is the last enzyme involved in proline synthesis from glutamate. (Deutch et al., 1982). However, this protein was also shown to be upregulated in DEPA (Section 3.4.1.3). As shown in Figure 3. 31, the protein was represented by two spots on the gel with slight pI changes. One of these spots showed higher expression while the other showed higher expression. It is likely that the protein has undergone PTM differentially in DEPA and NRRL 3585 strains.

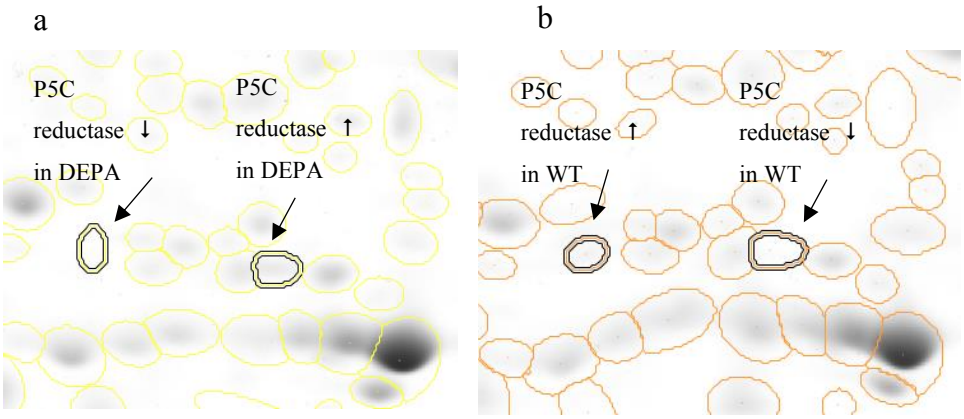


Figure 3. 31. Localized representation of protein spots detected on DEPA gel (a) and NRRL 3585 gel (b) by Delta2D software. Highlighted spots pointed with arrows were up- and down-regulated P5C reductase protein spots on DEPA (a) and NRRL 3585 (b) gel.

3-dehydroquinate synthase encoded by *aroB* and 3-phosphoshikimate 1-carboxyvinyltransferase (also known as 5-enolpyruvylshikimate-3-phosphate synthase) coded by *aroA* enzymes catalyze the reactions in shikimate pathway that leads to the production of aromatic compounds, aromatic amino acids and many aromatic secondary metabolites. Dehydroquinate synthase catalyzes the ring-closure

of 3-deoxy-D-arabino-heptulosonic acid 7-phosphate (DAHP) to form the saturated six-membered ring (of DHQ) (Bartlett and Satake, 1988; Schomburg, 1997; Herrmann and Weaver, 1999). 3-phosphoshikimate 1-carboxyvinyltransferase, on the other hand, catalyzes the transfer of carboxyvinyl moiety of phosphoenolpyruvate to shikimate 3-phosphate (S3P) producing inorganic phosphate and 5-enolpyruvylshikimate 3-phosphate (EPSP) (Huynh et al., 1988). Shikimate pathway can lead to production of aromatic polyketide antibiotic production. Several derivaties of shikimic acids such as cyclohexanecarboxylic acid and dihydroxycyclohexanecarboxylic acid are used in the production of ansatrienin in *Streptomyces collinus* and ascomycin in *Streptomyces hygroscopicus* var *ascomyceticus*, respectively (Wilson et al., 1998).

Alanine dehydrogenase catalyzes the reversible deaminating reaction of L-alanine to convert it to pyruvate. (Itoh and Morikawa, 1983*)*. This enzyme was detected on two spots differing in pI range, suggesting a PTM on the protein with a charge modification.

4-hydroxyphenylpyruvate dioxygenase involved in tyrosine catabolism was also shown to be downregulated in AK39 (Section 3.2.2.3).

Imidazole glycerol phosphate synthase (ImGP synthase) catalyzes imidazole ring closure in histidine biosynthesis by using glutamine and N'-[(5'phosphoribulosyl)formimino]-5-aminoimidazole-4-carboxamide-ribonucleotide) (PRFAR) as substrates and also provides 5-aminoimidazole-4-carboxamide ribotide (AICAR) to be exploited in purine biosynthesis (Douangamath et al., 2002). It is a key metabolic enzyme that links amino acid and nucleotide biosynthesis and consists of two subunits: HisF which converts (PRFAR) to imidazole glycerol phosphate (IGP) and AICAR, and HisH. HisF and HisH form a stable 1:1 dimeric complex which is called IGP synthase holoenzyme (Klem and Davisson, 1993). This protein was found as strongly downregulated although any role of this differential expression in CA production is unpredictable.

As already explained (Section 3.2.2.3), E1-alpha branched-chain alpha keto acid dehydrogenase, branched-chain alpha keto acid dehydrogenase E1 beta subunit and

dihydrolipoamide acyltransferase (E2) together comprise three of the four components of BCDH multienzyme complex and are involved in fatty acid synthesis, branched-chain amino acid catabolism and polyketide antibiotic biosynthesis (Denoya et al., 1995; Skinner et al, 1995; Harris et al., 1990). The decreased expression of these proteins might be attributed to cope with an increased production of CA since the suppression of competing metabolic activities would be to the advantage of cells

Urease subunit alpha (UreC) together with UreA and UreB subunits forms the complete urease enzyme. Urease can also require four accessory proteins coded by ureD, ureE, ureF, and ureG for its activity (Koper et al., 2004). It catalyzes the production of ammonia and CO_2 from urea which can be used as nitrogen source by soil bacteria (Contreras-Rodriguez et al., 2008). The downregulation of urease in DEPA strain might constitute an example to possible economical measures taken by DEPA to prevent extra metabolic burdens.

Phosphoserine phosphatase (SerB) is a subfamily of phosphoserine phosphatase (PSP) family (Burroughs et al., 2006). SerB performs the last step of a sequential serine biosynthesis pathway by dephosphorylating the 3-phospho-L-serine to produce L-serine. L-serine is used as a source of one-carbon units in purine, amino acid and thymidylate biosynthesis; moreover, it can also be used as an intermediate in phospholipid biosynthesis.

L-threonine 3-dehydrogenase converts threonine, which is produced through aspartate pathway, to acetaldehyde and glycine in *E. coli*. L-threonine exerts feedback inhibition on aspartate kinase, homoserine dehydrogenase and homoserine kinase enzymes which act through the aspartate pathway (Dong et al., 2011).

Putative ornithine cyclodeaminase encoded by *arcB* is responsible for the irreversible conversion of L-ornithine to L-proline by the removal of the ammonia (Goodman et al., 2004). Ornithine is the precursor of arginine amino acid. And arginine condensation with glyceraldehyde 3-phosphate initiates the CA biosynthesis pathway (Tahlan et al., 2004b). Given that DEPA strain is a CA overproducer, low abundance

of ornithine cyclodeaminase is as expected to decrease the conversion of ornithine to proline.

To summarize, methionine biosynthetic pathway addresses one of the very striking differences between wild type *S. clavuligerus* and its CA-overproducing mutant since among the downregulated proteins of amino acid metabolism, proteins involved in methionine biosynthesis were more populated when compared to the others. Methionine has a general role in protein elongation as well as as the initiator of peptide biosynthesis. However, over-representation of downregulated proteins involved in the methionine biosynthesis, especially some having probable PTMs suggests more on their function. For example, as in the case of S-adenosylmethionine (SAM) which is also involved in both secondary metabolism and differentiation, methionine derivatives can be the methyl donors in diverse biological processes. Therefore, downregulation of proteins related to methionine biosynthesis within the cell might suggest that they play important roles in the primary metabolic processes or production of secondary metabolites other than CA.

3.4.2.4. Lipid Metabolism

There were 19 proteins shown to be downregulated under this category.

Acetyl-CoA acetyltransferase, according to UniProtKB belongs to the thiolase family. Acetoacetyl-CoA thiolase is involved in the metabolism of acetoacetate and β-hydroxybutyrate (Igual et al., 1992). This protein encoded by *SCLAV 5209* was also found as downregulated in DEPA 2DE MALDI-TOF/MS experiments, but as encoded by a different gene, *SCLAV_4845*.

Acyl-CoA dehydrogenase enzymes comprise a large family and catalyzes the α, β-dehydrogenation of CoA-conjugated fatty acids that are produced during β-oxidation pathways (Ghisla and Thorpe, 2004). They can be involved in a large variety of biological processes. We have already demonstrated downregulation of acyl-CoA dehydrogenase encoded by *SCLAV_0960* in DEPA strain via 2DE MALDI-TOF/MS. Herein, we report downregulation of two other acyl-CoA dehydrogenases encoded by *SCLAV_04551* and *SCLAV_1946* in this strain, as shown by LC-MS/MS.

FabD, 3-oxoacyl-(acyl carrier protein [ACP]) synthase III, 3-Oxoacyl-[acyl-carrier-protein] reductase and enoyl-[acyl-carrier-protein] reductase (FabI) proteins are involved in fatty acid biosynthesis pathway in bacteria, which was explained in detail in Section 3.2.1.2. There are two 3-oxoacyl-(acyl carrier protein [ACP]) synthase III proteins identified as SCLAV_0454 and SCLAV_0484 which are located on the chromosome. There can be the alternative enzymes to FabH which is also 3-oxoacyl-(acyl carrier protein [ACP]) synthase, perhaps as a result of differential expression of these enzymes under certain conditions.

E1-alpha branched-chain alpha keto acid dehydrogenase, branched-chain alpha keto acid dehydrogenase E1 beta subunit and putative dihydrolipoamide acyltransferase (E2) are also categorized under the amino acid metabolism and explained in Section 3.2.2.3.

Acetyl-CoA acetyltransferase, according to UniProtKB database, belongs to the thiolase family. There are two types of thiolase enzymes; one of them is 3-oxoacyl-CoA thiolase and the other is acetoacetyl-CoA thiolase. Acetoacetyl-CoA thiolase is involved in the metabolism of acetoacetate and β-hydroxybutyrate, whereas 3-oxoacyl-CoA thiolase is a constituent of a multienzyme complex involved in fatty acid oxidation (Igual et al., 1992). This protein encoded by *SCLAV 5209* was also found as downregulated in DEPA 2DE MALDI-TOF/MS experiments, but as encoded by a different gene, *SCLAV_4845*.

Isobutyryl-CoA mutase (ICM) is one of the acyl-CoA mutases that play important roles in both primary and secondary metabolism (Jost et al., 2015). ICM catalyzes the reversible conversion of isobutyrylCoA to n-butyryl-CoA in a coenzyme B12-dependent manner. ICM isolated from *S. cinnamonensis* contains one large subunit (IcmA) and a smaller subunit (IcmB) which contains sequences of the cobalamin binding domains. ICMs are responsible for valine and fatty acid catabolism and also were shown to play key roles in producing building blocks for polyketide biosynthesis in several streptomycetes (Ratnatilleke et al., 1999). The decreased levels of this enzyme also suggests a decreased metabolic burden for secondary metabolites other than CA.

Enoyl-CoA hydratase/isomerase belongs to crotonase superfamily, each member of which has different functions such as dehalogenase, hydratase, and isomerase activities. However, generally they are involved in the stabilization of an enolate anion intermediate coming from an acyl-CoA substrate (Holden et al., 2001).

As explained previously in detail (Section 3.2.1.2), acetyl/propionyl CoA carboxylase enzyme is responsible for the production of malonyl-CoA used for both fatty acid biosynthesis as well as polyketide biosynthesis (Rodríguez and Gramajo, 1999; Gago et al., 2011).

Thioesterases form a large group of family whose substrates can differ depending on the family member such as coenzyme A (CoA), acyl carrier proteins, glutathione or its derivatives (Cantu et al., 2010.). Acyl-CoA thioesterases are responsible for the formation of free fatty acids and CoAs (Jones et al., 1999).

According to UniProtKB entry "transferase" enzyme coded by *plsC2* contains a phospholipid/glycerol acyltransferase domain and is involved in phospholipid biosynthesis in bacteria. PlsC was shown to be responsible for the production of a cell membrane component, phosphatidic acid (Zhang and Rock, 2008).

Acyl-CoA dehydrogenase enzymes comprise a large family and catalyzes the α, β-dehydrogenation of CoA-conjugated fatty acids that are produced during β-oxidation pathways (Ghisla and Thorpe, 2004). They can be involved in a large variety of biological processes. We have already demonstrated downregulation of acyl-CoA dehydrogenase encoded by *SCLAV_0960* in DEPA strain via 2DE MALDI TOF MS. Herein, we report downregulation of two other acyl-CoA dehydrogenases encoded by *SCLAV_04551* and *SCLAV_1946* in this strain, as shown by LC-MS/MS.

3.4.2.5. DNA Replication, Recombination, Repair, Transcripton

There were 5 proteins shown to be downregulated in this group: Two-component system response regulator (SCLAV_2102), two TetR-family transcriptional regulators (SCLAV_3146 and SCLAV_5246), MutT-like protein and LuxR family two-component response regulator (SCLAV_0205).

As explained earlier in Chapter I, two component systems (TCSs) are important mechanisms controlling the transcription of genes depending on changing environmental conditions. In *Streptomyces*, they regulate the secondary metabolism and morphological differentiation triggered by environmental stimuli such as nutritional changes or other stress factors (Martín and Liras, 2010; Rodríguez et al, 2013). SCLAV_2102 protein identified in this study showed 4.5 fold decreased levels suggesting that it is a negative regulator of CA biosynthesis according to 2DE MALDI-TOF/MS analysis, whereas LuxR family two-component response regulator (SCLAV_0205) identified by LC-MS/MS technique had an Rsc value of -2.460.

TetR family of proteins are transcriptional regulators that mostly function as repressors but sometimes as activators. These proteins form homodimers, each monomer consisting of one N-terminal DNA-binding domain (DNB domain) and one C-terminal ligand-binding domain. Homodimeric structure bind DNA through their DNB domains, thereby repressing the transcription of certain genes. When small ligand molecules bind their C-terminal ligand-binding domains, repression is relieved. More than 80% of the sequenced bacterial genomes were shown to contain at least one TetR family gene on average (Yu et al., 2010). Although these are widely distributed in bacterial genomes, function of only 85 members of this family have been elucidated (Ramos et al., 2005). Furthermore, more than 100 TetR family protein-coding genes were detected in some *Streptomyces* spp. probably owing to their complex morphological differentiation and secondary metabolism (Guo et al., 2013). Several processes that they regulate in the cell range from multidrug resistance, catabolic pathways, osmotic stress and antibiotic biosynthesis to pathogenicity (Ramos et al., 2005; Wei et al., 2014). As in the case of TCS response regulators, there are many TetR family of protein records from *S. clavuligerus* genome, one of which is SCLAV_3146 that was identified as 2.8 fold downregulated by 2DE analysis and the other being SCLAV_5246 with an Rsc value of -2.328.

MutT protein is involved in metabolism of 8-oxoguanine (8-oxoG) which is a mutagenic base analog produced during oxidative stress. In *E. coli*, three proteins are involved in this mechanism: MutT, MutY, and MutM. The function of MutT is to

convert 8-oxodGTP to 8-oxodGMP pyrophosphate so that DNA polymerase III cannot recognize the base during replication (Fowler et al., 2003).

3.4.2.6. Translation, Ribosomal Structures

There were 5 proteins categorized under this group: 30S ribosomal protein S1, elongation factor Tu, ribosomal RNA small subunit methyltransferase H, methionyl-tRNA formyltransferase and isoform 2 of Ribosomal protein S6 modification-like protein B.

30S ribosomal protein S1 is one of the ribosomal proteins; however, this protein is only loosely bound to ribosomes as it is not necessary for the constitution of 30S ribosomal structures. It is hypothesized to bind to the ribosomes and to unwind RNA for its entry to the ribosomes, hence it is not only involved in translation initiation but also in elongation. Nonetheless, even though most of the protein synthesis requires the presence of S1 proteins, it was also shown that translation of several mRNAs is also possible without them (Sørensen et al., 1998).

Elongation factor Tu is one of the three elongation factors that are involved in the elongation of the peptides on the ribosomes during translation. Elongation factor Tu (EF-Tu), in its GTP-bound form, recognizes its cognate amino acyl-tRNA (aa-tRNA) and helps it base-pair with the codons on mRNA located in the A-site of ribosomes. The base-pairing between mRNA codon and anti-codon of tRNA triggers the GTP hydrolysis and as a result, EF-Tu loses its affinity for its cognate tRNA and is released from the complex. Another elongation factor, nucleotide-exchange factor (EF-T), recycles EF-Tu back to its GTP form; meanwhile, the third elongation factor, EF-G, mediates the translocation of tRNA along the ribosome (Andersen et al., 2003; Nilsson and Nissen, 2005).

Ribosomal RNA small subunit methyltransferase H (RsmH) is a mehyltransferase responsible for the N^4-methylation of C1402 in *E. coli* 16S rRNA (Wei et al., 2012) as explained in Section 3.3.2.4.

Methionyl-tRNA formyltransferase is responsible for the formylation of initiatior tRNA so that it is targeted to the P site of the ribosomes by the help of iniatiator factor

IF2. The mutants lacking this enzyme were shown to be severely growth-deficient in *E. coli* (Schmitt et al., 1998).

Isoform 2 of ribosomal protein S6 modification-like protein B which is coded by *rimKLB* located on plasmid pSCL4 in *S. clavuligerus*. RimK protein isolated from *E. coli* and RimKLB protein found in mammals are categorized under ATP-grasp superfamily of enzymes (Fawaz et al., 2011; Collard et al; 2010). RimKLB protein is responsible for the production NAAG and β-citrylglutamate (Fawaz et al., 2011). RimK in *E. coli*, on the other hand, was first identified to carry out the post-translational modification of the 30S ribosomal protein S6 by adding Glu residues at the C-terminal of the protein (Kang et al., 1989). Moreover, it is suggested that the function of RimK protein is not limited to the modification of ribosomal protein, but might have diverse functions in bacteria identification (Zhou et al., 2013), yet unpredictable for DEPA.

Although certain extra-ribosomal functions of ribosomal proteins including EF-Tu and ribosomal protein S1 in bacteria are known (Warner and McIntosh, 2009), it is not yet possible to suggest a link between the downregulated proteins in this category and CA overproduction in the industrial strain

3.4.2.7. Stress-Related, Protein Turnover, Chaperones

This category contained 4 proteins: Trigger factor and DnaK protein, alkaline serine protease and secreted trypsin-like serine protease.

Trigger factor is a chaperone preventing misfolding or aggregation of the proteins during translation (Wong and Houry, 2004).

DnaK protein is a heat shock protein helping proteins assume the right conformation (Bucca et al., 1997). It was also explained in Section 3.3.1.5.

Trypsin like-serine proteases belong to serine protease family and the presence of them was found in *Streptomyces* sp. (Sinha et al., 1991). Trypsin hydrolyzes the lysil and arginyl bonds in the peptide (Yamashiro et al., 1997).

Alkaline serine protease has already been explained in Section 3.3.2.5.

3.4.2.8. Secondary Metabolism

There were 19 proteins categorized in this group. Of these, *avaA2* protein was shown to be the commom protein that was identified by both MALDI-TOF/MS and LC-MS/MS techniques.

Deacetoxycephalosporin C hydroxylase protein involved in cephamycin biosynthesis was explained in the upregulated proteins of AK39 and TB3585 strains in Sections 3.2.1.5 and 3.3.1.6, respectively. It was quite expected that deacetoxycephalosporin C hydroxylase is downregulated in DEPA strain. Since the mutagenesis processes that this strain has undergone is unknown, we can hypothesize that one or more of the pathways leading to the cephamycin C biosynthesis can be deficient. However, the production of CcaR protein located in cephamycin C gene cluster must be intact and expressed efficiently since CcaR is the regulator of both cephamycin and clavulanic acid.

Proclavaminate amidinohydrolase protein coded by *pah2* is involved in the early steps of clavulanic acid biosynthesis as well as clavam biosynthesis. It is known that *pah* gene have two paralogous genes: *pah1* and *pah2*. It was also shown that *pah1* mutants do not show considerable differences in the levels of CA production as compared to the parental strain (Jensen et al., 2004). The downregulation of *pah2* gene might suggest that industrial strain mostly prefers the transcriptional product of *pah1* gene over that of *pah2* gene for CA production.

Moenomycin biosynthesis protein (MoeA5) is involved in moenomycin antibiotic biosynthesis. Moenomycins are small phosphoglycolipid antibiotics produced by several *Streptomyces* species. Moenomycin A is composed of 3 main structures: a pentasaccharide, a phosphoglycerate, and a C25 isoprenyl (moenocinyl) lipid tail (McCranie and Bachmann, 2014). They inhibit the peptidoglycan glycosytransferases (PGTs) which carry out the process just before the addition of peptide chains to the peptidoglycan layers to form the peptidoglycan cell wall (Ostash et al., 2013). Moenomycins are the only natural products that can inhibit PGTs and were even shown to be effective against vancomycin-resistant pathogens (VRE) and methicillin-

resistant pathogens (MRSA) though they are not as effective against Gram-negative bacteria (Ostash et al., 2013; McCranie and Bachmann, 2014). On the other hand, moenomycins do not have application for human use but is just used as animal growth promoters. This is because they have such unfavorable physicochemical properties for human intake as low oral bioavailability and long half-life in the body. This must be a probable result of its structure containing long lipid chain (Ostash et al., 2009; Ostash et al., 2013). Moenomycin is produced in the late stages of *Streptomyces* life cycle. Furthermore, it was shown that *moeA5*, *moeO5*, *moeR5*, and *moeE5* genes, which are involved in moenomycin biosynthesis, contain rare TTA leucine codons and it is also known that Leu-tRNA^{UUA} is increased in the cell in the late stationary phase of *S. colicolor* life cycle, which suggests the presence of a control mechanism for the moenomyin biosynthesis (Ostash et al., 2007). Makitrynskyy et al. (2013) showed that *bldA, adpA* and *absB* gene products affect the moenomycin biosynthesis. *bldA* codes for rare Leu-tRNA^{UUA}. *absB* gene affects the expression of *adpA* gene, *bldA* regulates the translation of *adpA* mRNA which contains UUA-codon and AdpA, in turn, activates the transcripton of *bldA* and several *moe* genes (Makitrynskyy et al., 2013). However, interestingly, *moe* genes in *S. clavuligerus* do not contain TTA codon, suggesting different regulatory mechanism(s) are exploited for the biosynthesis of moenomycins. Since the organization and sequences of the *moe* genes show considerable similarity, it was suggested that these genes were transferred horizontally between streptomycetes (Ostash and Walker, 2010).

The downregulation of MoeA5 in DEPA strain indicates the possibility of the biosynthesis of a moenomycin derivative and also the downregulation of another antibiotic biosynthetic gene unrelated with CA

UniProtKB shows that γ-butyrolactone biosynthesis protein coded by *avaA2* has an A-factor biosynthesis hotdog domain. Niu et al. (2016) showed that this protein has 32% amino acid sequence identity to AfsA protein that is involved in the biosynthesis of a γ-butyrolactone called A-factor (Niu et al., 2016), which is important for the secondary metabolism and aerial mycelium formation in *S.griseus* (Ohnishi et al., 2005). In *S. clavuligerus* genome project of Medema et al. (2010), ScbA/AfsA-like putative

butyrolactone biosynthetic proteins were detected both on the chromosome and megaplasmid: three on the chromosome (SCLAV_0463, SCLAV_0471, SCLAV) and one on the megaplasmid (SCLAV_p0812). Two AfsA domain-containing butyrolactone biosynthetic proteins were shown within a PKS gene cluster of *S. clavuligerus* and suggested to be regulated by a γ-butyrolactone signaling pathway (Medema et al., 2010). Moreover, ScaR/Brp, a γ-butyrolactone receptor protein, was shown be involved in CC and CA production in *S. clavuligerus* (Kim et al. 2004; Santamarta et al. 2005). Medema et al. (2010) identified this gene as the megaplasmid-encoded SCLAV_p0894, which appeared to be the sole γ-butyrolactone receptor protein gene in the entire genome. In the present study, the 2.6 fold downregulated γ-butyrolactone biosynthetic protein was a chromosomal one.

Putative aminoglycoside 2-N-acetyltransferase encoded by *aac2*, with up to 23 fold downregulation, was the most drastically affected enzyme in DEPA strain. Aminoglycoside 2-N-acetyltransferases (AACs) are responsible for acetylating amino groups of aminoglycoside antibiotics that inhibit the protein synthesis by binding to the 16S rRNA of the 30S ribosome. The acetylation of the antibiotic confers bacterial resistance against these molecules. There are several types of AACs depending on their acetylation sites on the aminoglycosides: AAC(1), AAC(2'), AAC(3), and AAC(6'). Aminoglycoside 2-N-acetyltransferase are well-characterized in *Mycobacterium* (Vetting et al., 2002; Hegde et al., 2001). For example, AAC(2')-Ib enzyme coded by aac(2')-*Ib* gene in *M. fortuitum* catalyzes the acetylation of 2'-amino groups of such aminoglycosides as gentamicin, tobramycin, kanamycin B (Aínsa et al., 1997). There has not been revealed aminoglycoside gene clusters in the draft genome sequence of *S. clavuligerus* (Medema et al., 2010), thus the presence of such a resistance mechanism might suggest a horizontal gene transfer.

Polyketide synthases (PKSs) are responsible for the biosynthesis of polyketides (PKs) which can be used as commercially important molecules like antibiotics, immunosuppressants, anticancer agents, antiparasitic agents. (Katz and Donadio, 1993; Carreras and Santi, 1998). Furthermore, more than half of the biologically active, natural polyketides are produced by actinomycetes (Pfeifer and Khosla, 2001).

Generally, PKSs are divided into three categories in bacteria: type I PKSs, type II PKSs, type III PKSs (Shen, 2003). PKSs are multifunctional enzymes. Type I PKSs are comprised of polypeptides, each of which contain multicatalytic sites while type II PKSs are comprised of polypeptides, each of which contain only one active site. Both type I and type II PKSs contain one or more modular structures that usually contain three domains: AT domain, ketosynthase (KS), ACP domain. Type III PKSs, on the other hand, are comprised of monofunctional enzymes (Liou and Khosla 2003). They are homodimeric enzymes that perform the condensation reactions of the substrates (Shen, 2003). Each module on PKSs as described above carries out one cycle for the elongation of PK (Yoon et al., 2002). Modules can also contain domains other than described above such as ketoreductase (KR), dehydratase (DH), enoyl reductase that will modify the extending units. The number of extension modules, presence of modifying domains on the modules as well as starter units and extender units such as malonyl-, methylmalonyl-, and ethylmalonyl-coenzyme A have important impact on the structural and functional complexity and diversity of the polyketides (Bonnett et al., 2013). The biosynthesis of anthracyclines, like doxorubicin and its precursor, daunorubicin, for example, are produced through type II PKS system (Minotti et al., 2004; Castaldo et al., 2008). Both are used as antitumorigenic agents (Grimm et al., 1994) but have a drawback of causing cardiomyopathy (Minotti et al., 2004). During the synthesis of these antibiotics, propionyl-CoA is used as starter units and nine malonyl-CoA as extender units to produce decaketide intermediate which is then converted to the first enzyme-free intermediate aklanonic acid. The gene cluster for the daunorubicin biosynthesis consists of *dpsA*, -*B*, -*C*, -*D*, -*G*, -*E*, -*F*, and -*Y* genes in *S. peucetius* (Grimm et al., 1994; Castaldo et al., 2008).

Cytochrome P450 protein which was shown to be downregulated in TB3585 according to LC-MS/MS analysis was already explained in detail in Section 3.3.2.6,

cvm1, cvm2, cvm4, cvm5 and *cvm6* gene products are all involved in the biosynthesis of clavam biosynthesis. These proteins were also shown to be downregulated in LC-MS/MS results of AK39 and TB358 and extensively discussed in Section 3.2.2.8.

3'-hydroxymethylcephem-O-carbamoyltransferase coded by *cmcH* is required for the final steps of the CC biosynthesis and catalyzes the carbamoylation reaction of the DAC intermediate (Öster et al., 2006). It may be an important finding showing underexpression of an enzyme catalyzing a specific step in CC production, since this would reflect a slowdown in a parallel secondary pathway.

Goadsporin peptide in *Streptomyces* was shown to stimulate both the sporulation, the antibiotic and pigment production (Onaka et al., 2001). It is a linear peptide mostly composed of non-polar amino acids (Onaka, 2009). Furthermore, it also contains unusual amino acids such as oxazole and dehydroalanine derived from serine and thiazole derived from cysteine as well as methyoxazole derived from threonine. Goadsporin is synthesized on ribosomes and then subjected to the posttranslational modifications (Onaka et al, 2005; Onaka, 2009).

Although FkbH domain containing proteins have no clear function, they are suggested to play roles in polyketide synthesis. For example, *Streptomyces hygroscopicus* produces a secondary metabolite, FK520 (ascomycin). FK520 gene cluster contains *fkbGHIJK* genes. Of these, *fkbH* gene product is suggested to be responsible for transferring a glycerol-drived intermediate to an ACP protein, though the function of it could not be elucidated completely (Walton, et al., 2006).

Puromycin is an aminoacylnucleoside antibiotic produced by *Streptomyces alboniger*. Its main function is to cease the protein synthesis targeting the peptidyl-tRNA on 70S and 80S ribosomes thus forming a peptidyl-puromycin structure. *S. alboniger* is also sensitive to its own puromycin. As observed in many other antibiotic producing organisms, they have mechanisms to produce the self-resistance to their own products (Tercero et al., 1993). Puromycin N-acetyltransferase (PAC) provides resistance required against puromycin by acetylating the puromycin. Furthermore, it was also shown that *O*-demethylpuromycin, which is the penultimate product in the puromycin biosynthesis and also toxic to ribosomes, is acetylated by PAC enzyme, suggesting that acetylation is initiated just before the puromycin production (Vara et al., 1985; Cundliffe, 1989).

3.4.2.9. Energy Production/Electron-Iron Transfer

Ferredoxin is an iron-sulfur protein responsible for the transfer of electrons in diverse biological processes (Trower et al., 1990).

3.4.2.10. Cell Processes (Shape/Division/Motility)

FtsZ protein is a GTPase protein involved in bacterial cell division which includes septation, the constriction of the cell wall and cell membrane processes. During the septation process, FtsZ protein is accumulated at the division site forming a ring-shaped septum that is needed for cell membrane constriction (Löwe and Amos, 1998).

Teichoic acid biosynthesis protein encoded by *tagC* is involved in teichoic acid (TA) synthesis. Most of the Gram-positive bacteria are embellished with teichoic acids that are attached to either peptidoglycan layers or cell membrane. The ones attached to peptidoglycan cell walls are called wall teichoic acids (WTAs) and the ones anchored to bacterial membrane are lipoteichoic acids (LTAs). They are generally glycopolymers containing phosphodiester-linked polyol repeat units (Brown et al., 2013). However, their chemical structure differs greatly in detail not only between LTAs and WTAs but within the WTAs. WTAs can be composed of ribitol phosphate, glycerolphosphate (GroP) or more complex sugar-containing polymers (Reichmann and Gründling, 2011) which are covalently attached to peptidoglycan through a linkage unit that usually consists of a GlcNAc-ManNAc disaccharide (Xia and Peschel, 2008). LTAs have much simpler structures and are comprised of polyglycerolphosphate (PGP) chain which are linked to the bacterial cell membrane through a glycolipid anchor (Reichmann and Gründling, 2011). The presence of WTAs and LTAs seem to be quite important for the organism. For example, in the absence of WTAs, it was observed that cells lose their shapes, especially rod-shaped bacteria turned to have spherical shapes. Increase in the cell size or deficiencies in septation, unequal thickening along the peptidoglycan cell wall were other problems (Brown et al., 2013). As for LTA deficient mutants, morphological and growth defects were observed (Reichmann and Gründling, 2011). Moreover, WTAs were shown to bind to extracellular metal cations and even protons in the environment, thereby, affecting the pH distribution across the wall and hence regulating the activity of

enzymes. Their binding to ions may regulate the osmotic pressure inside and outside the cell so that cell is protected against unfavorable conditions and this might also help the cell integrity and shape be kept stabilized by decreasing the repulsion between the nearby phosphate groups. Furthermore, it was shown that blocking the WTA synthesis in MRSA or *B. subtilis* conferred these organisms sensitive to β-lactam antibiotics, further showing the protective importance of WTAs against penetrating molecules (Brown et al., 2013). Therefore, the presence of teichoic acids on the cell surface might be an additional protective layer for the microorganism.

UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase, which is also known as MurF, is involved in peptidoglycan cell biosynthesis. MurF with 3 other enzymes (MurC, MurD, MurE) is responsible sequentially catalyzing the addition of pentapeptide chain to UDP-N-acetylmuramic acid moiety in the cytoplasm (Shanmugam and Natarajan, 2012).

Alanine racemase domain protein is also involved in peptidoglycan biosynthesis. The pentapeptide chain that is added to UDP-N-acetylmuramic acid contain amino acids in D-configuration to form UDP-N-acetylmuramyl-LAla-D-Glu-meso-diaminopimelate-D-Ala-D-Ala required for peptidoglycan biosynthesis. However, it is not natural to find amino acids in D-amino acid configuration. Hence, D-alanines (D-Ala) in the structure of pentapeptide is formed by the activity of alanine racemases through the conversion of L-Ala to D-Ala. D-Ala is then ligated with another D-alanine by D-Ala: D-Ala ligase forming D-Ala dipeptide which is then incorporated into the UDP-N-acetylmuramyl-L-Ala-D-Glu-*meso*-diaminopimelate chain by D-Ala-D-Ala adding enzyme (Shaw et al., 1997).

As seen, all of the above-mentioned proteins are required for peptidoglycan biosynthesis, therefore their significant downregulation pointed to a general slowing down in cell division compared to the parental strain. Of interest was as many as 83 times decreased levels of TA biosynthetic protein and localization of the gene encoding this protein on pSCL4 since this megaplasmid did not seem to encode any functions essential to primary metabolism, as reported earlier (Medema et al., 2010).

3.7.2.11. Inorganic/Organic Molecule Transport

This category contained 6 proteins: ABC-Fe transporter, amino acid ABC transporter amino acid-binding protein, ABC transporter ATP-binding protein, sugar transporter sugar-binding protein, preprotein translocase subunit SecE and MMPL domain protein.

ABC transportors are important means of transporting molecules, either as exporters or importers, in bacteria (Rees et al., 2009) and were explained in Section 3.3.2.8. Importers require a specific protein attached to the cell membrane in Gram-positive bacteria for the recognition and hence uptake of the substrate. These proteins are called solute-binding proteins (Singh and Röhm, 2008). Iron is an important micronutrient source for all living organisms (Sebulsky et al., 2003). However, the stable form of the iron (ferric iron) is not soluble and thus has to be extracted from the environment (Hanks et al., 2005). Therefore, bacteria generated several systems to extract and transport the iron from the environment, one of which is the production of low-molecular weight proteins called siderophores. Siderophores can bind Fe (III) with high affinity and then they interact with the solute binding proteins that are part of ABC transporters to transport the acquired iron into the cell (Sebulsky et al., 2003). Solute-binding proteins are highly specific and bind their substrates with high affinity which might be the reason for their high selectivity towards the substrates. In *E. coli*, BtuF is the solute-binding component of BtuCD ABC transporter system responsible for the vitamin B12 transport (Borths et al., 2002) or FhuD component of an ABC transporter system in *E.coli* binds specifically to iron-hydroxamate siderophore complexes to transport them into the cell (Sebulsky et al., 2003).

The function of sugar transporter sugar-binding protein was not specified or the domains of the protein could not be predicted by UniProtKB database. However, it is known that this protein has a signal peptide sequence and an extracellular protein according to Gpos-mPLoc database. Moreover, STRING database predicts that this protein can have association with several ABC transporter permease proteins (data not shown). Nonetheless, type of the transporter and the substrate specificity of it needs to be elucidiated.

Sec is one of the two secretion pathways found in bacteria. Sec pathway mediates the translocation of unfolded proteins across the membrane or insertion of them into the cell membrane. SecA protein is a constituent of protein conducting channel (PCC) part of the system (Natale et al., 2008). Interestingly, although SecE protein was shown to be downregulated according to LC-MS/MS results, SecA protein was upregulated according to MALDI-TOF/MS results (3.4.1.10).

MMPL (*Mycobacterium* Membrane Protein Large) is suggested to be a subgroup of resistance nodulation division (RND) family transporters and was observed in *M. tuberculosis* (Pasca et al., 2005). RND proteins were thought to be present in only Gram-negative bacteria until the discovery of MMPL proteins in *M. tuberculosis*. RND proteins are usually associated with the multidrug efflux systems but involved in the efflux of a large variety substrates. There are 14 annotated MMPL proteins in *M. tuberculosis* (Sandhu and Akhter, 2015).Some of the *mmpL* genes were shown to be co-localized with lipid biosynthesis or polyketide synthesis genes, suggesting that they might be responsible for the lipid and polyketide transport. For example, MmpL8 transports a precursor required for the biosynthesis of sulfolipid 1 (Domenech et al., 2005) while MppL7 protein is involved in drug efflux, being responsible for the isoniazid efflux in *M. smegmatis* (Pasca et al., 2005; Sandhu and Akhter, 2015).

Proteins involved in transport mechanisms with diverse functions from sugar, iron and protein transport to drug efflux are less abundant in DEPA. Given the proteins in previous categories, it can be deduced that DEPA strain minimizes the metabolic processes that are related with growth or central metabolism.

3.4.2.12. Hypothetical/Unknown Proteins

There were 49 proteins identified in this category. Of these, putative transcriptional regulator AraC family, putative M28-family peptidase and putative alpha-glucosidase proteins appeared in two to three spots, suggesting that they have undergone PTMs. Moreover, putative transcriptional regulator AraC family was identified by both MALDI-TOF/MS and LC-MS/MS analyses. As AraC family proteins (Section 3.3.1.9), M28-family peptidase (3.2.2.12), TetR family of proteins (Section 3.2.2.5 and

3.4.2.5) and two-component systems (Chapter 1 and Section 3.2.1.3) have already been discussed earlier, only a couple of proteins will be mentioned from now on as follows:

α glucosidases are α-glycosidic *O*-linkage hydrolases and release the D-glucose moieties from the carbohydrate substrates (Chiba, 1997).

According to UniProt KB database, putative fatty acid oxidation complex alpha-subunit (SCLAV_4844) shown to be downregulated 71.4-fold in this study has a 3-hydroxyacyl-CoA dehydrogenase domain. It was shown in *E. coli* that 3-hydroxyacyl-CoA dehydrogenase is a part of a multienzyme complex involved in fatty acid oxidation. 3-hydroxyacyl-CoA dehydrogenase forms the α subunit of this complex together with enoyl-CoA hydratase, cis-A3-trans-A2-enoyl-CoA isomerase, and 3-hydroxyacyl-CoA epimerase and is coded from *fadAB* operon (Yang and Schulz, 1983).

DUF2587 domain-containing protein was searched through STRING database (data not shown). It has associations with at least 10 proteins, many of which are unknown. However, it was shown to interact/associate with three proteins related with proteolytic degradation of proteins; thus, this protein might be a chaperone-like protein or somewhat involved with protein degradation under unfavorable conditions.

Another DUF domain containing protein is DUF1254 multi-domain protein. STRING data gave an association network of this protein with proteins of unknown function except for one (Figure 3. 32). The only known protein that DUF1254 multi-domain protein is interacting is cephalosporin hydroxlyase (CmcI) which is involved in cephamycin C biosynthesis.

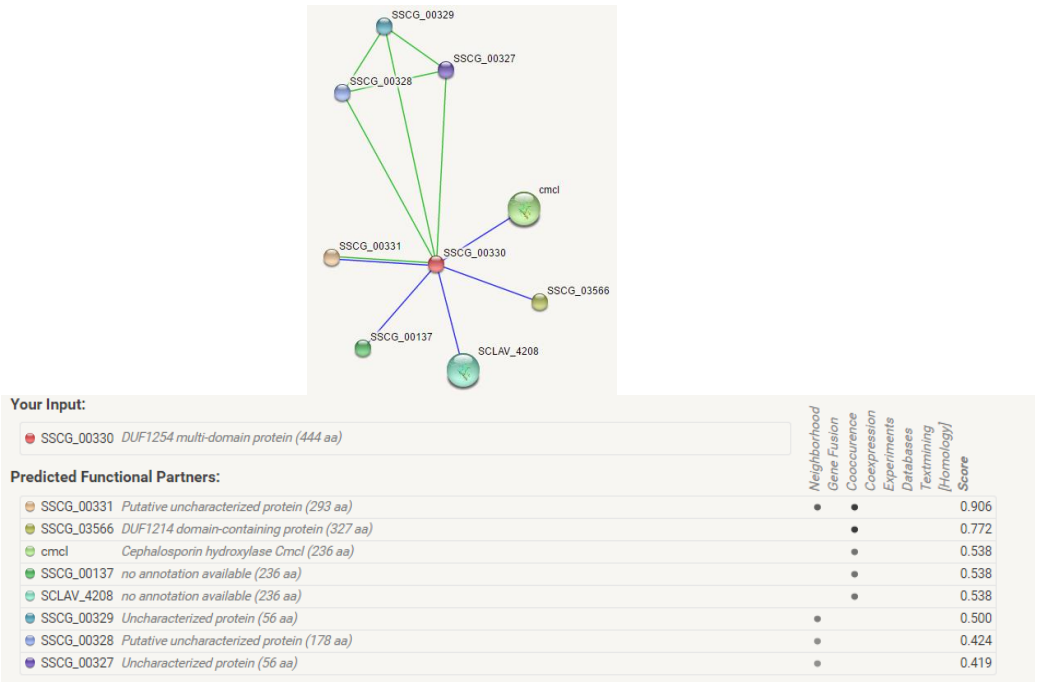


Figure 3. 32. The association network of DUF1254 multi-domain protein with other proteins on STRING database.

Putative epimerase/dehydratase protein was predicted to be a NAD-dependent epimerase/dehydratase according to UniProtKB and explained in Section 3.2.2.13.

Putative membrane protein was predicted to have an MMPL domain by UniProtKB database. MMPL domin proteins were explained in Section 3.4.2.11. These proteins were suggested to belong to RND family transporters and mainly involved in lipid and polyketide transport (Pasca et al., 2005; Domenech et al., 2005).

Interestingly, putative esterase protein was predicted to contain β-lactamase-like domain by UniProtKB database. β-lactamases were explained in both Chapter 1 and Section 3.4.1.11.

Ornithine aminotransferases are involved in the biosynthesis of L-glutamate (Yasuda et al., 1979), which is used in diverse processes as intermediate.

According to UniProtKB database, hydrolase superfamily dihydrolipoamide acyltransferase-like protein (SCLAV_0456) contains 2-oxoacid dehydrogenase acyltransferase domain and this domain is observed in the lipoamide acyltransferase component of the BCDH complex (Mattevi et al., 1993). Thus, this protein might be a counterpart of dihydrolipoamide acyltransferase (E2) of the BCDH complex. Nevertheless, while putative dihydrolipoamide acyltransferase has a length of 507 aa, hydrolase superfamily dihydrolipoamide acyltransferase-like protein is a 286 aa-long protein. EggNog database was also searched to find out the orthologues and thereby proposed function of the protein; the results showed that there were 19 orthologues from 18 species in the bacterial kingdom but no known function was predicted for this protein.

The function of the CHAD domain could not be characterized yet. However, it is known to contain charged amino acid residues that can produce a strong polar surface, thus may be acting as metal chelators or phosphoacceptors (Iyer and Aravind, 2002).

WGR domain is not characterized, as well and is known to contain conserved tryptophan, glycine, and arginine residues. However, it is suggested that it might have a nucleoside-binding function (Loseva et al., 2010).

When STRING database was searched for DUF1906 multi-domain protein, it was predicted to be associated with clavaminate synthase and clavaminic acid synthase 1 proteins as well as helicase protein. However, since this protein is completely unknown, it requires to be characterized to unravel its probable relationship with the secondary metabolism of the organism.

Putative uncharacterized protein (SCLAV_5506) which was predicted to be extracellular was searched through both the STRING and EggNog databases in order to understand the probable association network as well as orthologues of the protein. However, STRING data (data not shown) indicated that it interacts with a protein with no annotation. EggNog database found 10 orthologues of this protein in 10 species of the Bacterial kingdom, with no known function.

SMP-30/gluconolaconase/LRE domain-containing protein contains TolB-like domain aside from SMP-30/gluconolaconase/LRE domain according to UniProtKB database. TolB proteins are the periplasmic components of tol-dependent translocation system in *E. coli* that transports Group A colicins across the membrane. This translocation system also imports the DNA of filamentous phages. Association of TolB protein with the outer membrane is achieved through its interaction with the peptidoglycan-associated lipoprotein (Pal) (Carr et al., 2000). TolB proteins contain WD repeats. The proteins that harbor WD repeats were well-described in eukaryotes and shown to be involved in a large variety of processes such as signal transduction, transcription, splicing, cytoskeletal organisation and vesicular fusion (Neer et al., 1994; Carr et al., 2000); however, proteins with WD repeats identified in prokaryotes are not many in number; some of them that were identified include are serine/threonine protein kinase PkwA from *Thermomonospora curvata*, the HatA and HatR proteins from *Synechocystis* sp. PCC6803, and the prolyl oligopeptidase homologue from *Bacillus subtilis* (Ponting and Pallen, 1999; Carr et al., 2000). On the other hand, protein identified in this study was predicted to be a cytoplasmic protein, which contradicts the periplasmic property of a protein. Therefore, it was searched through both the STRING (Figure 3.33) and the EggNog databases for its association network and predicted functions of its orthologues. It was shown to be associated with considerable number of proteins; most of which were unknown. Interestingly, it was predicted to be interacting/associated with a β-lactamase domain-containing protein and glyoxalase/bleomycin resistance protein/dioxygenase, both of which also require to be identified. EggNog database retrieved several orthologues of this protein with no predicted function. Hence, this protein requires further characterization.

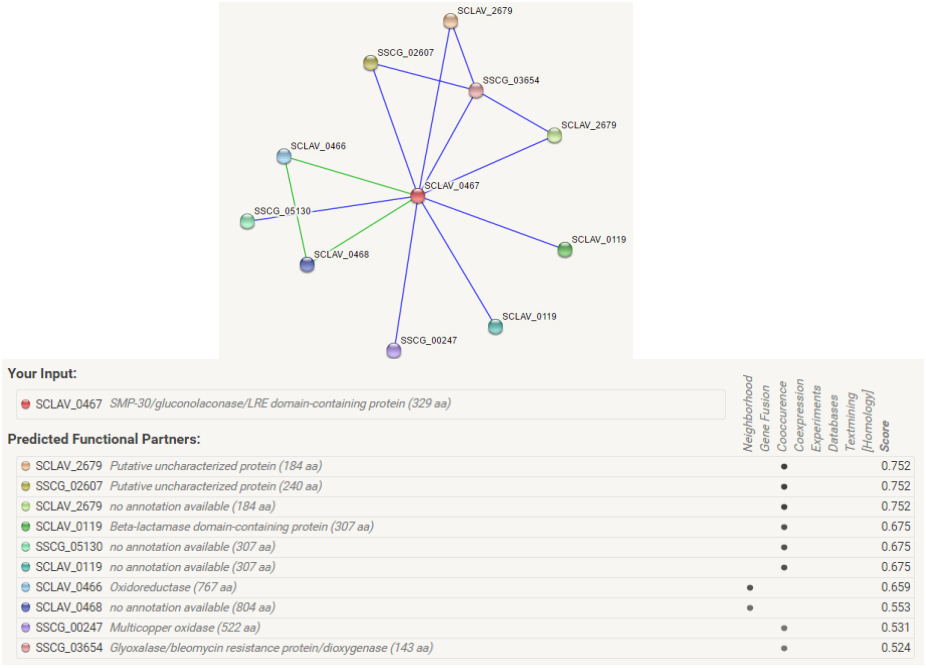


Figure 3. 33. STRING data showing the associaton of SMP-30/gluconolaconase/LRE domain-containing protein with other proteins.

Lipoprotein (SCLAV_4744) identified in this study was also searched through EggNog database and was shown to have orthologues with unknown functions. On the other hand, UniProtKB database predicts that it as a putative adhesion protein.

3.4.2.13. Others/General Function

There were 34 proteins shown to be downregulated in this category. Of these, according to 2DE MALDI-TOF/MS results, acetyltransferase GNAT family protein were shown to be located in two spots which suggests a PTM with a charge modification. Amidohydrolase was shown in 6 spots which must be a result of an extensive PTM on the protein. Moreover, bacterial luciferase domain-containing protein and amidohydrolase proteins were the common proteins that were identified by both MALDI-TOF/MS and LC-MS/MS analysis.

Phospholipid-binding protein was shown to be downregulated 22.7-fold in DEPA strain. UniProtKB database predicts that this protein belongs to phosphatidylethanolamine-binding protein (PEBP) family. Proteins in PEBP family

have conserved sequences and do not show considerable sequence similarity to proteins known to date (Banfield et al., 1998). They have been found in various organisms such as bacteria, yeast, nematodes, plants, drosophila and mammals. Their function is usually associated with signalling pathways; for example, they were shown to be inhibitors of several serine proteases or in MAP kinase and the NF-κB pathways (Vallée et al., 2003). Two proteins, YBHB and YBCL, which were isolated from *E. coli* were also shown to be the members of PEPB family and suggested to have multiple cellular partners (Serre et al., 2001). Protein identified in this study was predicted to be extracellular or localized in cellular membrane. This might suggest that it is involved in some cellular signalling mechanism. However, the function and importance of the protein needs further characterization.

GCN5-related N-acetyltransferase (GNAT) superfamily was already explained in Section 3.2.2.13.

According to UniProtKB database, oxygenase identified in this study with a 5.3 fold decreased level belongs to clavaminate synthase family which also includes VioC, MppO and AsnO proteins. VioC is produced by *Streptomyces vinaceus* and is alpha-ketoglutarate-dependent L-arginine hydroxylase that catalyzes Cβ-hydroxylation of L-arginine. It is involved in the biosynthesis of tuberactinomycin antibiotic viomycin which is produced through a nonribosomal peptide synthetase (NRPS) mechanism (Helmetag et al., 2009). AsnO produced by *Streptomyces coelicolor* is L-asparagine oxygenase that catalyzes the conversion of L-asparagine to (2S,3S)-3-hydroxyasparagine and is involved in the biosynthesis of calcium-dependent antibiotic (CDA) which is also synthesized by a NRPS system. AsnO enzyme produced by *Streptomyces coelicolor* shows almost 33% sequence identity to the clavaminate synthase protein which is coded by either *cas1* or *cas2 in S. clavuligerus* (Strieker et al 2007). Enduracididine β-hydroxylase (MppO) catalyzes β-hydroxylation of a nonproteinogenic amino acid and is involved in the biosynthesis of mannopeptimycins (MPPs) in *Streptomyces hygroscopicus* which are glycopeptide antibiotics (Haltli et al., 2005). Oxygenase found in this study showed 70% sequence coverage and 35% identity to VioC protein from by *Streptomyces vinaceus*, 90% sequence coverage and

38% identity to MppO from *Streptomyces hygroscopicus* and 85% sequence coverage and 34% identity to AsnO from *Streptomyces coelicolor*. Furthermore, the blastp results for oxygenase and CAS coded by *cas1* in *S. clavuligerus* showed 73% sequence coverage and 34% identity, for oxygenase and CAS coded by *cas2* in *S. clavuligerus* showed 73% sequence coverage and 33% identity. Downregulation of this protein in DEPA strain suggests that it might be involved in clavam biosynthesis other than CA and hence has decreased expression levels in order to compensate CA production.

Amidohydrolase superfamily is comprised of diverse members with diverse substrate specificities. Nevertheless, their main function is to hydrolyze the amide or ester groups located at carbon or phosphorus centers of the substrate (Seibert and Raushel, 2005). Since in this study, protein identified as amidohydrolase could not be characterized specifically, it requires further study to understand the function of it in *S. clavuligerus* and the reason for its downregulation in DEPA strain.

Aldehyde dehydrogenase (ALDH) comprises a large superfamily that contain enzymes catalyzing the oxidation of a variety of aromatic and aliphatic aldehydes. This protein was also explained in Section 3.2.1.9. In this study, four types of aldehyde dehydrogenases (one from MALDI-TOF/MS results and three from LC-MS/MS results) were found to be downregulated. These proteins are probably important for the central metabolism of the organism and hence downregulated in DEPA strain in order to compensate the overproduction of the antibiotics.

According to UniProtKB database, bacterial luciferase domain-containing protein belongs to luciferase-like, F_{420}-dependent oxidoreductase family. F_{420} is a derivative of FMN and is bound to enzymes non-covalently. In addition, it is mostly produced by archaeal methanogens and *Actinomycetes* (Taylor et al., 2013). In several *Streptomyces* spp., F_{420} -dependent oxidoreductases were shown to be involved in secondary metabolite production (Bown et al., 2016). For example, chlortetracycline produced by *Streptomyces scabies* (Nakano et al. 2004), oxytetracycline produced by *Streptomyces achromogenes* (Li et al., 2009) include the participation of such reductases (Bown et al., 2016).

2-hydroxyacid-family dehydrogenases catalyze the reversible conversion of 2-ketoacids to 2-hydroxy acids with the requirement of NAD(P) as the coenzyme. There are various 2-hydroxyacid-family dehydrogenases such as D-phosphoglycerate, D-lactate, D-hydroxyisocaproate, vancomycin-resistant protein H dehydrogenaases, having different substrate preferences as well as different functions (Wada et al., 2008).

Short-chain dehydrogenase/reductase SDR proteins are NAD(P)(H)-dependent oxidoreductases are involved in a large variety of biological processes (Kavanagh et al., 2008).

Oxidoreductase (SCLAV_2295) identified in this study was suggested to belong to the short-chain dehydrogenases/reductases family according to UniProtKB database.

GAF domains confer proteins diverse functional properties such as binding to small molecules or protein-protein interactions, etc. Nonethless, GAF domain protein described here awaits for its functional characterization (Heikaus et al., 2009).

Functions of an NAD-dependent epimerase/dehydratase protein were discussed in Section 3.4.2.10. In *Streptomyces*, NAD-dependent epimerase/dehydratase genes are usually found within secondary metabolite gene clusters.

NUDIX hydrolases are commonly found in organisms and responsible for the hydrolysis of several substrates such as nucleoside di- and triphosphates, dinucleoside and diphosphoinositol polyphosphates and nucleotide sugars. Members of this family show different substrate specificities; for example, MutT from *E. coli* is responsible for the degradation of oxidized nucleotides while others can control the accumulation of metabolic intermediates that are produced during biological processes. In actinomycetes and some *Streptomyces* spp., the number of Nudix genes can vary from 20 to 90 (McLennan et al., 2006), which indicates the probable diverse functions of the members of this family.

3.4.3. Summary for DEPA Upregulated and Downregulated Proteins

Having undergone numerous screening and mutagenesis processes, DEPA strain is a completely mysterious CA overproducer. Comparative proteome analysis of DEPA with the parental strain aimed to be able to shed some light on changes in its metabolic activities. There was a total of 80 upregulated and 174 downregulated proteins in DEPA identified by both MALDI-TOF/MS and LC-MS/MS techniques. Distribution of upregulated and downregulated proteins to their functional categories are shown in Figure 3. 34 and Figure 3. 35, respectively.

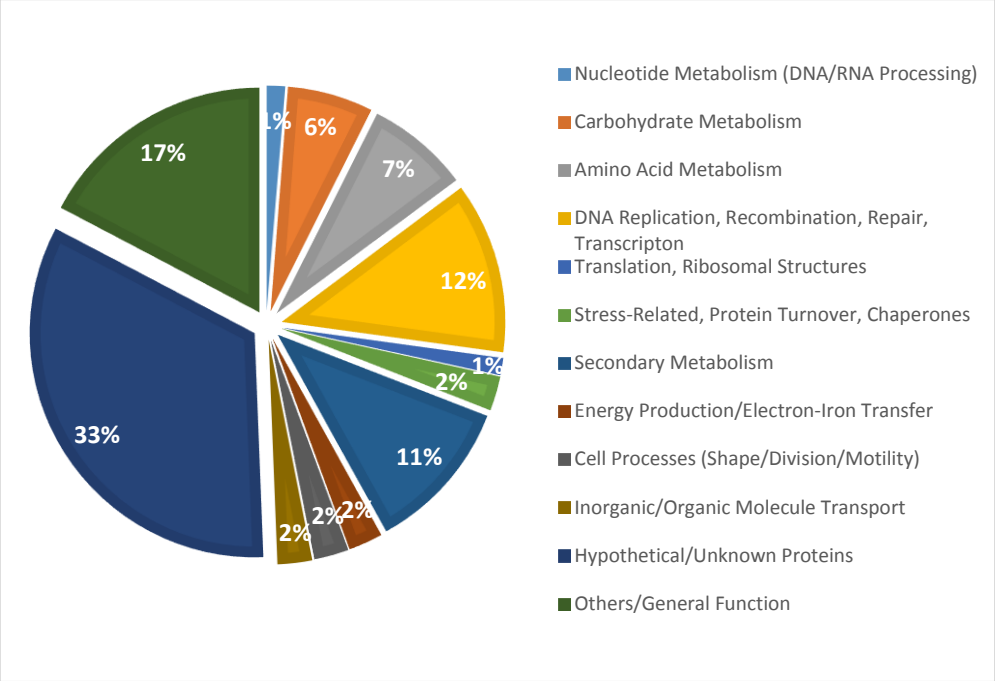


Figure 3. 34. Relative quantitative distribution of upregulated proteins identified in DEPA to the functional categories. Percentage of each category was calculated by dividing the number of proteins in each category to the total number of proteins in all categories.

"Hypothetical/Unknown Proteins" constituted the largest area in both groups. There were several putative regulatory proteins identified as upregulated or downregulated in this category, and they await for functional characterization in order to show their

relation with the CA overproduction. Upregulated proteins in "DNA replication, Recombination, Repair, Transcripton" category was higher than those in "Secondary Metabolism". LC-MS/MS results showed upregulation of several regulatory proteins in the former group such as two component transcriptional regulators, or DNA binding proteins, suggesting their probable participation in regulation of CA production. Interestingly, two proteins SCLAV_0037 and SCLAV_0043 shown to be upregulated in DEPA were proposed to be involved in T6SS just like SLAV_0035 protein upregulated in AK39. This finding is important in being the first report suggesting the presence of T6SS in *S. clavuligerus*. MALDI-TOF/MS analysis could identify proteins that are directly related with CA production while LC-MS/MS protein indicated the presence of other players in the secondary metabolism.

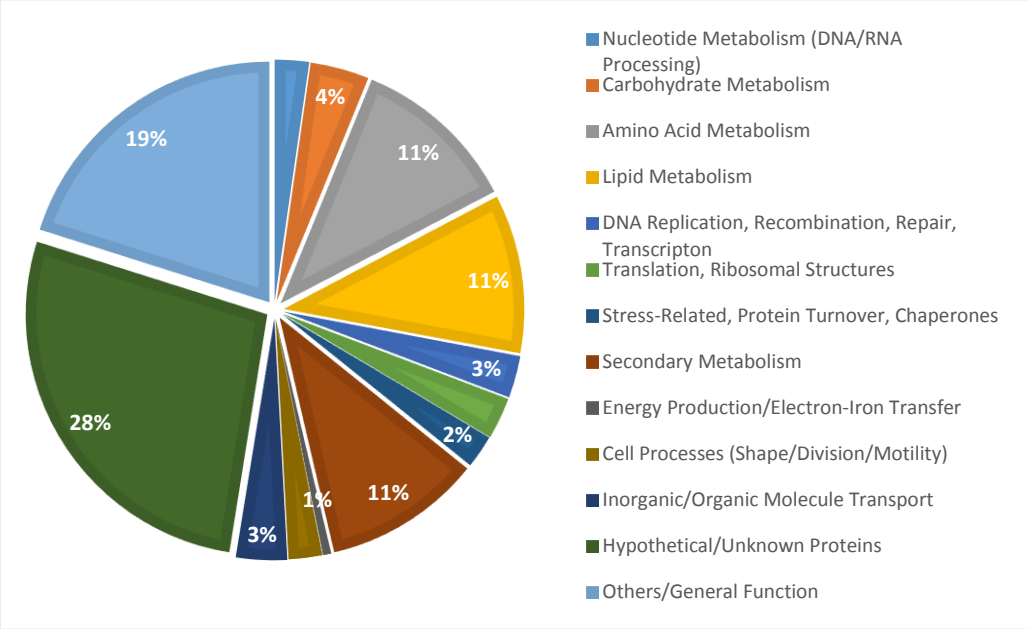


Figure 3. 35. Relative quantitative distribution of downregulated proteins identified in DEPA strain to the functional categories. Percentage of each category was calculated by dividing the number of proteins in each category to the total number of proteins in all categories.

As for the downregulated protein categories, “Amino Acid Metabolism”, “Lipid Metabolism” and “Secondary Metabolism” categories were the second most populated categories following the “Hypothetical/Unknown Proteins” category. Most of the proteins downregulated in “Amino Acid Metabolism” were involved in methionine biosynthesis suggesting its negative effect on CA production. Downregulated proteins in “Secondary Metabolism, on the other hand, showed considerable diversity ranging from ones involved in polyketide and clavam biosynthesis to γ butytolactone biosynthesis. In addition, a considerable portion of proteins in “Hypothetical/Unknown Proteins” was comprised of uncharacterized proteins and DUF-domain containing proteins, which proves the necessity for more studies to comprehend the physiological changes occurring in the mutant strain.

CHAPTER 4

CONCLUSION

- Proteomic analyses of two genetically/metabolically engineered *S. clavuligerus* strains AK39 and TB3585 as CC overproducers, and an industrial CA overproducing strain DEPA by combining (i) 2DE gel approach with MALDI-TOF/MS technique and (ii) FASP method with LC-MS/MS technique were conducted in order to reveal the protein profiles that have changed in response to genetic engineering and mutagenesis strategies and gain an insight on how the expression profiles of the proteins have changed to favor the production of one particular antibiotic among the many others, that is, pleiotropic and pathway-specific effects of manipulations on protein expression. With the identification of several over- and underexpressed proteins belonging to the major functional classes, which were either common to the strains or unique, the study almost completely reached its goals that we defined at the beginning.

- With inherent advantages and disadvantages of gel-based and gel-free proteomics, usually neither alone can provide a satisfactory picture of altered cellular physiology when used for comparative proteomic analysis. The combined approach employed in the present work allowed the identification of many more proteins accounting for β-lactam overproduction in *S. clavuligerus*. Moreover, by coupling the proteomics with the tools of metabolomics would illuminate many more critical alterations related with CC and CA overproduction.
- For both AK39 and TB3585, "Hypothetical/Unknown Proteins" category harbored higher numbers of upregulated proteins which was followed by "Secondary Metabolism". The third important cateogory was "DNA

Replication, Recombination, Repair, Transcription" in AK39 and "Others/General Function" in TB3585. Upregulated proteins of DEPA strain were ranked as "Hypothetical/Unknown", "Others/General Function", "DNA Replication, Recombination, Repair, Transcripton" and "Secondary Metabolism", respectively, according to their abundance in different categories.

- In AK39, "Others/General Function" category harbored highest number of the downregulated proteins which was followed by "Secondary Metabolism". "Hypothetical/Unknown Proteins".and "Amino Acid Metabolism" were the third most populated categories. As to the downregulated proteins of TB3585, there was a different pattern such that "Amino Acid Metabolism" was the most prominent while the secondary position was shared by "Secondary Metabolism" and "Hypothetical/Unknown Proteins". "Others/General Function" and "Lipid Metabolism" categories just followed them. Among downregulated proteins of DEPA strain, "Hypothetical/Unknown Proteins" was the most prominent group followed by "Others/General Function" category. The third position was shared by the proteins of "Amino Acid Metabolism", "Lipid Metabolism" and "Secondary Metabolism".

- Although the CC and CA overproducing strains of *S. clavuligerus* were obtained by different kinds of strain improvement approaches, it was quite impressive to find many proteins, especially those that are co-regulated in them. AK39 and TB3585 were constructed by different approaches, via *hom* deletion and multi-copy *ask* expression, respectively, that is, by targeting aspartic acid family genes at different levels of biosynthetic pathway, they had many up- and downregulated proteins in common. Because the industrial strain DEPA was improved to produce at least 100 folds more CA than the standard strain (compared to the targeted mutants producing only approximately 1.3 to 2.5 fold more CC than the standard strain) by successive cycles of random mutagenesis and screening which had taken several years of efforts of the DEPA Co. researchers, its differentially expressed proteins were much more in their number and more unique. Still, these three strains shared many of the

differentially expressed proteins related with overproduction of two different β-lactam antibiotics, CC and CA.

- Strongly elevated levels of CcaR (positive transcriptional regulator specific to both CC and CA pathways) had a key role in overproduction in all of the strains, although we do neither know the actual global actor governing its overexpression, nor the signal that motivates this global actor. There were many candidate global regulators commonly overexpressed in all overproducers such as two component system proteins or TetR family proteins.
- While there occurs to be the overexpression of proteins of primary and secondary metabolism, mainly the enzymes of CC and CA biosynthetic pathways and certain primary pathways (particularly carbohydrate and amino acid metabolism) supplying precursors to CC and CA biosynthesis and certain proteins coping with general stress, the key proteins of competing primary and secondary metabolic pathways were strongly downregulated. Downregulated ones included the enzymes negatively influencing precursor flow by competing for precursors and biosynthetic enzymes of other secondary metabolites like PK antibiotics, moenomycin, non-ribosomally synthesized peptides, etc., as well as certain important autoregulators of secondary metabolism.
- Besides the above mentioned differentially expressed proteins that could be directly linked to CC and CA overproduction, the present study could also successfully pointed to many hypothetical/unknown proteins with an altered expression pattern which would provoke researchers to conduct gene-level functional studies for a deep comprehension into their roles. In this way, the existing mysteries will be unraveled.

- Random mutagenesis is very slow and tedious for industrial fermentation researchers and might cause unintended changes to the entire system, but targeted approaches which do not have such limitations can be used to further improve the current high production strains as guided by current -omics technologies and other tools of systems biotechnology. Comparative proteomics is a powerful tool in industrial biotechnology since the information obtained by comparing two or more genetically different strains or the same

strain grown in different nutritional/environmental conditions can successfully lead to designing new strategies for strain improvement even when a limited number of protein spots could be identified. On the other hand, each x-ome alone is not sufficient since the levels of RNAs, proteins, metabolites and fluxes vary independently, but various regulatory circuits coordinate them in a highly orchestrated fashion (Lee et al., 2005). Especially for industrially important secondary metabolites, an integrated combined -omics for inspection of correlations among different x-omes is essential to better link the components of the primary and secondary metabolism, to define novel targets at gene and pathway levels and to design strategies for metabolic engineering of organisms for increased secondary metabolite titers. With some potentially crucial changes in the levels of certain proteins of primary and secondary metabolism, the present work shed light on proteome-wide changes that could contribute to increased production of CC and CA. Considering DEPA strain in particular, other mechanisms of CA overproduction could be specifically induced when it is grown in complex industrial culture medium rather than in defined SA. Moreover, coupling proteomics with the tools of especially metabolomics would illuminate many more critical alterations related with CC and CA overproduction.

REFERENCES

Abdallah, C., Dumas-Gaudot, E., Renaut, J., & Sergeant, K. (2012). Gel-based and gel-free quantitative proteomics approaches at a glance. Int J Plant Genomics. 2012.

Agarwal, A. K., & Auchus, R. J. (2005). Minireview: cellular redox state regulates hydroxysteroid dehydrogenase activity and intracellular hormone potency. Endocrinology. 146(6), 2531-2538.

Aínsa, J. A., Pérez, E., Pelicic, V., Berthet, F. X., Gicquel, B., & Martín, C. (1997). Aminoglycoside 2′-N-acetyltransferase genes are universally present in mycobacteria: characterization of the *aac (2')-Ic* gene from *Mycobacterium tuberculosis* and the *aac (2')-Id* gene from *Mycobacterium smegmatis*. Mol Microbiol. 24(2), 431-441.

Akanuma, G., Ueki, M., Ishizuka, M., Ohnishi, Y., & Horinouchi, S. (2011). Control of aerial mycelium formation by the BldK oligopeptide ABC transporter in *Streptomyces griseus*. FEMS Microbiol Lett. 315(1), 54-62.

Alexander, D. C., & Jensen, S. E. (1998). Investigation of the *Streptomyces clavuligerus* cephamycin C gene cluster and its regulation by the CcaR protein. J Bacteriol. 180(16), 4068-4079.

Alexander, D. C., Anders, C. L., Lee, L., & Jensen, S. E. (2007). *pcd* mutants of *Streptomyces clavuligerus* still produce cephamycin C. J Bacteriol. 189(16), 5867-5874.

Álvarez-Álvarez, R., Martínez-Burgo, Y., Pérez-Redondo, R., Braña, A. F., Martín, J. F., & Liras, P. (2013). Expression of the endogenous and heterologous clavulanic acid cluster in *Streptomyces flavogriseus*: why a silent cluster is sleeping. Appl Microbiol Biot. 97(21), 9451-9463.

Álvarez-Álvarez, R., Rodríguez-García, A., Santamarta, I., Pérez-Redondo, R., Prieto-Domínguez, A., Martínez-Burgo, Y., & Liras, P. (2014). Transcriptomic analysis of

Streptomyces clavuligerus ΔccaR::tsr: effects of the cephamycin C-clavulanic acid cluster regulator CcaR on global regulation. Microb Biotechnol. 7(3), 221-231.

Ammerman, J. W., & Azam, F. (1985). Bacterial 5'-nucleotidase in aquatic ecosystems: a novel mechanism of phosphorus regeneration. Science. 227(4692), 1338-1340.

Amundsen, S. K., Taylor, A. F., Chaudhury, A. M., & Smith, G. R. (1986). *recD*: the gene for an essential third subunit of exonuclease V. Proc Natl Acad Sci U S A. 83(15), 5558-5562.

Anantharaman, V. & Aravind, L. (2002). The PRC-barrel: A widespread, conserved domain shared by photosynthetic reaction center subunits and proteins of RNA metabolism. Genome Biol. 3(11), 1.

Andersen, G. R., Nissen, P., & Nyborg, J. (2003). Elongation factors in protein biosynthesis. Trends Biochem Sci. 28(8), 434-441.

Andersen, P. S., Smith, J. M., & Mygind, B. (1992). Characterization of the *upp* gene encoding uracil phosphoribosyltransferase of *Escherichia coli* K12. Eur J Biochem. 204(1), 51-56.

Anderson, A. S., Clark, D. J., Gibbons, P. H., & Sigmund, J. M. (2002). The detection of diverse aminoglycoside phosphotransferases within natural populations of actinomycetes. J Ind Microbiol Biotechnol. 29(2), 60-69.

Anderson, N. L., Esquer-Blasco, R., Hofmann, J. P., & Anderson, N. G. (1991). A two-dimensional gel database of rat liver proteins useful in gene regulation and drug effects studies. Electrophoresis. 12(11), 907-913.

Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M., & Iyer, L. M. (2005). The many faces of the helix-turn-helix domain: transcription regulation and beyond. FEMS Microbiol Rev. 29(2), 231-262.

Argyrou, A., & Blanchard, J. S. (2004). Flavoprotein disulfide reductases: advances in chemistry and function. Prog Nucleic Acid Res Mol Biol. 78, 89-142.

Argyrou, A., Vetting, M. W., & Blanchard, J. S. (2004). Characterization of a new member of the flavoprotein disulfide reductase family of enzymes from *Mycobacterium tuberculosis*. J Biol Chem. 279(50), 52694-52702.

Arima, J., Uesugi, Y., Iwabuchi, M., & Hatanaka, T. (2006). Study on peptide hydrolysis by aminopeptidases from *Streptomyces griseus*, *Streptomyces septatus* and *Aeromonas proteolytica*. Appl Microbiol Biot. 70(5), 541-547.

Armirotti, A., & Damonte, G. (2010). Achievements and perspectives of top-down proteomics. Proteomics. 10(20), 3566-3576.

Arsène-Ploetze, F., Nicoloff, H., Kammerer, B., Martinussen, J., & Bringel, F. (2006). Uracil salvage pathway in *Lactobacillus plantarum*: transcription and genetic studies. J Bacteriol.188(13), 4777-4786.

Arulanantham, H., Kershaw, N. J., Hewitson, K. S., Hughes, C. E., Thirkettle, J. E., & Schofield, C. J. (2006). ORF17 from the clavulanic acid biosynthesis gene cluster catalyzes the ATP-dependent formation of N-glycyl-clavaminic acid. J Biol Chem. 281(1), 279-287.

Asano Y., Ito H., Dairi T., Kato Y. (1996). An alkaline D -stereospecific endopeptidase with β-lactamase activity from *Bacillus cereus*. J Biol Chem. 271(47), 30256-30262

Ashby, M. N., & Edwards, P. A. (1990). Elucidation of the deficiency in two yeast coenzyme Q mutants. Characterization of the structural gene encoding hexaprenyl pyrophosphate synthetase. J Biol Chem. 265(22), 13157-13164.

Ashmarina, L. I., Rusnak, N., Miziorko, H. M., & Mitchell, G. A. (1994). 3-Hydroxy-3-methylglutaryl-CoA lyase is present in mouse and human liver peroxisomes. J Biol Chem. 269(50), 31929-31932.

Aufhammer, S. W., Warkentin, E., Ermler, U., Hagemeier, C. H., Thauer, R. K., & Shima, S. (2005). Crystal structure of methylenetetrahydromethanopterin reductase (Mer) in complex with coenzyme F420: Architecture of the F420/FMN binding site of

enzymes within the nonprolyl *cis*-peptide containing bacterial luciferase family. Protein Sci. 14(7), 1840-1849.

Azevedo, R. A., Lancien, M., & Lea, P. J. (2006). The aspartic acid metabolic pathway, an exciting and essential pathway in plants. Amino Acids. 30(2), 143-162.

Bachmann, B. O., Li, R., & Townsend, C. A. (1998). β-Lactam synthetase: a new biosynthetic enzyme. Proc Natl Acad Sci U S A. 95(16), 9082-9086.

Baggaley, K. H., Brown, A. G., & Schofield, C. J. (1997). Chemistry and biosynthesis of clavulanic acid and other clavams. Nat Prod Rep. 14(4), 309-333.

Baker, B. J., Dotzlaf, J. E., & Yeh, W. K. (1991). Deacetoxycephalosporin C hydroxylase of *Streptomyces clavuligerus*. Purification, characterization, bifunctionality, and evolutionary implication. J Biol Chem. 266(8), 5087-5093.

Banerjee, R. V., & Matthews, R. G. (1990). Cobalamin-dependent methionine synthase. FASEB J. 4(5), 1450-1459.

Banerjee, S., & Mazumdar, S. (2012). Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. Int J Anal Chem. 2012.

Banfield MJ, Barker JJ, Perry AC, Brady RL. (1998). Function from structure? The crystal structure of human phosphatidylethanolamine-binding protein suggests a role in membrane signal transduction. Structure, 6(10):1245-1254.

Banks, R. D., Blake, C. C. F., Evans, P. R., Haser, R., Rice, D. W., Hardy, G. W. et al. (1979). Sequence, structure and activity of phosphoglycerate kinase: a possible hinge-bending enzyme. Nature. 279(5716):773-7.

Bao, K., & Cohen, S. N. (2001). Terminal proteins essential for the replication of linear plasmids and chromosomes in *Streptomyces*. Gene Dev. 15(12), 1518-1527.

Barkovich, R., & Liao, J. C. (2001). Review: metabolic engineering of isoprenoids. Metab Eng. 3(1), 27-39.

Barnidge, D. R., Dratz, E. A., Martin, T., Bonilla, L. E., Moran, L. B., & Lindall, A. (2003). Absolute quantification of the G protein-coupled receptor rhodopsin by LC/MS/MS using proteolysis product peptides and synthetic peptide standards. Anal Chem. 75(3), 445-451.

Barnidge, D. R., Goodmanson, M. K., Klee, G. G., & Muddiman, D. C. (2004). Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-MS/MS using protein cleavage and isotope dilution mass spectrometry. J Proteome Res. 3(3), 644-652.

Barra, D., Martini, F., Angelaccio, S., Bossa, F., Gavilanes, F., Peterson, D., et al. (1983). Sequence homology between prokaryotic and eukaryotic forms of serine hydroxymethyltransferase. Biochem Bioph Res Co. 116(3), 1007-1012.

Bartlett, P. A., & Satake, K. (1988). Does dehydroquinate synthase synthesize dehydroquinate? J Am Chem Soc. 110(5), 1628-1630.

Bateman, A., & Rawlings, N. D. (2003). The CHAP domain: a large family of amidases including GSP amidase and peptidoglycan hydrolases. Trends Biochem Sci. 28(5), 234-237.

Bateman, A., Coggill, P., & Finn, R. D. (2010). DUFs: families in search of function. Acta Crystallogr Sect F Struct Biol Cryst Commun. *66*(Pt10), 1148-1152.

Bell, A. W., Ward, M. A., Blackstock, W. P., Freeman, H. N., Choudhary, J. S., et al. (2001). Proteomics characterization of abundant golgi membrane proteins. J Biol Chem. 276(7), 5152-5165.

Bellinzoni, M., Buroni, S., Pasca, M. R., Guglierame, P., Arcesi, F., De Rossi, E., & Riccardi, G. (2005). Glutamine amidotransferase activity of NAD^+ synthetase from *Mycobacterium tuberculosis* depends on an amino-terminal nitrilase domain. Res Microbiol. 156(2), 173-177.

Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., et al. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3 (2). Nature, 417(6885), 141-147.

Beranova-Giorgianni, S. (2003). Proteome analysis by two-dimensional gel electrophoresis and mass spectrometry: strengths and limitations. Trac-Trend Anal Chem. 22(5), 273-281.

Berdy, J. (1995). Are actinomycetes exhausted as source of secondary metabolites? In *Proceedings of the 9th International Symposium on the Biology of Actinomycetes* (pp. 13-34).

Bernhardt, R. (2006). Cytochromes P450 as versatile biocatalysts. J Biotechnol. 124(1), 128-145.

Bibb, M. J. (2005). Regulation of secondary metabolism in streptomycetes. Curr Opin Microbiol. 8(2), 208-215.

Bibb, M. J., Molle, V., & Buttner, M. J. (2000). σ^{BldN}, an extracytoplasmic function RNA polymerase sigma factor required for aerial mycelium formation in *Streptomyces coelicolor* A3 (2). J Bacteriol. 182(16), 4606-4616.

Bieganowski, P., Pace, H. C., & Brenner, C. (2003). Eukaryotic NAD^+ synthetase Qns1 contains an essential, obligate intramolecular thiol glutamine amidotransferase domain related to nitrilase. J Biol Chem. 278(35), 33049-33055.

Bignell, D. R., Tahlan, K., Colvin, K. R., Jensen, S. E., & Leskiw, B. K. (2005). Expression of *ccaR*, encoding the positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus*, is dependent on *bldG*. Antimicrob Agents Ch. *49*(4), 1529-1541.

Bignell, D. R., Warawa, J. L., Strap, J. L., Chater, K. F., & Leskiw, B. K. (2000). Study of the *bldG* locus suggests that an anti-anti-sigma factor and an anti-sigma factor may be involved in *Streptomyces coelicolor* antibiotic production and sporulation. Microbiology, 146(9), 2161-2173.

Blanco, J., Coque, J. J. R., & Martin, J. F. (1998). The folate branch of the methionine biosynthesis pathway in *Streptomyces lividans*: Disruption of the 5, 10-methylenetetrahydrofolate reductase gene leads to methionine auxotrophy. J Bacteriol. 180(6), 1586-1591.

Bloudoff, K., Alonzo, D. A., & Schmeing, T. M. (2016). Chemical probes allow structural insight into the condensation reaction of nonribosomal peptide synthetases. Cell Chem Biol. 23(3), 331-339.

Bonangelino, C. J., Chavez, E. M., & Bonifacino, J. S. (2002). Genomic screen for vacuolar protein sorting genes in *Saccharomyces cerevisiae*. Mol Biol Cell. 13(7), 2486-2501.

Bonnett, S. A., Whicher, J. R., Papireddy, K., Florova, G., Smith, J. L., & Reynolds, K. A. (2013). Structural and stereochemical analysis of a modular polyketide synthase ketoreductase domain required for the generation of a cis-alkene. Chem Biol. 20(6), 772-783.

Borths, E. L., Locher, K. P., Lee, A. T., & Rees, D. C. (2002). The structure of *Escherichia coli* BtuF and binding to its cognate ATP binding cassette transporter. Proc Natl Acad Sci U S A. 99(26), 16642-16647.

Bown, L., Altowairish, M. S., Fyans, J. K., & Bignell, D. R. (2016). Production of the *Streptomyces scabies* coronafacoyl phytotoxins involves a novel biosynthetic pathway with an F_{420}-dependent oxidoreductase and a short-chain dehydrogenase/reductase. Mol. Microbiol. 101(1):122-135.

Bralley, P., & Jones, G. H. (2004). Organization and expression of the polynucleotide phosphorylase gene (pnp) of Streptomyces: processing of pnp transcripts in Streptomyces antibioticus. J Bacteriol. 186(10), 3160-3172.

Braña, A. F., Paiva, N., & Demain, A. L. (1986). Pathways and regulation of ammonium assimilation in *Streptomyces clavuligerus*. Microbiology. 132(5), 1305-1317.

Briers, Y., Volckaert, G., Cornelissen, A., Lagaert, S., Michiels, C. W., Hertveldt, K., & Lavigne, R. (2007). Muralytic activity and modular structure of the endolysins of *Pseudomonas aeruginosa* bacteriophages φKZ and EL. Mol Microbiol. 65(5), 1334-1344.

Brown, A. G., Butterworth, D., Cole, M., Hanscomb, G., Hood, J. D., Reading, C., & Rolinson, G. N. (1976). Naturally-occurring β-lactamase inhibitors with antibacterial activity. J Antibiot (Tokyo). 29(6), 668-669.

Brown, S., Santa Maria Jr, J. P., & Walker, S. (2013). Wall teichoic acids of gram-positive bacteria. Annu Rev Microbiol. 67.

Brownlee, J. M., Johnson-Winters, K., Harrison, D. H., & Moran, G. R. (2004). Structure of the ferrous form of (4-hydroxyphenyl) pyruvate dioxygenase from *Streptomyces avermitilis* in complex with the therapeutic herbicide, NTBC. Biochemistry, 43(21), 6370-6377.

Bucca, G., Brassington, A. M. E., Hotchkiss, G., Mersinias, V. & Smith C. P. (2003). Negative feedback regulation of *dnaK*, *clpB* and *lon* expression by the DnaK chaperone machine in *Streptomyces coelicolor*, identified by transcriptome and *in vivo* DnaK-depletion analysis. Mol Microbiol. 50(1), 153–166

Bucca, G., Hindle, Z., & Smith, C. P. (1997). Regulation of the *dnaK* operon of *Streptomyces coelicolor* A3 (2) is governed by HspR, an autoregulatory repressor protein. J Bacteriol. 179(19), 5999-6004.

Bucca, G., Brassington, A. M., Schönfeld, H. J., & Smith, C. P. (2000). The HspR regulon of *Streptomyces coelicolor*: a role for the DnaK chaperone as a transcriptional corepressor. Mol Microbiol. 38(5), 1093-1103.

Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D., & Aravind, L. (2006). Evolutionary genomics of the HAD superfamily: understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. J Mol Biol. 361(5), 1003-1034.

Bursy, J., Pierik, A. J., Pica, N., & Bremer, E. (2007). Osmotically induced synthesis of the compatible solute hydroxyectoine is mediated by an evolutionarily conserved ectoine hydroxylase. J Biol Chem. 282(43), 31147-31155.

Bursy, J., Kuhlmann, A. U., Pittelkow, M., Hartmann, H., Jebbar, M., Pierik, A. J., & Bremer, E. (2008). Synthesis and uptake of the compatible solutes ectoine and 5-

hydroxyectoine by *Streptomyces coelicolor* A3 (2) in response to salt and heat stresses. Appl Environ Microb. 74(23), 7286-7296.

Bussari, B., Survase, S. A., Saudagar, P. S., & Singhal, R. S. (2009). An integrated approach for production of cephamycin C using *Streptomyces clavuligerus* NT4: Sequential optimization of production medium and effect of amino acids. Curr Trends Biotechnol Pharm. 3(4), 372-384.

Bzowska, A., Kulikowska, E., & Shugar, D. (2000). Purine nucleoside phosphorylases: properties, functions, and clinical aspects. Pharmacol Therapeut. 88(3), 349-425.

Calcott, M. J., Owen, J. G., Lamont, I. L., & Ackerley, D. F. (2014). Biosynthesis of novel pyoverdines by domain substitution in a nonribosomal peptide synthetase of *Pseudomonas aeruginosa*. Appl Environ Microbiol. 80(18), 5723-5731.

Campbell, J. A., Davies, G. J., Bulone, V., & Henrissat, B. (1997). A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochemical Journal*, *326*(Pt 3), 929.

Candiano, G., Bruschi, M., Musante, L., Santucci, L., Ghiggeri, G. M., Carnemolla, B., et al. (2004). Blue silver: A very sensitive colloidal Coomassie G-250 staining for proteome analysis. Electrophoresis. 25(9), 1327-1333.

Cantu, D. C., Chen, Y., & Reilly, P. J. (2010). Thioesterases: a new perspective based on their primary and tertiary structures. Protein Sci. 19(7), 1281-1295.

Capaldi, R. A. (1990). Structure and function of cytochrome c oxidase. Annu Rev Biochem. 59(1), 569-596.

Carr, S., Penfold, C. N., Bamford, V., James, R., & Hemmings, A. M. (2000). The structure of TolB, an essential component of the tol-dependent translocation system, and its protein–protein interaction with the translocation domain of colicin E9. Structure, 8(1), 57-66.

Carreras, C. W., & Santi, D. V. (1998). Engineering of modular polyketide synthases to produce novel polyketides. Curr Opin Biotech. 9(4), 403-411.

Cashel, M., & Kalbacher, B. (1970). The control of ribonucleic acid synthesis in *Escherichia coli* V. Characterization of a nucleotide associated with the stringent response. J Biol Chem. 245(9), 2309-2318.

Cashel, M., Gentry, D. R., Hernández, V. J. & Vinella, D. (1996). The stringent response. In *Escherichia coli* and *Salmonella typhimurium: Cellular and Molecular Biology*, pp. 1458–1496. Edited by F. C. Neidhardt and others. Washington, DC: American Society for Microbiology.

Castaldo, G., Zucko, J., Heidelberger, S., Vujaklija, D., Hranueli, D., Cullum, J. et al. (2008). Proposed arrangement of proteins forming a bacterial type II polyketide synthase. Chem Biol. 15(11), 1156-1165.

Çaydaşı A.K. (2006). Targeted disruption of homoserine dehydrogenase gene in *S. clavuligerus* and its effects on cephamycin C production. Middle East Technical University, Turkey.

Chait, B. T. (2006). Mass spectrometry: bottom-up or top-down? Science, 314(5796), 65-66.

Chakraburtty, R., & Bibb, M. (1997). The ppGpp synthetase gene (*relA*) of *Streptomyces coelicolor* A3 (2) plays a conditional role in antibiotic production and morphological differentiation. J Bacteriol. 179(18), 5854-5861.

Challis, G. L. & Hopwood, D. A. (2003). Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. Proc Natl Acad Sci U S A. 100(suppl 2), 14555-14561.

Challis, G. L., Ravel, J., & Townsend, C. A. (2000). Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem Biol. 7(3), 211-224.

Chan, M., & Sim, T. S. (1998). Malate synthase from *Streptomyces clavuligerus* NRRL3585: cloning, molecular characterization and its control by acetate. Microbiology, 144(11), 3229-3237.

Chandra, G., & Chater, K. F. (2008). Evolutionary flux of potentially *bldA*-dependent *Streptomyces* genes containing the rare leucine codon TTA. *Antonie Van Leeuwenhoek*, *94*(1), 111-126.

Chang, Z., & Vining, L. C. (2002). Biosynthesis of sulfur-containing amino acids in *Streptomyces venezuelae* ISP5230: roles for cystathionine β-synthase and transsulfuration. Microbiology, 148(7), 2135-2147.

Charusanti, P., Fong, N. L., Nagarajan, H., Pereira, A. R., Li, H. J., Abate, E. A., et al. (2012). Exploiting adaptive laboratory evolution of *Streptomyces clavuligerus* for antibiotic discovery and overproduction. PLoS One, 7(3), e33727.

Chater, K. F. (2001). Regulation of sporulation in *Streptomyces coelicolor* A3 (2): a checkpoint multiplex? Curr Opin Microbiol. 4(6), 667-673.

Chater, K. F., & Chandra, G. (2008). The use of the rare UUA codon to define "expression space" for genes involved in secondary metabolism, development and environmental adaptation in *Streptomyces*. J Microbiol. 46(1), 1-11.

Chater, K. F., & Kinashi, H. (2007). *Streptomyces* linear plasmids: their discovery, functions, interactions with other replicons, and evolutionary significance. In *Microbial Linear Plasmids* (pp. 1-31). Springer Berlin Heidelberg.

Chater, K. F., Biró, S., Lee, K. J., Palmer, T., & Schrempf, H. (2010). The complex extracellular biology of *Streptomyces*. FEMS Microbiol Rev. 34(2), 171-198.

Chávez, A., Forero, A., Sánchez, M., Rodríguez-Sanoja, R., Mendoza-Hernández, G., Servín-Gonzalez, L., et al. (2011). Interaction of SCO2127 with BldKB and its possible connection to carbon catabolite regulation of morphological differentiation in *Streptomyces coelicolor*. Appl Microbiol Biot. 89(3), 799-806.

Che, T., Bethel, C. R., Pusztai-Carey, M., Bonomo, R. A., & Carey, P. R. (2014). The different inhibition mechanisms of OXA-1 and OXA-24 β-lactamases are determined by the stability of active site carboxylated lysine. J Biol Chem. 289(9), 6152-6164.

Chen, L. & Helmann, J. D. (1995). *Bacillus subtilis* MrgA is a Dps (PexB) homologue: evidence for metalloregulation of an oxidative-stress gene. Mol Microbiol. 18(2), 295-300.

Chen, K. C., Lin, Y. H., Wu, J. Y., & Hwang, S. C. J. (2003). Enhancement of clavulanic acid production in *Streptomyces clavuligerus* with ornithine feeding. Enzyme Microb Tech. 32(1), 152-156.

Chen, C. W. (2007). *Streptomyces* linear plasmids: replication and telomeres. In *Microbial Linear Plasmids* (pp. 33-61). Springer Berlin Heidelberg.

Chevalier, F. (2010). Highlights on the capacities of "Gel-based" proteomics. Proteome Sci. 8(1), 1.

Chiba, S. (1997). Molecular mechanism in α-glucosidase and glucoamylase. Biosci Biotechnol Biochem. 61(8), 1233-1239.

Chiu, Y., Kuo, T. Y., Lin, C. C., & Chen, W. J. (2011). Proteomic analysis reveals responsive proteins of *Vibrio parahaemolyticus* on exposure to cationic antimicrobial peptides. J Appl Microbiol. 110(1), 80-89.

Chmara, H., Zähner, H., & Borowski, E. (1984). Anticapsin, an active-site directed irreversible inhibitor of glucosamine-6-phosphate synthetase from *Escherichia coli*. J Antibiot (Tokyo). 37(9), 1038-1043.

Choi, H., Fermin, D., & Nesvizhskii, A. I. (2008). Significance analysis of spectral count data in label-free shotgun proteomics. Mol Cell Proteomics. 7(12), 2373-2385.

Choi, S. U., Lee, C. K., Hwang, Y. I., Kinosita, H., & Nihira, T. (2003). γ-Butyrolactone autoregulators and receptor proteins in non-*Streptomyces* actinomycetes producing commercially important secondary metabolites. Arch Microbiol. 180(4), 303-307.

Chow, C. (2015). Mechanistic analysis of essential enzymes from *Mycobacterium tuberculosis* and other bacterial pathogens. Yeshiva University, New York, United States

Chowdhury, E. K., Nagata, S., & Misono, H. (1996). 3-Hydroxyisobutyrate dehydrogenase from *Pseudomonas putida* E23: purification and characterization. Biosci Biotechnol Biochem. 60(12), 2043-2047.

Claas, K. R., Parrish, J. R., Maggio-Hall, L. A., & Escalante-Semerena, J. C. (2010). Functional analysis of the nicotinate mononucleotide: 5, 6-dimethylbenzimidazole phosphoribosyltransferase (CobT) enzyme, involved in the late steps of coenzyme B12 biosynthesis in *Salmonella enterica*. J Bacteriol. 192(1), 145-154.

Cobb, J. A., Bjergbaek, L., & Gasser, S. M. (2002). RecQ helicases: at the heart of genetic stability. FEBS Let. 529(1), 43-48.

Collard, F., Stroobant, V., Lamosa, P., Kapanda, C. N., Lambert, D. M., Muccioli, G. G., et al. (2010). Molecular identification of N-acetylaspartylglutamate synthase and β-citrylglutamate synthase. J Biol Chem. 285(39), 29826-29833.

Contreras-Rodriguez, A., Quiroz-Limon, J., Martins, A. M., Peralta, H., Avila-Calderon, E., Sriranganathan, N., et al. (2008). Enzymatic, immunological and phylogenetic characterization of *Brucella suis* urease. BMC Microbiol. 8(1), 121.

Cooper, R. A., & Kornberg, H. L. (1967). The mechanism of the phosphoenolpyruvate synthase reaction. Biochim Biophys Acta. 141(1), 211-213.

Coque, J. J. R., Liras, P., Láiz, L. & Martín, J. F. (1991). A gene encoding lysine 6-aminotransferase which forms α-aminoadipic acid, a precursor of β-lactam antibiotics, is located in the cluster of cephamycin biosynthetic genes in *Nocardia lactamdurans*. J. Bacteriol. 173(19), 625–6264.

Cordin, O., Banroques, J., Tanner, N. K., & Linder, P. (2006). The DEAD-box protein family of RNA helicases. Gene. 367, 17-37.

Costa, T. R., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M., & Waksman, G. (2015). Secretion systems in Gram-negative bacteria: structural and mechanistic insights. Nat Rev Microbiol. 13(6), 343-359.

Cotter, R. J. (1999). Peer Reviewed: The New Time-of-Flight Mass Spectrometry. Anal Chem. 71(13), 445A-451A.

Cronan, J. E., & Thomas, J. (2009). Bacterial fatty acid synthesis and its relationships with polyketide synthetic pathways. Methods Enzymol. 459, 395-433.

Cundliffe, E. (1989). How antibiotic-producing organisms avoid suicide. Annu Rev Microbiol. 43(1), 207-233.

Curreem, S. O., Watt, R. M., Lau, S. K., & Woo, P. C. (2012). Two-dimensional gel electrophoresis in bacterial proteomics. Protein Cell. 3(5), 346-363.

Cuthbertson, L., & Nodwell, J. R. (2013). The TetR family of regulators. Microbiol Mol Biol R. 77(3), 440-475.

Dairi, T. (2005). Studies on biosynthetic genes and enzymes of isoprenoids produced by actinomycetes. J Antibiot (Tokyo). 58(4), 227-243.

Darwin, K. H., Ehrt, S., Gutierrez-Ramos, J. C., Weich, N., & Nathan, C. F. (2003). The proteasome of *Mycobacterium tuberculosis* is required for resistance to nitric oxide. Science, 302(5652), 1963-1966.

Davidsen, J. M., & Townsend, C. A. (2009). Identification and characterization of NocR as a positive transcriptional regulator of the β-lactam nocardicin A in *Nocardia uniformis*. J Bacteriol. 191(3), 1066-1077.

Davidsen, J. M., & Townsend, C. A. (2012). In vivo characterization of nonribosomal peptide synthetases NocA and NocB in the biosynthesis of nocardicin A. Chem Biol. 19(2), 297-306.

Davies, J., & Wright, G. D. (1997). Bacterial resistance to aminoglycoside antibiotics. Trends Microbiol. 5(6), 234-240.

de Crécy-Lagard, V., Servant-Moisson, P., Viala, J., Grandvalet, C., & Mazodier, P. (1999). Alteration of the synthesis of the Clp ATP-dependent protease affects morphological and physiological differentiation in *Streptomyces*. Mol Microbiol. 32(3), 505-517.

de la Fuente, A., Lorenzana, L. M., Martín, J. F., & Liras, P. (2002). Mutants of *Streptomyces clavuligerus* with disruptions in different genes for clavulanic acid biosynthesis produce large amounts of holomycin: possible cross-regulation of two unrelated secondary metabolic pathways. J Bacteriol. 184(23), 6559-6565.

De Mot, R., & Parret, A. H. (2002). A novel class of self-sufficient cytochrome P450 monooxygenases in prokaryotes. Trends Microbiol. 10(11), 502-508.

De Mot, R., Schoofs, G., & Nagy, I. (2007). Proteome analysis of *Streptomyces coelicolor* mutants affected in the proteasome system reveals changes in stress-responsive proteins. Arch Microbiol. 188(3), 257-271.

de Orué Lucana, D. O., Schaa, T., & Schrempf, H. (2004). The novel extracellular *Streptomyces reticuli* haem-binding protein HbpS influences the production of the catalase-peroxidase CpeB. Microbiology, 150(8), 2575-2585.

Dela Cruz, R. D., Gao, Y., Penumetcha, S., Sheplock, R., Weng, K., & Chander, M. (2010). Expression of the *Streptomyces coelicolor* SoxR regulon is intimately linked with actinorhodin production. J Bacteriol. 192(24), 6428-6438.

Deli, A., Koutsioulis, D., Fadouloglou, V. E., Spiliotopoulou, P., Balomenou, S., Arnaouteli, S. et al. (2010). LmbE proteins from *Bacillus cereus* are de-N-acetylases with broad substrate specificity and are highly similar to proteins in *Bacillus anthracis*. Febs J. 277(13), 2740-2753.

Demain, A. L., & Elander, R. P. (1999). The β-lactam antibiotics: past, present, and future. Antonie Van Leeuwenhoek, 75(1-2), 5-19.

den Hengst, C. D., Tran, N. T., Bibb, M. J., Chandra, G., Leskiw, B. K., & Buttner, M. J. (2010). Genes essential for morphological development and antibiotic

production in *Streptomyces coelicolor* are targets of BldD during vegetative growth. Mol Microbiol. 78(2), 361-379.

Denoya, C. D., Fedechko, R. W., Hafner, E. W., McArthur, H. A., Morgenstern, M. R., Skinner, D. D. et al. (1995). A second branched-chain alpha-keto acid dehydrogenase gene cluster (bkdFGH) from *Streptomyces avermitilis*: its relationship to avermectin biosynthesis and the construction of a *bkdF* mutant suitable for the production of novel antiparasitic avermectins. J Bacteriol. 177(12), 3504-3511.

Deutch, A. H., Smith, C. J., Rushlow, K. E., & Krelschmer, P. J. (1982). *Escherichia coli* Δ1-pyrroline-5-carboxylate reductase: gene sequence, protein overproduction and purification. Nucleic Acids Res. 10(23), 7701-7714.

Deutscher, J., Francke, C., & Postma, P. W. (2006). How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. Microbiol Mol Biol R. 70(4), 939-1031.

Deveau, H., Garneau, J. E., & Moineau, S. (2010). CRISPR/Cas system and its role in phage-bacteria interactions. Annu Rev Microbiol. 64, 475-493.

Devedjiev, Y., Symersky, J., Singh, R., Jedrzejas, M., Brouillette, C., Brouillette, W., et al. (2001). Stabilization of active-site loops in NH_3-dependent NAD^+ synthetase from *Bacillus subtilis*. Acta Crystallogr D Biol Crystallogr. 57(Pt6), 806-812.

Dhingra, V., Gupta, M., Andacht, T., & Fu, Z. F. (2005). New frontiers in proteomics research: a perspective. Int J Pharmaceut. 299(1), 1-18.

Dicker, L., Lin, X., & Ivanov, A. R. (2010). Increased power for the analysis of label-free LC-MS/MS proteomics data by combining spectral counts and peptide peak attributes. Mol Cell Proteomics. 9(12), 2704-2718.

Domain, F., Houot, L., Chauvat, F., & Cassier-Chauvat, C. (2004). Function and regulation of the cyanobacterial genes *lexA*, *recA* and *ruvB*: LexA is critical to the survival of cells facing inorganic carbon starvation. Mol Microbiol. 53(1), 65-80.

Domenech, P., Reed, M. B., & Barry, C. E. (2005). Contribution of the *Mycobacterium tuberculosis* MmpL protein family to virulence and drug resistance. Infect Immun. 73(6), 3492-3501.

Dong, X., Quinn, P. J., & Wang, X. (2011). Metabolic engineering of *Escherichia coli* and *Corynebacterium glutamicum* for the production of L-threonine. Biotechnol Adv. 29(1), 11-23.

Douangamath, A., Walker, M., Beismann-Driemeyer, S., Vega-Fernandez, M. C., Sterner, R., & Wilmanns, M. (2002). Structural evidence for ammonia tunneling across the (βα) 8 barrel of the imidazole glycerol phosphate synthase bienzyme complex. Structure, 10(2), 185-193.

Dörschug, M., Frank, R., Kalbitzer, H. R., Hengstenberg, W., & Deutscher, J. (1984). Phosphoenolpyruvate-dependent phosphorylation site in enzyme IIIglc of the *Escherichia coli* phosphotransferase system. Eur J Biochem. 144(1), 113-119.

Du, L., & Lou, L. (2010). PKS and NRPS release mechanisms. Nat Prod Rep. 27(2), 255-278.

Du, L., Sánchez, C., Chen, M., Edwards, D. J., & Shen, B. (2000). The biosynthetic gene cluster for the antitumor drug bleomycin from *Streptomyces verticillus* ATCC15003 supporting functional interactions between nonribosomal peptide synthetases and a polyketide synthase. Chem Biol. *7*(8), 623-642.

Dunn, M. F., Araíza, G., & Mora, J. (2004). Biochemical characterization of a *Rhizobium etli* monovalent cation-stimulated acyl-coenzyme A carboxylase with a high substrate specificity constant for propionyl-coenzyme A. Microbiology, 150(2), 399-406.

Durand, P., Golinelli-Pimpaneau, B., Mouilleron, S., Badet, B., & Badet-Denisot, M. A. (2008). Highlights of glucosamine-6P synthase catalysis. Arch Biochem Biophys. 474(2), 302-317.

Durocher, D., Henckel, J., Fersht, A. R., & Jackson, S. P. (1999). The FHA domain is a modular phosphopeptide recognition motif. Mol Cell. 4(3), 387-394.

Dyballa, N., & Metzger, S. (2009). Fast and sensitive colloidal coomassie G-250 staining for proteins in polyacrylamide gels. J Vis Exp, 30.

Dyda, F., Klein, D. C., & Hickman, A. B. (2000). GCN5-related N-acetyltransferases: a structural overview. Annu Rev Bioph Biom. 29(1), 81-103.

Ebersbach, G., & Gerdes, K. (2004). Bacterial mitosis: partitioning protein ParA oscillates in spiral-shaped structures and positions plasmids at mid-cell. Mol Microbiol. 52(2), 385-398.

Egan, L. A., Busby, R. W., Iwata-Reuyl, D., & Townsend, C. A. (1997). Probable role of clavaminic acid as the terminal intermediate in the common pathway to clavulanic acid and the antipodal clavam metabolites. J Am Chem Soc. 119(10), 2348-2355.

Eisaki, N., Tatsumi, H., Murakami, S., & Horiuchi, T. (1999). Pyruvate phosphate dikinase from a thermophilic actinomyces *Microbispora rosea* subsp. *aerata*: purification, characterization and molecular cloning of the gene. Biochim Biophys Acta. 1431(2), 363-373.

El Zoeiby, A., Sanschagrin, F., Havugimana, P. C., Garnier, A., & Levesque, R. C. (2001). In vitro reconstruction of the biosynthetic pathway of peptidoglycan cytoplasmic precursor in *Pseudomonas aeruginosa*. FEMS Microbiol Lett. 201(2), 229-235.

El-Aneed, A., Cohen, A., & Banoub, J. (2009). Mass spectrometry, review of the basics: electrospray, MALDI, and commonly used mass analyzers. Appl Spectrosc Rev. 44(3), 210-230.

Elgrably-Weiss, M., Schlosser-Silverman, E., Rosenshine, I., & Altuvia, S. (2006). DeoT, a DeoR-type transcriptional regulator of multiple target genes. FEMS Microbiol Lett. 254(1), 141-148.

Elliot, M. A., Buttner, M. J., & Nodwell, J. R. (2007). Multicellular development in *Streptomyces*. In: Whitworth DE, editor. *Myxobacteria: Multicellularity and Differentiation*, Washington D.C.: ASM Press. pp. 419–438.

Engels, V., & Wendisch, V. F. (2007). The DeoR-type regulator SugR represses expression of *ptsG* in *Corynebacterium glutamicum*. J Bacteriol. 189(8), 2955-2966.

Enguita, F. J., Liras, P., Leitão, A. L., & Martín, J. F. (1996). Interaction of the two proteins of the methoxylation system involved in cephamycin C biosynthesis Immunoaffinity, protein cross-linking, and fluorescence spectroscopy Studies. J Biol Chem. 271(52), 33225-33230.

Faurobert, M., Pelpoir, E., & Chaïb, J. (2007). Phenol extraction of proteins for proteomic studies of recalcitrant plant tissues. Methods Mol Biol. 355,9-14.

Fawaz, M. V., Topper, M. E., & Firestine, S. M. (2011). The ATP-grasp enzymes. Bioorg Chem. 39(5), 185-191.

Felnagle, E. A., Jackson, E. E., Chan, Y. A., Podevels, A. M., Berti, A. D., McMahon, M. D., & Thomas, M. G. (2008). Nonribosomal peptide synthetases involved in the production of medically relevant natural products. Mol Pharmaceut. 5(2), 191-211.

Feniouk, B. A., Suzuki, T., & Yoshida, M. (2007). Regulatory interplay between proton motive force, ADP, phosphate, and subunit ϵ in bacterial ATP synthase. J Biol Chem. 282(1), 764-772.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., & Whitehouse, C. M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. Science, 246(4926), 64-71.

Fenn, J. B. (2003). Electrospray wings for molecular elephants (Nobel lecture). Angewandte Chemie International Edition. 42(33), 3871-3894.

Ferguson, N. L., Peña-Castillo, L., Moore, M. A., Bignell, D. R., & Tahlan, K. (2016). Proteomics analysis of global regulatory cascades involved in clavulanic acid production and morphological development in *Streptomyces clavuligerus*. J Ind Microbiol Biot. 43(4):537-55.

Ferla, M. P., & Patrick, W. M. (2014). Bacterial methionine biosynthesis. Microbiology. 160(8), 1571-1584.

Fernández, E., Weißbach, U., Reillo, C. S., Braña, A. F., Méndez, C., Rohr, J., & Salas, J. A. (1998). Identification of two genes from *Streptomyces argillaceus* encoding glycosyltransferases involved in transfer of a disaccharide during biosynthesis of the antitumor drug mithramycin. J Bacteriol. 180(18), 4929-4937.

Ferrández, A., Miñambres, B., García, B., Olivera, E. R., Luengo, J. M., García, J. L., & Díaz, E. (1998). Catabolism of phenylacetic acid in *Escherichia coli* characterization of a new aerobic hybrid pathway. J Biol Chem. 273(40), 25974-25986.

Fisch, K. M. (2013). Biosynthesis of natural products by microbial iterative hybrid PKS–NRPS. RSC Adv. 3(40), 18228-18247.

Flärdh, K., & Buttner, M. J. (2009). *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. Nat Rev Microbiol. 7(1), 36-49.

Fowler, R. G., White, S. J., Koyama, C., Moore, S. C., Dunn, R. L., & Schaaper, R. M. (2003). Interactions among the *Escherichia coli mutT*, *mutM*, and *mutY* damage prevention pathways. DNA repair, 2(2), 159-173.

Fraser, M. E., James, M. N., Bridger, W. A., & Wolodko, W. T. (1999). A detailed structural description of *Escherichia coli* succinyl-CoA synthetase. J Mol Biol. 285(4), 1633-1653.

Freist, W., Verhey, J. F., Rühlmann, A., Gauss, D. H., & Arnez, J. G. (1999). Histidyl-tRNA synthetase. Biol Chem. 380(6), 623-646.

Fu, X., Gharib, S. A., Green, P. S., Aitken, M. L., Frazer, D. A., Park, D. R., et al. (2008). Spectral index for assessment of differential protein expression in shotgun proteomics. J Proteome Res. 7(3), 845-854.

Fulston, M., Davison, M., Elson, S. W., Nicholson, N. H., Tyler, J. W., & Woroniecki, S. R. (2001). Clavulanic acid biosynthesis; the final steps. J Chem Soc Perk T 1 1(9), 1122-1130.

Fundoiano-Hershcovitz, Y., Rabinovitch, L., Langut, Y., Reiland, V., Shoham, G., & Shoham, Y. (2004). Identification of the catalytic residues in the double-zinc aminopeptidase from *Streptomyces griseus*. FEBS Lett. 571(1-3), 192-196.

Gaballo, A., Abbrescia, A., Palese, L. L., Micelli, L., Di Summa, R., Alifano, P., & Papa, S. (2006). Structure and expression of the *atp* operon coding for F1 F0-ATP synthase from the antibiotic-producing actinomycete *Nonomuraea* sp. ATCC 39727. Res Microbiol. 157(7), 675-683.

Gago, G., Diacovich, L., Arabolaza, A., Tsai, S. C., & Gramajo, H. (2011). Fatty acid biosynthesis in actinomycetes. FEMS Microbiol Rev. 35(3), 475-497.

Galinski, E. A., Pfeiffer, H. P., & Trüper, H. G. (1985). 1, 4, 5, 6-Tetrahydro-2-methyl-4-pyrimidinecarboxylic acid. Eur J Biochem. 149(1), 135-139.

Gao, Y. G., Selmer, M., Dunham, C. M., Weixlbaumer, A., Kelley, A. C., & Ramakrishnan, V. (2009). The structure of the ribosome with elongation factor G trapped in the posttranslocational state. Science, 326(5953), 694-699.

Garcia, B. A., Siuti, N., Thomas, C. E., Mizzen, C. A., & Kelleher, N. L. (2007). Characterization of neurohistone variants and post-translational modifications by electron capture dissociation mass spectrometry. Int J Mass Spectrom. *259*(1), 184-196.

Garcia-Dominguez, M., Martin, J. F., & Liras, P. (1989). Characterization of sugar uptake in wild-type *Streptomyces clavuligerus*, which is impaired in glucose uptake, and in a glucose-utilizing mutant. J Bacteriol. 171(12), 6808-6814.

Gardner, J. G., Grundy, F. J., Henkin, T. M., & Escalante-Semerena, J. C. (2006). Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD^+ involvement in *Bacillus subtilis*. J Bacteriol. 188(15), 5460-5468.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31(13), 3784-3788.

Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., & Gygi, S. P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci U S A. 100(12), 6940-6945.

Ghimire, G. P., Koirala, N., & Sohng, J. K. (2015). Activation of cryptic *hop* genes from *Streptomyces peucetius* ATCC 27952 involved in hopanoid biosynthesis. J. Microbiol. Biotechnol, 25(5):658-61.

Ghisla, S., & Thorpe, C. (2004). Acyl-CoA dehydrogenases. Eur J Biochem. 271(3), 494-508.

Ghuysen, J. M., Lamotte-Brasseur, J., Joris, B., & Shockman, G. D. (1994). Binding site-shaped repeated sequences of bacterial wall peptidoglycan hydrolases. FEBS Let. *342*(1), 23-28.

Gibbs, K. A., Urbanowski, M. L., & Greenberg, E. P. (2008). Genetic determinants of self identity and social recognition in bacteria. Science, 321(5886), 256-259.

Givot, I. L., Smith, T. A., & Abeles, R. H. (1969). Studies on the mechanism of action and the structure of the electrophilic center of histidine ammonia lyase. J Biol Chem. 244(23), 6341-6353.

Gomez-Escribano, J. P., Martin, J. F., Hesketh, A., Bibb, M. J., & Liras, P. (2008). *Streptomyces clavuligerus relA*-null mutants overproduce clavulanic acid and cephamycin C: negative regulation of secondary metabolism by (p) ppGpp. Microbiology, 154(3), 744-755.

Goodacre, N. F., Gerloff, D. L., & Uetz, P. (2014). Protein domains of unknown function are essential in bacteria. MBio, 5(1), e00744-13.

Goodman, J. L., Wang, S., Alam, S., Ruzicka, F. J., Frey, P. A., & Wedekind, J. E. (2004). Ornithine cyclodeaminase: structure, mechanism of action, and implications for the μ-crystallin family. *Biochemistry*, *43*(44), 13883-13891.

Görg, A., Boguth, G., Obermaier, C., Posch, A., & Weiss, W. (1995). Two-dimensional polyacrylamide gel electrophoresis with immobilized pH gradients in the

first dimensin (IPG-Dalt): The state of the art and the controversy of vertical versus horizontal systems. Electrophoresis, 16(1), 1079-1086.

Görg, A., Weiss, W., & Dunn, M. J. (2004). Current two-dimensional electrophoresis technology for proteomics. Proteomics, 4(12), 3665-3685.

Görg, A., Drews, O., Lück, C., Weiland, F., & Weiss, W. (2009). 2-DE with IPGs. Electrophoresis, 30(S1), S122-S132.

Göthel, S. F., & Marahiel, M. A. (1999). Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. Cell Mol Life Sci. 55(3), 423-436.

Gray, M. J., Wholey, W. Y., Wagner, N. O., Cremers, C. M., Mueller-Schickert, A., Hock, N. T., et al. (2014). Polyphosphate is a primordial chaperone. Mol Cell. 53(5), 689-699.

Griffin, T. J. & Aebersold, R. (2001). Advances in proteome analysis by mass spectrometry. J Biol Chem. 276(49), 45497-45500.

Griffin, N. M., Yu, J., Long, F., Oh, P., Shore, S., Li, Y., et al. (2010). Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. Nat Biotechnol. 28(1), 83-89.

Grimm, A., Madduri, K., Ali, A., & Hutchinson, C. R. (1994). Characterization of the *Streptomyces peucetius* ATCC 29050 genes encoding doxorubicin polyketide synthase. Gene, 151(1-2), 1-10.

Grove, A. (2011). Functional evolution of bacterial histone-like HU proteins. Curr Issues Mol Biol. 13(1), 1.

Grundner, C., Gay, L. M., & Alber, T. (2005). *Mycobacterium tuberculosis* serine/threonine kinases PknB, PknD, PknE, and PknF phosphorylate multiple FHA domains. Protein Sci. 14(7), 1918-1921.

Gulick, A. M., Starai, V. J., Horswill, A. R., Homick, K. M., & Escalante-Semerena, J. C. (2003). The 1.75 Å crystal structure of acetyl-CoA synthetase bound to adenosine-5'-propylphosphate and coenzyme A. Biochemistry, 42(10), 2866-2873.

Gunsior, M., Breazeale, S. D., Lind, A. J., Ravel, J., Janc, J. W., & Townsend, C. A. (2004). The biosynthetic gene cluster for a monocyclic β-lactam antibiotic, nocardicin A. Chem Biol. 11(7), 927-938.

Guo, J., Zhang, X., Luo, S., He, F., Chen, Z., Wen, Y., & Li, J. (2013). A novel TetR family transcriptional regulator, SAV576, negatively controls avermectin biosynthesis in *Streptomyces avermitilis*. PLoS One, 8(8), e71330.

Hagelueken, G., Adams, T. M., Wiehlmann, L., Widow, U., Kolmar, H., Tümmler, B., et al. (2006). The crystal structure of SdsA1, an alkylsulfatase from *Pseudomonas aeruginosa*, defines a third class of sulfatases. Proc Natl Acad Sci U S A. 103(20), 7631-7636.

Hagen, K. D., & Meeks, J. C. (2001). The unique cyanobacterial protein OpcA is an allosteric effector of glucose-6-phosphate dehydrogenase in *Nostoc punctiforme* ATCC 29133. J Biol Chem. 276(15), 11477-11486.

Hale, C. R., Zhao, P., Olson, S., Duff, M. O., Graveley, B. R., Wells, L., et al. (2009). RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. Cell, 139(5), 945-956.

Haltli, B., Tan, Y., Magarvey, N. A., Wagenaar, M., Yin, X., Greenstein, M., et al. (2005). Investigating β-hydroxyenduracididine formation in the biosynthesis of the mannopeptimycins. Chem Biol. 12(11), 1163-1168.

Hamed, R. B., Gomez-Castellanos, J. R., Henry, L., Ducho, C., McDonough, M. A., & Schofield, C. J. (2013). The enzymes of β-lactam biosynthesis. Nat Prod Rep. 30(1), 21-107.

Han, X., Jin, M., Breuker, K., & McLafferty, F. W. (2006). Extending top-down mass spectrometry to proteins with masses greater than 200 kilodaltons. Science, 314(5796), 109-112.

Han, X., Aslanian, A., & Yates, J. R. (2008). Mass spectrometry for proteomics. Curr Opin Chem Biol. 12(5), 483-490.

Handa, N., Terada, T., Hirota, H., Tame, J. R., Park, S. Y., Kuramitsu, S., .et al. (2005). Crystal structure of a novel polyisoprenoid-binding protein from *Thermus thermophilus* HB8. Protein Sci. 14(4), 1004-1010.

Hanks, T. S., Liu, M., McClure, M. J., & Lei, B. (2005). ABC transporter FtsABCD of *Streptococcus pyogenes* mediates uptake of ferric ferrichrome. BMC microbial. 5(1), 1.

Hannemann, F., Bichet, A., Ewen, K. M., & Bernhardt, R. (2007). Cytochrome P450 systems - biological variations of electron transport chains. Biochim Biophys Acta. 1770(3), 330-344.

Harris, R. A., Zhang, B., Goodwin, G. W., Kuntz, M. J., Shimomura, Y., Rougraff, P. et al. (1990). Regulation of the branched-chain α-ketoacid dehydrogenase and elucidation of a molecular basis for maple syrup urine disease. Adv Enzyme Regul. 30, 245-256.

Haseltine, W. A., Block, R., Gilbert, W. & Weber, K. (1972). MSI and MSII are made on the ribosome in an idling reaction of protein synthesis. Nature. 238, 381–385.

Hashimoto, Y., Yamashita, M., & Murooka, Y. (1997). The *Propionibacterium freudenreichii hemYHBXRL* gene cluster, which encodes enzymes and a regulator involved in the biosynthetic pathway from glutamate to protoheme. Appl Microbiol Biot. 47(4), 385-392.

Hayashi, T., Tanaka, Y., Sakai, N., Watanabe, N., Tamura, T., Tanaka, I., & Yao, M. (2013). Structural and genomic DNA analysis of a putative transcription factor SCO5550 from *Streptomyces coelicolor* A3 (2): Regulating the expression of gene *sco5551* as a transcriptional activator with a novel dimer shape. Biochem Bioph Res Co. 435(1), 28-33.

Heath, R. J., Su, N., Murphy, C. K., & Rock, C. O. (2000). The enoyl-[acyl-carrier-protein] reductases FabI and FabL from *Bacillus subtilis*. J Biol Chem. 275(51), 40128-40133.

Heath, R. J., White, S. W., & Rock, C. O. (2001). Lipid biosynthesis as a target for antibacterial agents. Prog Lipid Res. 40(6), 467-497.

Hegde, S. S., Javid-Majd, F., & Blanchard, J. S. (2001). Overexpression and mechanistic analysis of chromosomally encoded aminoglycoside 2′-N-Acetyltransferase (AAC (2′)-Ic) from *Mycobacterium tuberculosis*. J Biol Chem. 276(49), 45876-45881.

Heikaus, C. C., Pandit, J., & Klevit, R. E. (2009). Cyclic nucleotide binding GAF domains from phosphodiesterases: structural and mechanistic insights. Structure, 17(12), 1551-1557.

Heinemann, I. U., Jahn, M., & Jahn, D. (2008). The biochemistry of heme biosynthesis. Arch Biochem Biophys. 474(2), 238-251.

Helmann, J. D. (1999). Anti-sigma factors. Curr Opin Microbiol. 2(2), 135-141.

Helmann, J. D. (2002). The extracytoplasmic function (ECF) sigma factors. Adv Microb Physiol. *46*, 47-110.

Helmetag, V., Samel, S. A., Thomas, M. G., Marahiel, M. A., & Essen, L. O. (2009). Structural basis for the erythro-stereospecificity of the l-arginine oxygenase VioC in viomycin biosynthesis. FEBS J. 276(13), 3669-3682.

Herrmann, K. M., & Weaver, L. M. (1999). The shikimate pathway. *Annual review of plant biology*, *50*(1), 473-503.

Higgens, C. E., & Kastner, R. E. (1971). *Streptomyces clavuligerus* sp. nov., a β-lactam antibiotic producer. Int J Syst Evol Micr. 21(4), 326-331.

Higo, A., Horinouchi, S., & Ohnishi, Y. (2011). Strict regulation of morphological differentiation and secondary metabolism by a positive feedback loop between two global regulators AdpA and BldA in *Streptomyces griseus*. Mol Microbiol. 81(6), 1607-1622.

Hill, C. W., Sandt, C. H., & Vlazny, D. A. (1994). Rhs elements of *Escherichia coli*: a family of genetic composites each encoding a large mosaic protein. Mol Microbiol. 12(6), 865-871.

Hill, C. W. (1999). Large genomic sequence repetitions in bacteria: lessons from rRNA operons and Rhs elements. Res Microbiol. 150(9), 665-674.

Hiller, K., Schobert, M., Hundertmark, C., Jahn, D., & Münch, R. (2003). JVirGel: calculation of virtual two-dimensional protein gels. Nucleic Acids Res. 31(13), 3862-3865.

Hiller, K., Grote, A., Maneck, M., Münch, R., & Jahn, D. (2006). JVirGel 2.0: computational prediction of proteomes separated via two-dimensional gel electrophoresis under consideration of membrane and secreted proteins. Bioinformatics, 22(19), 2441-2443.

Hofmann, K. & Bucher, P. (1995) The FHA domain: a putative nuclear signalling domain found in protein kinases and transcription factors. Trends Biochem Sci. 20, 347–349.

Hofstein, H. A., Feng, Y., Anderson, V. E., & Tonge, P. J. (1999). Role of glutamate 144 and glutamate 164 in the catalytic mechanism of enoyl-CoA hydratase. Biochemistry, 38(29), 9508-9516.

Holden, H. M., Benning, M. M., Haller, T., & Gerlt, J. A. (2001). The crotonase superfamily: divergently related enzymes that catalyze different reactions involving acyl coenzyme a thioesters. Accounts Chem Res. 34(2), 145-157.

Hong, B., Wang, L., Lammertyn, E., Geukens, N., Van Mellaert, L., Li, Y., & Anne, J. (2005). Inactivation of the 20S proteasome in *Streptomyces lividans* and its influence on the production of heterologous proteins. Microbiology. 151(9), 3137-3145.

Hopwood, D. A. (2006). Soil to genomics: the *Streptomyces* chromosome. Annu Rev Genet. 40, 1-23.

Hopwood, D. A. (2007). *Streptomyces* in nature and medicine: the antibiotic makers. Oxford University Press.

Hopwood, D. A., & Sherman, D. H. (1990). Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. Annu Rev Genet. 24(1), 37-62.

Hopwood, D. A., Bibb, M. J., & Chater, K. F. (1985). Kieser. *T*., Bruton. C.. J., Kieser HM, Lydiate DJ, Smith CP and Schrempf H. Genetic manipulation of *Streptomyces.* A Laboratory Manual. The John Innes Foundation, Norwich.

Horinouchi, S. (2003). AfsR as an integrator of signals that are sensed by multiple serine/threonine kinases in *Streptomyces coelicolor* A3 (2). J Ind Microbiol Biotechnol. 30(8), 462-467.

Horinouchi, S., & Beppu, T. (2007). Hormonal control by A-factor of morphological development and secondary metabolism in *Streptomyces*. Proc Jpn Acad Ser B Phys Biol Sci. 83(9-10), 277-295.

Hoving, S., Voshol, H., & van Oostrum, J. (2000). Towards high performance two-dimensional gel electrophoresis using ultrazoom gels. Electrophoresis. 21(13), 2617-2621.

Hsiao, N. H., Nakayama, S., Merlo, M. E., de Vries, M., Bunet, R., Kitani, S. et al. (2009). Analysis of two additional signaling molecules in *Streptomyces coelicolor* and the development of a butyrolactone-specific reporter system. Chem Biol. 16(9), 951-960.

Huang, S., Zhao, Y., Qin, Z., Wang, X., Onega, M., Chen, L. et al. (2011). Identification and heterologous expression of the biosynthetic gene cluster for holomycin produced by *Streptomyces clavuligerus*. Process Biochem. 46(3), 811-816.

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., et al. (2015). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res. 44(D1):D286-93.

Hughes, L. E., Beck, D. A., & O'Donovan, G. A. (2005). Pathways of pyrimidine salvage in *Streptomyces*. Curr Microbiol. 50(1), 8-10.

Hung, L. W., Wang, I. X., Nikaido, K., Liu, P. Q., Ames, G. F. L., & Kim, S. H. (1998). Crystal structure of the ATP-binding subunit of an ABC transporter. Nature, 396(6712), 703-707.

Hurley, J. H., Chen, R., & Dean, A. M. (1996). Determinants of cofactor specificity in isocitrate dehydrogenase: structure of an engineered $NADP^+ \rightarrow NAD^+$ specificity-reversal mutant. Biochemistry, 35(18), 5670-5678.

Hutchings, M. I., Hoskisson, P. A., Chandra, G., & Buttner, M. J. (2004). Sensing and responding to diverse extracellular signals? Analysis of the sensor kinases and response regulators of *Streptomyces coelicolor* A3 (2). Microbiology, 150(9), 2795-2806.

Huynh, Q. K., Bauer, S. C., Bild, G. S., Kishore, G. M., & Borgmeyer, J. R. (1988). Site-directed mutagenesis of *Petunia hybrida* 5-enolpyruvylshikimate-3-phosphate synthase: Lys-23 is essential for substrate binding. J Biol Chem. 263(24), 11636-11639.

Igual, J. C., Gonzalez-Bosch, C., Dopazo, J., & Perez-Ortin, J. E. (1992). Phylogenetic analysis of the thiolase family. Implications for the evolutionary origin of peroxisomes. J Mol Evol. 35(2), 147-155.

Ikeda, C., Hayashi, Y., Itoh, N., Seto, H., & Dairi, T. (2007). Functional analysis of eubacterial ent-copalyl diphosphate synthase and pimara-9 (11), 15-diene synthase with unique primary sequences. J Biochem. 141(1), 37-45.

Inbar, L., & Lapidot, A. (1988). The structure and biosynthesis of new tetrahydropyrimidine derivatives in actinomycin D producer *Streptomyces parvulus*. Use of 13C-and 15N-labeled L-glutamate and 13C and 15N NMR spectroscopy. J Biol Chem. 263(31), 16014-16022.

Iost, I., & Dreyfus, M. (2006). DEAD-box RNA helicases in *Escherichia coli*. Nucleic Acids Res. 34(15), 4189-4197.

Ishida, A., Ichikawa, M., Kobayashi, K., Hitomi, T., Kojima, S., & Toraya, T. (1993). Importance of the nucleotide loop moiety corrdinated to the cobalt atom of adenosylcobalamin for coenzymic function in the diol dehydrase reaction. J Nutr Sci Vitaminol. 39(2), 115-125.

Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., & Mann, M. (2005). Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics. 4(9), 1265-1272.

Itoh, N., & Morikawa, R. (1983). Crystallization and properties of L-alanine dehydrogenase from *Streptomyces phaeochromogenes*. Agr Biol Chem Tokyo. 47(11), 2511-2519.

Iyer, L. M., & Aravind, L. (2002). The catalytic domains of thiamine triphosphatase and CyaB-like adenylyl cyclase define a novel superfamily of domains that bind organic phosphates. BMC Genomics, 3, 33.

Jain, V., Kumar, M., & Chatterji, D. (2006). ppGpp: stringent response and survival. J Microbiol. 44(1), 1-10.

Jaroszewski, L., Li, Z., Krishna, S. S., Bakolitsa, C., Wooley, J., Deacon, A. M., et al. (2009). Exploration of uncharted regions of the protein universe. PLoS Biol. 7(9), e1000205.

Jensen, R. B., & Gerdes, K. (1999). Mechanism of DNA segregation in prokaryotes: ParM partitioning protein of plasmid R1 co-localizes with its replicon during the cell cycle. The EMBO J. 18(14), 4076-4084.

Jensen, S. E. & Paradkar, A. S. (1999). Biosynthesis and molecular genetics of clavulanic acid. Antonie Van Leeuwenhoek. 75(1-2), 125-133.

Jensen, S. E. (2012). Biosynthesis of clavam metabolites. J Ind Microbiol Biotechnol. 39(10), 1407-1419.

Jensen, S. E., Wong, A., Griffin, A., & Barton, B. (2004). *Streptomyces clavuligerus* has a second copy of the proclavaminate amidinohydrolase gene. Antimicrob Agents Ch. 48(2), 514-520.

Jin, W., Ryu, Y. G., Kang, S. G., Kim, S. K., Saito, N., Ochi, K. et al. (2004). Two *relA/spoT* homologous genes are involved in the morphological and physiological differentiation of *Streptomyces clavuligerus*. Microbiology, 150(5), 1485-1493.

Jing, H., Takagi, J., Liu, J. H., Lindgren, S., Zhang, R. G., Joachimiak, A., et al. (2002). Archaeal surface layer proteins contain β propeller, PKD, and β helix domains and are related to metazoan cell surface proteins. Structure. 10(10), 1453-1464.

Johnston, M., Jankowski, D., Marcotte, P., Tanaka, H., Esaki, N., Soda, K., & Walsh, C. (1979). Suicide inactivation of bacterial cystathionine γ-synthase and methionine γ-lyase during processing of L-propargylglycine. Biochemistry. 18(21), 4690-4701.

Johnstone, T. C., & Nolan, E. M. (2015). Beyond iron: non-classical biological functions of bacterial siderophores. Dalton T. 44(14), 6320-6339.

Jones, J. M., Nau, K., Geraghty, M. T., Erdmann, R., & Gould, S. J. (1999). Identification of peroxisomal acyl-CoA thioesterases in yeast and humans. J Biol Chem. 274(14), 9216-9223.

Jost, M., Born, D. A., Cracan, V., Banerjee, R., & Drennan, C. L. (2015). Structural basis for substrate specificity in adenosylcobalamin-dependent isobutyryl-CoA mutase and related Acyl-CoA mutases. J Biol Chem. 290(45), 26882-26898.

Kagliwal, L. D., Survase, S. A., & Singhal, R. S. (2009). A novel medium for the production of cephamycin C by *Nocardia lactamdurans* using solid-state fermentation. Bioresource Technol. 100(9), 2600-2606.

Kahel-Raifer, H., Jindou, S., Bahari, L., Nataf, Y., Shoham, Y., Bayer, E. A., et al. (2010). The unique set of putative membrane-associated anti-σ factors in *Clostridium thermocellum* suggests a novel extracellular carbohydrate-sensing mechanism involved in gene regulation. FEMS Microbial Lett. 308(1), 84-93.

Kang, W. K., Icho, T., Isono, S., Kitakawa, M., & Isono, K. (1989). Characterization of the gene *rimK* responsible for the addition of glutamic acid residues to the C-terminus of ribosomal protein S6 in *Escherichia coli K12*. Mol Gen Genet. 217(2-3), 281-288.

Kannenberg, E. L. & Poralla, K. (1999). Hopanoid biosynthesis and function in bacteria. *Naturwissenschaften*, *86*(4), 168-176.

Karas, M. & Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem. 60(20), 2299-2301.

Kasarenini, S., & Demain, A. L. (1994). A role for alanine in the ammonium regulation of cephalosporin biosynthesis in *Streptomyces clavuligerus*. J Ind Microbiol 13(4), 217-219.

Kashiwagi, K., & Igarashi, K. (1987). Nonspecific inhibition of *Escherichia coli* ornithine decarboxylase by various ribosomal proteins: detection of a new ribosomal protein possessing strong antizyme activity. Biochim Biophys Acta. 911(2), 180-190.

Kato, J. Y., Funa, N., Watanabe, H., Ohnishi, Y., & Horinouchi, S. (2007). Biosynthesis of γ-butyrolactone autoregulators that switch on secondary metabolism and morphological development in *Streptomyces*. Proc Natl Acad Sci U S A. 104(7), 2378-2383.

Katz, L., & Donadio, S. (1993). Polyketide synthesis: prospects for hybrid antibiotics. Annu Rev Microbiol. 47(1), 875-912.

Kaufmann, R. (1995). Matrix-assisted laser desorption ionization (MALDI) mass spectrometry: a novel analytical tool in molecular biology and biotechnology. J Biotechnol. 41(2), 155-175.

Kavanagh, K. L., Jörnvall, H., Persson, B., & Oppermann, U. (2008). The SDR superfamily: Functional and structural diversity within a family of metabolic and regulatory enzymes. Cell. Mol. Life Sci, 65, 3895-3906.

Kenig, M. & Reading, C. (1979). Holomycin and an antibiotic (MM 19290) related to tunicamycin, metabolites of *Streptomyces clavuligerus*. J Antibiot (Tokyo). 32(6), 549-554.

Kershaw, N. J., McNaughton, H. J., Hewitson, K. S., Hernández, H., Griffin, J., Hughes, C., et al. (2002). ORF6 from the clavulanic acid gene cluster of *Streptomyces clavuligerus* has ornithine acetyltransferase activity. Eur J Biochem. 269(8), 2052-2059.

Kim, D. J., Huh, J. H., Yang, Y. Y., Kang, C. M., Lee, I. H., Hyun, C. G., et al. (2003). Accumulation of S-adenosyl-L-methionine enhances production of actinorhodin but inhibits sporulation in *Streptomyces lividans* TK23. J Bacteriol. 185(2), 592-600.

Kim, H. S., Lee, Y. J., Lee, C. K., Choi, S. U., Yeo, S. H., Hwang, Y. I. et al. (2004). Cloning and characterization of a gene encoding the γ-butyrolactone autoregulator receptor from *Streptomyces clavuligerus.* Arch Microbiol. 182*(1), 44-50.*

Kim, H. S., & Park, Y. I. (2008). Isolation and identification of a novel microorganism producing the immunosuppressant tacrolimus. J Biosci Bioeng. 105(4), 418-421.

Kim, J. N., Kim, Y., Jeong, Y., Roe, J. H., Kim, B. G., & Cho, B. K. (2015a). Comparative genomics reveals the core and accessory genomes of *Streptomyces* species. J. Microbiol. Biotechnol, 25(10), 1599-1605.

Kim, P. B., Nelson, J. W., & Breaker, R. R. (2015b). An ancient riboswitch class in bacteria regulates purine biosynthesis and one-carbon metabolism. Mol Cell. 57(2), 317-328.

Kimura, S., & Suzuki, T. (2010). Fine-tuning of the ribosomal decoding center by conserved methyl-modifications in the *Escherichia coli* 16S rRNA. Nucleic Acids Res. 38(4), 1341-1352.

Kitagawa, H., Ozawa, T., Takahata, S., Iida, M., Saito, J., & Yamada, M. (2007). Phenylimidazole derivatives of 4-pyridone as dual inhibitors of bacterial enoyl-acyl carrier protein reductases FabI and FabK. J Med Chem. 50(19), 4710-4720.

Klem, T. J. & Davisson, V. J. (1993). Imidazole glycerol phosphate synthase: the glutamine amidotransferase in histidine biosynthesis. Biochemistry. 32(19), 5177-5186.

Knaus, T. Schober, M., Kepplinger, B., Faccinelli, M., Pitzer, J., Faber, K., et al. (2012). Structure and mechanism of an inverting alkylsulfatase from *Pseudomonas* sp. DSM6611 specific for secondary alkyl sulfates. Febs J. 279(23), 4374-4384.

Koetsier, M. J., Jekel, P. A., Van den Berg, M. A., Bovenberg, R. A., & Janssen, D. B. (2009). Characterization of a phenylacetate–CoA ligase from *Penicillium chrysogenum*. Biochem J. 417(2), 467-476.

Kofron, J. L., Kuzmic, P., Kishore, V., Colon-Bonilla, E., & Rich, D. H. (1991). Determination of kinetic constants for peptidyl prolyl cis-trans isomerases by an improved spectrophotometric assay. Biochemistry, 30(25), 6127-6134.

Koh, T. H., Yamaguchi, K., & Ishii, Y. (2008). Characterisation of the metallo-β-lactamase VIM-6 and its genetic support. *Int J Antimicrob Ag*. 32(5), 446-449.

Kohanski, M. A., Dwyer, D. J., & Collins, J. J. (2010). How antibiotics kill bacteria: from targets to networks. Nat Rev Microbiol. 8(6), 423-435.

Kohli, R. M., Trauger, J. W., Schwarzer, D., Marahiel, M. A., & Walsh, C. T. (2001). Generality of peptide cyclization catalyzed by isolated thioesterase domains of nonribosomal peptide synthetases. Biochemistry, 40(24), 7099-7108.

Koike, M., Reed, L. J., & Carroll, W. R. (1960). α-Keto acid dehydrogenation complexes I. Purification and properties of pyruvate and α-ketoglutarate dehydrogenation complexes of *Escherichia coli*. J Biol Chem. 235(7), 1924-1930.

Kojo, H., Mine, Y., Nishida, M., Goto, S., & Kuwahara, S. (1988). Nature of monocyclic β-lactam antibiotic nocardicin A to β-lactamases. Microbiol Immunol. *32*(2), 119-130.

Komatsu, M., Uchiyama, T., Ōmura, S., Cane, D. E., & Ikeda, H. (2010). Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. Proc Natl Acad Sci U S A. 107(6), 2646-2651.

Komeda, H., & Asano, Y. (2000). Gene cloning, nucleotide sequencing, and purification and characterization of the D-stereospecific amino-acid amidase from *Ochrobactrum anthropi* SV3. Eur J Biochem. 267(7), 2028-2035.

Komeda, H., & Asano, Y. (2003). Genes for an alkaline D-stereospecific endopeptidase and its homolog are located in tandem on *Bacillus cereus* genome. FEMS Microbial. Lett. 228(1), 1-9.

Konermann, L., Ahadi, E., Rodriguez, A. D., & Vahidi, S. (2012). Unraveling the mechanism of electrospray ionization. Anal Biochem. 85(1), 2-9.

Kong, R., Liu, X., Su, C., Ma, C., Qiu, R., & Tang, L. (2013). Elucidation of the biosynthetic gene cluster and the post-PKS modification mechanism for fostriecin in *Streptomyces pulveraceus*. Chem Biol. 20(1), 45-54.

Koper, T. E., El-Sheikh, A. F., Norton, J. M., & Klotz, M. G. (2004). Urease-encoding genes in ammonia-oxidizing bacteria. Appl Environ Microb. 70(4), 2342-2348.

Kovacevic, S. & Miller, J. R. (1991). Cloning and sequencing of the β-lactam hydroxylase gene (*cefF*) from *Streptomyces clavuligerus*: gene duplication may have led to separate hydroxylase and expandase activities in the actinomycetes. J Bacteriol. 173(1), 398-400.

Kreft, B. D., Townsend, A., Pohlenz, H. D., & Laber, B. (1994). Purification and properties of cystathionine γ-synthase from wheat (*Triticum aestivum* L.). Plant Physiol. 104(4), 1215-1220.

Kruger, N. J. & von Schaewen, A. (2003). The oxidative pentose phosphate pathway: structure and organisation. Curr Opin Plant Biol. 6(3), 236-246.

Kurt, A., Álvarez-Álvarez, R., Liras, P., & Özcengiz, G. (2013). Role of the *cmcH–ccaR* intergenic region and *ccaR* overexpression in cephamycin C biosynthesis in *Streptomyces clavuligerus*. Appl Microbiol Biot. 97(13), 5869-5880.

Kvittingen, E. A. (1985). Hereditary tyrosinemia type I--an overview. Scand J Clin Lab Invest Suppl. 184, 27-34.

Kwak, J., Kendrick, K.E. (1996). Bald mutants of *Streptomyces griseus* that prematurely undergo key events of sporulation. J Bacteriol. 178(15), 4643-4650.

Le Maréchal, P., Decottignies, P., Marchand, C. H., Degrouard, J., Jaillard, D., Dulermo, et al. (2013). Comparative proteomic analysis of *Streptomyces lividans* wild-type and *ppk* mutant strains reveals the importance of storage lipids for antibiotic biosynthesis. Appl Environ Microb. 79(19), 5907-5917.

le Roes-Hill, M., Khan, N., & Burton, S. G. (2011). Actinobacterial peroxidases: an unexplored resource for biocatalysis. Appl Biochem Biotech. 164(5), 681-713.

Lee, E. J., Karoonuthaisiri, N., Kim, H. S., Park, J. H., Cha, C. J., Kao, C. M., & Roe, J. H. (2005). A master regulator σ^B governs osmotic and oxidative response as well as differentiation via a network of sigma factors in *Streptomyces coelicolor*. Mol Microbiol. 57(5), 1252-1264.

Leiza, M. G., Perez-Diaz, J. C., Ayala, J., Casellas, J. M., Martinez-Beltran, J., Bush, K., & Baquero, F. (1994). Gene sequence and biochemical characterization of FOX-1 from *Klebsiella pneumoniae*, a new AmpC-type plasmid-mediated beta-lactamase with two molecular variants. Antimicrob Agents Ch. 38(9), 2150-2157.

Lenart, A., Dudkiewicz, M., Grynberg, M., & Pawłowski, K. (2013). CLCAs-a family of metalloproteases of intriguing phylogenetic distribution and with cases of substituted catalytic sites. PloS One, 8(5), e62272.

Leskiw, B. K., Lawlor, E. J., Fernandez-Abalos, J. M., & Chater, K. F. (1991). TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. Proc Natl Acad Sci U S A. 88(6), 2461-2465.

Li, L., Bannantine, J. P., Zhang, Q., Amonsin, A., May, B. J., Alt, D., et al. (2005a). The complete genome sequence of *Mycobacterium avium* subspecies *paratuberculosis*. Proc Natl Acad Sci U S A. 102(35), 12344-12349.

Li, J., Cross, J. B., Vreven, T., Meroueh, S. O., Mobashery, S., & Schlegel, H. B. (2005b). Lysine carboxylation in proteins: OXA-10 β-lactamase. Proteins. 61(2), 246-257.

Li, W., Ying, X., Guo, Y., Yu, Z., Zhou, X., Deng, Z., et al. (2006). Identification of a gene negatively affecting antibiotic production and morphological differentiation in *Streptomyces coelicolor* A3 (2). J Bacteriol. 188(24), 8368-8375.

Li, W., Chou, S., Khullar, A., & Gerratana, B. (2009). Cloning and characterization of the biosynthetic gene cluster for tomaymycin, an SJG-136 monomeric analog. Appl Environ Microb. 75(9), 2958-2963.

Li, B., & Walsh, C. T. (2010). Identification of the gene cluster for the dithiolopyrrolone antibiotic holomycin in *Streptomyces clavuligerus*. Proc Natl Acad Sci U S A. 107(46), 19731-19735.

Li, G., & Howard, S. P. (2010). ExeA binds to peptidoglycan and forms a multimer for assembly of the type II secretion apparatus in *Aeromonas hydrophila*. Mol Microbiol. 76(3), 772-781.

Li, B., & Walsh, C. T. (2011). *Streptomyces clavuligerus* HlmI is an intramolecular disulfide-forming dithiol oxidase in holomycin biosynthesis. Biochemistry, 50(21), 4615-4622.

Li, F., Liang, J., Wang, W., Zhou, X., Deng, Z., & Wang, Z. (2014). Analysis of *Streptomyces coelicolor* membrane proteome using two-dimensional native/native and native/sodium dodecyl sulfate gel electrophoresis. Anal Biochem. 465, 148-155.

Li, K., Chen, W. H., & Bruner, S. D. (2015). Structure and mechanism of the siderophore-interacting protein from the fuscachelin gene cluster of *Thermobifida fusca*. Biochemistry, 54(25), 3989-4000.

Liang, L., Liu, R., Wang, G., Gou, D., Ma, J., Chen, K., Jiang M. (2012). Regulation of NAD (H) pool and NADH/NAD^+ ratio by overexpression of nicotinic acid phosphoribosyltransferase for succinic acid production in *Escherichia coli* NZN111. Enzyme Microb Tech. 51(5), 286-293.

Liang, Y., Gao, Z., Dong, Y., & Liu, Q. (2014). Structural and functional analysis show that the *Escherichia coli* uncharacterized protein YjcS is likely an alkylsulfatase. Protein Sci. 23(10), 1442-1450.

Lindahl, R. (1992). Aldehyde dehydrogenases and their role in carcinogenesis. Crit Rev Biochem Mol. *27*(4-5), 283-335.

Lindner, S. N., Knebel, S., Wesseling, H., Schoberth, S. M., & Wendisch, V. F. (2009). Exopolyphosphatases PPX1 and PPX2 from *Corynebacterium glutamicum*. Appl Environ Microb. 75(10), 3161-3170.

Liou, G. F., & Khosla, C. (2003). Building-block selectivity of polyketide synthases. Curr Opin Chem Biol. 7(2), 279-284.

Liras, P. & Rodriguez-Garcia, A. (2000). Clavulanic acid, a β-lactamase inhibitor: biosynthesis and molecular genetics. Appl Microbiol Biot. 54(4), 467-475.

Liras, P. (1999). Biosynthesis and molecular genetics of cephamycins. *Antonie van Leeuwenhoek*, *75*(1-2), 109-124.

Liras, P., & Martín, J. F. (2006). Gene clusters for β-lactam antibiotics and control of their expression: why have clusters evolved, and from where did they originate? Int Microbiol. 9(1), 9.

Liras, P., Gomez-Escribano, J. P., & Santamarta, I. (2008). Regulatory mechanisms controlling antibiotic production in *Streptomyces clavuligerus*. J Ind Microbiol Biotechnol. 35(7), 667-676.

Liras, P. (2014). Holomycin, a dithiolopyrrolone compound produced by *Streptomyces clavuligerus*. Appl Microbiol Biot. 98(3), 1023-1030.

Liu, H., & Reynolds, K. A. (2001). Precursor supply for polyketide biosynthesis: the role of crotonyl-CoA reductase. Metab Eng. 3(1), 40-48.

Liu, W., Ahlert, J., Gao, Q., Wendt-Pienkowski, E., Shen, B., & Thorson, J. S. (2003). Rapid PCR amplification of minimal enediyne polyketide synthase cassettes leads to a predictive familial classification model. Proc Natl Acad Sci U S A. 100(21), 11959-11963.

Liu, W., Nonaka, K., Nie, L., Zhang, J., Christenson, S. D., Bae, J., et al. (2005). The neocarzinostatin biosynthetic gene cluster from *Streptomyces carzinostaticus* ATCC 15944 involving two iterative type I polyketide synthases. Chem Biol. 12(3), 293-302.

Lochlainn, L. N., & Caffrey, P. (2009). Phosphomannose isomerase and phosphomannomutase gene disruptions in *Streptomyces nodosus*: impact on amphotericin biosynthesis and implications for glycosylation engineering. Metab Eng. 11(1), 40-47.

Loeber, G., Infante, A. A., Maurer-Fogy, I., Krystek, E., & Dworkin, M. B. (1991). Human NAD (+)-dependent mitochondrial malic enzyme. cDNA cloning, primary structure, and expression in *Escherichia coli*. J Biol Chem. 266(5), 3016-3021.

Loessner, M. J. (2005). Bacteriophage endolysins—current state of research and applications. Curr Opin Microbiol. 8(4), 480-487.

Loke, P., Goh, L. L., Soh, B. S., Yeow, P., & Sim, T. S. (2002). Purification and characterization of recombinant malate synthase enzymes from *Streptomyces coelicolor* A3 (2) and *S. clavuligerus* NRRL3585. J Ind Microbiol Biotechnol. *28*(4), 239-243.

Lonetto, M. A., Brown, K. L., Rudd, K. E., & Buttner, M. J. (1994). Analysis of the *Streptomyces coelicolor sigE* gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. Proc Natl Acad Sci U S A. 91(16), 7573-7577.

López-García, M. T., Santamarta, I., & Liras, P. (2010). Morphological differentiation and clavulanic acid formation are affected in a *Streptomyces clavuligerus adpA*-deleted mutant. Microbiology, 156(8), 2354-2365.

Lorenzana, L. M., Pérez-Redondo, R., Santamarta, I., Martín, J. F., & Liras, P. (2004). Two oligopeptide-permease-encoding genes in the clavulanic acid cluster of *Streptomyces clavuligerus* are essential for production of the β-lactamase inhibitor. J Bacteriol. *186*(11), 3431-3438.

Loseva, O., Jemth, A. S., Bryant, H. E., Schüler, H., Lehtiö, L., Karlberg, T., & Helleday, T. (2010). PARP-3 is a mono-ADP-ribosylase that activates PARP-1 in the absence of DNA. J Biol Chem. 285(11), 8054-8060.

Löwe, J., & Amos, L. A. (1998). Crystal structure of the bacterial cell-division protein FtsZ. Nature, 391(6663), 203-206.

Lu, P., Vogel, C., Wang, R., Yao, X., & Marcotte, E. M. (2007). Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nat Biotechnol. 25(1), 117-124.

Luckner, M. (2013). Secondary metabolism in microorganisms, plants and animals. Springer Science & Business Media.

Mackenzie, A. (2007). Studies on the biosynthetic pathways of clavulanic acid and cephamycin C in *Streptomyces clavuligerus*. Swedish University of Agricultural Sciences, Uppsala, Sweden

Mackenzie, A. K., Valegård, K., Iqbal, A., Caines, M. E., Kershaw, N. J., Jensen, S. E., et al. (2010). Crystal structures of an oligopeptide-binding protein from the biosynthetic pathway of the β-Lactamase inhibitor clavulanic acid. J Mol Biol. 396(2), 332-344.

Magnusson, L. U., Farewell, A., & Nyström, T. (2005). ppGpp: a global regulator in *Escherichia coli*. Trends Microbiol. 13(5), 236-242.

Makitrynskyy, R., Ostash, B., Tsypik, O., Rebets, Y., Doud, E., Meredith, T. et al (2013). Pleiotropic regulatory genes *bldA*, *adpA* and *absB* are implicated in production of phosphoglycolipid antibiotic moenomycin. Open Biol. *3*(10), 130121.

Malin, G. & Lapidot, A. (1996). Induction of synthesis of tetrahydropyrimidine derivatives in *Streptomyces* strains and their effect on *Escherichia coli* in response to osmotic and heat stress. J Bacteriol. 178(2), 385-395.

Malmberg, L. H., Hu, W. S., & Sherman, D. H. (1993). Precursor flux control through targeted chromosomal insertion of the lysine epsilon-aminotransferase (*lat*) gene in cephamycin C biosynthesis. J Bacteriol. 175(21), 6916-6924.

Mann, M., Hendrickson, R. C., & Pandey, A. (2001). Analysis of proteins and proteomes by mass spectrometry. Annu Rev Biochem. 70(1), 437-473.

Mann, M. (2006). Functional and quantitative proteomics using SILAC. Nat Rev Mol Cell Bio. 7(12), 952-958.

Manteca, A., Alvarez, R., Salazar, N., Yagüe, P., & Sanchez, J. (2008). Mycelium differentiation and antibiotic production in submerged cultures of *Streptomyces coelicolor*. Appl Environ Microb. 74(12), 3877-3886.

Manteca, A., Claessen, D., Lopez-Iglesias, C., & Sanchez, J. (2007). Aerial hyphae in surface cultures of *Streptomyces lividans* and *Streptomyces coelicolor* originate from viable segments surviving an early programmed cell death event. FEMS Microbiol Lett. 274(1), 118-125.

Marahiel, M. A., & Essen, L. O. (2009). Nonribosomal peptide synthetases: mechanistic and structural aspects of essential domains. Method Enzymol. 458, 337-351.

Martín J.F., Gutiérrez S., Aparicio J.F. (2000) Secondary metabolites. In:Lederberg J (ed) Encyclopedia of Microbiology, vol. 4, 2nd ed. Academic Press, San Diego, CA, pp 213-236.

Martín, J. F. (1998). New aspects of genes and enzymes for β-lactam antibiotic biosynthesis. Appl Microbiol Biot. 50(1), 1-15.

Martín, J. F., & Liras, P. (2010). Engineering of regulatory cascades and networks controlling antibiotic biosynthesis in *Streptomyces*. Curr Opin Microbiol. 13(3), 263-273.

Martín, J. F., Ullán, R. V., & García-Estrada, C. (2010). Regulation and compartmentalization of β-lactam biosynthesis. Microb Biotechnol. 3(3), 285-299.

Martín, J. F., Sola-Landa, A., Santos-Beneit, F., Fernández-Martínez, L. T., Prieto, C., & Rodríguez-García, A. (2011). Cross-talk of global nutritional regulators in the control of primary and secondary metabolism in *Streptomyces*. Microb Biotechnol. 4(2), 165-174.

Martínez-Burgo, Y., Álvarez-Álvarez, R., Rodríguez-García, A., & Liras, P. (2015). The pathway-specific regulator ClaR of Streptomyces clavuligerus has a global effect on the expression of genes for secondary metabolism and differentiation. Appl Environ Microb. 81(19), 6637-6648.

Martínez-Costa, O. H., Arias, P., Romero, N. M., Parro, V., Mellado, R. P., & Malpartida, F. (1996). A *relA/spoT* homologous gene from *Streptomyces coelicolor* A3 (2) controls antibiotic biosynthetic genes. J Biol Chem. 271(18), 10627-10634.

Mascher, T., Helmann, J. D., & Unden, G. (2006). Stimulus perception in bacterial signal-transducing histidine kinases. Microbiol Mol Biol R. 70(4), 910-938.

Mattevi, A., Obmolova, G., Kalk, K. H., Westphal, A. H., de Kok, A., & Hol, W. G. (1993). Refined crystal structure of the catalytic domain of dihydrolipoyl transacetylase (E2p) from *Azotobacter vinelandii* at 2.6 Å resolution. J Mol Biol. 230(4), 1183-1199.

McCormick, J. R., & Flärdh, K. (2012). Signals and regulators that govern *Streptomyces* development. FEMS Microbiol Rev. *36*(1), 206-231.

McCranie, E. K., & Bachmann, B. O. (2014). Bioactive oligosaccharide natural products. Nat Prod Rep. 31(8), 1026-1042.

McLennan, A. G. (2006). The Nudix hydrolase superfamily. *Cellular and Molecular Life Sciences CMLS*, *63*(2), 123-143.

Medema, M. H., Alam, M. T., Heijne, W. H., van den Berg, M. A., Müller, U., Trefzer, A., et al. (2011). Genome-wide gene expression changes in an industrial clavulanic acid overproduction strain of *Streptomyces clavuligerus*. Microb Biotechnol. 4(2), 300-305.

Medema, M. H., Trefzer, A., Kovalchuk, A., van den Berg, M., Müller, U., Heijne, W. et al. (2010). The sequence of a 1.8-Mb bacterial linear plasmid reveals a rich evolutionary reservoir of secondary metabolic pathways. Genome Biol Evol. 2, 212-224.

Mehta, P. K., Hale, T. I., & Christen, P. (1993). Aminotransferases: demonstration of homology and division into evolutionary subgroups. Eur J Biochem. 214(2), 549-561.

Mehta, V. J., Thumar, J. T., & Singh, S. P. (2006). Production of alkaline protease from an alkaliphilic actinomycete. Bioresource Technol. 97(14), 1650-1654.

Mellado, E., Lorenzana, L. M., Rodríguez-Sáiz, M., Díez, B., Liras, P., & Barredo, J. L. (2002). The clavulanic acid biosynthetic cluster of *Streptomyces clavuligerus*: genetic organization of the region upstream of the car gene. Microbiology, 148(5), 1427-1438.

Mendelovitz, S., & Aharonowitz, Y. (1982). Regulation of cephamycin C synthesis, aspartokinase, dihydrodipicolinic acid synthetase, and homoserine dehydrogenase by aspartic acid family amino acids in *Streptomyces clavuligerus*. Antimicrob Agents Ch. 21(1), 74-84.

Mengin-Lecreulx, D., & van Heijenoort, J. (1996). Characterization of the essential gene *glmM* encoding phosphoglucosamine mutase in *Escherichia coli*. J Biol Chem. 271(1), 32-39.

Merchant, M., & Weinberger, S. R. (2000). Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry. Electrophoresis, 21(6), 1164-1177.

Merril, C. R., Switzer, R. C., & Van Keuren, M. L. (1979). Trace polypeptides in cellular extracts and human body fluids detected by two-dimensional electrophoresis and a highly sensitive silver stain. Proc Natl Acad Sci U S A. 76(9), 4335-4339.

Miguélez, E. M., Hardisson, C., & Manzanal, M. B. (1999). Hyphal death during colony development in *Streptomyces antibioticus*: morphological evidence for the existence of a process of cell deletion in a multicellular prokaryote. J Cell Biol. 145(3), 515-525.

Miller, M. T., Bachmann, B. O., Townsend, C. A., & Rosenzweig, A. C. (2002). The catalytic cycle of β-lactam synthetase observed by x-ray crystallographic snapshots. Proc Natl Acad Sci U S A. 99(23), 14752-14757.

Minotti, G., Menna, P., Salvatorelli, E., Cairo, G., & Gianni, L. (2004). Anthracyclines: molecular advances and pharmacologic developments in antitumor activity and cardiotoxicity. Pharmacol Rev. 56(2), 185-229.

Mliki, A., & Zimmermann, W. (1992). Purification and characterization of an intracellular peroxidase from *Streptomyces cyaneus*. Appl Environ Microb. 58(3), 916-919.

Mo, S., Ban, Y. H., Park, J. W., Yoo, Y. J., & Yoon, Y. J. (2009). Enhanced FK506 production in *Streptomyces clavuligerus* CKD1119 by engineering the supply of methylmalonyl-CoA precursor. J Ind Microbiol Biotechnol. 36(12), 1473-1482.

Moe Jr, O. A., & Hammes, G. G. (1974). Binding of thiamine diphosphate and thiochrome diphosphate to the pyruvate dehydrogenase multienzyme complex. Biochemistry, 13(12), 2547-2552.

Moe Jr, O. A., Lerner, D. A., & Hammes, G. G. (1974). Fluorescence energy transfer between the thiamine diphosphate and flavine adenine dinucleotide binding sites in the pyruvate dehydrogenase multienzyme complex. Biochemistry, 13(12), 2552-2557.

Mohanty, B. K., & Kushner, S. R. (2000). Polynucleotide phosphorylase functions both as a 3′→ 5′ exonuclease and a poly (A) polymerase in *Escherichia coli*. Proc Natl Acad Sci USA. 97(22), 11966-11971.

Molle, V., & Buttner, M. J. (2000). Different alleles of the response regulator gene *bldM* arrest *Streptomyces coelicolor* development at distinct stages. Mol Microbiol. 36(6), 1265-1278.

Molle, V., Kremer, L., Girard-Blanc, C., Besra, G. S., Cozzone, A. J., & Prost, J. F. (2003). An FHA phosphoprotein recognition domain mediates protein EmbR phosphorylation by PknH, a Ser/Thr protein kinase from *Mycobacterium tuberculosis*. Biochemistry, 42(51), 15300-15309.

Molle, V., Soulat, D., Jault, J. M., Grangeasse, C., Cozzone, A. J., & Prost, J. F. (2004). Two FHA domains on an ABC transporter, Rv1747, mediate its phosphorylation by PknF, a Ser/Thr protein kinase from *Mycobacterium tuberculosis*. FEMS Microbiol Lett. 234(2), 215-223.

Moore, B. S. (2008). Extending the biosynthetic repertoire in ribosomal peptide assembly. Angew Chem Int Ed Engl. 47(49), 9386-9388.

Moradian, A., Kalli, A., Sweredoski, M. J., & Hess, S. (2014). The top-down, middle-down, and bottom-up mass spectrometry approaches for characterization of histone variants and their post-translational modifications. Proteomics, 14(4-5), 489-497.

Moriguchi, M., & Ideta, K. (1988). Production of D-aminoacylase from *Alcaligenes denitrificans* subsp. *xylosoxydans* MI-4. *Appl Environ Microb. 54*(11), 2767-2770.

Mosher, R. H., Paradkar, A. S., Anders, C., Barton, B., & Jensen, S. E. (1999). Genes specific for the biosynthesis of clavam metabolites antipodal to clavulanic acid are clustered with the gene for clavaminate synthase 1 in *Streptomyces clavuligerus*. *Antimicrob Agents Ch. 43*(5), 1215-1224.

Mougous, J. D., Cuff, M. E., Raunser, S., Shen, A., Zhou, M., Gifford, C. A., et al. (2006). A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. Science, 312(5779), 1526-1530.

Muniz-Lozano, F. E., Domınguez-Sánchez, G., Dıaz-Viveros, Y., & Barradas-Dermitz, D. M. (1998). D-Aminoacylase from a novel producer: *Stenotrophomonas maltophilia* ITV-0595. J Ind Microbiol Biot. 21(6), 296-299.

Nagase, H. (2001). Metalloproteases. Curr Protoc. Protein Sci. 21-4.

Nagy, I., Tamura, T., Vanderleyden, J., Baumeister, W., & De Mot, R. (1998). The 20S proteasome of *Streptomyces coelicolor*. J Bacteriol. 180(20), 5448-5453.

Nakano, T., Miyake, K., Endo, H., Dairi, T., Mizukami, T., and Katsumata, R. (2004). Identification and cloning of the gene involved in the final step of chlortetracycline biosynthesis in *Streptomyces aureofaciens*. Biosci Biotechnol Biochem. 68: 1345–1352.

Nakamaru-Ogiso, E., Yano, T., Yagi, T., & Ohnishi, T. (2005). Characterization of the iron-sulfur cluster N7 (N1c) in the subunit NuoG of the proton-translocating NADH-quinone oxidoreductase from *Escherichia coli*. J Biol Chem. 280(1), 301-307.

Nandan, A., Pandey, A., & Nampoothiri, K. M. (2011). Proline-specific extracellular aminopeptidase purified from *Streptomyces lavendulae*. Appl Biochem Biotech. 163(8), 994-1001.

Nandan, A. S., & Nampoothiri, K. M. (2014). Unveiling aminopeptidase P from *Streptomyces lavendulae*: Molecular cloning, expression and biochemical characterization. Enzyme Microb Tech. 55, 7-13.

Nárdiz, N., Santamarta, I., Lorenzana, L. M., Martín, J. F., & Liras, P. (2011). A rhodanese-like protein is highly overrepresented in the mutant *S. clavuligerus oppA2:: aph:* effect on holomycin and other secondary metabolites production. Microb Biotechnol. 4(2), 216-225.

Natale, P., Brüser, T., & Driessen, A. J. (2008). Sec-and Tat-mediated protein secretion across the bacterial cytoplasmic membrane—distinct translocases and mechanisms. Bba-Biomembranes. 1778(9), 1735-1756.

Neer, E. J., Schmidt, C. J., Nambudripad, R., & Smith, T. F. (1994). The ancient regulatory-protein family of WD-repeat proteins. Nature, 371(6495), 297.

Neilson, K. A., Ali, N. A., Muralidharan, S., Mirzaei, M., Mariani, M., Assadourian, G., et al. (2011). Less label, more free: approaches in label-free quantitative mass spectrometry. Proteomics, 11(4), 535-553.

Netolitzky, D. J., Wu, X., Jensen, S. E., & Roy, K. L. (1995). Giant linear plasmids of β-lactam antibiotic producing *Streptomyces*. FEMS Microbiol Lett. *131*(1), 27-34.

Nett, M., Ikeda, H., & Moore, B. S. (2009). Genomic basis for natural product biosynthetic diversity in the actinomycetes. Nat Prod Rep. 26(11), 1362-1384.

Neuhoff, V., Stamm, R., & Eibl, H. (1985). Clear background and highly sensitive protein staining with Coomassie Blue dyes in polyacrylamide gels: a systematic analysis. Electrophoresis, 6(9), 427-448.

Neuhoff, V., Arold, N., Taube, D., & Ehrhardt, W. (1988). Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. Electrophoresis, 9(6), 255-262.

Newton, G. L., Arnold, K., Price, M. S., Sherrill, C., Delcardayre, S. B., Aharonowitz, Y., et al. (1996). Distribution of thiols in microorganisms: mycothiol is a major thiol in most actinomycetes. J Bacteriol. 178(7), 1990-1995.

Nguyen, K. T., Tenor, J., Stettler, H., Nguyen, L. T., Nguyen, L. D., & Thompson, C. J. (2003). Colonial differentiation in *Streptomyces coelicolor* depends on translation of a specific codon within the *adpA* gene. . J Bacteriol. 185(24), 7291-7296.

Nicholson, N. H., Baggaley, K. H., Cassels, R., Davison, M., Elson, S. W., Fulston, M. et al. (1994). Evidence that the immediate biosynthetic precursor of clavulanic acid is its N-aldehyde analogue. J. Chem. Soc., Chem. Commun. (11), 1281-1282.

Nilewar, S. S., & Kathiravan, M. K. (2014). Mycothiol: a promising antitubercular target. Bioorg Chem. 52, 62-68.

Nilsson, J., & Nissen, P. (2005). Elongation factors on the ribosome. Curr Opin Struc Biol. 15(3), 349-354.

Niu, G., Chater, K. F., Tian, Y., Zhang, J., & Tan, H. (2016). Specialised metabolites regulating antibiotic biosynthesis in *Streptomyces* spp. FEMS Microbiol Rev. 40(4), 554-73.

Nodwell JR, McGovern K, Losick R.(1996). An oligopeptide permease responsible for the import of an extracellular signal governing aerial mycelium formation in *Streptomyces coelicolor*. Mol Microbiol. 22(5), 881-893.

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. J Biol Chem. 250(10), 4007-4021.

Ohnishi, Y., Kameyama, S., Onaka, H., & Horinouchi, S. (1999). The A-factor regulatory cascade leading to streptomycin biosynthesis in *Streptomyces griseus*: identification of a target gene of the A-factor receptor. Mol Microbiol. 34(1), 102-111.

Ohnishi, Y., Seo, J. W., & Horinouchi, S. (2002). Deprogrammed sporulation in *Streptomyces*. FEMS Microbiol Rev. 216(1), 1-7.

Ohnishi, Y., Yamazaki, H., Kato, J. Y., Tomono, A., & Horinouchi, S. (2005). AdpA, a central transcriptional regulator in the A-factor regulatory cascade that leads to morphological development and secondary metabolism in *Streptomyces griseus*. Biosci Biotech Bioch. 69(3), 431-439.

Okamoto, S., Lezhava, A., Hosaka, T., Okamoto-Hosoya, Y., & Ochi, K. (2003). Enhanced expression of S-adenosylmethionine synthetase causes overproduction of actinorhodin in *Streptomyces coelicolor* A3 (2). J Bacteriol. 185(2), 601-609.

O'Keefe, D. P., & Harder, P. A. (1991). Occurrence and biological function of cytochrome P450 monooxygenases in the actinomycetes. Mol Microbiol. 5(9), 2099-2105.

Old, W. M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K. G., Mendoza, A., Sevinsky, J. R., et al. (2005). Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics. 4(10), 1487-1502.

Oliva, B., O'Neill, A., Wilson, J. M., O'Hanlon, P. J., & Chopra, I. (2001). Antimicrobial properties and mode of action of the pyrrothine holomycin. Antimicrob Agents Ch. 45(2), 532-539.

Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M. et al. (1992). The α/β hydrolase fold. Protein Eng. 5(3), 197-211.

Ōmura, S. (1992). The expanded horizon for microbial metabolites—a review. Gene, 115(1), 141-149.

Onaka H. (2009). Biosynthesis of indolocarbazole and goadsporin, two different heterocyclic antibiotics produced by actinomycetes. Biosci Biotech Bioch. 73(10):2149-55.

Onaka, H., Nakaho, M., Hayashi, K., Igarashi, Y., & Furumai, T. (2005). Cloning and characterization of the goadsporin biosynthetic gene cluster from *Streptomyces* sp. TP-A0584. Microbiology, 151(12), 3923-3933.

Onaka, H., Tabata, H., Igarashi, Y., Sato, Y., & Furumai, T. (2001). Goadsporin, a chemical substance which promotes secondary metabolism and morphogenesis in streptomycetes. I. Purification and characterization. J Antibiot (Tokyo). 54(12), 1036-1044.

Oppermann, U., Filling, C., Hult, M., Shafqat, N., Wu, X., Lindh, M. et al. (2003). Short-chain dehydrogenases/reductases (SDR): the 2002 update. Chem-Biol Interact. 143-144, 247-253.

Ostash, B., & Walker, S. (2010). Moenomycin family antibiotics: chemical synthesis, biosynthesis, and biological activity. Nat Prod Rep. 27(11), 1594-1617.

Ostash, B., Campbell, J., Luzhetskyy, A., & Walker, S. (2013). MoeH5: a natural glycorandomizer from the moenomycin biosynthetic pathway. Mol Microbiol. *90*(6), 1324-1338.

Ostash, B., Doud, E. H., Lin, C., Ostash, I., Perlstein, D. L., Fuse, S. et al. (2009). Complete characterization of the seventeen step moenomycin biosynthetic pathway. Biochemistry, 48(37), 8830-8841.

Ostash, B., Saghatelian, A., & Walker, S. (2007). A streamlined metabolic pathway for the biosynthesis of moenomycin A. Chem Biol. 14(3), 257-267.

Otto, A., Becher, D., & Schmidt, F. (2014). Quantitative proteomics in the field of microbiology. Proteomics. 14(4-5), 547-565.

Öster, L. M., Lester, D. R., Van Scheltinga, A. T., Svenda, M., Van Lun, M., Genereux, C., & Andersson, I. (2006). Insights into cephamycin biosynthesis: the crystal structure of CmcI from *Streptomyces clavuligerus*. J Mol Biol. 358(2), 546-558.

Özcengiz, G., Okay, S., Ünsaldı, E., Taşkın, B., Liras, P., & Piret, J. (2010). Homologous expression of aspartokinase (*ask*) gene in *Streptomyces clavuligerus* and its *hom*-deleted mutant: effects on cephamycin C production. Bioeng Bugs. 1(3), 191-197.

Özcengiz, G., & Demain, A. L. (2013). Recent advances in the biosynthesis of penicillins, cephalosporins and clavams and its regulation. Biotechnol Adv. 31(2), 287-311.

Page, M. G. (2012). β-lactam antibiotics. In *Antibiotic Discovery and Development* (pp. 79-117). Springer US.

Paget, M. S., & Helmann, J. D. (2003). The σ^{70} family of sigma factors. Genome Biol. *4*(1), 1.

Paget, M. S., Hong, H. J., Bibb, M. J., & Buttner, M. J. (2002, January). The ECF sigma factors of *Streptomyces coelicolor* A3 (2). In *Symposıa-society for general microbiology* (pp. 105-126). Cambridge; Cambridge University Press; 1999.

Palaniappan, N., Ayers, S., Gupta, S., Habib, E. S., & Reynolds, K. A. (2006). Production of hygromycin A analogs in *Streptomyces hygroscopicus* NRRL 2388 through identification and manipulation of the biosynthetic gene cluster. *Chem Biol.* 13(7), 753-764.

Pallen, M., Chaudhuri, R., & Khan, A. (2002). Bacterial FHA domains: neglected players in the phospho-threonine signalling game?. *Trends Microbiol*. 10(12), 556-563.

Pan, Y., Liu, G., Yang, H., Tian, Y., & Tan, H. (2009). The pleiotropic regulator AdpA-L directly controls the pathway-specific activator of nikkomycin biosynthesis in *Streptomyces ansochromogenes*. Mol Microbiol. 72(3), 710-723.

Paradkar, A. S., Aidoo, K. A., & Jensen, S. E. (1998). A pathway-specific transcriptional activator regulates late steps of clavulanic acid biosynthesis in *Streptomyces clavuligerus*. Mol Microbiol. 27(4), 831-843

Paradkar, A., Trefzer, A., Chakraburtty, R., & Stassi, D. (2003). *Streptomyces* genetics: a genomic perspective. Crit Rev Biotechnol. 23(1), 1-27.

Park, S. J., & Lee, S. Y. (2003). Identification and characterization of a new enoyl coenzyme A hydratase involved in biosynthesis of medium-chain-length polyhydroxyalkanoates in recombinant *Escherichia coli*. J Bacteriol. 185(18), 5391-5397.

Pasca, M. R., Guglierame, P., De Rossi, E., Zara, F., & Riccardi, G. (2005). *mmpL7* gene of *Mycobacterium tuberculosis* is responsible for isoniazid efflux in *Mycobacterium smegmatis*. Antimicrob Agents Ch. 49(11), 4775-4777.

Patel, P., Song, L., & Challis, G. L. (2010). Distinct extracytoplasmic siderophore binding proteins recognize ferrioxamines and ferricoelichelin in *Streptomyces coelicolor* A3 (2). Biochemistry, 49(37), 8033-8042.

Patnaik, R. (2008). Engineering complex phenotypes in industrial strains. Biotechnol Progr. 24(1), 38-47.

Paulsen, I. T., Brown, M. H., & Skurray, R. A. (1996). Proton-dependent multidrug efflux systems. Microbiol Rev. 60(4), 575-608.

Pérez-Llarena, F. J., Rodríguez-García, A., Enguita, F. J., Martín, J. F., & Liras, P. (1998). The *pcd* gene encoding piperideine-6-carboxylate dehydrogenase involved in biosynthesis of α-aminoadipic acid is located in the cephamycin cluster of *Streptomyces clavuligerus*. J Bacteriol. 180(17), 4753-4756.

Pérez-Redondo, R., Rodríguez-García, A., Martín, J. F., & Liras, P. (1998). The *claR* gene of *Streptomyces clavuligerus*, encoding a LysR-type regulatory protein controlling clavulanic acid biosynthesis, is linked to the clavulanate-9-aldehyde reductase (*car*) gene. Gene, 211(2), 311-321.

Pérez-Redondo, R., Santamarta, I., Bovenberg, R., Martín, J. F., & Liras, P. (2010). The enigmatic lack of glucose utilization in *Streptomyces clavuligerus* is due to inefficient expression of the glucose permease gene. Microbiology, 156(5), 1527-1537.

Petříček, M., Petříčková, K., Havlíček, L., & Felsberg, J. (2006). Occurrence of two 5-aminolevulinate biosynthetic pathways in *Streptomyces nodosus* subsp. *asukaensis* is linked with the production of asukamycin. J Bacteriol. 188(14), 5113-5123.

Pettis, G. S., & Cohen, S. N. (1996). Plasmid transfer and expression of the transfer (*tra*) gene product of plasmid pIJ101 are temporally regulated during the *Streptomyces lividans* life cycle. Mol Microbiol. 19(5), 1127-1135.

Pfeifer, B. A. & Khosla, C. (2001). Biosynthesis of polyketides in heterologous hosts. Microbiol Mol Biol R. 65(1), 106-118.

Pinho, M. G., Kjos, M., & Veening, J. W. (2013). How to get (a) round: mechanisms controlling growth and division of coccoid bacteria. Nat Rev Microbiol. 11(9), 601-614.

Ponting, C. P., & Pallen, M. J. (1999). A β-propeller domain within TolB. Mol Microbiol. *31*(2), 739-740.

Porankiewicz, J., Wang, J., & Clarke, A. K. (1999). New insights into the ATP-dependent Clp protease: *Escherichia coli* and beyond. Mol Microbiol. 32(3), 449-458.

Pukatzki, S., Ma, A. T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W. C., Heidelberg J.F. & Mekalanos, J. J. (2006). Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. Proc Natl Acad Sci U S A. 103(5), 1528-1533.

Raghavan, V., & Groisman, E. A. (2010). Orphan and hybrid two-component system proteins in health and disease. Curr Opin Microbiol. 13(2), 226-231.

Raju, R. M., Jedrychowski, M. P., Wei, J. R., Pinkham, J. T., Park, A. S., O'Brien, K., et al. (2014). Post-translational regulation via Clp protease is critical for survival of *Mycobacterium tuberculosis*. PLoS Pathog. 10(3), e1003994.

Ramagli, L. S., & Rodriguez, L. V. (1985). Quantitation of microgram amounts of protein in two-dimensional polyacrylamide gel electrophoresis sample buffer. Electrophoresis, 6(11), 559-563.

Ramesh, S., Rajesh, M., & Mathivanan, N. (2009). Characterization of a thermostable alkaline protease produced by marine *Streptomyces fungicidicus* MML1614. Bioproc Biosyst Eng. 32(6), 791-800.

Ramirez, J. M., Del Campo, F. F., Paneque, A., & Losada, M. (1966). Ferredoxin-nitrite reductase from spinach. Biochim. Biophys. Acta. 118(1), 58-71.

Ramos, A., Honrubia, M. P., Vega, D., Ayala, J. A., Bouhss, A., Mengin-Lecreulx, D., & Gil, J. A. (2004). Characterization and chromosomal organization of the *murD–murC–ftsQ* region of *Corynebacterium glutamicum* ATCC 13869. Res Microbiol. 155(3), 174-184.

Ramos, J. L., Martínez-Bueno, M., Molina-Henares, A. J., Terán, W., Watanabe, K., Zhang, X., et al. (2005). The TetR family of transcriptional repressors. Microbiol Mol Biol R. ,69(2), 326-356.

Rappsilber, J., Ryder, U., Lamond, A. I., & Mann, M. (2002). Large-scale proteomic analysis of the human spliceosome. Genome Res. 12(8), 1231-1245.

Rasmussen, U. B., Mygind, B., & Per, N. (1986). Purification and some properties of uracil phosphoribosyltransferase from *Escherichia coli* K12. Biochim Biophys Acta. *881*(2), 268-275.

Ratnatilleke, A., Vrijbloed, J. W., & Robinson, J. A. (1999). Cloning and sequencing of the coenzyme B12-binding domain of isobutyryl-CoA mutase from *Streptomyces cinnamonensis*, reconstitution of mutase activity, and characterization of the recombinant enzyme produced in *Escherichia coli*. J Biol Chem. 274(44), 31679-31685.

Rawlings, N. D., & Barrett, A. J. (1993). Evolutionary families of peptidases. Biochem J. 290(1), 205-218.

Rawlings, N. D., & Barrett, A. J. (1997). Structure of membrane glutamate carboxypeptidase. Biochim Biophys Acta. 1339(2), 247-252.

Raymond, S., & Weintraub, L. (1959). Acrylamide gel as a supporting medium for zone electrophoresis. Science, 130(3377), 711-711.

Rees, D. C., Johnson, E., & Lewinson, O. (2009). ABC transporters: the power to change. Nat Rev Mol Cell Bio. 10(3), 218-227.

Reichmann, N. T., & Gründling, A. (2011). Location, synthesis and function of glycolipids and polyglycerolphosphate lipoteichoic acid in Gram-positive bacteria of the phylum *Firmicutes*. FEMS Microbiol Lett. 319(2), 97-105.

Reid, M. F., & Fewson, C. A. (1994). Molecular characterization of microbial alcohol dehydrogenases. Crit Rev Microbiol. 20(1), 13-56.

Rhee, S., Silva, M. M., Hyde, C. C., Rogers, P. H., Metzler, C. M., Metzler, D. E., & Arnone, A. (1997). Refinement and comparisons of the crystal structures of pig cytosolic aspartate aminotransferase and its complex with 2-methylaspartate. Crit Rev Microbiol. 272(28), 17293-17302.

Ricoux, R., Dubuc, R., Dupont, C., Marechal, J. D., Martin, A., Sellier, M., & Mahy, J. P. (2008). Hemozymes peroxidase activity of artificial hemoproteins constructed from the *Streptomyces lividans* xylanase A and iron (III)-carboxy-substituted porphyrins. Bioconjugate Chem. 19(4), 899-910.

Riesenberg, D., Bergter, F., & Kari, C. (1984). Effect of serine hydroxamate and methyl a-D-glucopyranoside treatment on nucleoside polyphosphate pools, RNA and protein accumulation in *Streptomyces hygroscopicus*. Microbiology, 130(10), 2549-2558.

Rigali, S., Titgemeyer, F., Barends, S., Mulder, S., Thomae, A. W., Hopwood, D. A., & van Wezel, G. P. (2008). Feast or famine: the global regulator DasR links nutrient stress to antibiotic production by *Streptomyces*. Embo Rep. 9(7), 670-675.

Rigden, D. J., Jedrzejas, M. J., & Galperin, M. Y. (2003). Amidase domains from bacterial and phage autolysins define a family of γ-D, L-glutamate-specific amidohydrolases. Trends Biochem Sci. 28(5), 230-234.

Rizzi, M., Nessi, C., Mattevi, A., Coda, A., Bolognesi, M., & Galizzi, A. (1996). Crystal structure of NH_3-dependent NAD^+ synthetase from *Bacillus subtilis*. EMBO J. *15*(19), 5125.

Rodríguez, A. M., Olano, C., Méndez, C., Hutchinson, C. R., & Salas, J. (1995). A cytochrome P450-like gene possibly involved in oleandomycin biosynthesis by *Streptomyces antibioticus*. FEMS Microbiol Lett. 127(1-2), 117-120.

Rodríguez, E., & Gramajo, H. (1999). Genetic and biochemical characterization of the α and β components of a propionyl-CoA carboxylase complex of *Streptomyces coelicolor* A3 (2). Microbiology, 145(11), 3109-3119.

Rodríguez, E., Banchio, C., Diacovich, L., Bibb, M. J., & Gramajo, H. (2001). Role of an essential acyl coenzyme A carboxylase in the primary and secondary metabolism of *Streptomyces coelicolor* A3 (2). Appl Environ Microb. 67(9), 4166-4176.

Rodríguez, H., Rico, S., Díaz, M., & Santamaría, R. I. (2013). Two-component systems in *Streptomyces*: key regulators of antibiotic complex pathways. Microb Cell Fact. 12(1), 1.

Rogowska-Wrzesinska, A., Le Bihan, M. C., Thaysen-Andersen, M., & Roepstorff, P. (2013). 2D gels still have a niche in proteomics. J Proteomics. 88, 4-13.

Román-Hernández, G., Hou, J. Y., Grant, R. A., Sauer, R. T., & Baker, T. A. (2011). The ClpS adaptor mediates staged delivery of N-end rule substrates to the AAA+ ClpAP protease. Mol Cell. 43(2), 217-228.

Romero-Rodríguez, A., Robledo-Casados, I., & Sánchez, S. (2015). An overview on transcriptional regulators in *Streptomyces*. Biochim Biophys Acta. 1849(8), 1017-1039.

Roth, J. R., Lawrence, J. G., & Bobik, T. A. (1996). Cobalamin (coenzyme B12): synthesis and biological significance. Annu Rev Microbiol. 50(1), 137-181.

Rouch, D. A., Byrne, M. E., Kong, Y. C., & Skurray, R. A. (1987). The *aacA-aphD* gentamicin and kanamycin resistance determinant of Tn4001 from *Staphylococcus aureus*: expression and nucleotide sequence analysis. Microbiology, 133(11), 3039-3052.

Saito, K., Kurosawa, M., Tatsuguchi, K., Takagi, Y., & Murakoshi, I. (1994). Modulation of cysteine biosynthesis in chloroplasts of transgenic tobacco overexpressing cysteine synthase [O-acetylserine (thiol)-Iyase]. Plant Physiol. 106(3), 887-895.

Salas, J. A., & Méndez, C. (2009). Indolocarbazole antitumour compounds by combinatorial biosynthesis. Curr Opin Chem Biol. 13(2), 152-160.

Salie, M. J., & Thelen, J. J. (2016). Regulation and structure of the heteromeric acetyl-CoA carboxylase. Biochim Biophys Acta. 1861(9 Pt B):1207-13.

Sandhu, P., & Akhter, Y. (2015). The internal gene duplication and interrupted coding sequences in the MmpL genes of *Mycobacterium tuberculosis*: Towards understanding the multidrug transport in an evolutionary perspective. Int J Med Microbiol. 305(3), 413-423.

Santamarta, I., López-García, M. T., Kurt, A., Nárdiz, N., Álvarez-Álvarez, R., Pérez-Redondo, R. et al. (2011). Characterization of DNA-binding sequences for CcaR in the cephamycin–clavulanic acid supercluster of *Streptomyces clavuligerus*. Mol Microbiol. 81(4), 968-981.

Santamarta, I., López-García, M. T., Pérez-Redondo, R., Koekman, B., Martín, J. F., & Liras, P. (2007). Connecting primary and secondary metabolism: AreB, an IclR-like protein, binds the ARE_{ccaR} sequence of *S. clavuligerus* and modulates leucine biosynthesis and cephamycin C and clavulanic acid production. Mol Microbiol. 66(2), 511-524.

Santamarta, I., Pérez-Redondo, R., Lorenzana, L. M., Martìn, J. F., & Liras, P. (2005). Different proteins bind to the butyrolactone receptor protein ARE sequence located upstream of the regulatory *ccaR* gene of *Streptomyces clavuligerus*. Mol Microbiol. 56(3), 824-835.

Santamarta, I., Rodríguez-García, A., Pérez-Redondo, R., Martín, J. F., & Liras, P. (2002). CcaR is an autoregulatory protein that binds to the *ccaR* and *cefD-cmcI* promoters of the cephamycin C-clavulanic acid cluster in *Streptomyces clavuligerus*. J Bacteriol. 184(11), 3106-3113.

Santos-Beneit, F., Barriuso-Iglesias, M., Fernández-Martínez, L. T., Martínez-Castro, M., Sola-Landa, A., Rodríguez-García, A., & Martín, J. F. (2011a). The RNA polymerase omega factor RpoZ is regulated by PhoP and has an important role in antibiotic biosynthesis and morphological differentiation in *Streptomyces coelicolor*. Appl Environ Microb. 77(21), 7586-7594.

Santos-Beneit, F., Rodríguez-García, A., & Martín, J. F. (2011b). Complex transcriptional control of the antibiotic regulator *afsS* in *Streptomyces*: PhoP and AfsR are overlapping, competitive activators. J Bacteriol. 193(9), 2242-2251.

Saudagar, P. S., Survase, S. A., & Singhal, R. S. (2008). Clavulanic acid: a review. Biotechnol Adv. 26(4), 335-351.

Schägger H. (1994). Chapter 4: Native electrophoresis. In: A practical guide to membrane protein purification (Von Jagow G, Schägger H, eds.) Academic Press, New York, pp. 81–103

Schägger, H., & von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. Anal Biochem. 199(2), 223-231.

Schatschneider, S., Vorhölter, F. J., Rückert, C., Becker, A., Eisenreich, W., Pühler, A., & Niehaus, K. (2011). Genome-enabled determination of amino acid biosynthesis in *Xanthomonas campestris* pv. campestris and identification of biosynthetic pathways for alanine, glycine, and isoleucine by 13C-isotopologue profiling. Mol Genet Genomics. 286(3-4), 247-259.

Schelin, J., Lindmark, F., & Clarke, A. K. (2002). The *clpP* multigene family for the ATP-dependent Clp protease in the cyanobacterium *Synechococcusc*. Microbiology. 148(7), 2255-2265.

Schirch, V. (1998). Mechanism of folate-requiring enzymes in one-carbon metabolism. In *Comprehensive Biological Catalysis* (Vol. 2, pp. 211-252). New York, NY, USA: Academic Press.

Schmitt, E., Panvert, M., Blanquet, S., & Mechulam, Y. (1998). Crystal structure of methionyl-tRNAfMet transformylase complexed with the initiator formyl-methionyl-tRNAfMet. EMBO J. 17(23), 6819-6826.

Schnoor, M., Voß, P., Cullen, P., Böking, T., Galla, H. J., Galinski, E. A., & Lorkowski, S. (2004). Characterization of the synthetic compatible solute homoectoine as a potent PCR enhancer. Biochem Bioph Res Co. 322(3), 867-872.

Schomburg, D. 1997. Enzyme Handbook: This collection of datasheets was generated from the Database" BRENDA". Class 2.5-EC 2.7. 1.104 Transferases. Springer.

Schulze, W. X., & Usadel, B. (2010). Quantitation in mass-spectrometry-based proteomics. Annu Rev Plant Biol. 61, 491-516.

Schwarz, S., Hood, R. D., & Mougous, J. D. (2010). What is type VI secretion doing in all those bugs? Trends Microbiol. 18(12), 531-537.

Schwede, T. F., Rétey, J., & Schulz, G. E. (1999). Crystal structure of histidine ammonia-lyase revealing a novel polypeptide modification as the catalytic electrophile. Biochemistry, 38(17), 5355-5361.

Sebulsky, M. T., Shilton, B. H., Speziali, C. D., & Heinrichs, D. E. (2003). The role of FhuD2 in iron (III)-hydroxamate transport in *Staphylococcus aureus* demonstration that FhuD2 binds iron (III)-hydroxamates but with minimal conformational change and implication of mutations on transport. J Biol Chem. 278(50), 49890-49900.

Seibert, C. M., & Raushel, F. M. (2005). Structural and catalytic diversity within the amidohydrolase superfamily. Biochemistry. 44(17), 6383-6391.

Serre, L., de Jesus, K. P., Zelwer, C., Bureaud, N., Schoentgen, F., & Bénédetti, H. (2001). Crystal structures of YBHB and YBCL from *Escherichia coli*, two bacterial homologues to a Raf kinase inhibitor protein. J Mol Biol. 310(3), 617-634.

Sevcikova, B., Rezuchova, B., Homerova, D., & Kormanec, J. (2010). The anti-anti-sigma factor BldG is involved in activation of the stress response sigma factor σH in *Streptomyces coelicolor* A3 (2). J Bacteriol. 192(21), 5674-5681.

Shanmugam, A., & Natarajan, J. (2012). Comparative modeling of UDP-N-acetylmuramoyl-glycyl-D-glutamate-2, 6-diaminopimelate ligase from Mycobacterium leprae and analysis of its binding features through molecular docking studies. J Mol Model. 18(1), 115-125.

Sharp, J. D., Cruz, J. W., Raman, S., Inouye, M., Husson, R. N., & Woychik, N. A. (2012). Growth and translation inhibition through sequence-specific RNA binding by *Mycobacterium tuberculosis* VapC toxin. J Biol Chem. 287(16), 12835-12847.

Shaw, J. P., Petsko, G. A., & Ringe, D. (1997). Determination of the structure of alanine racemase from *Bacillus stearothermophilus* at 1.9-Å resolution. Biochemistry, 36(6), 1329-1342.

Shen, B. (2003). Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. Curr Opin Chem Biol. 7(2), 285-295.

Shevchenko, A., Wilm, M., Vorm, O., & Mann, M. (1996). Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Anal Chem. 68(5), 850-858.

Shima, S., Warkentin, E., Grabarse, W., Sordel, M., Wicke, M., Thauer, R. K., & Ermler, U. (2000). Structure of coenzyme F420 dependent methylenetetrahydromethanopterin reductase from two methanogenic archaea. J Mol Biol. 300(4), 935-950.

Shin, S. K., Xu, D., Kwon, H. J., & Suh, J. W. (2006). S-adenosylmethionine activates *adpA* transcription and promotes streptomycin biosynthesis in *Streptomyces griseus*. FEMS Microbiol Lett. 259(1), 53-59.

Shu, D., Chen, L., Wang, W., Yu, Z., Ren, C., Zhang, W., et al. (2009). *afsQ1-Q2-sigQ* is a pleiotropic but conditionally required signal transduction system for both secondary metabolism and morphological development in *Streptomyces coelicolor*. Appl Microbiol Biot. 81(6), 1149-1160.

Sieber, S. A., & Marahiel, M. A. (2005). Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. Chem Rev. 105(2), 715-738.

Siedenburg, G., & Jendrossek, D. (2011). Squalene-hopene cyclases. Appl Environ Microb. 77(12), 3905-3915.

Singh, B., & Röhm, K. H. (2008). Characterization of a *Pseudomonas putida* ABC transporter (AatJMQP) required for acidic amino acid uptake: biochemical properties and regulation by the Aau two-component system. Microbiology, 154(3), 797-809.

Singh, D., Seo, M. J., Kwon, H. J., Rajkarnikar, A., Kim, K. R., Kim, S. O., & Suh, J. W. (2006). Genetic localization and heterologous expression of validamycin biosynthetic gene cluster isolated from *Streptomyces hygroscopicus* var. *limoneus* KCCM 11405 (IFO 12704). Gene, 376(1), 13-23.

Sinha, U., Wolz, S. A., & Lad, P. J. (1991). Two new extracellular serine proteases from *Streptomyces fradiae*. Int J Biochem Cell B. 23(10), 979-984.

Sisinni, L., Cendron, L., Favaro, G., & Zanotti, G. (2010). *Helicobacter pylori* acidic stress response factor HP1286 is a YceI homolog with new binding specificity. FEBS J. *277*(8), 1896-1905.

Skinner, D. D., Morgenstern, M. R., Fedechko, R. W., & Denoya, C. D. (1995). Cloning and sequencing of a cluster of genes encoding branched-chain alpha-keto acid dehydrogenase from *Streptomyces avermitilis* and the production of a functional E1 [alpha beta] component in *Escherichia coli*. J Bacteriol. 177(1), 183-190.

Smith, D. D. S., Wood, N. J., & Hodgson, D. A. (1995). Interaction between primary and secondary metabolism in *Streptomyces coelicolor* A3 (2): role of pyrroline-5-carboxylate dehydrogenase. Microbiology, 141(7), 1739-1744.

Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F., & Miller, N. E. (2001). Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. Nucleic Acids Res. 29(5), 1097-1106.

Sohlberg, B., Huang, J., & Cohen, S. N. (2003). The *Streptomyces coelicolor* polynucleotide phosphorylase homologue, and not the putative poly (A) polymerase, can polyadenylate RNA. J Bacteriol. 185(24), 7273-7278.

Song, J. Y., Jeong, H., Yu, D. S., Fischbach, M. A., Park, H. S., Kim, J. J. et al. (2010a). Draft genome sequence of *Streptomyces clavuligerus* NRRL 3585, a producer of diverse secondary metabolites. J Bacteriol. 192(23), 6317-6318.

Song, J. Y., Jensen, S. E., & Lee, K. J. (2010b). Clavulanic acid biosynthesis and genetic manipulation for its overproduction. Appl Microbiol Biot. 88(3), 659-669.

Sophos, N. A., & Vasiliou, V. (2003). Aldehyde dehydrogenase gene superfamily: the 2002 update. Chem Biol Interact. 143-144, 5-22.

Sørensen, M. A., Fricke, J., & Pedersen, S. (1998). Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in *Escherichia coli in vivo*. J Mol Biol. 280(4), 561-569.

Sperisen, P., Schmid, C. D., Bucher, P., & Zilian, O. (2005). Stealth proteins: in silico identification of a novel protein family rendering bacterial pathogens invisible to host immune defense. PLoS Comput Biol, 1(6), e63.

Spiridonov, N. A., & Wilson, D. B. (2001). Cloning and biochemical characterization of BglC, a β-glucosidase from the cellulolytic actinomycete *Thermobifida fusca*. Curr Microbiol. 42(4), 295-301.

Sprenger, G. A. (1995). Genetics of pentose-phosphate pathway enzymes of *Escherichia coli* K-12. Arch Microbiol. 164(5), 324-330.

Sprusansky, O., Zhou, L., Jordan, S., White, J., & Westpheling, J. (2003). Identification of three new genes involved in morphogenesis and antibiotic production in *Streptomyces coelicolor*. J Bacteriol. 185(20), 6147-6157.

Staunton, J., & Weissman, K. J. (2001). Polyketide biosynthesis: a millennium review. Nat Prod Rep. 18(4), 380-416.

Štefanková, P., Perečko, D., Barák, I., & Kollárová, M. (2006). The thioredoxin system from *Streptomyces coelicolor*. J Basic Microb. 46(1), 47-55.

Steffek, M., Newton, G. L., Av-Gay, Y., & Fahey, R. C. (2003). Characterization of *Mycobacterium tuberculosis* mycothiol S-conjugate amidase. Biochemistry, 42(41), 12067-12076.

Stock, D., Leslie, A. G., & Walker, J. E. (1999). Molecular architecture of the rotary motor in ATP synthase. Science, 286(5445), 1700-1705.

Stoller, G., Rücknagel, K. P., Nierhaus, K. H., Schmid, F. X., Fischer, G., & Rahfeld, J. U. (1995). A ribosome-associated peptidyl-prolyl cis/trans isomerase identified as the trigger factor. EMBO J. 14(20), 4939.

Stover, P. J. (2009). One-carbon metabolism–genome interactions in folate-associated pathologies. J Nutr. 139(12), 2402-2405.

Strieker, M., Kopp, F., Mahlert, C., Essen, L. O., & Marahiel, M. A. (2007). Mechanistic and structural basis of stereospecific Cβ-hydroxylation in calcium-dependent antibiotic, a daptomycin-type lipopeptide. ACS Chem Biol. 2(3), 187-196.

Strieker, M., Tanović, A., & Marahiel, M. A. (2010). Nonribosomal peptide synthetases: structures and dynamics. Curr Opin Struc Biol. *20*(2), 234-240.

Sugawara, K., Dohmae, N., Kasai, K., Saido-Sakanaka, H., Okamoto, S., Takio, K., & Ochi, K. (2002). Isolation and identification of novel ADP-ribosylated proteins from *Streptomyces coelicolor* A3 (2). Biosci Biotech Bioch. 66(10), 2292-2296.

Sugie M., and Suzuki H. (1978). Purification and properties of D-aminoacylase of *Streptomyces olivaceus*. Agric. Biol. Chem. 42:107-113.

Sun, Q. L., Wang, L. Y., Shan, J. J., Jiang, R., Guo, L. H., Zhang, Y., et al. (2007). Knockout of the gene (*ste15*) encoding a glycosyltransferase and its function in biosynthesis of exopolysaccharide in *Streptomyces* sp. 139. Arch Microbiol. 188(4), 333-340.

Sundaram, S., Karakaya, H., Scanlan, D. J., & Mann, N. H. (1998). Multiple oligomieric forms of glucoses-6-phosphate dehydrogenase in cyanobacteria and the role of OpeA in the assembly process. Microbiology, 144(6), 1549-1556.

Swiatek, M. A. (2012). Global control of development and antibiotic production by nutrient-responsive signalling pathways in *Streptomyces* (Doctoral dissertation, Department of Molecular Biotechnology, Leiden Institute of Chemistry (LIC), Faculty of Science, Leiden University).

Świątek, M. A., Urem, M., Tenconi, E., Rigali, S., & van Wezel, G. P. (2012). Engineering of N-acetylglucosamine metabolism for improved antibiotic production in *Streptomyces coelicolor* A3 (2) and an unsuspected role of NagA in glucosamine metabolism. Bioengineered. 3(5), 280-285.

Sy, J. (1977). In vitro degradation of guanosine 5'-diphosphate, 3'-diphosphate. Proc Natl Acad Sci U S A. 74(12), 5529-5533.

Taguchi, F., Ogawa, Y., Takeuchi, K., Suzuki, T., Toyoda, K., Shiraishi, T., & Ichinose, Y. (2006). A homologue of the 3-oxoacyl-(acyl carrier protein) synthase III gene located in the glycosylation island of *Pseudomonas syringae* pv. tabaci regulates virulence factors via N-acyl homoserine lactone and fatty acid synthesis. J Bacteriol. 188(24), 8376-8384.

Tahlan, K., Anders, C., & Jensen, S. E. (2004a). The paralogous pairs of genes involved in clavulanic acid and clavam metabolite biosynthesis are differently regulated in *Streptomyces clavuligerus*. J Bacteriol. 186(18), 6286-6297.

Tahlan, K., Park, H. U., & Jensen, S. E. (2004b). Three unlinked gene clusters are involved in clavam metabolite biosynthesis in *Streptomyces clavuligerus*. Can J Microbiol. 50(10), 803-810.

Tahlan, K., Anders, C., Wong, A., Mosher, R. H., Beatty, P. H., Brumlik, M. J. (2007). 5S clavam biosynthetic genes are located in both the clavam and paralog gene clusters in *Streptomyces clavuligerus*. Chem Biol. 14(2), 131-142.

Takahashi-Iñiguez, T., Cruz-Rabadán, S., Miguel Burciaga-Cifuentes, L., & Elena Flores, M. (2014). Molecular cloning, purification, and biochemical characterization of recombinant isocitrate dehydrogenase from *Streptomyces coelicolor* M-145. Biosci Biotech Bioch. 78(9), 1490-1494.

Takano, E. (2006). γ-Butyrolactones: *Streptomyces* signalling molecules regulating antibiotic production and differentiation. Curr Opin Microbiol. 9(3), 287-294.

Takano, E., Tao, M., Long, F., Bibb, M. J., Wang, L., Li, W. et al. (2003). A rare leucine codon in *adpA* is implicated in the morphological defect of *bldA* mutants of *Streptomyces coelicolor*. Mol Microbiol. 50(2), 475-486.

Takano, H., Hagiwara, K., & Ueda, K. (2015). Fundamental role of cobalamin biosynthesis in the developmental growth of *Streptomyces coelicolor* A3 (2). Appl Microbiol Biot. 99(5), 2329-2337.

Tanaka, T., Fukui, T., Fujiwara, S., Atomi, H., & Imanaka, T. (2004). Concerted action of diacetylchitobiose deacetylase and exo-β-D-glucosaminidase in a novel chitinolytic pathway in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. J Biol Chem. 279(29), 30021-30027.

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

Taylor, M., Scott, C., & Grogan, G. (2013). F420-dependent enzymes-potential for applications in biotechnology. Trends Biotechnol. 31, 63-4.

Tercero, J. A., Lacalle, R. A., & Jimenez, A. (1993). The *pur8* gene from the pur cluster of *Streptomyces alboniger* encodes a highly hydrophobic polypeptide which confers resistance to puromycin. Eur J Biochem. 218(3), 963-971.

Teufel, R., Mascaraque, V., Ismail, W., Voss, M., Perera, J., Eisenreich, W., et al. (2010). Bacterial phenylalanine and phenylacetate catabolic pathway revealed. Proc Natl Acad Sci U S A. 107(32), 14390-14395.

Thauer, R. K., & Kunow, J. (1995). Sulfate-reducing archaea. In *Sulfate-Reducing Bacteria* (pp. 33-48). Springer US.

Thoden, J. B., Frey, P. A., & Holden, H. M. (1996). Molecular structure of the NADH/UDP-glucose abortive complex of UDP-galactose 4-epimerase from

Escherichia coli: implications for the catalytic mechanism. Biochemistry, 35(16), 5137-5144.

Tiffany, K. A., Roberts, D. L., Wang, M., Paschke, R., Mohsen, A. W. A., Vockley, J., & Kim, J. J. P. (1997). Structure of human isovaleryl-CoA dehydrogenase at 2.6 Å resolution: Structural basis for substrate specificity. Biochemistry, 36(28), 8455-8464.

Tipper, D. J. & Strominger, J. L. (1965). Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. Proc Natl Acad Sci U S A. 54(4), 1133-1141.

Torre, F., Santis, L. D., Suárez, M. F., Crespillo, R., & Cánovas, F. M. (2006). Identification and functional analysis of a prokaryotic-type aspartate aminotransferase: implications for plant amino acid metabolism. Plant J. 46(3), 414-425.

Trepanier, N. K., Jensen, S. E., Alexander, D. C., & Leskiw, B. K. (2002). The positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus* is mistranslated in a *bldA* mutant. Microbiology. 148(3), 643-656.

Tripathi, L. P., & Sowdhamini, R. (2008). Genome-wide survey of prokaryotic serine proteases: analysis of distribution and domain architectures of five serine protease families in prokaryotes. *BMC genomics*, *9*(1), 1.

Trivedi, V., Gupta, A., Jala, V. R., Saravanan, P., Rao, G. J., Rao, N. A., et al. (2002). Crystal structure of binary and ternary complexes of serine hydroxymethyltransferase from *Bacillus stearothermophilus*: insights into the catalytic mechanism. J Biol Chem. 277(19), 17161-17169.

Tsai, Y. C., Tseng, C. P., Hsiao, K. M., & Chen, L. Y. (1988). Production and purification of D-aminoacylase from *Alcaligenes denitrificans* and taxonomic study of the strain. Appl Environ Microb. 54(4), 984-989.

Tu, J., Lu, F., Miao, S., Ni, X., Jiang, P., Yu, H., et al. (2014). The siderophore-interacting protein is involved in iron acquisition and virulence of *Riemerella anatipestifer* strain CH3. Vet Microbiol. 168(2), 395-402.

Tucker, A. C., & Escalante-Semerena, J. C. (2013). Acetoacetyl-CoA synthetase activity is controlled by a protein acetyltransferase with unique domain organization in *Streptomyces lividans*. Mol Microbiol. 87(1), 152-167.

Tunca S., Yılmaz E.I., Piret J., Liras P., Özcengiz G. (2004). Cloning, characterization and heterologous expression of the aspartokinase and aspartate semialdehyde dehydrogenase genes of cephamycin C-producer *Streptomyces clavuligerus*. Res Microbiol. 155(7):525-534

Ung, K. S. & Av-Gay, Y. (2006). Mycothiol-dependent mycobacterial response to oxidative stress. FEBS Lett. 580(11), 2712-2716.

Ünligil, U. M., & Rini, J. M. (2000). Glycosyltransferase structure and mechanism. Curr Opin Struc Biol. 10(5), 510-517.

Valiña, A. L. B., Mazumder-Shivakumar, D., & Bruice, T. C. (2004). Probing the Ser-Ser-Lys catalytic triad mechanism of peptide amidase: computational studies of the ground state, transition state, and intermediate. Biochemistry. 43(50), 15657-15672.

Vallée, B. S., Coadou, G., Labbe, H., Sy, D., Vovelle, F., & Schoentgen, F. (2003). Peptides corresponding to the N-and C-terminal parts of PEBP are well-structured in solution: new insights into their possible interaction with partners *in vivo*. J Pept Res. 61(2), 47-57.

van der Oost, J., Huynen, M. A., & Verhees, C. H. (2002). Molecular characterization of phosphoglycerate mutase in archaea. FEMS Microbial. Lett. *212*(1), 111-120.

van Wezel, G. P., and McDowall, K. J. (2011). The regulation of the secondary metabolism of *Streptomyces*: new links and experimental advances. Nat Prod Rep. 28(7), 1311-1333.

Vannini, A., Volpari, C., Gargioli, C., Muraglia, E., Cortese, R., De Francesco, R., et al. (2002). The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. EMBO J. 21(17), 4393-4401.

Vara, J., Perez-Gonzalez, J. A., & Jimenez, A. (1985). Biosynthesis of puromycin by *Streptomyces alboniger*: characterization of puromycin N-acetyltransferase. Biochemistry. 24(27), 8074-8081.

Vaupel, M., & Thauer, R. K. (1995). Coenzyme F420-dependent N5, N10-methylenetetrahydromethanopterin reductase (Mer) from *Methanobacterium thermoautotrophicum* strain Marburg. Eur J Biochem. 231(3), 773-778.

Vetting, M. W., de Carvalho, L. P. S., Yu, M., Hegde, S. S., Magnet, S., Roderick, S. L., & Blanchard, J. S. (2005). Structure and functions of the GNAT superfamily of acetyltransferases. Arch Biochem Biophys. 433(1), 212-226.

Vetting, M. W., Hegde, S. S., Javid-Majd, F., Blanchard, J. S., & Roderick, S. L. (2002). Aminoglycoside 2′-N-acetyltransferase from *Mycobacterium tuberculosis* in complex with coenzyme A and aminoglycoside substrates. Nat Struct Mol Biol. 9(9), 653-658.

Viars, S., Valentine, J., & Hernick, M. (2014). Structure and function of the LmbE-like superfamily. Biomolecules. 4(2), 527-545.

Vierling, S., Weber, T., Wohlleben, W., & Muth, G. (2001). Evidence that an additional mutation is required to tolerate insertional inactivation of the *Streptomyces lividans recA* gene. J Bacteriol. 183(14), 4374-4381.

Viollier, P. H., Nguyen, K. T., Minas, W., Folcher, M., Dale, G. E., & Thompson, C. J. (2001). Roles of aconitase in growth, metabolism, and morphological differentiation of *Streptomyces coelicolor*. J Bacteriol. 183(10), 3193-3203.

Volff, J. N., & Altenbuchner, J. (1998). Genetic instability of the *Streptomyces* chromosome. Mol Microbiol. 27(2), 239-246.

von Mering, C., Huynen, M., Jaeggi, D., Schmidt, S., Bork, P., & Snel, B. (2003). STRING: a database of predicted functional associations between proteins. Nucleic Acids Res. 31(1), 258-261.

Vuong, C., Kocianova, S., Voyich, J. M., Yao, Y., Fischer, E. R., DeLeo, F. R., & Otto, M. (2004). A crucial role for exopolysaccharide modification in bacterial biofilm formation, immune evasion, and virulence. J Biol Chem. 279(52), 54881-54886.

Wada, Y., Iwai, S., Tamura, Y., Ando, T., Shinoda, T., Arai, K., & Taguchi, H. (2008). A new family of D-2-hydroxyacid dehydrogenases that comprises D-mandelate dehydrogenases and 2-ketopantoate reductases. Biosci Biotech Bioch. 72(4), 1087-1094.

Walker, J. E. (1992). The NADH: ubiquinone oxidoreductase (complex I) of respiratory chains. Q Rev Biophys. 25(03), 253-324.

Walton, L. J., Corre, C., & Challis, G. L. (2006). Mechanisms for incorporation of glycerol-derived precursors into polyketide metabolites. J Ind Microbiol Biotechnol. 33(2), 105-120.

Wang, B., Weng, J., Fan, K., & Wang, W. (2011). Elastic network model-based normal mode analysis reveals the conformational couplings in the tripartite AcrAB-TolC multidrug efflux complex. Proteins. 79(10), 2936-2945.

Wang, L. Y., Li, S. T., & Li, Y. (2003). Identification and characterization of a new exopolysaccharide biosynthesis gene cluster from *Streptomyces*. FEMS Microbiol. Lett. *220*(1), 21-27.

Wang, L., Tahlan, K., Kaziuk, T. L., Alexander, D. C., & Jensen, S. E. (2004). Transcriptional and translational analysis of the *ccaR* gene from *Streptomyces clavuligerus*. Microbiology, 150(12), 4137-4145.

Wang, R., Mast, Y., Wang, J., Zhang, W., Zhao, G., Wohlleben, W., et al. (2013). Identification of two-component system AfsQ1/Q2 regulon and its cross-regulation with GlnR in *Streptomyces coelicolor*. Mol Microbiol. 87(1), 30-48.

Wang, X. J., Guo, S. L., Guo, W. Q., Xi, D., & Xiang, W. S. (2009). Role of *nsdA* in negative regulation of antibiotic production and morphological differentiation in *Streptomyces bingchengensis*. J Antibiot (Tokyo). 62(6), 309-313.

Wang, X., Li, X., & Li, Y. (2007). A modified Coomassie Brilliant Blue staining method at nanogram sensitivity compatible with proteomic analysis. Biotechnol Lett. 29(10), 1599-1603.

Wang, Y. D., Zhao, S., & Hill, C. W. (1998). Rhs elements comprise three subfamilies which diverged prior to acquisition by *Escherichia coli*. J Bacteriol. 180(16), 4102-4110.

Warner, J. R., & McIntosh, K. B. (2009). How common are extraribosomal functions of ribosomal proteins? Mol Cell. 34(1), 3-11.

Wei, J., Tian, Y., Niu, G., & Tan, H. (2014). GouR, a TetR family transcriptional regulator, coordinates the biosynthesis and export of gougerotin in *Streptomyces graminearus*. Appl Environ Microb. 80(2), 714-722.

Wei, Y., Zhang, H., Gao, Z. Q., Wang, W. J., Shtykova, E. V., Xu, J. H., et al. (2012). Crystal and solution structures of methyltransferase RsmH provide basis for methylation of C1402 in 16S rRNA. J Struct Biol. 179(1), 29-40.

Weinig, S., Hecht, H. J., Mahmud, T., & Müller, R. (2003). Melithiazol biosynthesis: further insights into myxobacterial PKS/NRPS systems and evidence for a new subclass of methyl transferases. Chem Biol. 10(10), 939-952.

Wendrich, T. M., & Marahiel, M. A. (1997). Cloning and characterization of a *relA/spoT* homologue from *Bacillus subtilis*. Mol Microbiol. 26(1), 65-79.

Westermeier, R. & Marouga, R. (2005). Protein detection methods in proteomics research. Bioscience Rep. 25(1-2), 19-32.

Westermeier, R. (2006). Electrophoresis in practice. John Wiley & Sons.

Westermeier, R., Naven, T., & Höpker, H. R. (2008). Proteomics in practice: a guide to successful experimental design. John Wiley & Sons.

Westermeier, R. (2016). 2D Gel-based proteomics: There's life in the old dog yet. Arch Physiol Biochem. (just-accepted), 1-6.

Whitfield, C. D., Steers, E. J., & Weissbach, H. (1970). Purification and properties of 5-methyltetrahydropteroyltriglutamate-homocysteine transmethylase. J Biol Chem. 245(2), 390-401.

Wilderman, P. J., Vasil, A. I., Johnson, Z., & Vasil, M. L. (2001). Genetic and biochemical analyses of a eukaryotic-like phospholipase D of *Pseudomonas aeruginosa* suggest horizontal acquisition and a role for persistence in a chronic pulmonary infection model. Mol Microbiol. 39(2), 291-304.

Wildermuth, H. (1970). Development and organization of the aerial mycelium in *Streptomyces coelicolor*. Microbiology, 60(1), 43-50.

Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D., Humphery-Smith, I., Hochstrasser, D. F., & Williams, K. L. (1996). Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. Biotechnol Genet Eng Rev**.** 13(1), 19-50.

Williams, C. H., Arscott, L. D., Müller, S., Lennon, B. W., Ludwig, M. L., Wang, P. F., et al. (2000). Thioredoxin reductase. Eur J Biochem. 267(20), 6110-6117.

Williams, G. J. (2013). Engineering polyketide synthases and nonribosomal peptide synthetases. Curr Opin Struc Biol. 23(4), 603-612.

Wilson, D. J., Patton, S., Florova, G., Hale, V., & Reynolds, K. A. (1998). The shikimic acid pathway and polyketide biosynthesis. J Ind Microbiol Biotechnol, 20(5), 299-303.

Wiśniewski J.R., Zougman A., Nagaraj N., Mann M. (2009). Universal sample preparation method for proteome analysis. Nat Methods. 6(5):359-62.

Wiśniewski, J. R. & Mann, M. (2012). Consecutive proteolytic digestion in an enzyme reactor increases depth of proteomic and phosphoproteomic analysis. Anal Chem. 84(6), 2631-2637.

Wong, L. L. (1998). Cytochrome P450 monooxygenases. Curr Opin Chem Biol.

2(2), 263-268.

Wong, P. & Houry, W. A. (2004). Chaperone networks in bacteria: analysis of protein homeostasis in minimal cells. J Struct Biol. 146(1), 79-89.

Wöhlbrand, L., Trautwein, K., & Rabus, R. (2013). Proteomic tools for environmental microbiology—a roadmap from sample preparation to protein identification and quantification. Proteomics. 13(18-19), 2700-2730.

Wright, G. D. & Thompson, P. R. (1999). Aminoglycoside phosphotransferases: proteins, structure, and mechanism. Front Biosci. 4, D9-D21.

Wu, J., Jiang, Z., Liu, M., Gong, X., Wu, S., Burns, C. M., & Li, Z. (2009). Polynucleotide phosphorylase protects *Escherichia coli* against oxidative stress. Biochemistry. 48(9), 2012-2020.

Wu, W., Leblanc, S. K. D., Piktel, J., Jensen, S. E., & Roy, K. L. (2006). Prediction and functional analysis of the replication origin of the linear plasmid pSCL2 in *Streptomyces clavuligerus*. Can J Microbiol. 52(4), 293-300.

Wu, X. & Roy, K. L. (1993). Complete nucleotide sequence of a linear plasmid from *Streptomyces clavuligerus* and characterization of its RNA transcripts. J Bacteriol. 175(1), 37-52.

Wysocki, V. H., Resing, K. A., Zhang, Q., & Cheng, G. (2005). Mass spectrometry of peptides and proteins. Methods. 35(3), 211-222.

Wyszynski, F. J., Lee, S. S., Yabe, T., Wang, H., Gomez-Escribano, J. P., Bibb, M. J., et al. (2012). Biosynthesis of the tunicamycin antibiotics proceeds via unique exo-glycal intermediates. Nat Chem. 4(7), 539-546.

Xia, G. & Peschel, A. (2008). Toward the pathway of *S. aureus* WTA biosynthesis. Chem Biol. 15(2), 95-96.

Xia, X., Wu, C., Cui, Y., Kang, M., Li, X., Ding, S., & Shen, J. (2015). Proteomic analysis of tylosin-resistant *Mycoplasma gallisepticum* reveals enzymatic activities associated with resistance. Sci Rep. 5:17077.

Xiao, M., Xu, P., Zhao, J., Wang, Z., Zuo, F., Zhang, J., et al. (2011). Oxidative stress-related responses of *Bifidobacterium longum* subsp. *longum* BBMN68 at the proteomic level after exposure to oxygen. Microbiology. 157(6), 1573-1588.

Yamashiro, K., Tsuruoka, N., Kodama, S., Tsujimoto, M., Yamamura, Y., Tanaka, T., et al. (1997). Molecular cloning of a novel trypsin-like serine protease (neurosin) preferentially expressed in brain. Biochim Biophys Acta. 1350(1), 11-14.

Yan, D., Kang, J., & Liu, D. Q. (2009). Genomic analysis of the aromatic catabolic pathways from *Silicibacter pomeroyi* DSS-3. Ann Microbiol. 59(4), 789-800.

Yang, S. Y., & Schulz, H. (1983). The large subunit of the fatty acid oxidation complex from *Escherichia coli* is a multifunctional polypeptide. Evidence for the existence of a fatty acid oxidation operon (*fadAB*) in *Escherichia coli*. J Biol Chem. 258(16), 9780-9785.

Yang, H., Wang, L., Xie, Z., Tian, Y., Liu, G., & Tan, H. (2007). The tyrosine degradation gene *hppD* is transcriptionally activated by HpdA and repressed by HpdR in *Streptomyces coelicolor*, while *hpdA* is negatively autoregulated and repressed by HpdR. Mol Microbiol. 65(4), 1064-1077.

Yano, T. & Ohnishi, T. (2001). The origin of cluster N2 of the energy-transducing NADH–quinone oxidoreductase: comparisons of phylogenetically related enzymes. J Bioenerg Biomembr. *33*(3), 213-222.

Yasuda, M., Misono, H., Soda, K., Yonaha, K., & Toyama, S. (1979). Purification and crystallization of L-ornithine: α-ketoglutarate δ-aminotransferase from *Bacillus sphaericus*. FEBS Lett. 105(2), 209-212.

Yates, J. R. (1998). Mass spectrometry and the age of the proteome. J Mass Spectrom. 33(1), 1-19.

Yates III, J. R. (2004). Mass spectral analysis in proteomics. Annu Rev Biophys. Biomol Struct. 33, 297-316.

Yates, J. R., Ruse, C. I., & Nakorchevsky, A. (2009). Proteomics by mass spectrometry: approaches, advances, and applications. Annu Rev Biomed Eng. 11, 49-79.

Yılmaz, E. I., Çaydasi, A. K., & Özcengiz, G. (2008). Targeted disruption of homoserine dehydrogenase gene and its effect on cephamycin C production in *Streptomyces clavuligerus*. J Ind Microbiol Biotechnol, 35(1), 1-7.

Yin, P., Wang, Y. H., Zhang, S. L., Chu, J., Zhuang, Y. P., Wang, M. L., & Zhou, J. (2008). Isolation of soluble proteins from an industrial strain *Streptomyces avermitilis* in complex culture medium for two-dimensional gel electrophoresis. J Microbiol Meth. 73(2), 105-110.

Yon, J. M., Desmadril, M., Betton, J. M., Minard, P., Ballery, N., Missiakas, D., et al. (1990). Flexibility and folding of phosphoglycerate kinase. *Biochimie*, *72*(6), 417-429.

Yoo, C., Ahn, K., Park, J. E., Kim, M. J., & Jo, S. A. (2010). An aminopeptidase from *Streptomyces* sp. KK565 degrades beta amyloid monomers, oligomers and fibrils. FEBS Lett. 584(19), 4157-4162.

Yoon, G. S., Ko, K. H., Kang, H. W., Suh, J. W., Kim, Y. S., & Ryu, Y. W. (2006). Characterization of S-adenosylmethionine synthetase from *Streptomyces avermitilis* NRRL8165 and its effect on antibiotic production. Enzyme Microb Tech. 39(3), 466-473.

Yoon, Y. J., Beck, B. J., Kim, B. S., Kang, H. Y., Reynolds, K. A., & Sherman, D. H. (2002). Generation of multiple bioactive macrolides by hybrid modular polyketide synthases in *Streptomyces venezuelae*. Chem Biol. 9(2), 203-214.

Youn, H. D., Yim, Y. I., Kim, K., Hah, Y. C., & Kang, S. O. (1995). Spectral characterization and chemical modification of catalase-peroxidase from *Streptomyces* sp. J Biol Chem. 270(23), 13740-13747.

Yu, Z., Reichheld, S. E., Savchenko, A., Parkinson, J., & Davidson, A. R. (2010). A comprehensive analysis of structural and sequence conservation in the TetR family transcriptional regulators. J Mol Biol. 400(4), 847-864.

Yu, Z., Zhu, H., Dang, F., Zhang, W., Qin, Z., Yang, S., et al. (2012). Differential regulation of antibiotic biosynthesis by DraR-K, a novel two-component system in *Streptomyces coelicolor*. Mol Microbiol. 85(3), 535-556.

Yum, D. Y., Chung, H. C., Bai, D. H., Oh, D. H., & Yu, J. H. (1994). Purification and characterization of alkaline serine protease from an alkalophilic *Streptomyces* sp. Biosci Biotech Bioch. 58(3), 470-474.

Zammit, C. M., Mangold, S., Rao Jonna, V., Mutch, L. A., Watling, H. R., Dopson, M., & Watkin, E. L. (2012). Bioleaching in brackish waters - effect of chloride ions on the acidophile population and proteomes of model species. Appl Microbiol Biot. 93(1), 319-329.

Zámocký, M., Hallberg, M., Ludwig, R., Divne, C., & Haltrich, D. (2004). Ancestral gene fusion in cellobiose dehydrogenases reflects a specific evolution of GMC oxidoreductases in fungi. Gene, 338(1), 1-14.

Zappa, S., Li, K., & Bauer, C. E. (2010). The tetrapyrrole biosynthetic pathway and its regulation in *Rhodobacter capsulatus*. In recent advances in phototrophic prokaryotes (pp. 229-250). Springer New York.

Zelyas, N. J., Cai, H., Kwong, T., & Jensen, S. E. (2008). Alanylclavam biosynthetic genes are clustered together with one group of clavulanic acid biosynthetic genes in *Streptomyces clavuligerus*. J Bacteriol. 190(24), 7957-7965.

Zhang, Y. M., Lu, Y. J., & Rock, C. O. (2004). The reductase steps of the type II fatty acid synthase as antimicrobial targets. Lipids, 39(11), 1055-1060.

Zhang, Y. M., & Rock, C. O. (2008). Thematic review series: glycerolipids. Acyltransferases in bacterial glycerophospholipid synthesis. J Lipid Res. 49(9), 1867-1874.

Zhang, B., Wang, B., Wang, P., Cao, Z., Huang, E., Hao, J., et al. (2009). Enzymatic characterization of a monomeric isocitrate dehydrogenase from *Streptomyces lividans* TK54. Biochimie. 91(11), 1405-1410.

Zhang J., Kasciukovic T., White M.F. (2012). The CRISPR associated protein Cas4 Is a 5′ to 3′ DNA exonuclease with an iron-sulfur cluster. PLoS One. 7,10, e47232.

Zhang, Y., Fonslow, B. R., Shan, B., Baek, M. C., & Yates III, J. R. (2013a). Protein analysis by shotgun/bottom-up proteomics. Chem Rev. 113(4), 2343-2394.

Zhang, B. B., Wang, P., Wang, A., Wang, W. C., Tang, W. G., & Zhu, G. P. (2013b). Expression and characterization of a novel isocitrate dehydrogenase from *Streptomyces diastaticus* No. 7 strain M1033. Mol Biol Rep. 40(2), 1615-1623.

Zhang, Z., Wu, S., Stenoien, D. L., & Paša-Tolic, L. (2014). High-throughput proteomics. Annu Rev Anal Chem. 7, 427-454.

Zhao, S., & Hill, C. W. (1995). Reshuffling of Rhs components to create a new element. J Bacteriol. 177(5), 1393-1398.

Zhou, M., & Veenstra, T. (2008). Mass spectrometry: m/z 1983-2008. Biotechniques, 44(5), 667-8.

Zhou, Z. S., Yang, S. N., Li, H., Zhu, C. C., Liu, Z. P., & Yang, Z. M. (2013). Molecular dissection of mercury-responsive transcriptome and sense/antisense genes in *Medicago truncatula*. J Hazard Mater. 252, 123-131.

Zhu, D., He, X., Zhou, X., & Deng, Z. (2005). Expression of the *melC* operon in several *Streptomyces strains* is positively regulated by AdpA, an AraC family transcriptional regulator involved in morphological development in *Streptomyces coelicolor*. J Bacteriol. 187(9), 3180-3187.

Zhu, J., & Winans, S. C. (1998). Activity of the quorum-sensing regulator TraR of *Agrobacterium tumefaciens* is inhibited by a truncated, dominant defective TraR-like protein. Mol Microbiol. 27(2), 289-297.

Zhu, J., Oger, P. M., Schrammeijer, B., Hooykaas, P. J., Farrand, S. K., & Winans, S. C. (2000). The bases of crown gall tumorigenesis. J Bacteriol. 182(14), 3885-3895.

Zhu, W., Smith, J. W., & Huang, C. M. (2009). Mass spectrometry-based label-free quantitative proteomics. Biomed Res Int. 2010.

Zimmermann, H. (1992). 5'-Nucleotidase: molecular structure and functional aspects. Biochem J. *285*(Pt 2), 345.

Zmudka, M. W., Thoden, J. B., & Holden, H. M. (2013). The structure of DesR from *Streptomyces venezuelae*, a β-glucosidase involved in macrolide activation. Protein Sci. 22(7), 883-892.

Zou, P. J., Borovok, I., de Orué Lucana, D. O., Müller, D., & Schrempf, H. (1999). The mycelium-associated *Streptomyces reticuli* catalase-peroxidase, its gene and regulation by FurS. Microbiology, 145(3), 549-559.

Zybailov, B., Mosley, A. L., Sardiu, M. E., Coleman, M. K., Florens, L., & Washburn, M. P. (2006). Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*. J Proteome Res. 5(9), 2339-2347.

APPENDIX A

A.1. Trypticase Soy Broth (TSB)

	g/L
Trypticase Soy Broth	30

Sterilized at 121 °C for 15 min

A.2. Starch-Asparagine (SA) Medium (Aharonowitz and Demain, 1979)

600 ml of dH_2O was boiled and the following components were added to the water by continuous stirring:

Components	g/L
Starch	10
MOPS	21
K_2HPO_4	4.4

After complete dissolution of the components, the solution was cooled to RT, pH is adjusted to 6.8 and volume is completed to 800 ml. It was sterilized at 121 °C for 15 min, after sterilization, the following sterile components were added to the medium:

Components	ml/L
L-Asparagine (10 g/L)	200
$MgSO_4$ .7 H_2O (0.6 g/ml)	2
Trace element solution*	1

*Trace element solution	g/L
$FeSO_4.7H_2O$	1
$MnCl_2.4H_2O$	1
$ZnSO_4.7H_2O$	1
$CaCl_2.3H_2O$	1.3

A.3. Protein Extraction Buffer

500 mM Tris-HCl

50 mM EDTA

700 mM Sucrose

100 mM KCl

Adjust pH to 8.0 (It can be stored at +4°C for a week.)

Freshly add;

2% β-mercaptoethanol

1 mM Phenylmethylsulfonyl fluoride (PMSF)

A.4. Precipitation Solution

0.1 M Ammonium acetate in cold acetone (It is stored at -20°C)

A.5. Rehydration Buffer

8 M urea

2% (w/v) CHAPS

50mM DTT

1% ampholytes

A.6. 5X Bradford Dye

100 mg of Coomassie Brilliant Blue G-250 was dissolved in 50 mL of 95% ethanol in brown bottle. Afterwards, 100 mL of 85% ortho-phosphoric acid was added and the volume was completed to 200 ml with dH2O. Solution was filtered through Whatman No.1 filter paper and stored at +4°C. Before use for Bradford assay, 5X Bradford dye was diluted to 1X with dH_2O

A.7. Equilibration Buffers:

A.7.1. Solution I

50 mM Tris-Cl, pH 6.8

6 M urea

30% (v/v) glycerol

1% SDS

2% DTT (freshly added)

A.7.2. Solution II

50 mM Tris-Cl, pH 6.8

6 M urea

30% (v/v) glycerol

1% SDS

2.5% iodoacetamide (IAA) (freshly added)

3.5 μM bromophenol blue (freshly added)

A.8. Separating and Stacking Gel

	Separating Gel (0.375 M Tris, pH 8.8)	**Stacking Gel (0.125 M Tris, pH 6.8)**
Monomer Concentration (% T, 2.67 % C)	*12 %*	*4 %*
Acrylamide/bis (30 % T, 2.67 % C stock)	40 ml	1.3 ml
Distilled water	33.5 ml	6.1 ml
1.5 M Tris-HCl, pH 8.8	25.0 ml	2.5 ml
0.5 M Tris-HCl, pH 6.8	—	100 µl
10 % (w/v) SDS	1.0 ml	50 µl
10% Ammonium persulfate (stored at 4°C)	500 µl	10 µl
TEMED	50 µl	10 ml
TOTAL MONOMER	100 ml	

A.9. FASP Digestion Kit Solutions:

A.9.1. Urea Sample Solution (Prepared as explained in the kit manual)

It should be prepared fresh before digestion. 1 mL of Tris-hydrochloride solution provided with the FASP Kit was added to one tube of Urea, also provided with the FASP Kit. The mixture was vortexed until all the powder was dissolved.

A.9.2. 10X Iodoacetamide Solution (It should be prepared fresh before digestion).

10X Iodoacetamide Solution was prepared by adding 100 µL Urea Sample Solution to one tube of Iodoacetamide provided with the FASP Kit. The mixture was mixed and dissolved by pipetting for 15 times and then transferred to a clean, dry microcentrifuge tube.

A.9.2.1. 1X Iodoacetamide Solution (Prepared as explained in the kit manual

1X Iodoacetamide Solution was prepared by mixing 10 µl of 10X Iodoacetamide Solution with 90 µl of Urea Sample Solution

A.9.3. Digestion Solution I

Final concentration of Lys-C stock solution was adjusted to 1µg of LysC enzyme / 65 µl Ammonium Bicarbonate Solution provided by FASP kit for each sample.

A.9.4. Digestion Solution II

100X Trypsin Solution (0.2 µg/µl) was diluted to a final concentration so that each sample contained ~ 0.6-0.7 µg of Trypsin enzyme. Each 15 µl of Digestion Solution II added to the Spin Filter contained 3.2 µl of 100X Trypsin Solution and 11.8 µl of Ammonium Bicarbonate Solution provided by FASP kit.

A.9.5. ZipTip Purification Solutions

All the solvents are MS-grade

A.9.5.1. Wetting Solution (70% ACN)

Acetonitrile (ACN)	700 µl
dH_2O	300 µl

A.9.5.2. Equilibration Solution (3% ACN + 0.1% TFA)

Acetonitrile	30 µl
dH_2O	969 µl
Trifluoroacetic acid (TFA)	1 µl

A.9.5.3. Washing Solution (5% MeOH + 0.1% TFA)

Methanol (MeOH)	50 µl
dH_2O	949 µl
Trifluoroacetic acid (TFA)	1 µl

A.9.5.4. Elution Solution (60% ACN + 0.1% TFA)

Acetonitrile	60 µl
dH_2O	939 µl
Trifluoroacetic acid (TFA)	1 µl

APPENDIX B

CHEMICALS AND THEIR SUPPLIES

Chemicals	Supplier
Acetic acid	Merck
Acetone	Merck
Acetonitrile (ACN)	Merck
Acrylamide	Sigma
Ammonium acetate	Sigma
Ammonium persulfate	Applichem
Ampholytes pH (4-10)	Fluka
L-Asparagine	Sigma
Bromophenol blue	Merck
Bovine Serum Albumin (BSA)	Sigma
$CaCl_2.3H_2O$	Sigma
CHAPS	Sigma
Comassie Brillant Blue G250	Sigma
DTT	Sigma
EDTA	Sigma
Ethanol	Merck
$FeSO_4 .7H_2O$	Merck
Glycerol	Merck
HCI	Merck
Immersion Oil	Sigma
Iodoacetamide (IAA)	Sigma
IPG strips	BioRad
K_2HPO_4	Merck

KCl	Merck
β-mercaptoethanol	Merck
Methanol	Merck
N,*N*′-Methylenebis(acrylamide)	Sigma
$MgSO_4.7H_2O$	Merck
$MnCl_2 .4H_2O$	Merck
MOPS	Sigma
Ortho-phosphoric acid	Merck
Phenylmethylsulfonyl fluoride (PMSF)	Merck
SDS	Sigma
Starch	Merck
Sucrose	Merck
TEMED	Merck
Trypticase Soy Broth (TSB)	Oxoid
Tris-buffered phenol (pH 8.0)	Sigma
Tris-HCl	Sigma
Trifluoroacetic acid (TFA)	Merck
Urea	Merck
$ZnSO_4.7H_2O$	Merck

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Ünsaldı Eser

Nationality: Turkish (TC)

Date and Place of Birth: 2 November 1984, Elazığ

Marital Status: Single

Phone: +90 312 210 51 90

Fax: +90 312 210 79 76

email: eunsaldi@metu.edu.tr

EDUCATION

Degree	Institution	Year of Graduation
MS	METU Biology	2010
BS	METU Molecular Biology and Genetics	2007
High School	Balakgazi High School, Elazığ	2002

WORK EXPERIENCE

Year	Place	Enrollment
2007-2016	METU Department of Biology	Research Assistant

FOREIGN LANGUAGES

English (Advanced), Japanese (Beginner)

PUBLICATIONS

Papers

Özcengiz G., Okay S., Ünsaldı E., Taşkın B., Liras P. and Piret J. (2010). Homologous expression of aspartokinase (*ask*) gene in *Streptomyces clavuligerus* and its *hom*-deleted mutant: Effects on cephamycin C production. Bioeng Bugs. 1(3):191-197.

Posters

Ünsaldı, E., Okay, S. and Özcengiz, G. (2008). Cephamycin C overproduction upon releasing relaxing precursor flux by genetic engineering and medium formulation. XII. International Congress of Bacteriology and Applied Microbiology (5-9 August 2008, İstanbul, Turkey) Abstract Book p. 98.

Sezer Okay, Eser Ünsaldı, Bilgin Taşkın,Aslıhan Kurt, Jacqueline Piret, Paloma Liras, Gülay Özcengiz. Metabolic engineering of aspartate pathway for increased production of cephamycin C in *Streptomyces clavuligerus*. BIOTECH METU 2009, International Symposium on Biotechnology: Developments and Trends (September, 27-30, 2009 Ankara, Turkey).

Oral Presentations

Ünsaldı E., Özcengiz G. Differential proteomics of wild type and engineered strains of *S. clavuligerus,* ISBA'17, XII. International Symposium on the Biology of *Actinomycetes* (October, 8-12 2014, Kuşadası-Aydın, Turkey) Abstract Book p. 109

Workshops

Proteomics Bioinformatics Workshop during the X Annual Congress of the European Proteomics Association (22-25 June 2016, İstanbul, Turkey)

AWARDS/GRANTS

2211-C Doctorate Scholorship Programme of The Scientific and Technological Research Council of Turkey (TUBITAK).

2224-B Support for Participation to National Scientific Events by The Scientific and Technological Research Council of Turkey (TUBITAK).