Hes·so VALAIS WALLIS
Haute Ecole Spécialisée de Suisse occidentale
Fachhochschule Westschweiz
University of Applied Sciences Western Switzerland

Università degli Studi di Palermo
Dipartimento di Biologia Cellulare e dello Sviluppo "A. Monroy"

Degree Course Life Technologies

Option Biotechnology

Diploma 2007

Mirko Ledda

Strategies for the study of the *trpX* role in *Streptomyces coelicolor*

Professor Sergio Schmid

Expert Anna—Maria Puglia

Palermo, 04.03.2008

HES-SO Valais

SI	TV	EE	IG	EST
X	X	X	X	

Données du travail de diplôme
Daten der Diplomarbeit

FO.0.2.02.07.DB
DD/18/05/2006

Filière / *Studiengang* : Technologies du vivant

Confidentiel / *Vertraulich* ☐

Etudiant / *Student*	**Année scolaire / *Schuljahr***	**No TD / *Nr. DA***
Mirko Ledda	2006/07	TV/2007/68

Proposé par / *vorgeschlagen von*
Prof Anna-Maria Puglia
Università degli Studi, Palermo
ampuglia@unipa.it

Lieu d'exécution / *Ausführungsort*
Università degli Studi, Palermo

Expert / *Experte*
Anna-Maria Puglia

Titre / *Titel*:

Amino acid metabolism and its impact on antibiotic production

Description / *Beschreibung*:

Amino acids, either natural or non-proteinogenic, are frequently precursors of many antibiotics.

Although in *Streptomyces* extensive researches have been focused on the molecular genetics of antibiotic biosynthesis, little is known about the regulation of primary metabolic pathways that provide amino acids. Better knowledge of amino acid metabolism and its regulation can throw light on the interconnection with regulons affecting antibiotic production and could be used to enhance the yield of specific antibiotics.

The aim of our work is to analyse amino acid biosynthetic pathways and their interplay with regulons affecting antibiotic synthesis. In particular we intend to identify *S. coelicolor* key genes involved in amino acid biosynthesis. To this goal the collaboration with partners 10 and 11 will be crucial. When key genes will be identified, our group will create up-regulated mutants using inducible promoters and, in collaboration with partners 2 and 8, knock-out mutant strains. In the meantime we will create mutants for genes known as global regulators of amino acid metabolism, such as argR.

The strains of this "mutant collection" will be tested for the effects on *S. coelicolor* antibiotic production.
Gene expression studies will be carried out by transcriptional analysis and 2D-PAGE followed by computer analysis and characterisation of the proteins.

The suitable mutant strains will be provided for Activity 3 partners.

Objectifs / *Ziele:*

- identify *Streptomyces coelicolor* key genes involved in aromatic amino acid biosynthesis.
- create up-regulated and knock-out mutant strains.
- create mutants for genes known as global regulators of amino acid metabolism.
- test antibiotic production of mutant strains.
- characterize interesting mutant strains by transcriptional (qRT-PCR) and proteomic analysis (2D-PAGE, followed by computer analysis and identification of proteins by MALDI TOF).

Signature ou visa / *Unterschrift oder Visum*	**Délais / *Termine***
Leiter Vertiefungsrichtung Biotechnologie	Attribution du thème / Ausgabe des Auftrags: 03.09.2007
Professeur/*Dozent*: Sergio Schmid	Remise du rapport / Abgabe des Schlussberichts:
Etudiant/*Student*:	Exposition publique / Ausstellung Diplomarbeiten: -
	Défenses orales / Mündliche Verfechtungen -

Rapport reçu le / *Schlussbericht erhalten am* ..... Visa du secrétariat / *Visum des Sekretariats*: .....

Ledda Mirko 2007

Strategies for the study of *trpX* role in *Streptomyces coelicolor*

Specific objective

The *trpX* gene is a little *Streptomyces coelicolor* orf, coding for a hypothetical 63-amino acids protein, TrpX , having, to date, unknown function. This gene is present in the *trpA/B/X/C1* operon, coding for important enzymes involved in the tryptphan biosynthesis pathway.
The specific objective of this work is to start the study of the *trpX* gene role in *Streptomyces coelicolor*. In particular to:

1) study the *trpX* expression in the strain *S. coelicolor* A3(2) M145, in various growth phases and media by qRT-PCR.

2) perform in silico analysis of protein structure using the Robetta server.

3) carry out EMSA of intergenic regions of *trpA/B/X/C1* operon with crude extract of wt and mutant strains.

4) overexpress and purify TrpX protein from *E. coli*.

5) overexpress the *trpX* gene in *S. coelicolor*

Results

Real-Time RT-PCR analysis showed that *trpX* is not costitutively expressed during the growth in minimal medium, but is growth-phase dependent.
In silico approaches and previous experiments have shown that TrpX protein is a putative regulator of the *trpB* gene, a gene involved in the tryptophan biosynthesis and a tridimentional structure prediction showed that TrpX protein could have a DNA-binding protein structure. The most probable hypothesis of the TrpX protein function in *S. coelicolor*, is that it binds to intergenic regions of the *trpA/B/X/C1* operon, regulating the *trpB* gene expression. This hypothesis has not been confirmed by DNA binding assays carried out in this work, thus, further DNA or RNA binding assays are necessary. The His_6-TrpX protein was successfully overproduced in *E. coli*, and its purification at present is in progress, it will be used in DNA or RNA binding assays, in order to confirm the previous hypothesis.
The *trpX* gene overexpression in *S. coelicolor*, indicating ways to understand *trpX* function has not been obtained, but new strategies at present are in progess.

Key words

ActinoGEN project*, Streptomyces coelicolor*, antibiotic, tryptophan, pathway, cluster, operon, regulation, *trpX*, *SCO2038*, electrophoretic mobility shift assay, qRT-PCR, cloning

Ledda Mirko 2007

Stratégies pour l'étude du rôle du gène *trpX* dans la souche *Streptomyces coelicolor*

Objectifs spécifiques

Le gène trpX est une petite orf de la souche *Streptomyces coelicolor*, qui code pour une protéine hypothétique de 63 acides aminés, TrpX, ayant à ce jour, aucune fonction connue. Ce gène est présent dans l'opéron *trpA/B/X/C1*, codant pour d'importantes protéines impliquées dans le pathway de la biosynthèse du tryptophane, un acide aminé aromatique.
L'objectif spécifique de ce travail est l'étude du gène *trpX* dans la souche *S. coelicolor*. En particulier :

1) l'étude de l'expression du gène *trpX* dans la souche *S. coelicolor*, dans différentes phase de croissance et milieux de culture, par qRT-PCR.

2) l'analyse de la structure de la protéine TrpX, in silico, avec le programme informatique Robetta server.

3) effectuer des EMSAs sur les régions intergéniques de l'opéron *trpA/B/X/C1*, avec des extraits protéiques brutes des souche de *S. coelicolor* wt et de mutants *trpX*.

4) surexprimer et purifier la protéine TrpX produite dans une souche *E. coli*.

5) surexprimer le gène *trpX* dans la souche *S. coelicolor*

Résultats

Une analyse par Real-Time RT-PCR a montré que le gène *trpX* n'est pas exprimer constitutivement mais est dépendant de la phase de croissance.
Une approche in silico et des analyses précédentes à ce travail, ont montrés que la protéine TrpX est un possible régulateur du gène *trpB*, un gène impliqué dans la biosynthèse du tryptophane et une prédiction de la stucture tridimensionnelle de TrpX a montré qu'il s'agit d'une protéine liant l'ADN. L'hypothèse la plus probable du rôle de la protéine TrpX dans *S. coelicolor*, est qu'elle lie une région intergénique de l'opéron *trpA/B/X/C1*, régulant l'expression génique du gène *trpB*. Cette hypothèse n'a pas été confirmée par les analyses effectuées durant ce travail (DNA binding assays), donc d'autres essais sont nécessaires, notamment des DNA or RNA binding assays avec des conditions différentes. La protéine His_6-TrpX a été surproduite, avec succès, dans *E. coli*, et sa purification est, à l'heure actuelle, en cours. Cette protéine purifiée pourra être utilisée dans des
DNA or RNA binding assays, afin de confirmer l'hypothèse émise précédemment.
La surexpression du gène *trpX* dans *S. coelicolor*, permettant d'indiquer des voies, afin de comprendre la fonction de trpX n'as pas été obtenues, mais de nouvelles stratégies ont été étudiées et sont désormais en cours.

Mot-clés

Projet ActinoGEN*, Streptomyces coelicolor*, antibiotique, tryptophane, pathway, cluster, opéron, régulation, *trpX*, *SCO2038*, electrophoretic mobility shift assay, qRT-PCR, clonage

Hes·so VALAIS WALLIS
Haute Ecole Spécialisée de Suisse occidentale
Fachhochschule Westschweiz
University of Applied Sciences Western Switzerland

Università degli Studi di Palermo
Dipartimento di Biologia Cellulare e dello Sviluppo "A. Monroy"

Degree Course Life Technologies

Option Biotechnology

Diploma 2007

Strategies for the study of *trpX* role in *Streptomyces coelicolor*

03 September 2007 – 27 February 2008

Professor : **Schmid Sergio**
HES-SO Valais/Wallis
Sion, Switzerland

Expert : **Puglia Anna-Maria**
Università degli studi di Palermo
Palermo, Italy

Key words: ActinoGEN project, *Streptomyces coelicolor*, antibiotic, tryptophan, pathway, cluster, operon, regulation, *trpX*, *SCO2038*, electrophoretic mobility shift assay, qRT-PCR, cloning

Abstract

The emergence of several pathogenic bacteria and the adaptation of part of those resistant to antibiotics currently used, in clinically treatments, needs the research for new antibiotic development and the creation of a "Superhost", allowing a large scale production of antibiotics. *Streptomyces* belongs to Actinomycetes , a group of gram-positive bacteria, producing more than 50% of known antibiotics. To increase antibiotic level production in these organisms, the pathway of antibiotic precursors, like tryptophan, must be studied.
The *trpX* gene is a little *Streptomyces coelicolor* orf, coding for a hypothetical 63-amino acids protein, TrpX , having, to date, unknown function. This gene is present in the *trpA/B/X/C1* operon, coding for important enzymes involved in the tryptphan biosynthesis pathway, a aromatic amino acid.
Real-Time RT-PCR analysis showed that *trpX* is not costitutively expressed during the growth in minimal medium, but is growth-phase dependent.
In silico approaches and previous experiments have shown that TrpX protein is a putative regulator of the *trpB* gene, a gene involved in the tryptophan biosynthesis and a tridimentional structure prediction showed that TrpX protein could have a DNA-binding protein structure. The most probable hypothesis of the TrpX protein function in *S. coelicolor*, is that it binds to intergenic regions of the *trpA/B/X/C1* operon, regulating the *trpB* gene expression. This hypothesis has not been confirmed by DNA binding assays carried out in this work, thus, further DNA or RNA binding assays are necessary. The His_6-TrpX protein was successfully overproduced in *E. coli*, and its purification at present is in progress, it will be used in DNA or RNA binding assays, in order to confirm the previous hypothesis.
The *trpX* gene overexpression in *S. coelicolor*, indicating ways to understand *trpX* function, has not been obtained, but new strategies at present are in progess.

Contents

1 INTRODUCTION

1.1 The ActinoGEN project

Title : Integrating genomics-based applications to exploit actinomycetes as a resource for new antibiotics
Acronym : ActinoGEN
Coordinator : Dr. Paul Dyson, University of Wales Swansea, UK

ActinoGEN is an European integrated project aimed at developing novel genomics-based approaches to exploit hitherto overlooked genetic resources for new antibiotics. Drug discovery will focus on accessing new antibiotic biosynthetic pathways from diverse actinomycetes that have yet cultured; activating cryptic pathways from well-characterized actinomycetes; and engineering novel hybrid antibiotics by combinatorial biosynthesis.
To accelerate drug discovery, another parallel strategy could be to engineer generic hosts optimized to produce high antibiotic yields, these hosts are called “Superhosts”. Indeed, a rate-limiting step to developing new antibiotics is yield improvement. Post-genomic analyses permit for the first time, a concerted and holistic approach to engineering generic Superhosts for use in the production of high yields of a wide variety of antibiotics.
This synergy of discovery linked to overproduction will place the European biotechnology sector at the forefront of developing much-needed new antibiotics for fighting multi-drug resistant pathogens.
ActinoGEN is a consortium made up of groups of 14 universities and 4 SMEs from 9 different countries, UK, Czech Republic, Danemark, Germany, Italy, Netherlands, France, Switzerland and Spain.
This work is in the fields of research aiming at creating a Superhosts, for high yields production of antibiotics.

1.2 The Streptomyces genus

1.2.1 General information on *Streptomyces* genus

Streptomyces is the largest genus of Actinobacteria, a group of gram-positive bacteria. Streptomycetes are organisms widely spread in nature and in particular in the soil and the decaying vegetation where they act in the natural decomposition process and thus in the carbon cycle.

The *Streptomyces* genus is characterized by a genome with high G+C content (69-78%), a large linear chromosome (approximately 8.7 Mb) and a complex life cycle. In addition to a morphological differentiation, *Streptomyces* undergoes a physiological differentiation on the basis of which, in given conditions of growth, they produce a vast range of secondary metabolites. These have many applications in the medical field, due to their employment like antibiotics, immunosuppressors and antitumoral agents, and in the agricultural field as weedkillers and pesticides. It could be estimated than more of 50% of 11,900 known antibiotics, are produced by *Streptomyces* (Kieser T, 2000). Each year, about 500 new antibiotics are identified and a lot of these come from *Streptomyces* genus. For this reason, streptomycetes have a great importance for pharmaceutical industry. (Hodgson DA, 2000)

1.2.2 Natural environment and life cycle of *Streptomyces* genus

The soil, *Streptomyces* natural environment, is rich in carbohydrates and poor in phosphate and nitrogen, in fact, the greater source of nutriment derives from dead plants. Moreover due to the presence of many competitor microorganisms, the ground is an oligotrophic environment. Streptomycetes are saprophytes and produced extracellular enzymes, used in wall digestion of plant cells and insoluble polymers. They behave like facultative oligotrophic, because they have possibility to growth in a poor (oligotrophic) or a rich (copiotrophic) nourishing medium. (Hodgson DA, 2000)

The growth of *Streptomyces* colony begins with spore germination in answer to easily assimilable soluble nutrients, which are used to form the vegetative mycelium. When these nutrients start to lack, the colony begins to produce lytic enzymes (e.g. cellulases and xylanases) for insoluble polymer degradation, generally dead plant tissues. During this same phase, the antibiotic production starts because it becomes crucial for the colony to protect, from other competitor microorganisms, nutrients which were previously solubilised (fig. 1).

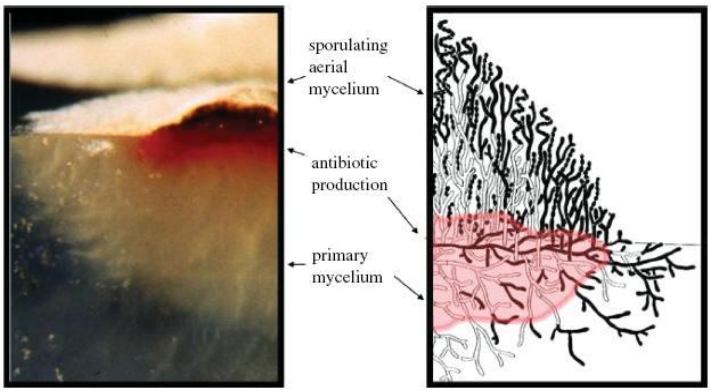


Figure 1 Vertical section of a colony of *Streptomyces* (Chater KF, 2006).

Moreover, in response to lacking nutrients, the colony starts producing an aerial mycelium. (fig. 2) (Chater KF, 2006).

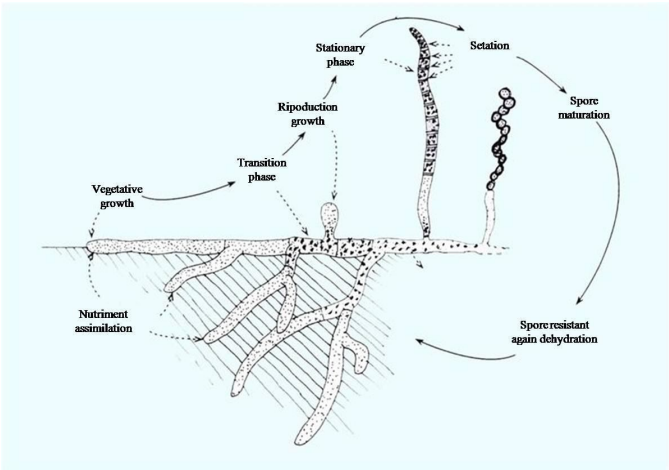


Figure 2 Life cycle of a *Streptomyces* colony (Chater KC, 1998).

The aerial mycelium septate to form spores, which allow the colonization of neighboring zones by the bacterium. This colonization is carried out by spore transport performed by the wind, by arthropods, or by organisms which nourish from microbes present on the decaying plants.
The growth of a *Streptomyces* colony is diauxic. On solid medium, in the first rapide growth phase (RG1), the colony produces the vegetative mycelium and in the second rapid growth phase (RG2), the aerial mycelium. In the transition phase (T), the colony growth is slowed down and the secondary metabolite production initiates and continues in the RG2 and stationary (S) phases (fig. 3). In liquid medium *Streptomyces* does not differentiate in aerial mycelium and spores.

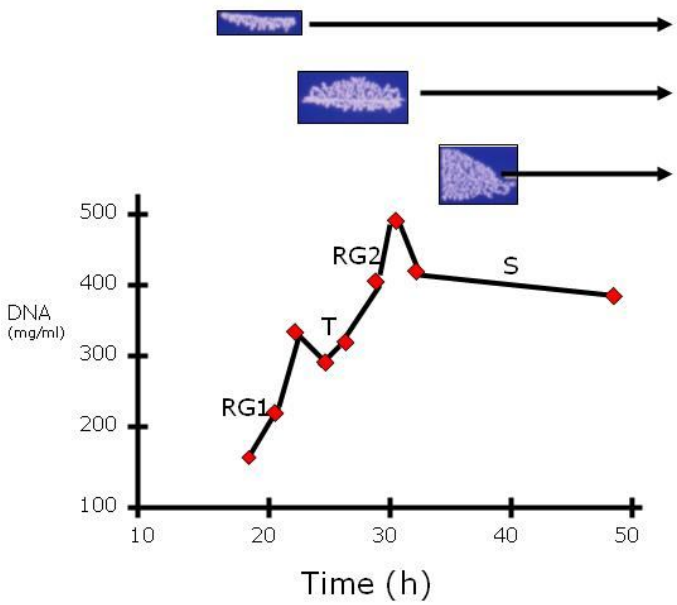


Figure 3 Growth curve of *Streptomyces coelicolor* (Granozzi C, 1990).

1.3 The Streptomyces coelicolor A3(2) genome

The *Streptomyces* mycelium is cenocytic, each hypha containing multiple chromosome copy, spores, instead contain only one copy of the chromosome.
The chromosome of *Streptomyces coelicolor* is linear. This characteristic confers a certain instability to the genome and thus a strong tendency to undergo, spontaneously, to deletions of DNA fragments, especially in the telomeric region. These deletions seem to be caused by transposition of DNA mobile elements (Volff JN, 1998). *S. coelicolor* genome, is the most characterized of all streptomycetes, the chromosome having been sequenced completely. This chromosome has a size of 8.7 Mb.
The origin of replication, *oriC*, is located in the chromosome center (Bentley SD, 2002). At the two ends of the chromosome, there are terminal inverted repeated sequences (TIRs) which carry proteins bounded to their free 5'-ends. The replication of the genome is carried out bidirectionally starting from the *oriC* and leaves a gap after the last RNA primer elimination. This gap is then filled by a DNA synthesis, starting from proteins being at the end. This mechanism has been called "end-patching". (Bao K, 2001)

With 7'825 identified genes, the chromosome of *S. coelicolor* has an enormous codifying potential. The gene distribution appears uniform along the chromosome, with a c.a. 4.9 Mb central core and two arms of 1.5 Mb for the left arm and 2.3 Mb for the right arm (fig. 4).

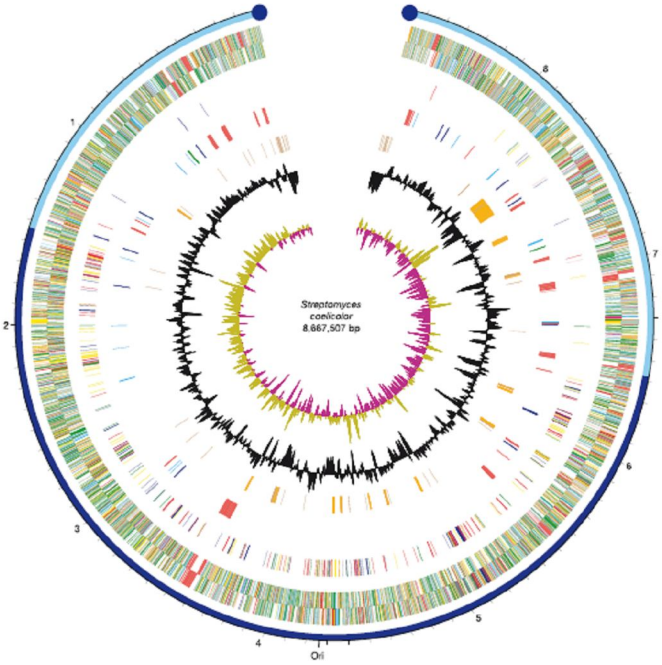


Figure 4 Circular map of the genome of *S. coelicolor* (Bentley SD, 2002).

Almost all genes present in the core are essential for the cellular division, DNA replication, transcription and translation and amino acids biosynthesis. Instead, contingency genes, codifying for non essential functions, like secondary metabolites and hydrolytic enzymes, reside in the arms. It has been found a correspondence between the core of the genome of *S. coelicolor* and the genome of the pathogenic bacterium *Mycobacterium tuberculosis*, indicating that the two microorganisms can have common ancestors, and *S. coelicolor* genome arms correspond to DNA acquired successively. (Bentley SD, 2002)
Through analysis of global gene expression, using DNA microarray, it has been pointed out that during the vegetative growth, genes involved in the primary metabolism, are mainly expressed, both in rich or minimal medium. When the stationary growth phase starts, the gene expression in the core decreases, contrary to genes present in arms, which are more expressed (Karoonuthaisiri N, 2005). The proteins having regulatory function are approximately 965. *S. coelicolor* has, moreover, numerous regulatory two components systems (Bentley SD, 2002).

1.4 Primary metabolism

The primary metabolism includes all catabolic and anabolic reactions or pathways, which carry to a cell growth and thus an increase of the biomass (Hodgson DA, 2000).
Many carbohydrate catabolic pathways are inducible. In substrate absence, a repressor binds to the promoter region of the catabolic operon and blocks the expression of the operon. For example, in *S. coelicolor* the operon of the glycerol catabolism is activated through an inducer, the glycerol-3-phosphate, produced by a kinase, activated when high level of glycerol is present in the medium (Seno ET, 1983).

When the natural environment of streptomycetes, the soil, contains nitrogen sources, like ammonium, glutamate, glutamine or nitrate, cells start using these compounds for their growth.
When these substances start to lack, the cell begin the degradation of some amino acids and nucleotides, in order to have nitrogen sources. The soil, is generally poor in nitrogen, this is the reason why these pathways are for half constitutive and for half inducible. (Hodgson DA, 2000)
The arginine catabolism needs the presence of the ornitine aminotransferase enzyme. This enzyme is constitutive but is repressed by ammonium and aspartate (Bascaran V, 1989). The catabolic pathway of tryptophan has been identified, for the first time, in *S.parvulus* (Katz E, 1984). The first enzyme of the catabolic pathway is the tryptophan dioxygenase, which is not induced and not repressed by tryptophan. On the contrary its activity is increased during the antibiotic actinomycin D, biosynthesis, thanks to a glutamate presence into the medium,. In bacteria, the majority of amino acid biosynthetic pathways are repressed by feedback repression, the final product repressing biosynthetic genes. Streptomycetes, seems to have lost or then not to have never evolved mechanisms of repression of amino acids biosynthesis. This is probably due to the fact that the soil is generally poor in amino acids. Indeed, there exist only three examples of feedback repression system : aromatic amino acids biosynthesis (Murphy MF, 1980); arginine biosynthesis (Flett F, 1987); aliphatic amino acids biosynthesis (Potter CA, 1996). It thus seems that streptomycetes have a reduced capacity to modify their primary metabolism, because of their oligotrofic natural environment. (Hodgson DA, 2000)

1.5 Tryptophan biosynthetic pathway

Studies carried out on auxotrophic strains of *Streptomyces coelicolor* allowed the identification of genes involved in the tryptophan biosynthesis, which are *trpA*, *trpB*, *trpC*, *trpD*, *trpE*, *trpG* and *hisA*. Some of these genes, like *trpC*, *trpD* and *trpE* have paralogue genes.
The genes *trpC2*, *trpD2*, *trpG* and *trpE2* belong to an operon internal to the cluster of the calcium dependent antibiotic (CDA) biosynthesis. The other *trp* gene are in three different clusters. The cluster which contains the s *trpA*, *trpB* and *trpC* genes, contains also the *trpX* (*SCO2038*) gene, a small orf, which would code for a 63 amino-acids peptide with unknown function (fig. 5). (Hu DSJ, 1999)

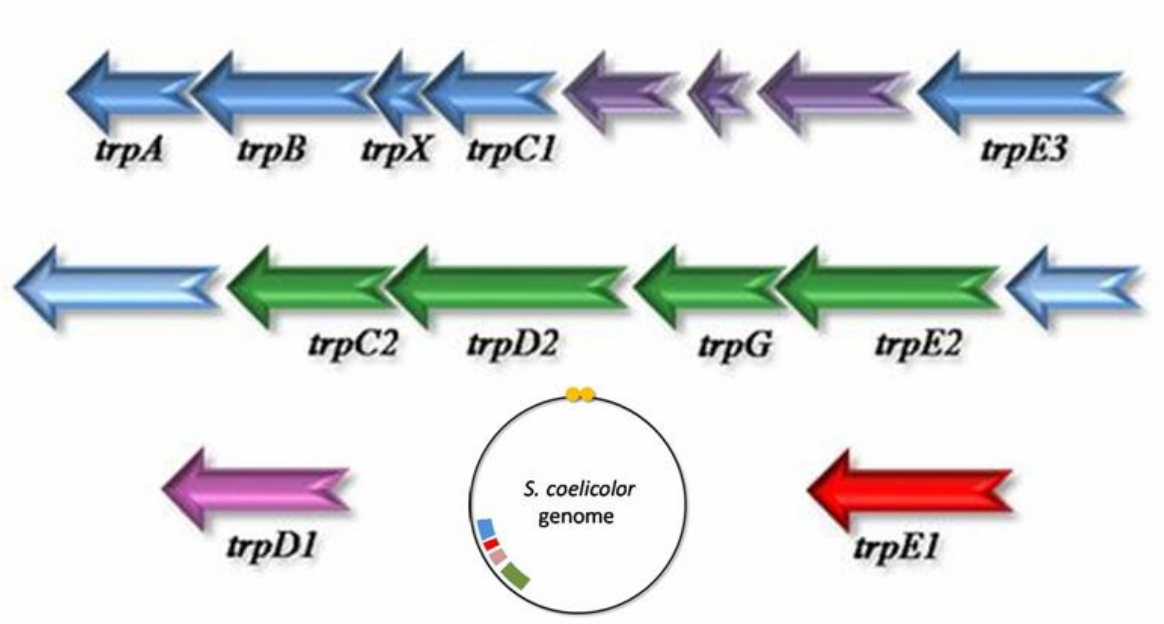


Figure 5 Schematization of genes for tryptophan biosynthesis in *S. coelicolor.*

The roles of various genes implied in the tryptophan biosynthesis are described hereafter :

- *trpE1/2/3* and *trpG* code for the anthranilate synthase. The *trpG* gene codes for a glutamine amidotransferase, its function being, in physiologic condition, to provide ammonium from glutamine. This reaction occurs, with lowest efficiency, without amidotransferase at alkaline pH, so that *trpG* is not essential. (Zalkin H, 1998)
- *trpD1/2* code for the anthranilate phosphoribosyl transferase, which catalyzes the formation of phosphoribosyl anthranilate, in the presence of anthranilate and 5-phosphoribosyl-1-pyrophosphate.
- *hisA* codes for the N-(5'-phosphoribosyl)-anthranilate isomerase, which catalyzes the tranformation of phosphoribosyl anthranilate in 1-(O-carboxyphenylamino)-1-deoxyribulose-5-phosphate. The *hisA* gene is also involved in the histidine biosynthesis.
- *trpC1/2* code for the indole-3-glycerol-phosphate synthetase, an enzyme catalyzing the synthesis of indole-3-glycerol phosphate from 1-(O-carboxyphenylamino)-1-deoxyribulose-5-phosphate.
- *trpA* and *trpB* code for the tryptophan synthase. The *trpA* gene codes for the α-subunit of the enzyme, responsible for the conversion of indole-3-glycerol phosphate in indole and glyceraldehydes-3-phosphate. The α-subunit forms a complex with β-subunit encoded by *trpB*, catalyzing the formation of L-tryptophan, from indole and L-serine.

Aromatic amino acids (tryptophan, phenylalanine, tyrosine) are synthesized by a common precursor, chorismate, produced by the shikimate pathway, identified in all streptomycetes analysed. (Hodgson,2000)

The tryptophan pathway requires initially the synthesis of anthranilate from chorismate by anthranilate synthase. The mechanism of this reaction is an enolpyruvic elimination accompanied with an ammonium transfer of a glutamine. Then, there is the addition of the phosphorybosyl functional group of 5-phosphoribosyl-1-pyrophosphate in 3' position of anthranilate, catalyzed by anthranilate phosphoribosyl transferase. This product is transformed in 1-(O-carboxyphenylamino)-1-deoxyribulose-5-phosphate by N-(5'-phosphoribosyl)-anthranilate isomerase by rearrangement of phosphoribosyl-anthranilate. The indole-3-glycerol-phosphate synthetase then synthesizes indole-3-glycerol phosphate, by decarboxylation and closing of the ring. The last reaction is the suppression of the lateral chain of glycerophophate of indole-3-glycerol phosphate and its substitution with the alaninic group of L-serine, by tryptophan synthase, leading to the final product, the tryptophan (fig. 6).

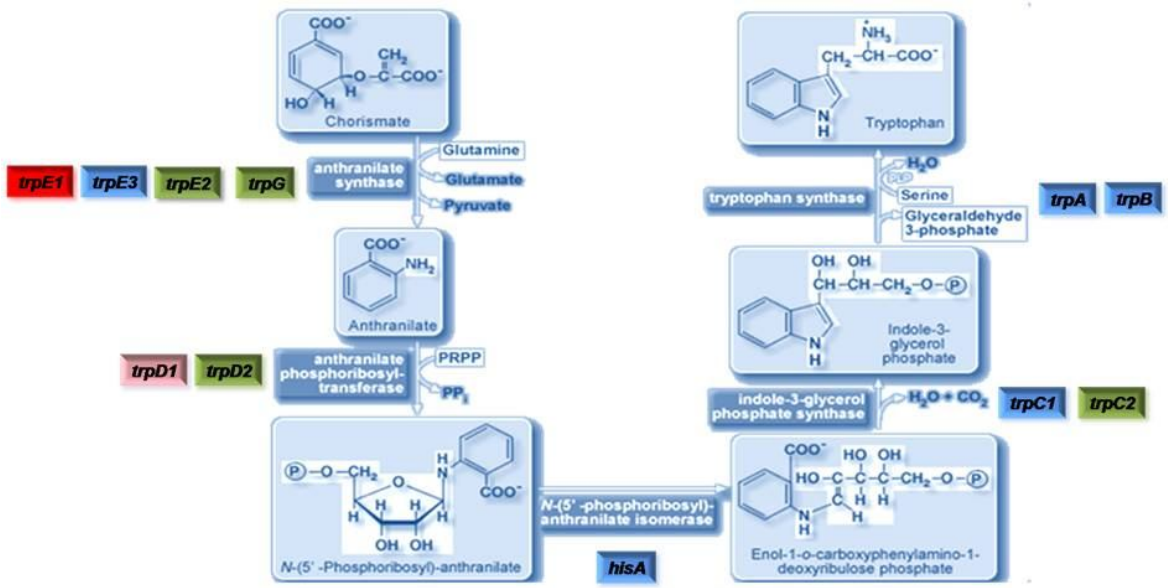


Figure 6 Schematization of the tryptophan biosynthetic pathway.

1.6 Regulation of tryptophan biosynthetic genes

The tryptophan biosynthesis requires a great amount of energy, so in bacteria it is finely regulated. Generally, in bacteria, the regulation of tryptophan biosynthetic gene is performed by attenuation mechanism and is based on the possibility to form hairpin structures, using a peptide-leader region, downstream of *trp* operon, that can form secondary structures, during the translation of the tryptophan operon. The regulation mechanism of the tryptophan biosynthesis has been extensively studied in various organisms, such as *Escherichia coli* and *Bacillus subtilis* (Yanofsky C, 2000).

Expression of tryptophan biosynthetic genes in *E. coli* is regulated by two mechanisms. The first one is regulated *via* TrpR protein, encoded by *trpR* gene. This protein can have two different conformations, depending on the binding of tryptophan (the corepressor) with this protein. When TrpR is bound to tryptophan, it acts as a repressor for tryptophan biosynthesis, binding to the *trp* operon operator-promoter region. This regulation is thus dependent on the tryptophan concentration in the cell. When the concentration of tryptophan is high, it can bind to TrpR (Zurawski G, 1978).

The second mechanism has been called attenuation and it is due to a peptide-leader region transcribed downstream of *trp* operon and containing two codons for the tryptophan amino acid. The transcriptional regulation of the *trp* operon is thus dependent on the leader region translation, regulated by tryptophan into the cell (Yanofsky C, 2004).

The *Bacillus subtilis trpEDCFBA* operon contains six of the seven genes required for the biosynthesis of tryptophan from chorismate, the common aromatic amino acid precursor. *pabA* (*trpG*), the remaining tryptophan biosynthetic gene, is present in an operon whose genes are primarily involved in folic acid biosynthesis. Expression of the *B. subtilis* tryptophan biosynthetic genes (*trpEDCFBA*) is regulated in response to tryptophan by the *trp* RNA-binding attenuation protein (TRAP), encoded by the *mtrB* gene. In presence of high concentration of tryptophan, TRAP binds to the leader sequence of *trp* mRNA and induce the formation of the transcription termination structure by hairspin structure formation. When concentration of tryptophan in the cell is low, TRAP does not bind to the leader sequence and an antitermination structure forms. This mechanism is relatively similar to that present in *E. coli*.
TRAP-mediated regulation of the tryptophan biosynthetic genes includes also another distinct translation control mechanisms. TRAP regulates translation of *trpP* (*yhaG*), a single-gene operon that encodes a putative tryptophan transporter, and *pabA* (*trpG*) gene. TRAP binds specifically to a segment of the *trpP* and *pabA* transcript that includes the untranslated leader and translation initiation region. The tryptophan-activated TRAP regulates translation of *trpP* and *pabA* by preventing binding of a 30S ribosomal subunit or by blocking ribosome binding on the *trpP* and *pabA* mRNAs. (Yakhnin H, 2004)
The regulation mechanisms of the *trp* operon in *Streptomyces coelicolor* are still unknown. Two potential sites of transcription initiation of the *trpCXBA* operon, however have been identified, the first upstream of *trpC* gene and the other upstream of *trpX* gene. The regulation of tryptophan biosynthetic genes in *S. coelicolor* does not seem to be due to feedback mechanism, but rather growth phase-dependent and growth rate-dependent. In fact, the *trp* operon is more massively expressed during the first exponential phase (RG1) and decreases during the second exponential phase (RG2) (Hogdson DA, 2000 ; Hu DSJ, 1999).

1.7 Secondary metabolism and antibiotic production

1.7.1 Antibiotic production in *S. coelicolor*

The secondary metabolism contains all reactions carrying to the biosynthesis of non-essential, but useful, substances for the cell. Generally secondary metabolites are produced in large amount during the stationary phase or slow growth phases.
Antibiotics are secondary metabolites and one can estimate than more of 50% of 11,900 antibiotics today known, are produced by *Streptomyces* (Kieser T, 2000). *S. coelicolor* produces several antibiotics, the most characterized being the actinorhodin, the undecylprodigiosin and the calcium dependent antibiotic (CDA). The actinorhodin polyketide (ACT) gives a characteristic blue color to *Streptomyces* colonies and to the surrounding medium (from where the name *coelicolor*). The undecylprodigiosin (RED) is a non-diffusible antibiotic which gives a red pigmentation to the mycelium. (Hopwood, 1999) The CDA belongs to the group of lipopeptidic acid antibiotics.

Through microarray analysis, the relation between the growth phase and the antibiotic production in the strain *S. coelicolor* M145, has been evaluated (fig. 7). (Huang, 2001)

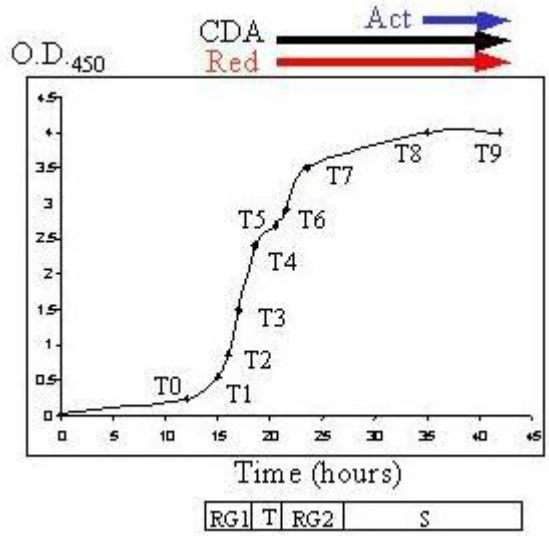


Figure 7 Growth of *Streptomyces coelicolor* A3(2) strain M145 in liquid R5–medium and production of antibiotics (Huang, 2001).

CDA and RED antibiotic production starts during the transition phase (T), whereas the ACT antibiotic is not produced before the stationary phase (S).
The biosynthesis of secondary metabolites is influenced by a huge variety of signals, as much extracellular than intracellular. The expression of their biosynthetic cluster is controlled by various families of regulatory proteins, a great part of them belonging to the SARP family (Bibb MJ, 2005). An example of this phenomenon is RelA, the synthetase of highly phosphoryl guanilic nucleotide (ppGpp), associated to the ribosome, which is used for the antibiotic production under nitrogen limiting conditions in *S. coelicolor* (Chakraburtty R, 1997). The secondary metabolites biosynthesis can also be influenced by the availability of primary metabolism products and is triggered by several stressful conditions.

1.7.2 Calcium dependent antibiotic (CDA)

The calcium dependent antibiotic (CDA) belongs to the class of the natural products called non-ribosomal polypeptides (Kim HB, 2004). It is a lipopeptide formed by a cyclic peptide of eleven amino-acids, with the fatty acid 2,3-ipoxyexanol in the N-terminal position of the side chains (fig. 8). (Hojati Z, 2002)

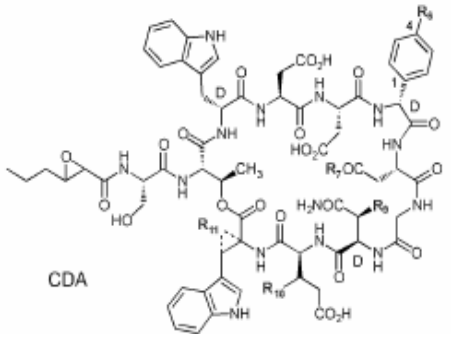


Figure 8 Structure of CDA isolated from *S. coelicolor* (Hojati Z, 2002).

The CDA acts against a vast range of Gram positive bacteria, but only in presence of calcium ions (for this reason is called calcium dependent). Calcium ions (Ca^{2+}) allow an aggregation of several CDA molecules, through interactions between acid residues. This brings to the formation of a transmembrane canal in the bacterial membrane (Hopwood DA, 1983; Grunewald J, 2003).
Recently the metabolic network responsible for the production of CDA, which includes more than 400 enzymatic reactions has been built (Kim HB, 2004).
Genes codifying these enzymes are organized in a 82 Kb cluster, also containing genes implied in the CDA regulation and transport, for a total of 40 orf. This cluster is localised in the central region (core) of the *Streptomyces coelicolor* chromosome (Chong PP, 1998; Bentley SD, 2002).
Inside the CDA cluster there is an operon containing some genes potentially involved in tryptophan biosynthesis. The *trpC2*, *trpD2* and *trpE2* genes are analogues to other genes present in the chromosome, outside the cluster, whereas the *trpG* gene is present in a single copy. In the center of the CDA cluster there is three long genes (*cdaPS1/2/3*), codifying for CdaPS1, 2 and 3, non-ribosomal peptide synthetases (NPRS), allowing the CDA biosynthesis (fig. 9) (Hojati Z, 2002).

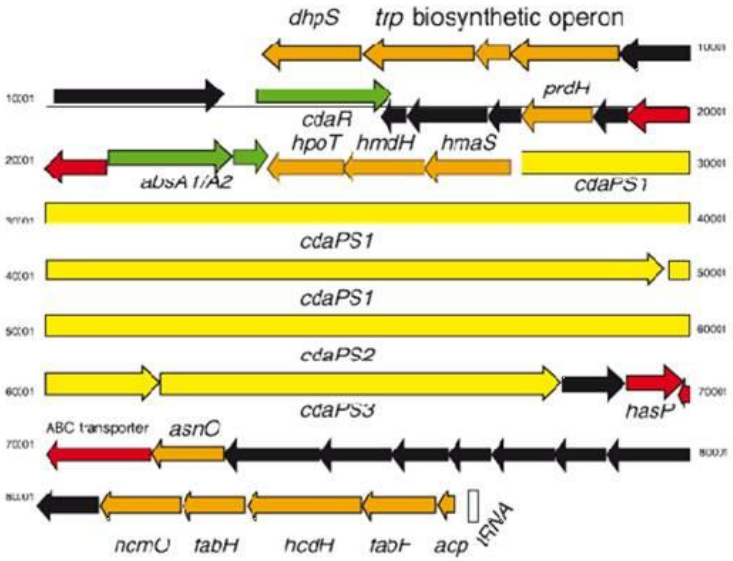


Figure 9 CDA biosynthetic cluster in *S. coelicolor* (Hojati Z, 2002).

The NPRS biosynthesis requires some primary metabolites, necessary for the growth; in fact, these precursors are (Kim HB, 2004) :

- Coenzyme-A deriving from fatty acids metabolism
- Mevalonate deriving from Acetyl-CoA
- Nucleotides
- Carbon sources (sugars)
- Shikimate
- Amino acids (mainly tyrosine, aspartate, tryptophan, threonine, glycine, serine, glutamate and oxoglutarate)

The CDA synthesis rate has been analyzed studying the effect of mutations in genes which code for enzymes involved in the production of primary metabolites, precursors of CDA. The CDA production was increased intervening on the assimilation way of the nitrogen, on the pentose phosphate pathway and also on oxoglutarate flow. (Kim HB, 2004)

2 MATERIALS AND METHODS

2.1 Strains and plasmids

Table 1 Strains and plasmids used in this study.

Strains and plasmids	Relevant genotype	Source or reference	Acronym
***E. coli* strains**			
DH5α	F⁻Φ80dlacZΔM15Δ(lacZY A-argF)U169 endA1 recA1 hdsR17 deoR thi-1 supE44 λ⁻ gyrA96 rel1A1	(Hanahan D, 1983)	DH5α
BL21(DE3) pLysS	F⁻ ompT hsdSB (rB⁻ mB⁻) gal dcm (DE3) pLysS (CamR) araB::T7RNAP-tetA	(Studier FW, 1986)	BL21
***S. coelicolor* strains**			
M145	Wild type, SCP1⁻, SCP2⁻	(Kieser T, 2000)	M145
M145 trpX::Tn5062	M145 trpX::Tn5062	UNIPA	ΔX
M145 trpX::Tn5062 /pKCHyg796	M145 trpX::Tn5062 pKC796Hyg	UNIPA	ΔX/pKC
M145 trpX::Tn5062 /pKCHyg796trpX	M145 trpX::Tn5062 pKCHyg796trpX	UNIPA	ΔX/X
Plasmids			
pGEM®-T easy	lacZ, Amp^r	Promega	pGEM
pGEM-trpX	pGEM®-T carrying trpX	UNIPA	pGEM-X
pRSETB	T7 promotor, N-terminal His6 tag, Amp^r	Invitrogen	pRSETB
pRSETB-trpX	pRSETB carrying trpX	This study	pRSETB-X
pIJ8600	tipAp, tsr, apr^r, int/xis oriT, attP	John Innes Center	pIJ8600
pIJ8600-trpX	pIJ8600 carrying trpX	This study	pIJ-X

2.1.1 E. coli

2.1.1.1 *E. coli* DH5α and *E. coli* DH5α pGEM-X

The strain *E. coli* DH5α is a well known strain used for cloning. Its genotype is F⁻ Φ80d*lac*ZΔM15Δ(*lac*ZY A-argF)U169 *end*A1 *rec*A1 *hds*R17 *deo*R *thi*-1 *sup*E44 λ⁻ *gyr*A96 *rel*1A1 (Hanahan D, 1983).
The strain *E. coli* DH5α pGEM-X has been provided by UNIPA, Palermo, Italy. The pGEM-X plasmid has been construct by insertion of *trpX* in pGEM-T Easy™ (Promega), using *Bam*HI and *Nde*I restriction sites. The *trpX* correct insertion has been guaranteed by DNA sequencing.

2.1.1.2 *E. coli* BL21(DE3) pLysS

The genotype of *E. coli* BL21(DE3) pLysS is F^- ompT hsdSB (rB^- mB^-) gal dcm (DE3) pLysS (CamR) araB::T7RNAP-tetA. (Studier FW, 1986)
E. coli BL21 have a T7 bacteriophage RNA polymerase-based expression system, for high yield expression of recombinant proteins. This strain is a lisogen of λDE3 virus, it thus contains a copy of the gene encoding for the T7 RNA polymerase. The recombinant protein expression can be induced by IPTG.
Another characteristic is a mutation in the gene *rne131* coding for the RNase E enzyme, an essential endonuclease. This enzyme is involved in rRNA maturation and mRNA degradation, thus mRNAs expressed in *E. coli* BL21, have increased stability.
The plasmid pLysS contains a gene encoding for the T7 lysozyme (*lysS*) and also a chloramphenicol resistance gene.

2.1.2 Streptomyces coelicolor A3(2)

The strain *S. coelicolor* A3(2) M145 trpX::Tn5062 has been constructed, on the basis of *S. coelicolor* A(3)2 M145, by *trpX* gene interruption using the 5062 transposon. The complementation of *trpX* in this mutant has been made using the recombinant plasmid pKCHyg796trpX (*S. coelicolor* A3(2) M145 trpX::Tn5062/pKCHyg796trpX). The strain *S. coelicolor* A3(2) M145 trpX::Tn5062/pKCHyg796 has been also constructed.

2.1.3 pGEM®-T easy

The pGEM®-T easy vector (A1360, Promega, Madison WI, USA) is a high copy number plasmid and it has been used for *E. coli* cloning. It also contains a resistance to ampicillin (*amp*^{r}) gene (fig. 10).

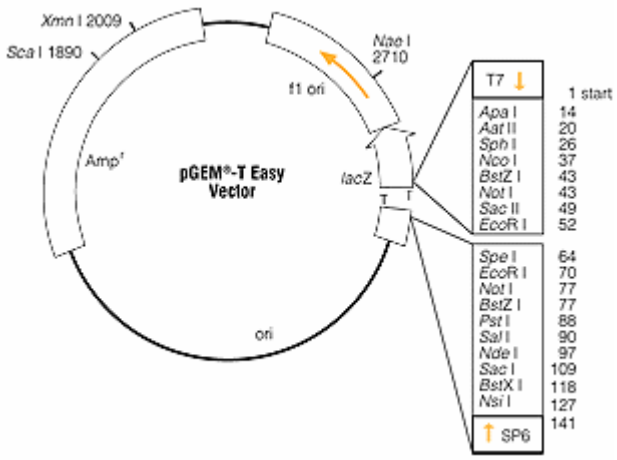


Figure 10 pGEM®-T Easy vector map (Promega, Cat. N° TM042).

2.1.4 pIJ8600

The pIJ8600 plasmid has been built by John Innes Centre, Norwich, UK. It can be replicated in *E. coli* and it can be used for integration in the *S. coelicolor* chromosome and overexpression of a gene. It is characterised by the following element (fig. 11) :

- A thiostreptone inducible promoter (*tipAp*)
- A thiostreptone resistance gene (*tsr*)
- A apramycin resistance gene (apr^r)
- A multiple cloning site (MCS)
- *int/xis* gene coding for an integrase, useful for plasmid integration in *S. coelicolor* chromosome
- *oriT* origine site, used for the conjugation
- *attP* site homologous to *attB* sequence in *S. coelicolor* chromosome, used for a specific integration of the DNA insert

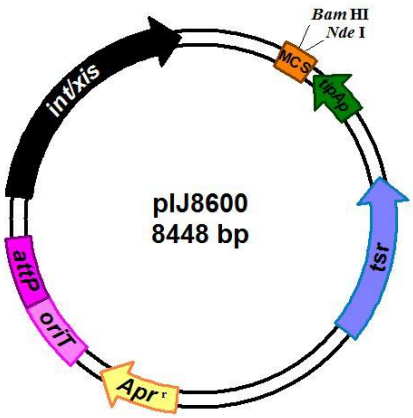


Figure 11 pIJ8600 plasmid map

2.1.5 pRSETB

The pRSETB plasmid (V35120, Invitrogen, Carlsbad CA, USA) comes from a vector expression family derived from pUC and has been developed for recombinant protein overexpression in *E. coli*. The recombinant protein overexpression is possible due to a T7 promoter. Moreover inserts are positioned downstream and in the reading frame of a N-terminal fusion peptide coding sequence. This peptide sequence includes a ATG codon for the translation initiation and a polyhistidine tag. His-tag allows a simple purification of recombinant protein with a nickel containing resin (fig. 12).

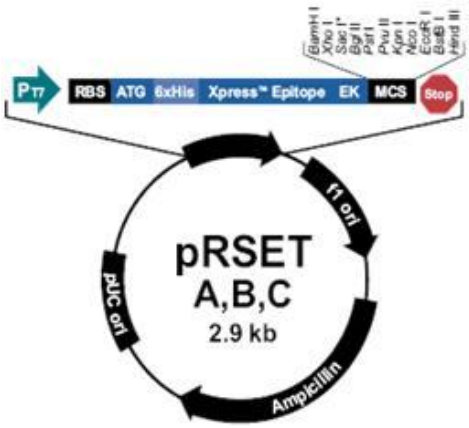


Figure 12 pRSETB plasmid map

2.2 Culture medium

2.2.1 Solid culture medium

LB Agar Medium (also called LA Medium)

Put 25 g of Luria Broth Base (Miller's LB Broth Base, 12795, Invitrogen, Carlsbad CA, USA) in a 1 L bottle and add 10 g of Technical Agar. Complete to 1 L with demineralised water, mix and sterilise at 121°C for 20 min.

2.2.2 Liquid culture medium

LB Medium

Put 25 g of Luria Broth Base (Miller's LB Broth Base, 12795, Invitrogen, Carlsbad CA, USA) in a 1 L bottle and complete to 1 L with demineralised water, mix and sterilize at 121°C for 20 min.

2×YT Medium

Product	Mass [g]
Difco Bacto tryptone	16
Yeast Extract	10
Sodium chloride (NaCl)	5

- Complete to 1L with demineralised water
- Sterilize at 121°C for 20 min

SMMS Medium

Product	Quantity
Casaminoacid	2 g
TES 1M pH 7.6	50 ml
NaH_2PO_4 50 mM	10 ml
K_2HPO_4 50 mM	10 ml
$MgSO_4$ 1 M	5 ml
D-Glucose 50 %	18 ml
Trace elements**	1 ml
Tryptophan 7.5 mg/ml	7.5 ml

** Trace elements for SMMS Medium : **0.1 g/l** of $ZnSO_4 \cdot 7\ H_2O$, $FeSO_4 \cdot 7\ H_2O$, $MnCl_2 \cdot 4\ H_2O$, $CaCl_2 \cdot 6\ H_2O$ and NaCl

- Complete to 1 L with demineralised water
- Sterilize at 121°C for 20 min

SpMR Medium

Product	Quantity
Sucrose	103 g
Magnesium chloride ($MgCl_2$)	10 g
D-Glucose	5 g
Yeast Extract	5 g
TES 1M pH 7.6	50 ml
Trace elements*	2 ml

* Trace elements for SpMR Medium : **0.6 mM** $ZnCl_2$, **0.12 mM** $CuCl_2$, **0.10 mM** $MnCl_2$, **0.05** mM $Na_2B_4O_7$, **0.016** mM $(NH_4)_6Mo_7O_{24}$

- Complete to 1 L with demineralised water
- Sterilize at 121°C for 20 min
- Add 2 ml of sterile $CaCl_2$ 5M

2.3 DNA Manipulation

2.3.1 Agarose Gel Electrophoresis

The DNA Agarose Gel Electrophoresis is used for DNA fragment separation using an electric current applied to a gel matrix. Indeed, the DNA (negatively charged) migrates inside the agarose gel towards the anode (positively charged). A small DNA fragment migrate more quickly than a larger one. The fragment separation is thus made by size and topology. The agarose has been solubilised in TAE buffer 1X and the gel electrophoresis has been made in a bath of TAE 1X.
Ethidium bromide has been added to the gel (0.001% of the final gel volume) because it allows the DNA visualisation. The ethidium bromide have a planar circular structure and link to DNA by intercalation between nucleic bases. When ethidium bromide is bound to DNA, it can be excite with UV light, and it emits fluorescence. A DNA fragment size is determinated with a DNA ladder. DNA fragments can also be quantified by confronting the fluorescence intensity emitted by the ladder and the DNA sample. The ladder used has been GeneRuler™ DNA Ladder Mix (SM0332, Fermentas, Hanover MD, USA). A 6X loading buffer has been added to the samples, in a ratio of 1:5 (loading buffer : sample).

TAE buffer 50X (Tris-Acetate-EDTA)

242 g Tris BASE, **57.1 ml** Acetic acid glacial and **100 ml** EDTA (0.5 M, pH8.0)

- Complete to 1 L with demineralised water
- Adjust the pH to 8.5

Loading buffer 6X

30% Glycerol, **60 mM** EDTA (ph 8.0) and **0.06%** Bromophenol Blue and **0.06%** Xylene cyanol

2.3.2 PCR

The Polymerase chain reaction (PCR) is a technique of *in vitro* targeted DNA amplification. It allows to obtain, important amount of a specific DNA fragment. The amplification of DNA is carried out at the time of three steps repeated recurringly, after a first step of denaturation (initial denaturation) :

- **Denaturation** at 95°C of dsDNA in ssDNA.
- **Annealing**, step allowing to primers to bind the ssDNA. The temperature of annealing is determined by primer sequences.
- **Extension** of DNA, starting from primers, by the Taq polymerase, at 72°C.

The PCR has been made in the thermocycler Tpersonal® (050-550, Biometra, Goettingen, Germany).
PCR amplification of *trpX* gene has been carried out in the master mix and with PCR conditions shown thereafter:

Master mix (for 1 sample)

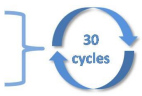
Product	Concentration	Volume [µl]
Buffer Taq	10X	5
Magnesium chloride ($MgCl_2$)	50 mM	1.5
dNTP	10 mM	1
Primer forward	-	1
Primer reverse	-	1
Taq Polymerase	5 U/µl	0.5
DMSO	-	5
H_2O demineralised	-	33
Final volume	-	**48**

➔ Add 2 µl of DNA

Temperature and times of PCR steps

5 min at 95°C → Initial denaturation

N°	Temp. [°C]	Time [min]	N°	Temp. [°C]	Time [min]	N°	Temp. [°C]	Time [min]
1	95°C	1	7	95°C	1	13	95°C	1
2	64°C	1	8	60°C	1	14	56°C	1
3	72°C	1	9	72°C	1	15	72°C	1
4	95°C	1	10	95°C	1	**16**	**95°C**	**1**
5	62°C	1	11	58°C	1	**17**	**54°C**	**1**
6	72°C	1	12	72°C	1	**18**	**72°C**	**1**



2.3.3 Colony pool PCR

The colony pool PCR is a DNA amplification technique from a microorganism colony. Some cells taken from few colonies on a Petri dish are suspended in the PCR master mix. The cell lysis, and thus the release of DNA, are carried out during PCR steps.
This method is typically used for the screening after a cloning, allowing to analyze several colonies in a single PCR tube. The colony pool PCR has been made in the thermocycler Tpersonal® (050-550, Biometra, Goettingen, Germany). Colony pool PCR amplification of *trpX* has been made with the conditions described in the PCR chapter.

2.3.4 Extraction of plasmid DNA

2.3.4.1 MiniPrep extraction

The MiniPrep protocol allows us to extract plasmid DNA from a liquid culture. Two methods have been used for this extraction. First method has been used for a high purity of plasmid DNA after extraction and has been made with the Kit GenElute™ Plasmid Miniprep (cat. n° PLN-10, Sigma-Aldrich, Steinheim, Germany). Handlings carried out are described in the protocol provided with the Kit. Second method allows to extract plasmid DNA, with a less cost compared to the commercial kit. The disadvantage is the poor quality of the plasmid DNA extracted. In fact, it contains, generally, a not negligible part of chromosomal DNA and proteins. The protocol used is shown hereafter :

1. Pour 1.5 ml of culture in one Eppendorf tube and spin 10 min at 12'000 rpm.
2. Discard the supernatant
3. Suspend the pellet in 100 µl of Solution I and wait 5 min at room temperature (RT)

This step weakens the cellular membrane

4. Add 200 µl of Solution II, mix and keep on ice during 5 min

This step permits the cell lysis and it is important to maintain the temperature at 4°C for not damage the DNA

5. Add 150 µl of Solution III, mix gently and keep on ice during 5 min

During this step, the pH of the solution returns to pH7 abruptly. The chromosomal DNA is not correctly renatured and thus precipitates with cell debris while most of plasmid DNA renatures remaining in solution

6. Spin during 1-2 min at 4°C, take 400 µl of the supernatant and pour in another Eppendorf tube
7. Add 2 volumes (800 µl) of cold ethanol 95%, mix and wait 10 min at RT
8. Spin 5 min, discard the ethanol and dry the tube in a Speed Vacuum for 30 min

It is possible air dry the tube on a piece of paper at room conditions

9. Suspend the pellet in 30 µl of demineralised water

Sometimes it is not possible to see a pellet, in this case wash well the wall of the eppendorf tube with the pipette

Solution I	
Glucose	50 mM
Tris-HCl (pH8)	25 mM
SDS	1%

Sterilize and stock at 4°C

Solution II	
NaOH	0.2 N
SDS	1%

Prepare a fresh solution

Solution III	
Potassium acetate	5 M
Acetic acid glacial	28.75%

Sterilize and stock at 4°C

2.3.4.2 MaxiPrep extraction

The MaxiPrep extraction uses the same principle as the MiniPrep extraction, but the amount of plasmid DNA extracted is much higher. This extraction has been made with the Kit QIAGEN® Plasmid Maxi (12163, QIAGEN Sciences, Germantown MD, USA). Extraction has been carried out according manufacturer's protocol.

2.3.5 Phenol-chloroform extraction

The removal of proteins, from a DNA sample, is important because cells contain enzymes degrading nucleic acids or DNA binding proteins and can interfere with subsequent enzymatic treatments. The classic protocol uses a mixture of phenol and chloroform. The protocol used is shown below :

1. If the volume is less than 50 µl, add demineralised water to arrive at this volume
2. Add one volume of phenol-chloroform and mix vigorously vortexing

The phenol permit to denaturate proteins which will precipitate now at the interphase

3. Spin for 5-7 min at 13'000 rpm
4. Recover the aqueous phase (the upper phase)

Do not touch the interphase with the point of the pipette

5. Add a identical volume of chloroform and mix
6. Spin for 5-7 min at 13'000 rpm
7. Recover the aqueous phase like as at the time of step 4
8. Add 1/10 volume of sodium acetate 3M and 2 volumes of ethanol absolute
9. Keep the sample at -80°C for 1 hour

During this step the DNA precipitates

10. Spin at 13'000 rpm for 20 min at 4°C and discard the supernatant
11. Add 400 µl of ethanol 70%
12. Spin at 13'000 rpm for 10 min at room temperature and discard the supernatant
13. Dry the tube in a Speed Vacuum for 30 min

It is possible air dry the tube on a piece of paper at room conditions

14. Suspend the pellet in a appropriate volume of H_20 demineralised

Sometimes it is not possible to see a pellet, in this case wash well the wall of the eppendorf tube with the pipette

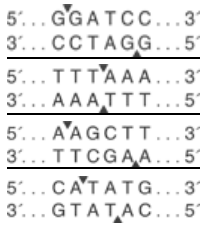
2.3.6 Drop dialysis

The drop dialysis is a simple and fast method to eliminate salts and other small molecules from a small DNA or protein sample volume. The elimination of salts in the sample is achieved by diffusion through the membrane, from the sample to an aqueous phase.
It has been carried out by filling a container (typically a 100 mm Petri dish) with demineralised water, on which a floating filter disc Millipore "V" Series Membranes 0.05µm (VMWP02500, Millipore, Billerica MA,USA) has been delicately deposited. When the filter has been wet, the sample, of a typical volume between 20 and 400 µl, has been deposited on the surface of the disc. The drop dialysis duration changes according to the sample volume and composition, but it generally lasted between 30 min and 1h. It is possible to add a marker dye to the sample to follow the purification process.

2.3.7 Digestion with restriction enzymes

Restriction enzyme amount used during a DNA digestion has been calculated in relation to the quantity of DNA to digest. One unit (1U) of enzyme allows, generally, to cut 1 µg of DNA in one hour. These amounts are efficient for a digestion at the optimal temperature of the enzyme (generally 37°C) and in the buffer provided with the enzyme (see the note provided during the purchase). Moreover, it is important to remember that the enzyme is stored in a solution containing glycerol, because problems of enzymatic activity can appear when the glycerol concentration during the digestion is higher than 4%.

Restriction enzymes used



- *Bam*HI (10 U/µl), 15201-031, Invitrogen, Carlsbad CA, USA
- *Dra*I (10 U/µl), 526S, New England Biolabs, Beverly MA, USA
- *Hind*III (10 U/µl), 15222-011, Invitrogen, Carlsbad CA, USA
- *Nde*I (5U/µl), 15426-018, Invitrogen, Carlsbad CA, USA

2.3.8 Alkaline phosphatase treatment

During a ligation between a DNA insert and a vector, it is important to prevent the vector recircularization (self-ligation). The ligation asks for the presence of a 5'-ends phosphate group at least on one of the two DNA fragments to bind. During the alkaline phosphatase treatment, phosphate groups at 5'-ends of DNA fragments are removed, avoiding vector self-ligation. The alkaline phosphatase treatment has been carried out with Alkaline phosphatase (AP) (11097075001, Roche Diagnostics GmbH, Indianapolis IN, USA) during 1h at 37°C.

To calculate the right quantity of alkaline phosphatase to use for a DNA dephosphorylation, it is necessary to know the DNA extremity quantity in the sample. In fact, 1 U AP dephosphoryls 1 pmol DNA extremity in 1 h at 37°C. To calcultate the DNA extremity quantity in the sample, the following two calculations have been used :

1) $\mathbf{MW_{DNA} = S_{DNA} \times 330\ [g/mol] \times Strand_{Nb}}$ **2)** $\mathbf{DNA_{extr} = (m_{DNA} / MW_{DNA}) \times 2}$

$\mathbf{DNA_{extr}}$: DNA extremity quantity in the sample in [pmol]
$\mathbf{m_{DNA}}$: DNA quantity in the sample in [pg]
$\mathbf{MW_{DNA}}$: Masse weight of DNA in the sample in [g/mol]
$\mathbf{S_{DNA}}$: DNA size in [bp]
$\mathbf{Strand_{Nb}}$: Number of DNA strand (1 = ssDNA and 2 = dsDNA)
330 [g/mol] = Average masse weight of a nucleotide

2.3.9 Ligation

The ligation is a reaction linking two DNA fragments. It allows, typically, to link an DNA insert to a vector. This reaction is carried out at room temperature, using an DNA ligase (T4 DNA Ligase, 15224-017, Invitrogen, Carlsbad CA, USA), during 1h.

To obtain a correct ligation the presence of contaminant in the reaction solution should be avoid and the concentration of the two DNA fragments to be bound accurately calculated. For a ligation insert-vector, it is generally used a mass ratio insert : vector of 3:1. This mass ratio makes it possible to minimize the formation of vector-vector and insert-insert dimers, but can nevertheless be modified to optimize the reaction.

2.3.10 Preparation of calcium competent cells

For the preparation of calcium competent cells, it is important to remember that **the solutions and handling must be sterile**.

1. Prepare 3 ml of pre-inoculum in LB Medium with *E. coli* cells. Incubate over-night at 37°C
2. Add the 3 ml of pre-inoculum to 200 ml of LB Medium in a 500 ml flask. Incubate at 37°C until the culture reaches an $OD_{650nm} = 0.6$
3. Pour the culture in 4 falcon tubes and keep 1-2h on ice
4. Spin at 3'700 rpm during 7 min at room temperature
5. Discard the supernatant and suspend the pellet in 25 ml sterile $CaCl_2$ 70mM (Pre-cooled on ice)
6. Put 1h on ice
7. Spin at 3'400 rpm during 7 min at room temperature
8. Discard the supernatant and suspend the pellet in 5 ml of freezing solution
9. Join together the samples, and prepare 8 eppendorf tubes with 200 µl of the cells, 16 eppendorf with 300 µl and 24 eppendorf with 400 µl
10. Store at -80°C

Freezing solution (for 20 ml)

14.28 ml $CaCl_2$ 70mM**, 4 ml** Glycerol 100% and **1.72 ml** sterile H_2O

2.3.11 Transformation of calcium competent cells

Calcium cations (Ca^{2+}) neutralize negative charges of phospholipids on the *E. coli* cellular membrane and allow the DNA molecules (Negatively charged too) to get closer to cell membrane. An abrupt change of the temperature (heat-shock) causes the formation of pores increasing thus, the permeability of the cell, allowing the plasmid to enter in the cell. The transformation of calcium competent cells protocol is presented to follow :

1. Freeze sterile tubes
2. Defrost calcium competent cells on ice

Calcium competent cells are sensible, a too fast defrosting will risk to damage cells

3. Add 200 µl of competent cells in a sterile tube
4. Add 20 µl of the DNA after ligation
5. Mix and keep tubes 40 min on ice
6. Heat shock of cells at 37°C for 3 min

The plasmid enters during this step. It is a critical step because cells are now very fragile, and a heat-shock prolonged could destroy the cellular membrane

7. Add 1 ml of LB medium
8. Keep tubes at 37°C for 1h
9. Put 400 µl of LB containing cells on selective plates (3 selective plates for each ligation)
10. Incubate at 37°C over night

Not more than 20h! After this time, antibiotics do not have any more the same effectiveness, and it's possible observe wrong clones

2.3.12 Preparation of electro-competent cells

1. Prepare 5 ml of pre-inoculum in LB Medium with *E. coli* cells. Incubate over-night at 37°C
2. Add the 3 ml of pre-inoculum to 500 ml of LB Medium in a 2l flask. Incubate at 37°C until the culture reaches an OD_{650nm} = 0.5
3. Pour the culture flask 10-15 min on ice
4. Pour the culture in a centrifuge bottle, spin 20min. at 2°C and 2500 g (Beckman JA-14 rotor), and discard the supernatant
5. Resuspend cells in 5 ml HEPES (1 mM, pH 7) and add other 500 ml cold HEPES
6. Well mix, spin at 2500 g during 20 min and discard the supernatant
7. Resuspend cells in 500 ml cold demineralised water, spin spin at 2500 g during 20 min and discard the supernatant
8. Resuspend cells in the remaining liquid and pour 10 ml of suspension in four 50 ml Falcon tube
9. Add 40 ml cold 10% glycerol and well mix
10. Spin 10 min at 2°C and 2500 g (Beckman JA-20 rotor) and discard the supernatant
11. Resuspend cells in the remaining liquid
12. Aliquot 40 µl cells in 1.5 ml Eppendorf tube
13. Freeze aliquots at -20°C and pour it 10 min in dry ice
14. Stock cell aliquots at -80°C

2.3.13 Transformation of electro-competent cells

During the transformation of electro-competent *E. coli* cells, the electric field, created by a high voltage impulse (12.5-15 $kV \cdot cm^{-1}$), induces the pore formation in the cell membrane. DNA molecules can enter in cells, using cell membrane pores.

1. Defrost electro-competent cells on ice

Electro-competent cells are sensible, a too fast defrosting will risk to damage cells

2. Add 2 µl of DNA after ligation to cells
3. Set the electroporator MicroPulser™ (411BR, Bio-Rad, Hercules CA, USA) at 25 µF, 2.5 kV and 200 ohm
4. Pour DNA-cell suspension in a cold electroporation cuvette, insert it in the electroporator and give the electric impulse

If too much salts are present in the cell suspension, that can generate an electric arc, the electric current passing through salts. If an electric arc is formed, cells are not affected, and it is necessary to repeat the transformation

5. Immediately after the electric impulse, remove the electroporation cuvette and add 1 ml of rich medium (like LB medium) to cells.

It is necessary to add the rich medium as soon as possible, for avoid the diminution of cell viability

6. Recover the cell suspension, pour it in a 1.5 ml sterile Eppendorf tube and keep at 37°C during 1h
7. Pour LB containing cells on selective plates (3 selective plates for each transformation)
8. Incubate at 37°C over night

Not more than 20h! After this time, antibiotics do not have any more the same effectiveness, and it's possible observe wrong clones

2.4 Protein methods

2.4.1 Protein structure prediction with Robetta server

The **Robetta server** (http://robetta.bakerlab.org) provides automated tools for protein structure prediction and analysis. For structure prediction, sequences submitted to the server are parsed into putative domains and structural models are generated using either comparative modeling or *de novo* structure prediction methods. If a confident match to a protein of known structure is found using BLAST, PSI-BLAST, FFAS03 or 3D-Jury, it is used as a template for comparative modeling. If no match is found, structure predictions are made using the *de novo* Rosetta fragment insertion method. The detailed operation of the Robetta server is described by Kim DE, 2004. The result of the Robetta protein structure prediction is a series of possible models, where for each model, a series of known proteins, having a similar structure, is proposed.

2.4.2 SDS-PAGE

The SDS-PAGE (Sodium Dodecylsulphate – Polyacrylamide Gel Electrophoresis) is a protein electrophoresis on a polyacrylamide gel. This technique have the same principle of the DNA electrophoresis on agarose gel. It allows the protein separation according to their molecular weight. The gel is constituted from a three-dimensional polymer of polyacrylamide, obtained from the copolymerization of acrylamide and methylen-bis-acrylamide, initiated from ammonium-persulphate (APS, initiator of the chain reaction) and TEMED (catalysator of the polymerization reaction). The SDS is a detergent and a strong ionic surfactant. It removes the non-covalent bond of proteins, allowing their denaturations, and thus the loss of their native folding. Moreover, SDS anions bind to proteins conferring a constant density of negative charges, independently of the molecular weight and amino acids composition. These conditions allows to obtain a separation only parameterized by the molecular weight of proteins.
The β-mercaptoethanol is a reducer of disulfur bonds. The disulfur bonds are important for the tertiary structure of proteins, and the cleavage of these bonds allow the linearization of proteins.
Before the SDS-PAGE analysis, the samples must be prepared. The samples are prepared in a 3X sample buffer (or loading buffer), in a ratio of 1:2 (loading buffer : sample), put during 5 min into a boiling water bath, followed by 5 min on ice. The high temperature facilitates the proteomic denaturation.
The polyacrylamide gel is discontinuous :

- Stacking Gel (wide meshes, in pH 6.8 buffer)
  The Stacking Gel allows the formation of a tight protein bands
- Resolving Gel (tight meshes, in pH 6.8 buffer)
  The Resolving Gel allows the separation of proteins

The SDS-PAGE analysis has been carried out with a polyacrylamide concentration of 12%. 15 µl of BenchMark™ Pre-Stained Protein Ladder (10748-010, Invitrogen, Carlsbad CA, USA) has been charged in the gel. The run of the SDS-PAGE analysis has been made, in the Tris-Glycine buffer 1X, at 30 mA until appropriates separation of protein bands.

Sample buffer 3X

150 mM Tris-HCl (pH 6.8), **300 mM** DTT, **6%** SDS, **0.3%** bromophenol blue and **30%** glycerol

Tris-Glycine buffer 5X

15 g Tris BASE, **72 g** Glycine and **10 ml** SDS 10%

- Complete to 1 L with demineralised water

Stacking Gel 5% (for 4 ml)

Product	Volume [ml]
H_20 demineralised	2.7
Acrylamide Mix 30%	0.67
Tris 1M (pH 6.8)	0.5
SDS 10%	0.04
Ammonium persulphate 10%	0.04
TEMED	0.004

Resolving Gel 12% (for 10 ml)

Product	Volume [ml]
H_20 demineralised	3.3
Acrylamide Mix 30%	4
Tris 1.5M (pH 8.8)	2.5
SDS 10%	0.1
Ammonium persulphate 10%	0.1
TEMED	0.004

The gel coloration has been made for minimum 1h (generally 1h30-2h), in Coomassie staining solution. After this step, the staining solution has been removed (it's possible to re-use the Coomassie staining solution as long as the color is sharp) and the gel has been washed with demineralised water. Afterward, the gel has been put in a bath of destaining solution and quietly shaken until discoloration of the gel (generally 3h or more).

Coomassie staining solution (for 1 l)

Product	Volume [ml]
Methanol	225
Acetic acid glacial	50
Coomassie Blue G250	1

Destaining solution (for 1 l)

Product	Volume [ml]
Methanol	225
Acetic acid glacial	50

- Complete to 1 l with demineralised water

2.4.3 Protein quantification with Bradford assay

The **Bradford Protein Assay**, is a spectroscopic analytical procedure used to measure the protein concentration in a solution. The Bradford assay, a colorimetric protein assay, is based on the absorbance shift of the Coomassie dye. Proteins bind to the green and red Coomassie and forms the blue Coomassie form. The blue form have a maximal absorbance spectrum at 595 nm. The absorbance increase, at 595 nm, is proportional to the amount of bounded dye, and thus to the protein amount (concentration) in the sample. The analysis has been done with the following protocol :

1. Build a calibration curve with BSA at known concentration on the desired scale (generally a scale between 0 and 200 µg/ml). To do that, add 200 µl of the sample to 800 µl of Bradford reagent, mix, wait 5 min. and read the OD_{595nm}.
2. Add 200 µl of the protein sample to analysis to 800 µl of Bradford reagent, mix, wait 5 min. and read the OD_{595nm}.

If the OD_{595nm} is out of range (c.a. OD_{595nm} > 2), adjust the concentration of the sample by dilution with demineralised H_20 or concentration

3. To know the protein concentration in the sample, defer the OD_{595nm} obtained on the calibration curve previously built

Bradford reagent (for 1 l)

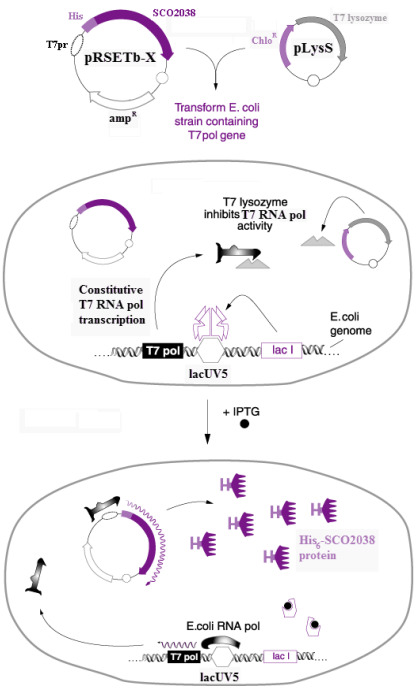
Product	Volume [ml]
Coomassie Blue G250	0.1
Ethanol 95 %	50
Phosphoric acid 85 %	100

- Complete to 1 l with demineralised water
- Filter before use with Whatman no.1 filter paper (1001-325, Whatman International Ltd, Oxon, UK)

2.4.4 Overexpression of His_6-TrpX protein in *E. coli* BL21 pRSETB-X

2.4.4.1 Overexpression of recombinant protein in *E. coli* BL21

E. coli BL21 has a T7 bacteriophage RNA polymerase-based expression system for recombinant protein. The plasmid used for the overexpression has been the pRSETB containing the coding sequence *trpX* (pRSETB-X), for the recombinant protein His_6-TrpX, downstream of the T7 promoter (*T7pr*). The *E. coli* strain used contains a genomic copy of the T7 RNA polymerase (*T7pol*) under the control of the *lacUV5* promoter. There is also a T7 lysozyme encoding plasmid (pLysS), which provides a means to inhibit the small amount of T7 RNA polymerase that is expressed constitutively, in absence of IPTG. pLysS contains also a chloramphenicol-resistance ($chlo^R$),



High expression level of the His_6-TrpX protein is obtained by blocking *lacI* repressor function by the IPTG addition to the media. When IPTG is added to the medium, the inhibitor, synthesized by *lacI*, binds to the IPTG and cannot fix any more on the *lacUV5* promoter. The RNA polymerase produced naturally by *E. coli* (*E. coli* RNA pol), can transcribe the gene encoding for the T7 RNA polymerase. The T7 RNA polymerase binds to the T7 promoter (*T7pr*), the gene *trpX* is transcribed and then translated in the His_6-TrpX protein (fig. 13).

Figure 13 Schematization of genes for tryptophan biosynthesis in *S. coelicolor*.

2.4.4.2 Preparation of inocula and induction with IPTG

Pre-inocula have been prepared with *E. coli* BL21 w.t., pRSETB and pRSETB-X strains in 3ml LB medium containing 0.1 g/l ampicillin and 0.025 g/l chloramphenicol. Cultures have been put over-night at 37°C.

For each culture, 2 ml of pre-inoculum has been added to 50 ml LB medium containing 0.1 g/l ampicillin and 0.025 g/l chloramphenicol. The culture has been put at 37°C until a $OD_{590nm} \sim 0.4$.

Cultures has been aliquoted in 5x10ml. The overexpression induction of the recombinant protein has been made, for culture with *E. coli* pRSETB-X. Three aliquots have been induced with 1 mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG). Two last aliquots, of this culture, have been not induced. All aliquots have been cultivated at 30°C during 3h. Afterwards, samples have been spin at 4000 g during 10 min. The supernatant (the culture medium) has been discarded. If pellets must be stored, freeze pellet at -80°C.

2.4.5 Purification of His_6-TrpX protein

The cell lysis and purification of the His_6-TrpX protein has been carried out using the kit Ni-NTA Purification System (K950-01, Invitrogen, Carlsbad CA, USA). The Ni-NTA kit use a chromatographic column, containing an agarose resin, where the nitriltriactetic acid (NTA) bind Ni^{2+} ions through to four coordination sites. The purification of His-tagged protein is possible because Ni^{2+} ions (bound to the matrix) also bind imidazolic rings of the histidine of the fusion protein. The purification is carried out in three steps, the step of connection between the matrix and the his-tagged protein, the washing step and finally the elution step (fig. 14).

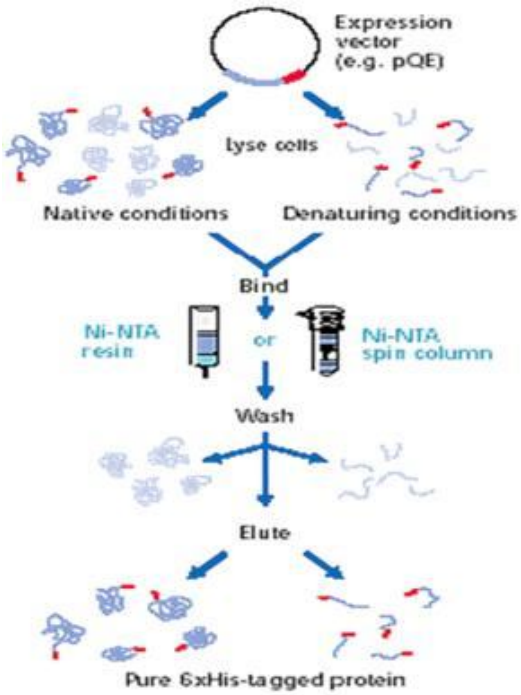


Figure 14 Schematization of genes for tryptophan biosynthesis in *S. coelicolor*.

The manipulation has been carried out following the protocol provided by manufacturer. The differences between the manipulation carried out and the original protocol have been the use of a 15 ml Falcon tube, instead of a purification column and the elution of proteins made in a different way.

2.4.5.1 Elution of proteins in native conditions

Elution in native conditions has been not carried out in a continuous, but in a discontinuous way. Elution has been carried out in 5 steps, adding every time 1.5 ml of native elution buffer pH 8.0, respectively containing 50,100,150,200 and 250 mM Imidazol (from stock solution 3M pH 6.0), to the purification matrix, mix and eluate recovery.

2.4.5.2 Elution of proteins in denaturing conditions

Elution in denaturing conditions has not been carried out in a continuous, but in a discontinuous way. Elution has been carried out in 4 steps, by addition each time of 1 ml of denaturing elution buffer pH 4.0, to the purification matrix, mix and eluate recovery.

2.4.6 Electrophoretic mobility shift assay (EMSA)

An electrophoretic mobility shift assay (EMSA), also referred as a gel shift assay, gel mobility shift assay, band shift assay, or gel retardation assay, is a common technique used to study protein-DNA or protein-RNA interactions *in-vitro*.
An EMSA generally involves a electrophoretic separation of a protein-DNA or protein-RNA mixture, on a polyacrylamide or agarose gel. The rate at which different molecules (and combinations thereof) move through the gel is determined by their size and charge, and to a lesser extent, their shape. The free (unbound) DNA probe is more fast, than a protein-DNA complex. It then appears a shift-up on the gel. For avoid bounds of aspecific proteins with the labelled probe, an unlabelled oligonucleotides are added, in excess, to the mix (fig. 15).
The EMSA gel is non denaturing for proteins (native conditions). For visualization purposes, the nucleic acid fragment is usually labelled with a radioactive, fluorescent or biotin label.

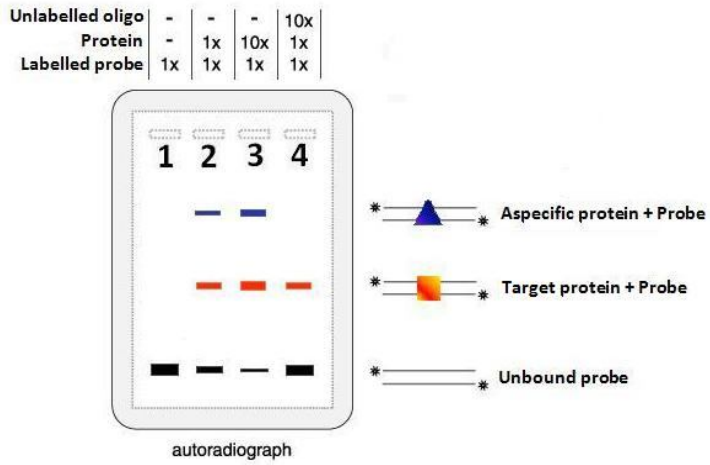


Figure 15 Schematization of a EMSA result.

1 : Control line (free labelled probe) ; **2 and 3** : Mix of proteins with labelled probe ; **4** : Mix of proteins with labelled probe and unlabelled oligonucleotides

2.4.6.1 Preparation of an EMSA

After preparation and casting of the EMSA gel, it must be left over night at room temperature, to guarantee a complete gel polymerization.

Non denaturing EMSA Gel 5% (for 50 ml)

Product	Volume [ml]
H_20 demineralised	36.3
Acrylamide Mix 30%	8.3
TBE 5X	5
Ammonium persulphate 10%	0.4
TEMED	0.02

TBE buffer 5X (Tris-Borate-EDTA)

54 g Tris BASE, **27.5 g** Boric acid and **32 ml** EDTA (0.5 M, pH8.0)

- Complete to 1 L with demineralised water
- Adjust the pH to 8.3

2.4.6.1.1 Preparation of protein extracts

The crude extracts has been obtained by lysis of cells in the CRACK buffer, cell pellets have been ressuspended in 2.5 ml of CRACK buffer and lysis has been made by sonication, with 10 cycles of 15s impulses followed by 15s on ice. Lysates obtained have been centrifuged 10 min at 3000 g and the supernatant has been recovered. The supernatant has been then dialysed, in order to desalt it, in a dialysis tube (Neflex membrane 45, 358-263-610, Union Carbide Corporation, Chicago IL, USA), which have been put in a demineralised water bath for 2h at 4°C. After 2h, demineralised water has been changed and the second dialysis has been carried out over night at 4°C. A third and last dialysis has been made for 2h at 4°C in fresh demineralised water. Thereafter, samples in the dialysis tubes have been recovered.
Proteins have been quantified by Bradford Assay and for all samples, a 100 µl aliquot containing 30 µg/ml proteins has been prepared.

CRACK Buffer

10 mM Tris HCl, **0.5 mM** EDTA pH 8 and **0.3 mM** DTT

2.4.6.1.2 Preparation of the labelled probe

The targeted DNA, on which to carry out the analysis, is amplified and labelled by PCR, with the radioactive, EasyTides® Deoxycytidine 5'.triphosphate [α-^{32}P]-(NEG513Z250UC, PerkinElmer, Boston MA, USA). Master mix composition and PCR conditions are shown thereafter :

Master mix (for 1 sample)

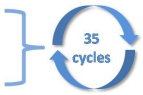
Product	Concentration	Volume [µl]
Buffer Taq	10X	5
Magnesium chloride ($MgCl_2$)	50 mM	2
dNTP	10 mM	1
Primer forward	-	1
Primer reverse	-	1
Taq Polymerase	5 U/µl	0.5
DMSO	-	5
[α-^{32}P]-dCTP	222 TBq/mmol	3
H_2O demineralised	-	30.5
Final volume	**-**	**49**

➔ Add 1 µl of *S. coelicolor* genome DNA

Temperature and times of PCR steps

5 min at 95°C → Initial denaturation

N°	Temp. [°C]	Time [s]	N°	Temp. [°C]	Time [s]	N°	Temp. [°C]	Time [s]
1	95°C	45	7	95°C	45	13	95°C	45
2	72°C	45	8	68°C	45	14	64°C	45
3	72°C	45	9	72°C	45	15	72°C	45
4	95°C	45	10	95°C	45	**16**	**95°C**	**60**
5	70°C	45	11	66°C	45	**17**	**62°C**	**60**
6	72°C	45	12	72°C	45	**18**	**72°C**	**60**



To eliminate not used radioactive and primers in the PCR products, a G25 column (illustra™ MicroSpin™ G-25 Columns, 27-5325-01, GE Healthcare, Bukinghamshire, UK) has been used, following the manufacturer's protocol.

2.4.6.1.3 Preparation of samples

To allow protein-DNA bounds, a binding reaction mix has been carried out, according to the following composition (for 1 sample) :

Product	Volume [µl]
$MgCl_2$ 0.1M	1
Glycerol 50%	1
10X BB	2
Salmon Sperm (SS)	1
H_20 demineralised	3

10X BB

Product	Concentration
Tris HCl pH 7.5	12.5 mM
Glycerol	10%
Potassium Chloride (KCl)	62.5 mM
Dithiothreitol (DTT)	0.75 mM

- Salmon sperm, D-1626, Sigma, St-Louis, USA.

10 µl of proteins extracted, at a concentration of 30 µg/ml (determined by Bradford Assay) has been added to the reaction mix and samples have been kept 15 min on ice. Afterwards, 2 µl of the radioactive labelled probe has been added to each samples, for a final volume of 20 µl and a sample radioactivity emission between 30 $Bq \cdot cm^{-2}$ and 80 $Bq \cdot cm^{-2}$ (if the radioactivity of the sample are too much or not sufficiency, it's possible to adjust the volume of labelled probe, adjusting the volume of H_20 demineralised into the initial reaction mix). Sample have been kept on ice during 20 min. After this step, 1 µl of Dye has been added to each samples, to show the progress of the electrophoretic run. Now samples are ready to use for the electroporetic run.

2.4.6.2 EMSA gel electrophoretic run

The EMSA gel electrophoretic run has been carried out using as running buffer, TBE 0.5x. Before the electrophoresis run, a pre-run during 20 min, without sample charged, has been carried out, to permit a good repartition of the running buffer (TBE 0.5X) and a uniform composition of the EMSA gel.
Afterwards samples have been charged on the gel :

- **21 µl** of protein samples prepared as shown previously
- The labelled probe :

  16 µl H20 demineralised + **2 µl** glycerol 50% + **2 µl** labelled probe + **1 µl** Dye

The gel run has been made at 40 mA.

2.4.6.3 Revelation of the EMSA gel

After the EMSA gel run, the gel has been removed from the support. For that, glass plates framing have been separated and a Whatman 3MM paper has been applied on the gel. It thus remains "stuck" to the Whatman 3MM paper and can now be moved. The gel then has been dried in a vacuum gel dryer (GD2000, 80-6428-46, Amersham Pharmacia Biotech AB, San Francisco CA, USA) at 60°C during 1h, directly on the sheet of Whatman 3MM paper which has been used to remove it from the glass support. Thereafter, the Whatman 3MM paper has been fixed, using adhesive tape, in a leaded frame and a radiosensitive film (Amersham Hyperfilm™ MP, 28906850, GE Healthcare, Buckinghamshire, UK) has been placed on the top. The exposure is done at -80°C during a variable time, according to the radioactivity emission.
After exposure of radiosensitive film, it has been developed in a development bath until clear appearance of various bands and then wash in a water bath. Afterwards, it has been put 15 min in a fixing bath. A final washing in a water bath has been finally carried out. The radiosensitive film has been dried at 37°C.

- Development solution : Kodak® Processing chemicals for autoradiography films, GBX developer/replenisher, P7042, Sigma-Aldrich, St. Louis MO, USA.
- Fixer solution : Kodak® Processing chemicals for autoradiography films, GBX fixer/replenisher, P7167, Sigma-Aldrich, St. Louis MO, USA.

2.5 RNA methods

2.5.1 RNA extraction from *E. coli*

2.5.1.1 RNA extraction from *E. coli* with hot SDS/Phenol protocol

1. Warm 1 volume (based on final volume of cells, media and lysis solution) of phenol/chloroform to 65°C
2. When culture has reached desired O.D., bring 0.5 volume of SDS lysis solution (i.e. 5ml of lysis solution for 10ml of culture) to 100°C in a Falcon tube for 5 minutes
3. Quickly add the appropriate volume of cells (typically 10-15ml) directly into boiling lysis solution
   Instant lysis of cells using hot SDS quickly inactivates endogenous RNases. This can be very important due to the extremely short half-life of many E. coli mRNAs and is of particular importance when time-course studies are performing
4. Keep at 100°C for 5 min. with occasional mixing
5. Pour this sample directly into 65°C phenol/chloroform seal and parafilm cap (to avoid phenol escaping tube during mixing) and mix well by vortexing at high speed.
6. Keep at 65°C for 10 min. with periodic vortexing
7. Spin the sample at 2500 g during 15 min.
8. Carefully transfer aqueous phase (avoiding any of the white interface) to fresh 50ml tube and add an equal volume of phenol/chloroform
9. Seal cap with parafilm and mix well by vortexing at high speed
10. Spin at 2500 g during 15 min.
11. Carefully transfer aqueous phase to fresh 50 ml Falcon tube and add an equal volume of chloroform/isoamyl alcohol
12. Seal cap and mix well by vortexing
13. Spin at 2500 g during 15 min. and put aqueous phase in a 30 ml corex tube and add an equal volume of isopropanol
14. Keep at -20°C at least 2 hrs or overnight
15. Spin down RNA at 15-20000 g during 20 min. at room temperature
16. Dry the tube in a Speed Vacuum for 30 min

It is possible air dry the tube on a piece of paper at room conditions

<u>SDS lysis solution :</u>

2% SDS and **16 mM** EDTA pH 8

- **Optional** : add 200 mM NaCl if cells are grown in low salt medium like LB Medium
- Sterilise at 121°C for 20 min.

2.5.1.2 DNase treatment of RNA extracts

The DNase treatment of total RNA extracts is critical remove contaminating genome or plasmid DNA from RNA samples. In fact, genome or plasmid DNA can be used as template by Taq polymerase during a RT-PCR or a Real-Time PCR.
Solutions used for the DNase treatment are provided with the kit DNA-*free*™ (1906, Ambion INC., Austin TX, USA). The reaction mix was the following :

180 µl RNA, **20 µl** 10X DNase 1 Buffer and **2 µl** DNAse

The reaction mix has been keep at 37°C during 2h30. After this step, RNA extracts have been purified, from DNAse enzyme, by phenol-chloroform extraction.

2.5.2 Real-Time RT-PCR

The real-time polymerase chain reaction (Real-Time PCR), also called quantitative real time polymerase chain reaction (qRT-PCR) or kinetic polymerase chain reaction, is a laboratory technique based on polymerase chain reaction, which is used to amplify and simultaneously quantify a targeted DNA molecule. It enables both detection and quantification (as absolute number of copies or relative amount when normalized to DNA input or additional normalizing genes) of a specific sequence in a DNA sample.
The procedure follows the general principle of polymerase chain reaction; its key feature is that the amplified DNA is quantified as it accumulates in the reaction in *real time* after each amplification cycle. Two common methods of quantification are the use of fluorescent dyes that intercalate with dsDNA, and modified DNA oligonucleotide probes that fluoresce when hybridized with a complementary DNA.
Frequently, real-time polymerase chain reaction is combined with reverse transcription polymerase chain reaction to quantify low abundance mRNA, enabling a researcher to quantify relative gene expression at a particular time, or in a particular cell or tissue type.

2.5.2.1 cDNA preparation from RNA extracts

RNA extracts have been reverse transcripted in DNA, for easiest sample manipulations and better storage. cDNAs have been obtained by PCR, using the High Capacity cDNA Reverse Transcription Kit (P/N 4368814, Applied Biosystems, Foster City CA, USA). The PCR has been carried out by 10 min at 25°C (annealing), 120 min at 37°C (DNA extension) and 5 min at 85°C (RT enzyme inactivation), using the thermocycler Tpersonal® (050-550, Biometra, Goettingen, Germany), in the following master mix :

Product	Volume [µl]
Buffer 10X	5
dNTP 10 mM	8.3
Random primers	5
Reverse transcriptase	2.5
RNA inhibitor	2.5
RNA 300 ng/µl	2
Demineralised water	31

2.5.2.2 Real-Time RT-PCR

The Real-Time PCR has been carried out using *Power* SYBR® Green PCR Master Mix (P/N 4309155, Applied Biosystems, Foster City CA, USA), in the following master mix :

Product	Volume [µl]
Primer forward 100 pM	2.5
Primer reverse	2.5
SYBR® Green	12.5
cDNA 5 ng/µl	2.5
Demineralised water	2.5

The Real-Time PCR has been made in the 7300 Real Time PCR system (4351103, Applied Biosystem, Foster City CA, USA).
The thermic profile used begin with a hot start (initial denaturation) at 95°C during 10 min, and 40 cycles of 15s at 95°C (denaturation) and 1 min at 60°C (primers annealing and DNA extension).
The analysis of amplified DNA fragments is based on the fluorescence variation in each samples. The fluorescence detected correspond to the SYBR® Green emitted fluorescence. The SYBR® Green bind stoichiometrically to dsDNA only, and it emit fluorescence only after bound to nucleic acid.
The detector subtracts the passive emitted fluorescence (Baseline) in the sample, from the total fluorescence detected. The fluorescence augmentation (ΔRn), correlated to amplified fragment amount, in function of amplification cycles represents the “amplification plot”, a logarithmic curve.
Amplified sample analysis is made using the Sequence Detection Software (SDS), Version 1.2.3 (7000 System SDS Software, Applied Biosystems, Foster City CA, USA).

3 Aims of this work

3.1 General purpose of the work

3.1.1 The increase of antibiotic production in *S. coelicolor* A3(2)

The emergence of several pathogenic bacteria and the adaptation of part of those resistant to antibiotics currently used, in clinically treatments, needs the research for new antibiotic development. ActinoGEN is an European project for the discovery of new antibiotics and the construction of a "Superhost" for their industrial scale production. Organisms, on which this research is based, are the actinomycetes, and more specifically *Streptomyces* genus. Actinomycetes are known to produce approximately 6000 different antibiotics (half of those currently known).

The *Streptomyces coelicolor* genome having been completely sequenced (Bentley SD, 2002), allowed the identification of many genes involved in the secondary metabolite biosynthesis. This allowed the discovery as well as the regulation knowledge of various biosynthetic clusters for secondary metabolites. In particular, the CDA biosynthetic cluster in *Streptomyces coelicolor* has been well studied and represent about 1% of total genome. (Hojati Z, 2002).

The CDA is one of a natural peptide family produced by non-ribosomal way, like the cyclosporin (immunosupressor), the bleomycin (antitumor) and the vancomycin (antibiotic). Antibiotics of this family are used against gram-positive pathogenic bacteria resistant to common antibiotic treatments.

The CDA cluster contains some genes involved in the tryptophan pathway (*trpC2*, *trpD2*, *trpE2* and *trpG*), tryptophan being a CDA precursor (Kim HB, 2004). Two strategies, aimed to increase the CDA production, are the overexpression of genes involved in the tryptophan biosynthesis and the mutation of genes involved in its catabolism. These two strategies could allow an accumulation of tryptophan in the cell, carrying to a major production of CDA (fig. 16).

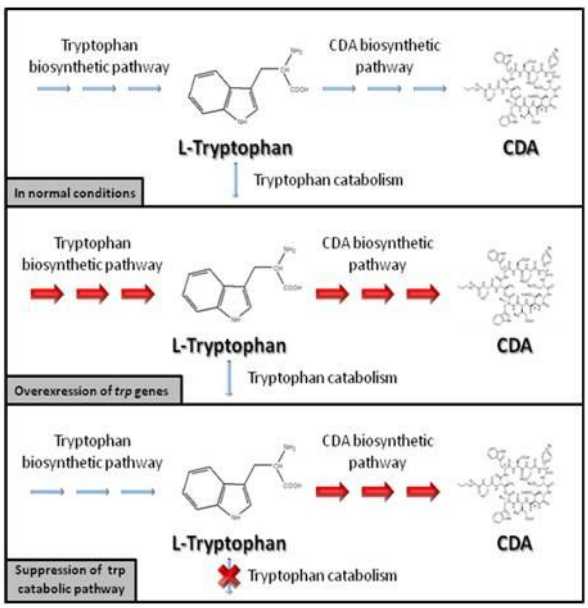


Figure 16 Strategies to increase CDA production in *S. coelicolor*.

3.1.2 The *trpX* orf in *S. coelicolor* A3(2)

One of fundamental operons for the tryptophan biosynthesis is the *trpA/B/X/C1* operon (fig. 17).

Figure 17 *trpA/B/X/C1* operon in *S. coelicolor*.

This cluster contains the **small *trpX* orf, having to date any known function**. The *trpX* orf (also called *SCO2038*), codes for a hypothetical 63 amino acids protein, with a mass weight of 7167 Daltons and a pI of 11.754. The small orf presence among important genes implied in a amino acids biosynthesis does not constitute an isolated case. Indeed, small orfs are present also between *proA* and *proB* genes, involved in the proline synthesis (Hood DW, 1992), and in principal operons involved in the histidine synthesis (Limauro D, 1990). Moreover, to date, the role of these small orf remains still unknown. This shows the importance of elucidating the *trpX* gene role. This will imply a better comprehension of the amino acids synthesis regulation in the *S. coelicolor*, and more mainly in actinomycetes, still poorly understood.

Previous experiments showed that the *S. coelicolor* A3(2) ΔX mutant is not able to grow on a minimal medium (SMMS), whereas it grows on this same medium after tryptophan addition. The mutant *S. coelicolor* A3(2) ΔX/X grows on a minimal medium, like the strain *S. coelicolor* A3(2) M145 (w.t.). This effect cannot be explained by a simple polar effect, due to the *trpX* gene interruption, which does not allow the *trpB* gene expression. Indeed, it is possible to complement the *trpX* mutation with the pKCHyg796trpX plasmid, containing a w.t. copy of this gene.

These results show that *trpX* gene is essential for the *S. coelicolor* growth in a medium where tryptophan lacks, and thus it plays a key role in the tryptophan biosynthesis.

Real-Time RT-PCR, of the global expression of *S. coelicolor* A3(2) genes, has shown that *trpA* gene is expressed at the same level, in *S. coelicolor* M145 and *trpX* mutants, indicating that this gene is not influenced or regulated by *trpX* gene.

Microarrays analysis, confirmed by Real-Time RT-PCR, has shown that the *trpB* gene is less expressed in *S. coelicolor* ΔX and ΔX/pKC mutants, than in M145 strain. When *trpX* mutation is complemented with pKCHyg796trpX plasmid, the *trpB* shows normal expression level.

These results suggest that the *trpB* gene is regulated by *trpX*, and that the *trpB-trpX* intergenic region could contain a promoter allowing *trpA* and *trpB* expression.

3.2 Specific objective of this work

The specific objective of this work is to start the study of the *trpX* gene role in *Streptomyces coelicolor*. In particular to:

1) study the *trpX* expression in the strain *S. coelicolor* A3(2) M145, in various growth phases and media by qRT-PCR.

2) perform in silico analysis of protein structure using the Robetta server.

3) carry out EMSA of intergenic regions of *trpA/B/X/C1* operon with crude extract of wt and mutant strains.

4) overexpress and purify TrpX protein from *E. coli*.

5) overexpress the *trpX* gene in *S. coelicolor*.

4 RESULTS

4.1 Expression of *trpA*, *trpB* and *trpX* in *S. coelicolor* A3(2) M145

The expression analysis of *trpA*, *trpB* and *trpX* genes, in *S. coelicolor* A3(2) M145 have been carried out by Real-Time RT-PCR, from total RNA extracts. RNAs have been extracted from cultures grown in R5 medium, in the RG1 and RG2 growth phases and in SMMS medium, in RG, T and S phases (in SMMS medium RG2 phase is absent) (fig. 18).

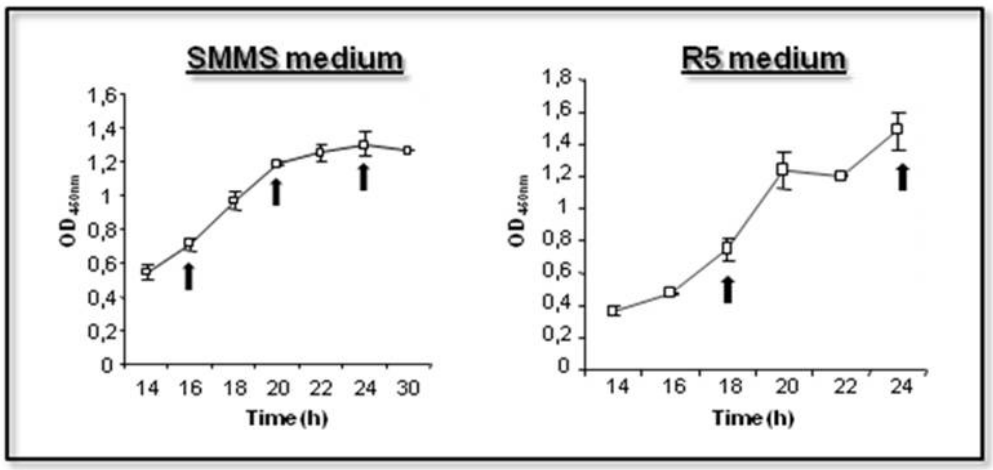


Figure 18 *S. coelicolor* growth curve in SMMS and R5 medium. Black arrows indicate the time of global RNA extraction.

RNA extracts, after DNA removing, using DNase, and purification, by phenol-chloroform protocol, have been reverse transcribed in cDNA, using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems).
The Real-Time RT-PCR has been carried for *trpA*, *trpB* and *trpX* genes, using three primer pairs (table 2).

Table 2 Real-Time PCR primers

Targeted gene	Primer direction	Primer sequences
trpA	Forward	5'-TCCGCGAGGCACACGAGGC-3'
	Reverse	5'-CGGGCAGGTCGGGCAGGAT-3'
trpB	Forward	5'-GCGCGAGGACCTGAACCACAC-3
	Reverse	5'-GGCGGCGGGTGTCGATCT-3
trpX	Forward	5'-CGCTCCCGCTCGTCCC-3
	Reverse	5'-CCTGATGGGGCGCTTGA-3

The qRT-PCR has been carried out using absolute quantification of the transcripts, and the quantification of targeted genes in samples has been obtained using 5 standards, containing 10^2, 10^3, 10^4, 10^5 and 10^6 molecules of *S. coelicolor* chromosome DNA respectively. The cDNA concentration, in each analysed sample was 5 ng/µl.
Real-Time RT-PCR results, analysed, using the Sequence Detection Software (SDS), Version 1.2.3 (Applied Biosystems), are shown in fig. 19.

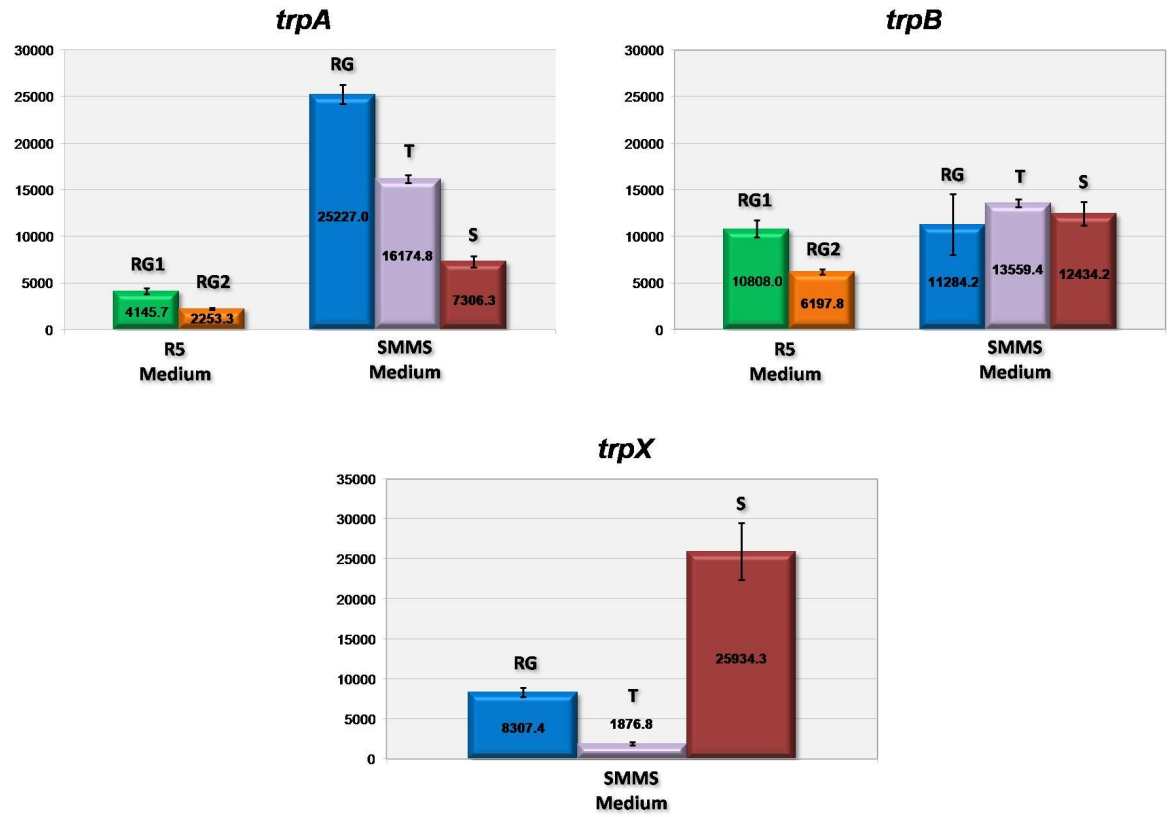


Figure 19 Expression level of *trpA, trpB and trpX* genes of *S. coelicolor* A3(2) M145, in R5 medium, during the RG1 and RG2 growth phases, and in SMMS medium, during RG, T and S phases, quantified by qRT-PCR and analysed using Sequence Detection Software (SDS), Version 1.2.3 (Applied Biosystems).

The *trpA* expression level decreases during the *S. coelicolor* A3(2) growth, independently of the culture medium used. Nevertheless, it seems sensible to culture medium used. In fact, it is less expressed in R5 than in SMMS medium.
In R5 medium, the *trpB* expression level decreases during the *S. coelicolor* A3(2) growth, but it remains constant in SMMS medium. The *trpB* expression level is not influenced by the medium used.
The expression level of *trpX* gene is showed only for SMMS medium, because of an aspecific fragment amplified in R5 medium samples. *trpX* is more expressed in rapid growth (RG) and stationary (S) phases than in transition (T) phase.

4.2 Tridimensional structure prediction of TrpX with Robetta server

The TrpX amino acidic sequence has been submitted to the Robetta server (http://robetta.bakerlab.org), to have its structure predictions and thus putative functions.
The amino acid sequence submitted was:

MTLPLVPARDPYARLARGCRPRGCRAPARRVHGRRVRYVIGDEPGQVNGMRWLKRPIRGAGLY*

Several protein structure predictions have been found and for each prediction, several known protein models have been proposed. Robetta server results are shown, in fig. 20.

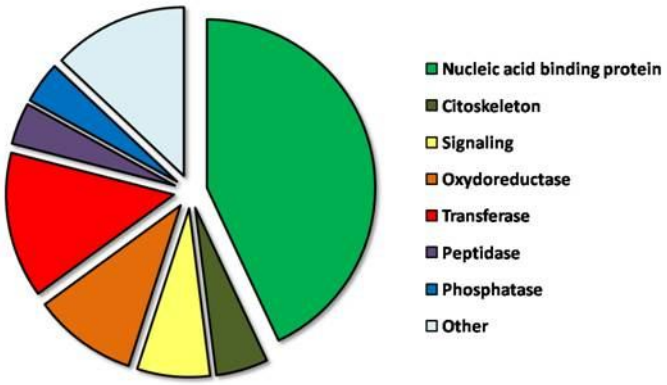


Figure 20 Representation of TrpX model families, obtained by the Robetta server.

Robetta server protein structure prediction show clearly that a nucleic acid binding protein is the most probable model family for TrpX, with 45% of total models obtained. Informatic structure prediction cannot be considered as reliable, and thus this gives simply some TrpX hypothesis function in *S. coelicolor*.

4.3 Gel mobility shift assays of ^{32}P-end-labeled fragments spanning the intergenic region of *trpA/B/X/C1* operon.

The TrpX structure prediction, using Robetta server, suggested a DNA-binding role.
Study on global gene expression carried out by qRT-PCR, in *S. coelicolor* A3(2), using *trpX* mutants, suggested that *trpX* could be involved in the *trpB* gene regulation and also that *trpB-trpX* intergenic region could contain a promoter of *trpA/B* gene.
Study on the *trpA/B/X/C1* operon have evidenced that *trpX-trpC1* intergenic region is a promoter of this operon (Hu DSJ, 1999).
Gel mobility shift assay have been carried out, on *trpB-trpX* and *trpX-trpC1* intergenic regions (fig. 21), using crude extracts of *S. coelicolor* A3(2) wt and *trpX* mutant strains.

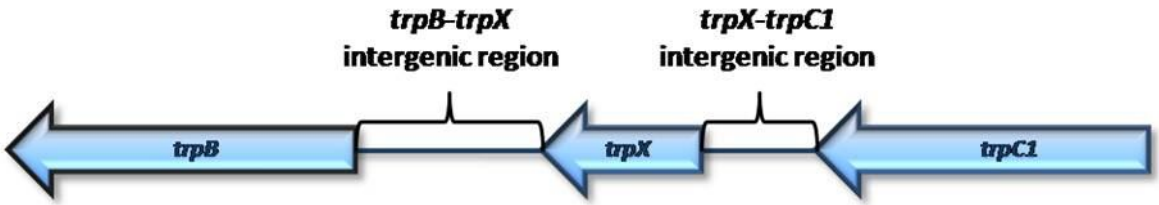
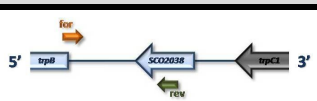


Figure 21 Schematic representation of *trpB/X/C1* intergenic regions.

Protein extracts have been prepared from *S. coelicolor* A3(2) M145, ΔX, ΔX/X and ΔX/pKC strains grown in SMMS and SpMR media. Two samples have been collected, after 16h and 42h of growth. *S. coelicolor* cells have been harvested and resuspended in 2.5 ml of CRACK buffer, and cell lysis has beencarried out using a sonicator. Proteins have been separated from cell debris, by centrifugation and recovery of the supernatant. Then, to desalt protein extracts, a dialysis of samples has been carried out.
Proteins in solution have been quantified by the Bradford assay, and 100 µl aliquots containing 30 µg/ml proteins has been prepared, for each strain.
Radioactive labelled probes have been prepared by PCR amplification of *trpB-trpX* and *trpX-trpC1* intergenic regions, from *S. coelicolor* M145 genome DNA, using primer pairs shown in table 3:

Table 3 Primers used for PCR PCR amplification of *trpB-trpX* and *trpX-trpC1* intergenic regions

Targeted region	Primer direction	Primer sequences	Primers bound schematization
trpB-trpX	Forward	5'-GGATCGGGGATGAAGAAGTT-3'	
	Reverse	5'-GGTACGTCATCGGAGACGAG-3'	
trpX-trpC1	Forward	5'-CTCGTCTCCGATGACGTACC-3	
	Reverse	5'-GTCGGCGAGTCCCTGGTC-3'	

For each amplification, three samples have been prepared, the negative control without DNA (by addition of demineralised water), two samples with *S. coelicolor* genome DNA, with and without dC^{32}TP). Non radioactive samples, after PCR, have been analysed by gel electrophoresis 1% (fig. 22).

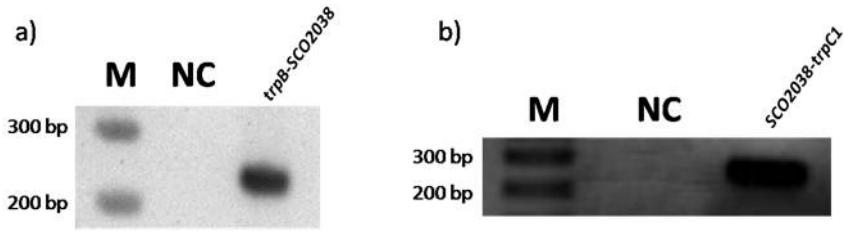


Figure 22 PCR amplification of *trpB-trpX* (a) and *trpX-trpC1* (b) intergenic regions. The negative control (NC) corresponds to PCR master mix without DNA template. The size of analysed DNA fragments has been determined using the Fermentas O'GeneRuler™ DNA Ladder Mix (M).

Amplified DNA fragments have been extracted from the gel using the illustra GFX™ PCR DNA and Gel Band Purification Kit (28903466, Ge Healthcare, Buckinghamshire, UK). Fragments extracted have been sequenced by BMR Genomics, Padova, Italy using same primers used for PCR amplification. Chromatograms received have been analysed using Sequence Scanner V1.0 (Applied Biosystems). Sequenced fragment sequences have been aligned on the *S. coelicolor* M145 sequences, using the ClustalX V2.0.3 software (fig. 23).

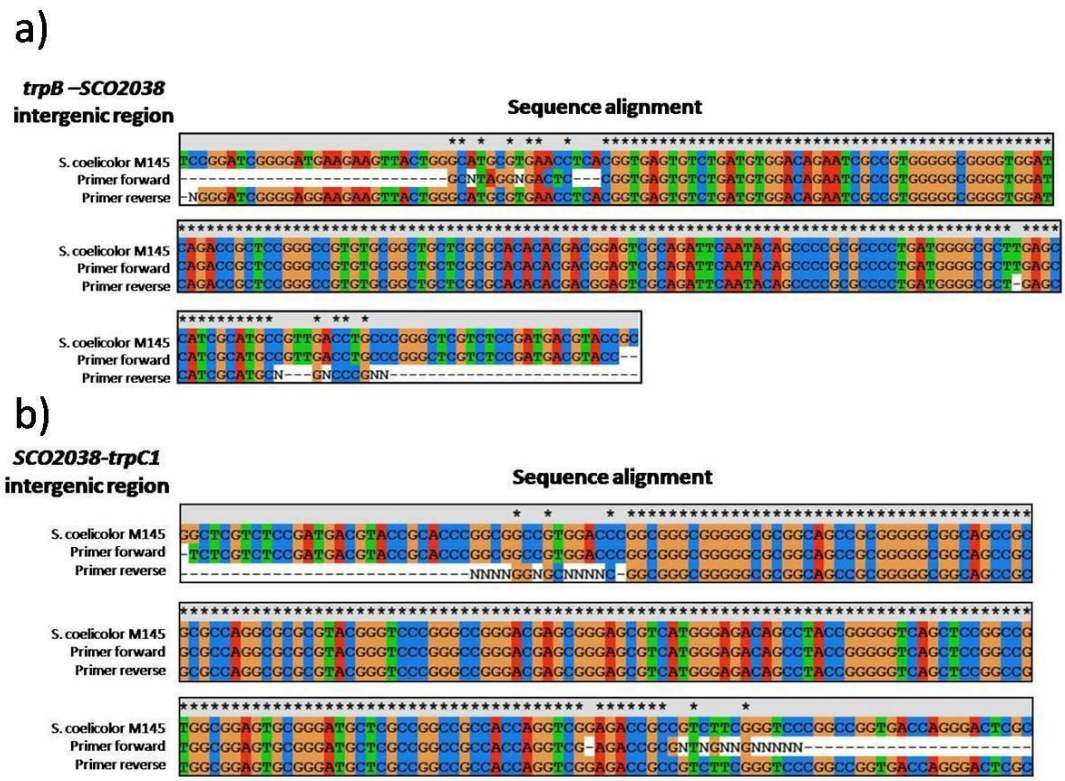


Figure 23 Sequence alignment, using the ClustalX 2.0.3 software, of sequenced *trpB-trpX* intergenic region amplified by PCR (a) and sequenced *trpX-trpC1* intergenic region amplified by PCR (b). Sequenced sequences have been aligned on the *S. coelicolor* M145 sequences.

The sequence alignment shows that amplified *trpB-trpX* and *trpX-trpC1* fragments has the correct sequence (fig. 23).

Radioactive labelled probes have been purified from primers and free nucleotides through a G-25 column.
EMSA analysis of the *trpB-trpX* and *trpX-trpC1* intergenic region did not revealed any shift different from those present in the labeled probe without protein crude extract (fig. 24).

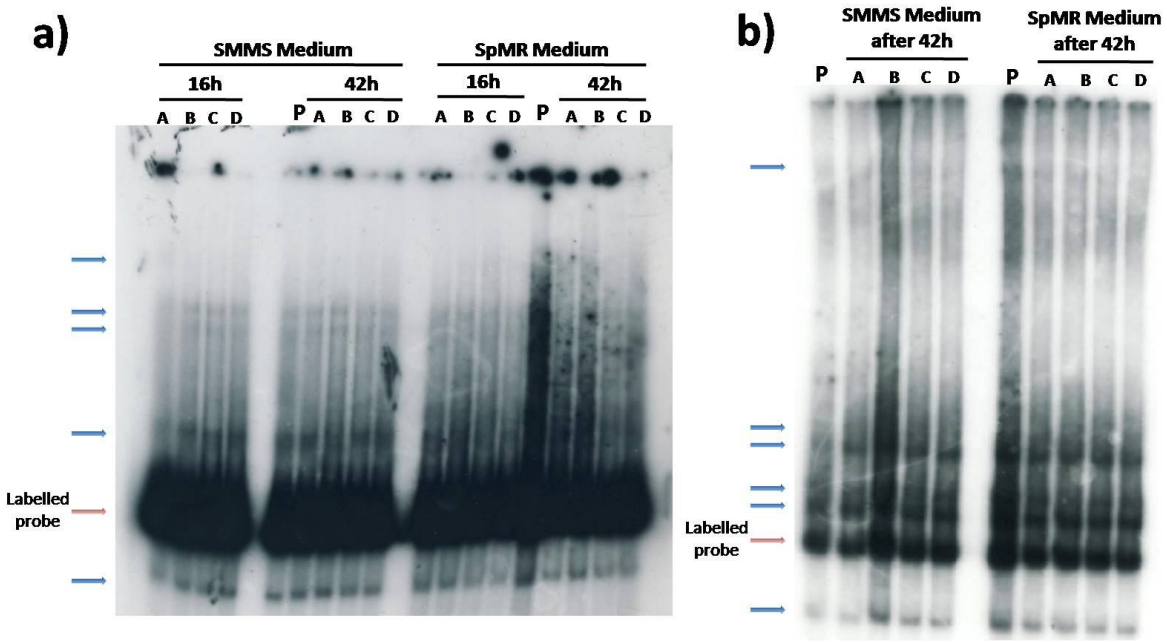


Figure 24 Gel mobility shift assay of the 32P labelled DNA fragments spanning the intergenic regions *trpB-trpX (a) and trpX-trpC1 (b)* EMSA was performed using crude extracts from strains *S. coelicolor* M145 (A), ΔX (B), ΔX/X (C) and ΔX/pKC (D) grown in SMMS or SpMR medium. P lanes correspond to free probe.

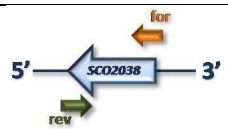
This results obtained showed that, in the used conditions, there are no proteins bound to the intergenic regions analysed.

4.4 Overexpression of TrpX in *E. coli* BL21

One objective of this work was to obtain the purified TrpX protein, in order to use it in EMSA assay.
The strategy choosed for the TrpX protein overexpression involves the use of pRSETB plasmid. The pRSETB plasmid has been constructed for the recombinant protein overexpression in *E. coli* BL21, allowing its control and a easy purification of the overexpressed protein.
The first step of this strategy is the construction of the pRSETB-X plasmid in *E. coli* DH5α, followed by the *E. coli* BL21 transformation with the constructed plasmid.

4.4.1 Cloning of *SCO2038* in pRSETB

The *trpX* gene has been amplified, from *S. coelicolor* A3(2) M145 genome, by PCR with following primer pair :

SCO2038	Forward	5'- AAAGGATCCTATGACGCTCCCGCTCGTCC -3' *Bam*HI	
	Reverse	5'- AAAAAGCTTTCAATACAGCCCCGCGCCCCT -3' *Hind*III	

This primer pairs add *Hind*III and *Bam*HI restriction sites and polyA tags at the 5'-end of the amplified fragment. PolyA tag allows an easier binding of the restriction enzyme. The PCR product has been analysed by gel electrophoresis 1% (fig. 25)

Figure 25 Gel electrophoresis 1% of PCR product (1). The size of analysed DNA fragments has been determined using the Fermentas O'GeneRuler™ DNA Ladder Mix (M).

The 210 bp amplified fragment, corresponding to *trpX* gene, has been purified, from the gel matrix, using the illustra GFX™ PCR DNA and Gel Band Purification Kit (28903466, Ge Healthcare, Buckinghamshire, UK).
The pRSETB plasmid has been extracted, from a *E. coli* BL21 pRSETB culture in LB medium, containing 0.025 g/l chloramphenicol. After extraction, it has been digested with *HindIII* and *BamHI*.
Amplified and purified *trpX* and pRSETB have been digested with *Bam*HI and *Hind*III restriction enzymes. After digestion of pRSETB, an alkaline phosphatase treatment has been carried out in order to avoid self-ligation of the vector. The *trpX* cloning in pRSETB has been carried out by ligation reaction, using an insert-vector ratio of 3:1.

4.4.2 *E. coli* DH5α transformation with pRSETB-X

E. coli DH5α electrocompetent cells have been transformed with the ligation mixture. After electroporation cells have been selected on plates containing 0.1 g/l ampicillin. Twentyfour clones have been selected for further analysis. Plasmids from these selected clones have been extracted by MiniPrep. The restriction pattern of extracted plasmids has been carried out using *Dra*I and *BamHI* restriction enzymes. Fragments has been analysed by gel electrophoresis 1% (fig. 26).

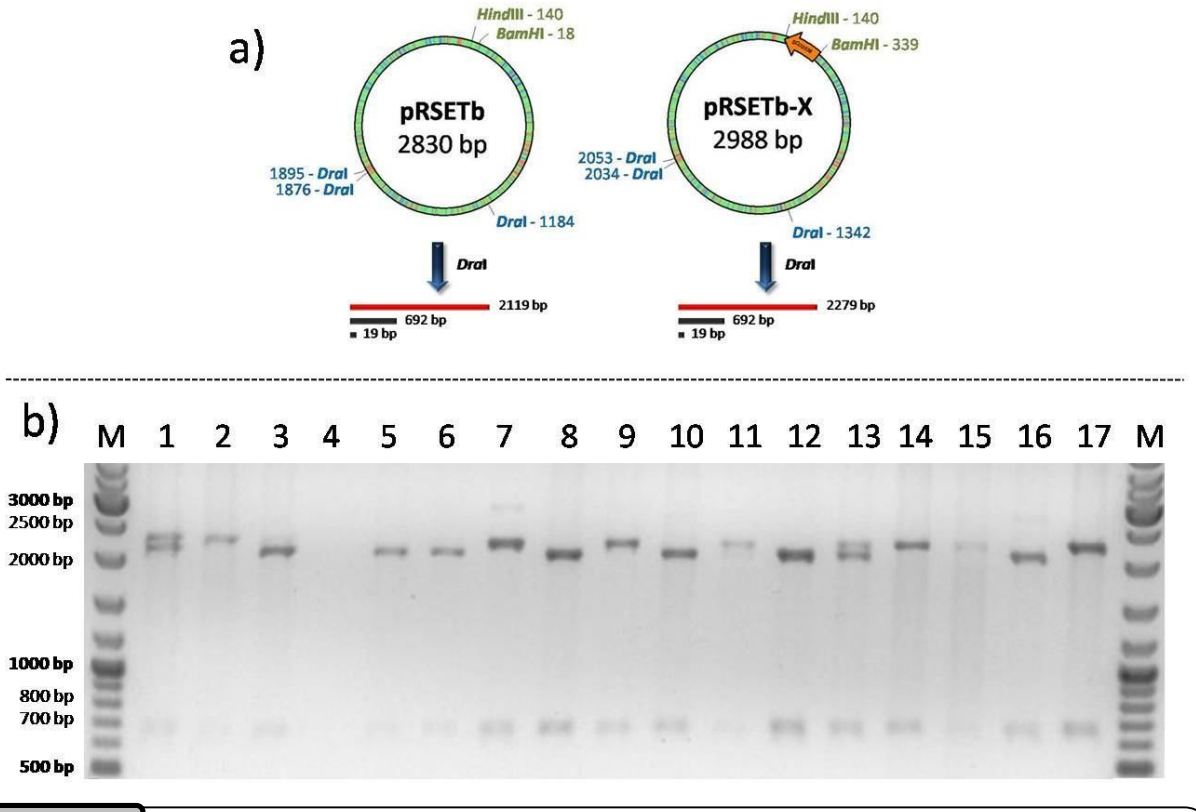


Figure 26 Expected fragments after *Dra*I digestion of pRSETB and pRSETB-X plasmids (a). Gel electrophoresis 1% of extracted plasmids after digestion with *Kpn*I (b) (lanes 1-17). The size of analysed DNA fragments has been determined using the Fermentas O'GeneRuler™ DNA Ladder Mix (M).

*Dra*I digestion of extracted plasmid, after gel electrophoresis, shows the two expected fragments (692 bp and 2119/2279 bp), the third one being not visible due to its small size (19 bp) (fig. 26). Seven *E. coli* DH5α clones, containing a plasmid with expected fragments (fig. 31, lanes 2, 7, 9, 11, 14, 15 and 17), have been selected. From this selected clones, a plasmid extraction with the GenElute™ Plasmid Miniprep Kit (Invitrogen), has been carried out. These plasmids have been sequenced by BMR Genomics, Padova, Italy, using primers which allow the amplification of DNA cloned in pRSET the *trpX* gene. Chromatograms received have been analysed using Sequence Scanner V1.0 (Applied Biosystems).

The sequence alignment shows that three out of four *E. coli* DH5α pRSETB-X clones, contain the pRSETB-X plasmid with the correct *trpX* gene sequence. Two out of three clones (called pRSETB-X2 and pRSETB-X24) has been selected to transform *E. coli* BL21 cells, after plasmid extraction.

4.4.3 Transformation of *E. coli* BL21 with pRSETB-X plasmid

E. coli BL21 electrocompetent cells have been transformed with the pRSETB-X2 and pRSETB-X24 plasmids, extracted from *E. coli* DH5α, by electroporation. Cells have been cultivated on selective plates containing 0.025 g/l chloramphenicol and 0.1 g/l ampicillin, for clone screening.

4.4.4 TrpX purification from *E. coli* cells

In order to purify His_6-TrpX, an *E. coli* BL21 pRSETB-X24 culture has been grown at 37°C until a OD_{590nm} =0.4 has been reached. Later, the culture has been splitted in three subcultures and grown

1) 37°C, 3h
2) 37°C, 3h; 0,1 mM IPTG
3) 37°C, 3h; 1 mM IPTG

Then, the medium has been removed from cells by centrifugation at 3000 g.

4.4.4.1 Cell lysis and His_6-TrpX purification in native conditions

Cells have been thereafter lysed by sonication in the native buffer (see Ni-NTA purification system, Invitrogen). A centrifugation has been carried out to separate soluble proteins from cells debris and insoluble proteins. SDS-PAGE analysis has been carried out with soluble proteins of *E. coli* BL21 pRSETB-X24 strain (fig. 27).

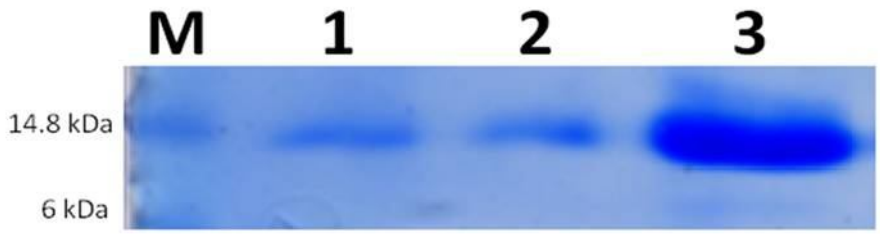


Figure 27 SDS-PAGE of native and soluble proteins extracted from *E. coli* BL21 pRSETB-X24 not induced (1-2) or induced with 1 mM IPTG (3).

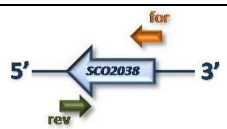
SDS-PAGE analysis of native and soluble protein extracts (fig. 34) revealed the induction of a 8-9 kDa protein. This protein is present in fact only in strain harbouring the pRSET-X24 plasmid. Purification of the protein is in progress.

4.5 *trpX* gene overexpression in *S. coelicolor*

One objective of this work was to overexpress the *trpX* gene in *S. coelicolor*, in order to study differences in the global gene level expression and antibiotic level production, between mutant and wt strains, indicating ways to understand the *trpX* gene role.
The strategy choosed for the *trpX* gene overexpression involves the use of pIJ8600 plasmid. The pIJ8600 plasmid has been constructed for the integration of a gene in *S. coelicolor* genome, followed by its overexpression.
The first step of this strategy is the construction of the pIJ8600-X plasmid in *E. coli* DH5α.

The pGEM-X plasmid has been extracted, from a *E. coli* DH5α pGEM-X culture in LB medium, containing 0.1 g/l ampicillin. The *trpX* gene has been obtained by digestion of the extracted pGEM-X plasmid with *Bam*HI and *Nde*I restriction enzymes.
The pIJ8600 plasmid has been extracted, from a *E. coli* DH5α pIJ8600 culture in LB medium, containing 0.05 g/l apramycin. After extraction, it has been digested with *BamHI* and *NdeI*. After digestion of pIJ8600, an alkaline phosphatase treatment has been carried out in order to avoid self-ligation of the vector.
The *trpX* cloning in pRSETB has been carried out by ligation reaction, using an insert-vector ratio of 3:1.

E. coli DH5α electrocompetent cells have been transformed with the ligation mixture. After electroporation cells have been selected on plates containing 0.05 g/l apramycin.
Two hundred clones have been selected for further analysis. Clones have been used in colony pool PCR, using 10 clones per PCR tubes, and amplified fragments have been analysed by gel electrophoresis. The colony pool PCR has been carried out using the following primer pair :

SCO2038	Forward	5'-AAA<u>CATATG</u>ACGCTCCCGCTCGTCC-3' *Nde*I	
	Reverse	5'-AAA<u>CCTAGG</u>TCAATACAGCCCCGCGCCCCT-3' *Bam*HI	

The gel electrophoresis of colony pool PCR products have not shown a 210 bp amplified fragment, corresponding to the amplified *trpX* gene. Clones obtained contains only the pIJ8600 without *trpX* insert.

Three other *E. coli* DH5α transformations, changing insert-vector ratio or using *E. coli* DH5α calcium-competent cells, transformed with the corresponding protocol, has been made. Clones have been screened but no positive *E. coli* DH5α pIJ8600-X clone has been found.

5 DISCUSSION

5.1 Expression of *trpX* gene in *S. coelicolor*

Real-Time RT-PCR analysis of *S. coelicolor* M145 gene expression, carried out during this work, showed that *trpX* is not costitutively expressed during the growth in SMMS medium (minimal medium). In fact, this gene is less expressed in the transition phase than in rapid growth and stationary phases. These results are in agreement with those reported by Hodgson DA, 2000. In fact the expression of many *S. coelicolor* genes involved in amino acid biosynthesis is growth-phase dependent.

5.2 Putative function of TrpX protein in *S. coelicolor*

In silico approaches and previous experiments have shown that TrpX protein is a putative regulator of the *trpB* gene, a gene involved in the tryptophan biosynthesis. Microarray analysis, confirmed by Real-Time RT-PCR, of *S. coelicolor trpX* mutants showed that *trpB* gene expression in *trpX* mutant is lower than the one observed in wt strain. When mutation is complemented, *trpB* expression level is similar to that of wt strain.
A tridimentional structure prediction, using a predictional software has been carried out, and it showed that TrpX protein has a DNA-binding protein structure.
These in silico results suggested that TrpX could bind to DNA.

5.3 Gel mobility shift assays (EMSA) of ^{32}P end labelled fragments spanning intergenic regions of trpA-B-X-C1 operon.

The most probable hypothesis of the TrpX protein function in *S. coelicolor*, is that it binds to intergenic regions of the *trpA/B/X/C1* operon, such as *trpB-trpX* and *trpX-trpC1* intergenic regions, regulating the *trpB* gene expression.
Two EMSA have been carried out, during this work, using protein extracts of *S. coelicolor* M145, ΔX, ΔX/X and ΔX/pKC strains. Results showed that no proteins are bound to the two regions, in the used conditions.
Further EMSA analysis is necessary, changing some parameters of the assay.

5.4 TrpX protein overexpression in *E. coli*

To understand TrpX protein function in *S. coelicolor*, its overexpression and purification were carried out using a *E. coli* BL21 pRSETB-X strain.
The strategy choosed for the SCO2038 protein overexpression involves the use of pRSETB plasmid. The pRSETB plasmid has been constructed for the recombinant protein overexpression in *E. coli* BL21, allowing the control of its expression and an easy purification of the overexpressed protein.
The *E. coli* DH5α pRSETB-X, and successively the *E. coli* BL21 pRSETB-X24 strains have been successfully constructed and controlled. The His_6-TrpX protein was successfully overproduced and its purification at present is in progress. The protein will be used in DNA or RNA binding assays.

5.5 *trpX* gene overexpression in *S. coelicolor*

One objective of this work was to overexpress the *trpX* gene in *S. coelicolor*, in order to study differences in the global gene level expression and antibiotic level production, between mutant and wt strains, indicating ways to understand the *trpX* gene role.
To understand *trpX* gene function in *S. coelicolor*, its overexpression in *S. coelicolor* was carried out using the pIJ8600-X plasmid. The strategy choosed for the *trpX* gene overexpression involves the use of pIJ8600 plasmid. The pIJ8600 plasmid has been constructed for the integration of a gene in *S. coelicolor* genome, followed by its overexpression.
The first step of this strategy is the construction of the pIJ8600-X plasmid in *E. coli* DH5α.

E. coli DH5α cells has been transformed with the pIJ8600-X plasmid, but no positive clone has been found, also changing cell transformation conditions.

Other strategies to obtain the *trpX* gene overexpression in *S. coelicolor* have been studied. Two new strategies have been selected, and each involved the construction of the *E. coli* DH5α pIJ8600-pGEM-X stain. The advantage of these strategies is the clone screening using plates containing two antibiotics, apramycin (gene resistance present in pIJ8600) and ampicillin (gene resistance present in pGEM), avoiding the growth of strains containing only the pIJ8600 plasmid without *trpX* insert.
First one involved the pIJ8600-pGEM-X plasmid direct integration in the *S. coelicolor* genome.
Second one involved the suppression of the pGEM fragment in the pIJ8600-pGEM-X plasmid, using restriction enzymes, to obtain the targeted *E. coli* DH5α pIJ8600-X stain.
This two strategies at present are in progress.

6 CONCLUSIONS AND PERPECTIVES

Real-Time RT-PCR analysis showed that *trpX* is not costitutively expressed during the growth in minimal medium, but is growth-phase dependent.
In silico approaches and previous experiments have shown that TrpX protein is a putative regulator of the *trpB* gene, a gene involved in the tryptophan biosynthesis and a tridimentional structure prediction showed that TrpX protein could have a DNA-binding protein structure.
The most probable hypothesis of the TrpX protein function in *S. coelicolor*, is that it binds to intergenic regions of the *trpA/B/X/C1* operon, regulating the *trpB* gene expression.
This hypothesis has not been confirmed by DNA binding assays carried out in this work, thus, further DNA or RNA binding assays, are necessary.
The His_6-TrpX protein was successfully overproduced in *E. coli*, and its purification at present is in progress, it will be used in DNA or RNA binding assays, in order to confirm the previous hypothesis.
To understand *trpX* gene role in *S. coelicolor*, it has been overexpressed in *S. coelicolor*, in order to study differences in the global gene level expression and antibiotic level production, between obtained mutant and wt strains. This strategy indicating ways to understand the *trpX* gene role in *S. coelicolor*. Nevertheless, the *trpX* gene overexpression in *S. coelicolor* was unsuccessfully. The pIJ8600-X plasmid used for the *trpX* gene overexpression, has not been constructed in *E. coli* DH5α.
New strategies has been studied and at present are in progess.

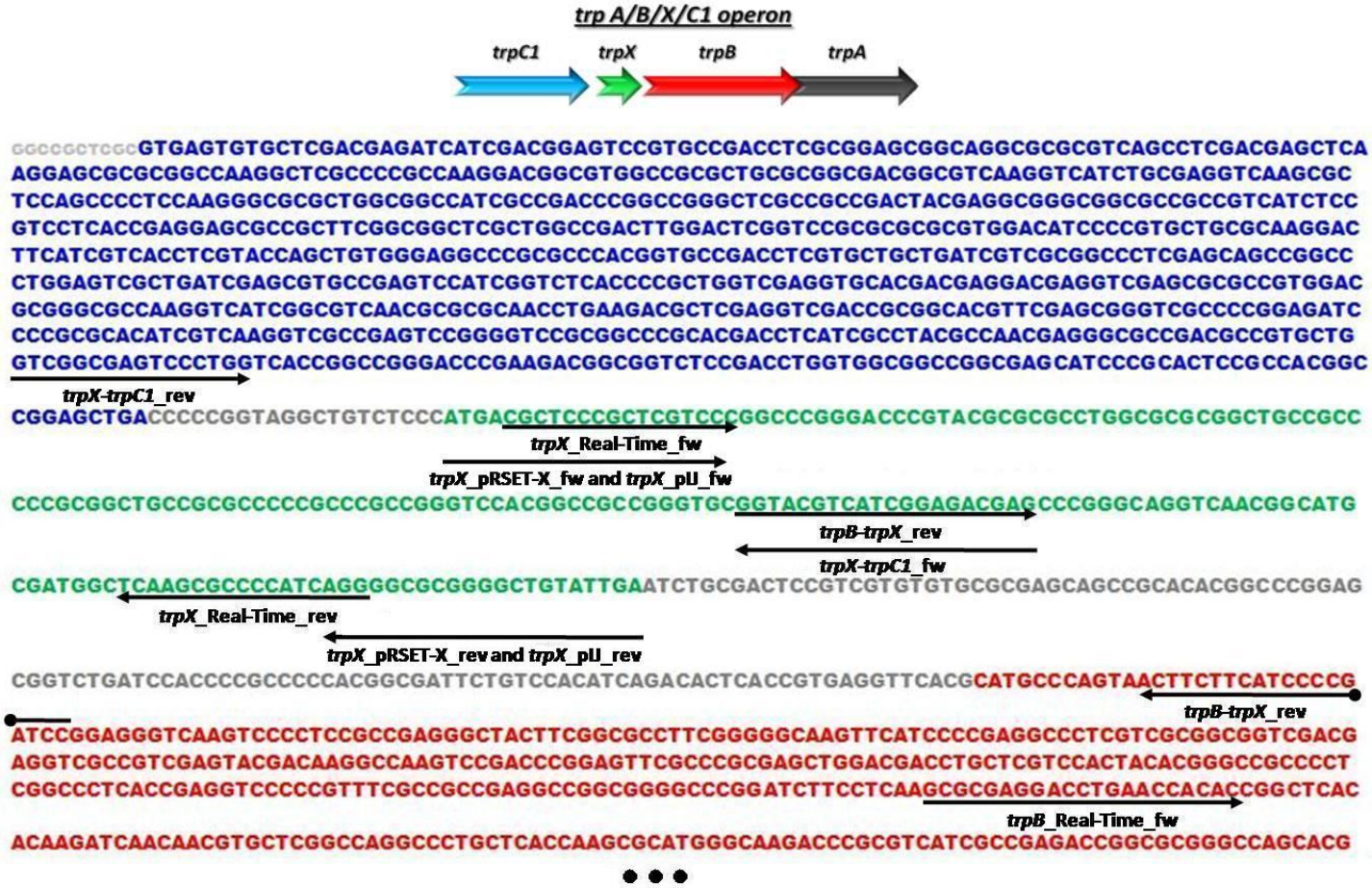
7 REFERENCES

- **Bao K. & Cohen S.N. (2001).** Terminal proteins essential for the replication of linear plasmids and chromosomes in *Streptomyces*, Stanford University School of Medicine, CA, USA, *Genes & Dev.* **15**, 1518-1527
- **Bascaran V. & Al. (1989).** Regulation of nitrogen catabolic enzymes in *Streptomyces clavuligerus*, Universidad de Oviedo, Spain, *J. Gen. Microbiol.* **135(9),** 2465-74
- **Bentley S.D. & Al. (2002)**. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2), The Wellcome Trust Sanger Institute , UK, *Nature* **417**, 141-147
- **Bibb M.J. (2005)**. Regulation of secondary metabolism in streptomycetes, John Innes Centre, UK, *Curr. Opin. Microbiol.* **8**, 208-215
- **Chakraburtty R. & Bibb M.J. (1997)**. The ppGpp synthetase gene (relA) of Streptomyces coelicolor A3(2) plays a conditional role in antibiotic production and morphological differentiation, John Innes Centre, UK, *J Bacteriol.* **179**, 5854-5861
- **Chater, K.C. (1998)**. Taking a genetic scalpel to the Streptomyces colony, *Microbiology* **144**, 1465-1478
- **Chater K.F. (2006)**. *Streptomyces* inside-out: a new perspective on the bacteria that provide us with antibiotics, Norwich Research Park, UK, *Phil.Trans.R.Soc. B* **361**, 761-768
- **Chong P.P. & Al. (1998)**. Physical identification of a chromosomal locus encoding biosynthetic genes for the lipopeptide calcium-dependent antibiotic (CDA) of *Streptomyces coelicolor* A3(2), *Microbiology* **144(1)**, 193-199
- **Flett F. & Al. (1987)**. DNA rearrangements associated with instability of an arginine gene in *Streptomyces coelicolor* A3(2), University of Manchester, *J. Basic Microbiol.* **27**, 3-10
- **Granozzi C. & Al. (1990)**. A breakdown in macromolecular synthesis preceding differentiation in Streptomyces coelicolor A3(2), *J. Gen. Microbiol.* **136(4)**, 713-6
- **Grunewald J. & Al. (2004)**. Synthesis and derivatization of daptomycin: a chemoenzymatic route to acidic lipopeptide antibiotics, Philipps-Universität Marburg, Germany, *J Am. Chem. Soc.* **126(51)**, 17025-17031
- **Hanahan D. (1983)**. Studies on transformation of *Escherichia coli* with plasmids, *J. Mol. Biol.* **166,** 557–580
- **Hodgson D. A. (2000)**. Primary Metabolism and its Control in Streptomycetes : A Most Unusual Group of Bacteria, University of Warwick, UK, *Adv Microb Physiol.* **42**, 47-238
- **Hojati Z. & Al. (2002)**. Structure, biosynthetic origin, and engineered biosynthesis of calcium-dependent antibiotics from *Streptomyces coelicolor*, UMIST, UK, *Chem. Biol.* **9(11)**, 1175-1187
- **Hood D.W. & Al. (1992)**. Molecular genetic analysis of proline and tryptophan biosynthesis in *Streptomyces coelicolor* A3(2): interaction between primary and secondary metabolism, *Gene* **115**, 5-12
- **Hopwood D.A. & Wright H.M. (1983)**. CDA is a new chromosomally-determined antibiotic from *Streptomyces coelicolor* A3(2), John Innes Centre, UK, *J Gen. Microbiol.* **129(12)**, 3575-3579
- **Hopwood D.A. (1999)**. Forty years of genetics with Streptomyces : from in vivo throught in virto to in silico, John Innes Centre, UK, *Microbiology* **145**, 2183-2202
- **Hu D.S.-J. & Al. (1999)**. The expression of the *trpD*, *trpC* and *trpBA* genes of *Streptomyces coelicolor* A3(2) is regulated by growth rate and growth phase but not by feedback repression, University of Warwick, UK, *Molecular Microbiology* **32(4)**, 869-880
- **Huang J. & Al. (2001)**. Global analysis of growth phase responsive gene expression and regulation of antibiotic biosynthetic pathways in Streptomyces coelicolor using DNA microarrays, Stanford University School of Medicine, CA, USA, *Genes Dev.* **15**, 3183-3192
- **Karoonuthaisiri N. & Al. (2005)**. Regional organization of gene expression in *Streptomyces coelicolor*, Stanford University, CA, USA, *Gene* **353**, 53-66
- **Katz E. & Al. (1984)**. Biological, Biochemical and Biomedical aspects of Actinomycetes, 325-342
- **Kieser T., Bibb, M.J., Buttner, M.J., Chater, K.F. & Hopwood, D.A. (2000)**. General introduction to actinomycete biology, Practical *Streptomyces* genetics, p.1-42.
- **Kim D.E & Al. (2004)**. Protein structure prediction and analysis using the Robetta server, University of Washington, Seattle, WA, USA, *Nucleic Acid Research* **32**, W526-W531

- **Kim H.B. & Al. (2004)**. Metabolic flux analysis for calcium dependent antibiotic (CDA) production in *Streptomyces coelicolor*, UMIST, UK, *Metab. Eng.* **6**, 313-325
- **Limauro D. & Al. (1990)**. Cloning and characterization of the histidine biosynthetic gene cluster of *Streptomyces coelicolor* A3(2), Universitá di Napoli, Italy, *Gene* **90**, 31-41
- **Murphy M.F. & Katz E. (1980)**. Regulatory control of 3-deoxy-D-arabino-heptulosonic acid 7-phosphate synthetase in *Streptomyces antibioticus*, *Can J. Microbiol.* **26**, 874-880
- **Potter C.A. & Baumberg S. (1996)**. End-product control of enzymes of branched-chain amino acid biosynthesis in *Streptomyces coelicolor*, University of Leeds, UK, *Microbiology* **142**, 1945-1952
- **Seno E.T. & Chater K.F. (1983)**. Glycerol catabolic enzymes and their regulation in wild-type and mutant strains of *Streptomyces coelicolor* A3(2), *J.Gen.Microbiol.* **129(5)**, 1403-1413
- **Studier F.W. & Moffatt B.A. (1986)**. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes, *J Mol Biol*, **189**, 113–130
- Volff J.N. & Altenbuchner J. (1998). Genetic instability of the *Streptomyces* chromosome, *Molecular Microbiology* **27**, 239-246(8)
- **Yakhnin H. & Al. (2004)**. The *trp* RNA-Binding Attenuation Protein of *Bacillus subtilis* Regulates Translation of the Tryptophan Transport Gene *trpP* (*yhaG*) by Blocking Ribosome Binding, The Pennsylvania State University, PA, USA, *J Bacteriol.* **186**, 278-286
- **Yanofsky C. (2000)**. Transcription Attenuation: Once Viewed as a Novel Regulatory Strategy, Stanford University, CA, USA, *J Bacteriol.* **182(1),** 1-8
- **Yanofsky C. (2004)**. The different roles of tryptophan transfer RNA in regulating *trp* operon expression in *E. coli* versus *B. subtilis*, Stanford University, CA, USA, *Trend genet.* **20,** 367-374
- **Zalkin H. & Smith J.L. (1998)**. Enzymes utilizing glutamine as an amide donor, *Adv. Enzymol. Relat. Areas. Mol. Bio.* **72**, 87-144
- **Zurawski G. & Al. (1978)**. Translational control of transcription termination at the attenuator of the *Escherichia coli* tryptophan operon, Stanford University, CA, USA, *Proc. Natl. Acad. Sc4. USA* **75(12)**, 5988-5992

8 APPENDICES

Appendix I *trpA/B/X/C1* operon sequence and primers used in this work




trpB_Real-Time_rev
trpC1_Real-Time_fw
trpC1_Real-Time_rev

ACKNOWLEDGMENTS

This diploma thesis would never have been completed without the support and help of many friends, colleagues and family.
First of all I would like to thank Anna-Maria Puglia, who gave me the opportunity to do my diploma work in Palermo, who gave me a big support during my work and who gave me a warm welcome and Letizia Vittorelli who have found me a home in Palermo and for her warm welcome.
I would also like to thanks Sandra Marineo and Anna Giardina, my supervisors, who have always supported my work and have always been available for my questions.
Next persons I want to thank is Sergio Schmid, who gave me the possibility to go in Palermo.
I would also like to thanks Hebert, from Perou, who gave me a big support when I arrived in Palermo.
I must acknowledge all (ex)co-workers who have all taught me something new and contributed towards the pleasant and friendly working environment in Palermo. I would like to thanks Franscesco, Maria-Grazia, Turi, Valentina, Erica, Giuseppe, Luca, Ruggero, Daniele, Eleonora, Calogero, Salvo, Giovanna, Michele, Fulvio, Sergio and everyone from AMP group.
Many thanks go to Maurizio Noto, for its tecnical support.
Special thanks go to my roommates Julius, Alexandra and Christian, from Basel, and Jana from Prague.
A big big special thanks go to Adele Caneppa, she knows why.
I wish to thank my family, my mother Ilide, my brothers Alain and Dario, my cousin Alberto, for their loving support.
Last but not least, I would like to thank all my friends in Palermo, without whom the past six month would not have been the same! So big thanks to Laura, Daniella and friends, Franscesca, Valentina, Patrizia, Claudia, Marco, Julia, Silvia, Martin, Xavier, Gabriele, Elisa, Alessia, Leandro, Lorenzo, Gemma, Roberto and many more…