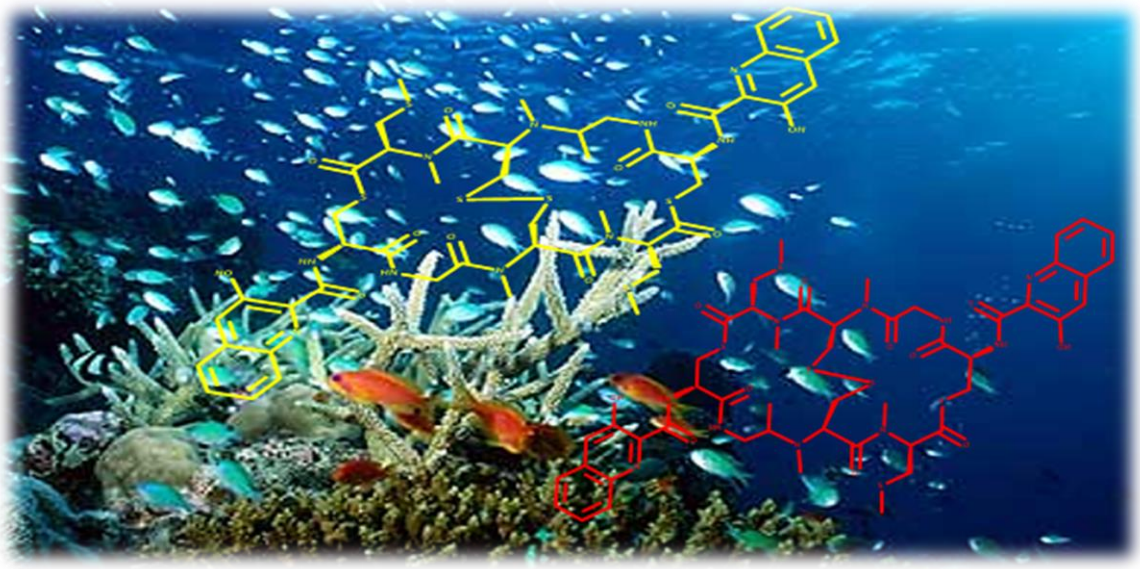
Deciphering the biosynthetic pathway for the starter unit of the antitumoral thiocoraline

Serena Galiè

Master in Biotechnology of Environment and Health

University of Oviedo-University of Bologna



Index

1. Summary

The growing interest of pharmaceutical companies towards more and more effective antitumor compounds opened up the process of "searching for a natural compound" in all sorts of ecosystems, from terrestrial to marine ones. To this regard, recent advances in modern technologies of screening have increased the efforts for the discovery of newly and effective marine drugs. Against this backdrop, the thiodepsipeptide thiocoraline represents an antitumor compound of great potential as chemotherapeutic. It is produced by two *Micromonospora* sp. strains isolated from two marine invertebrates found in the coast of Mozambique. Previous studies have identified the biosynthesis gene cluster of thiocoraline which allowed its genetic expression in the more versatile bacterial platform of *Streptomyces albus*. The present work is focused on the generation of three different mutant forms of thiocoraline producer strain *S. albus* pFL1049, in order to obtain a better understanding of the initial stages involved in the biosynthesis of thiocoraline starter unit, the 3-hydroxyquinaldic acid 3HQA. In particular, the attention has been paid on *tioL* and *tioM* genes in relation to their supposed involvement in one of last stages of 3 HQA biosynthesis. The optimization of thiocoraline production and its comparison between control and mutant strains have been carried out together with the identification of 3HQA intermediates by means of HPLC-MS analysis.

A hypothesis of *tioL* and *tioM* functions has been proposed, based on the results obtained during this work about their probable involvement in the regulation of thiocoraline biosynthesis.

2. Introduction

2.1 Anticancer drugs from Nature

Over the centuries, plant extracts have been used to treat various diseases. Until now, natural products have played an important role in anticancer therapy as there are more than 500 compounds from terrestrial and marine plants or microorganisms, which have antioxidant, antiproliferative, or antiangiogenic properties and are therefore able to reduce tumor growth. The recent discovery of new natural products has been accelerated by novel technologies (high throughput screening of natural products in plants, animals, marine organisms, and microorganisms). Vincristine, irinotecan, etoposide, and paclitaxel are examples of compounds derived from plants that are used in cancer treatment. Similarly, actinomycin D, mitomycin C, bleomycin, doxorubicin, and L-asparaginase are drugs derived from microorganisms (Orlikova et al., 2014).

The definition of a natural compound is very complex. Traditionally, a natural product is a chemical compound produced by living organisms and possessing biological or pharmacological activity. Moreover, a natural product can also be synthesized and thus be chemically identical to its natural counterpart.

In order to elucidate the contribution of natural products in chemotherapeutic drug discovery and development, Newman and Cragg generated a drug classification after evaluation of all approved anticancer drugs between 1940 and 2010. In their analysis, 206 approved anticancer agents were classified into clearly defined groups. While compounds belonging to the first 8 categories were classified as naturally derived or inspired compounds, the last category lacks any natural product inspiration and is therefore considered as the only truly synthetic class of compounds (Orlikova et al., 2014).

Accordingly, 78.6 % of all approved anticancer agents are either natural products or based thereon, or agents that mimic them in one form or another, and only 21.4 % were fully synthetic with no prototype or conception from nature. If we consider these data, it becomes clear that Mother Nature plays a predominant role in modern therapy. Moreover, this classification emphasizes the importance of the development of drugs from natural origins. However, the use of natural products has limitations since living organisms sometimes synthesize only trace quantities of otherwise interesting bioactive compounds and natural products have limited bioactivity. Here, the polarity of the molecule often complicates its cellular uptake, thus leading to a reduced activity.

Nevertheless, natural products can be optimized, on the one hand, by removal, introduction or modification of functional groups of active natural product scaffolds in order to improve their bioactivity and, on the other hand, by natural product-inspired combinatorial synthesis thereby providing large libraries of compounds in a short time. These are the promising strategies of the future to obtain powerful drug leads (Molinski et al., 2008).

2.2 The great resource of marine environment

Since ancient times, nature has been an important source of medicines: a fact illustrated by the large number of natural products currently in use in medical practice. Although marine compounds are under-represented in current pharmacopoeia, it is anticipated that the aquatic environment will become an invaluable source of novel compounds in the future, as it could be demonstrated by the increasing amount of marine products discovered in last decade (Figure 1).

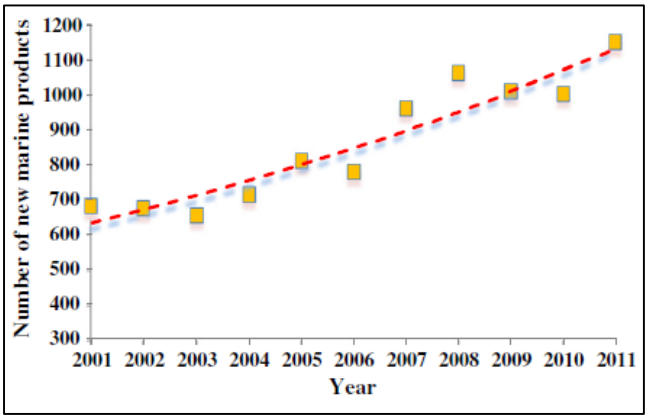


Figure 1: Number of new marine products discovered per year (Hartmann et al., 2014).

The marine ecosystem represents 95% of the biosphere, and all except one of the 33 animal phyla are represented in aquatic environments. Most sessile marine invertebrates contain a primitive immune system and produce toxic chemicals as a form of defense. Many of these products act as regulators of specific biological functions. Some of them have pharmacological activity due to their specific interactions with receptors and enzymes. Because these substances become immediately diluted by large volumes of seawater, they need to be highly potent on a molar basis, and also have to retain a relatively low solubility. The development of marine compounds as therapeutic agents is still in its infancy due to the lack of an

analogous ethno-medical history as compared with terrestrial habitats, together with the relative technical difficulties in collecting marine organisms. Over the last few decades significant efforts have been made, by both pharmaceutical companies and academic institutions, to isolate and identify new marine-derived, natural products (Vinothkumar et al., 2013).

Molecular approaches offer particularly promising alternatives not only to the supply of known natural products (e.g., through the identification, isolation, cloning, and heterologous expression of genes involved in the production of the chemicals) but also to the discovery of novel sources of molecular diversity (e.g., through the identification of genes and biosynthetic pathways from uncultured microorganisms) (Amador et al., 2003).

A cursory review of the literature indicates that more than 70% of marine metabolites are obtained from marine sponges, corals and microorganisms. The contribution from other organisms like mollusks, ascidians and algae adds up to only 30%. The importance of marine metabolites in modern day drug research is revealed by the fact that around 50% of the USA FDA approved drugs during 1981–2002 consist of either marine metabolites or their synthetic analogs. Their low effective dosage, better selectivity against target malignant tissues and relative non-vulnerability to resistance development as compared to compounds of terrestrial origin, render them useful target molecules. According to USA National Cancer Institute estimates, more than 1% of marine natural products show antitumor properties as against only 0.01% amongst their terrestrial counterparts (Vinothkumar et al., 2013).

2.3 Marine actinobacteria as producers of bioactive compounds

The phylum Actinobacteria is one of the largest taxonomic groups within the domain Bacteria. Actinomycetes are Gram-positive bacteria belonging to this phylum. Filamentous actinomycetes are characterized by a complex life cycle of morphological differentiation that starts with the germination of a spore to give a substrate mycelium and later an aerial mycelium, followed then by a sporulation stage (Olano et al., 2014).

Actinobacteria are widely distributed in terrestrial and aquatic ecosystem with a ubiquitous presence in the marine environment, playing an important ecological role in the recycling of refractory biomaterials and producing novel natural products with pharmacological applications.

Actinobacteria have been detected or isolated from different marine creatures such as sponges, corals, mollusks, ascidians, seaweeds, and seagrass. The high number of

marine organism-associated actinobacterial 16S rRNA gene sequences, deposited in the NCBI database clearly reveal enormous numbers of actinobacteria associated with marine organisms. For example, RDP classification of these sequences showed that 112 and 62 actinobacterial genera were associated with the sponges and corals, respectively. In most cases, it is expected that these actinobacteria protect the host against pathogens by producing bioactive compounds. Natural products investigation and functional gene screening of the actinobacteria associated with the marine organisms revealed that they can synthesize numerous natural products including polyketides, isoprenoids, phenazines, peptides, indolocarbazoles, sterols, and others. These compounds showed anticancer, antimicrobial, antiparasitic, neurological, antioxidant, and anti-HIV activities. Therefore, marine organism-associated actinobacteria represent an important resource for marine drugs.

To give an example, it has been estimated that 58 % of natural products are derived from the marine sponge-associated actinobacteria, while the remaining are from the actinobacteria associated with mollusks, mangroves, corals, ascidians, and seaweeds (Figure 2).

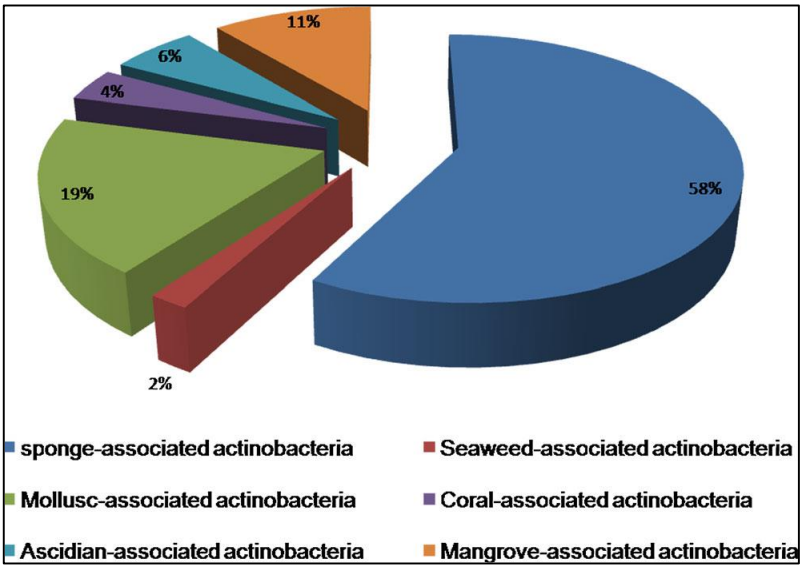


Figure 2: Marine organisms distribution of actinobacterial natural products (Valliappan et al., 2014).

2.4 Thiocoraline: from the discovery to the application

Thiocoraline is a twofold-symmetric bicyclic nonribosomal octathiodepsipeptide produced by two actinomycete strains isolated from two marine invertebrates from the

Indian Ocean coast of Mozambique: *Micromonospora* sp. ML1 from a marine mollusc and *Micromonospora* ACM2-092 from a soft coral. This compound shows antibacterial activity against Gram-positive bacteria and is a potent antitumor compound active against a variety of human cancer cell lines, due to its DNA bisintercalative properties and its DNA polymerase α inhibition (Romero et al., 1997).

2.4.1 Molecular insights into thiocoraline

From a structural standpoint, thiocoraline belongs to a class of naturally occurring 2-fold symmetric or pseudosymmetric bicyclic octadepsipeptides that include the well-known quinoxaline antibiotics triostin A and echinomycin.

It contains two moieties of the chromophore 3-hydroxy-quinaldic acid (3HQA), which probably acts as a starter unit during biosynthesis. Each of the two symmetrical halves is constituted by 3HQA-D-Cys-Gly-*N*-methyl-LCys-N,S-dimethyl-L-Cys (Figura 3). The halves are linked together through thioester bonds between the first D-Cys residue of one half and the last *N,S*-dimethyl-L-Cys residue of the other half, and by one disulfide cross-bridge, to generate the final bicyclic thiolactone.

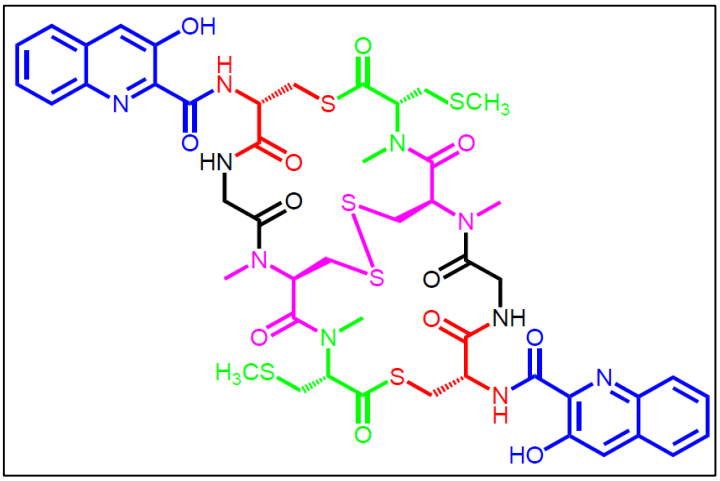


Figure 3: Thiocoraline chemical structure (Fernández et al., 2014).

The X-ray crystal structure of thiocoraline has revealed uniquely stacked arrays of cross-docked pairs of staple-shaped molecules that are highly preorganized for DNA binding through intercalation, although they still need to adapt to the geometric restrictions imposed by the bisintercalation binding mode. The 3-hydroxyquinaldic system present in thiocoraline ensures a tricyclic hydrogen-bonded conformation that

allows the antibiotic to intercalate into DNA and establish four good hydrogen bonds between its thiodepsipeptide and the minor groove of GC-rich sequences and particularly those encompassing a central CpG step (Figure 4).

Thiocoraline is then active as antitumor compound against human breast cancer, non-small-cell lung carcinoma, colon, renal and melanoma cancer cell lines in vitro and in in vivo xenografts. Together with its strong DNA minor groove binding, this antitumor compound is also a strong inhibitor of DNA polymerase α subunit. It inhibits a wide array of Gram positive pathogenic bacteria (Erba et al., 1999).

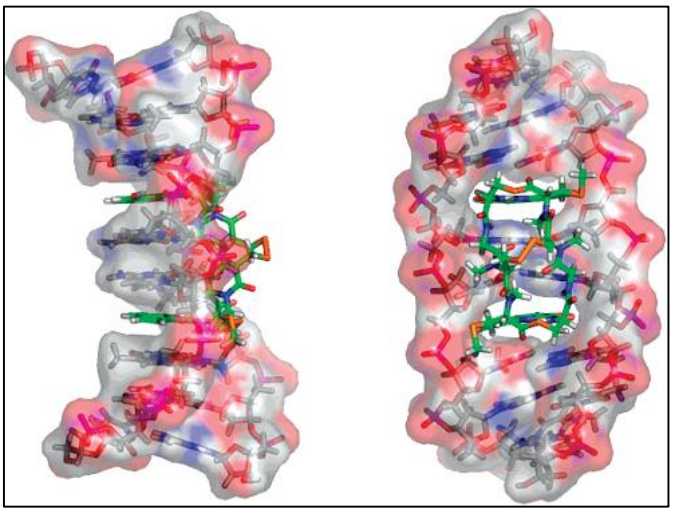


Figure 4: Side and front views of the bisintercalated thiocoraline complex (Negri et al., 2007).

2.4.2 Deciphering the biosynthetic pathway of thiocoraline

The biosynthesis of the family of antitumor compounds is carried out by nonribosomal peptide synthetases (NRPSs), which are large multifunctional enzymes with a modular organization of catalytic domains. A minimal module is composed of three domains: 1) an adenylation domain (A, approximately 550 residues) which selects a specific amino acid and generates its aminoacyl adenylated version by using adenosine triphosphate (ATP); 2) a peptidyl-carrier protein domain (P, approximately 80 residues) which contains a 4'-phosphopantetheine (PP) prosthetic group for covalent thioester binding of the activated adenylated amino acid; and 3) a condensation domain (C, approximately 450 residues) which generates a new peptide bond between the two aminoacyl adenylated residues located on consecutive P domains (Lombó et al, 2006).

Thiocoraline is structurally related to other actinomycete-produced chromopeptides which contain 3HQA or quinoxaline-2-carboxylic acid (QX) as the starter unit, both of which are derived from L-tryptophan. Their common scaffold structure has suggested a putative conserved biosynthetic pathway for the starter moieties and similar modular architecture for the corresponding NRPSs, although with differences in the specificity of some A domains (Lombó et al, 2006).

Relating to this assumption, a combinatorial study has been carried out by means of bioinformatics tools which allow designing different oligonucleotide primers that were used to amplify 6 PCR fragments in *Micromonospora sp. ML1* and construct a cosmid library for a hybridization protocol. Positive clones for hybridization were tested for the involvement in thiocoraline biosynthesis thanks to trans-conjugation of inactivated genes in *Micromonospora* up to the identification of a defined sequence region of 64150 bp that coded for 36 open reading frames (ORFs) (Lombó et al, 2006).

Heterologous expression in several *Streptomyces* species of the ORF found with the procedure described above allowed the identification of all the genes necessary for thiocoraline production, within a sequence of about 50 kbp that goes from *tio A* to *tio Z* (Figure 5). Therefore, this region has been successfully integrated into *S. albus* J1074 chromosome in order to obtain the bacterial platform *S. albus*-pFL1049 that has been used throughout this work (Lombó et al, 2006).

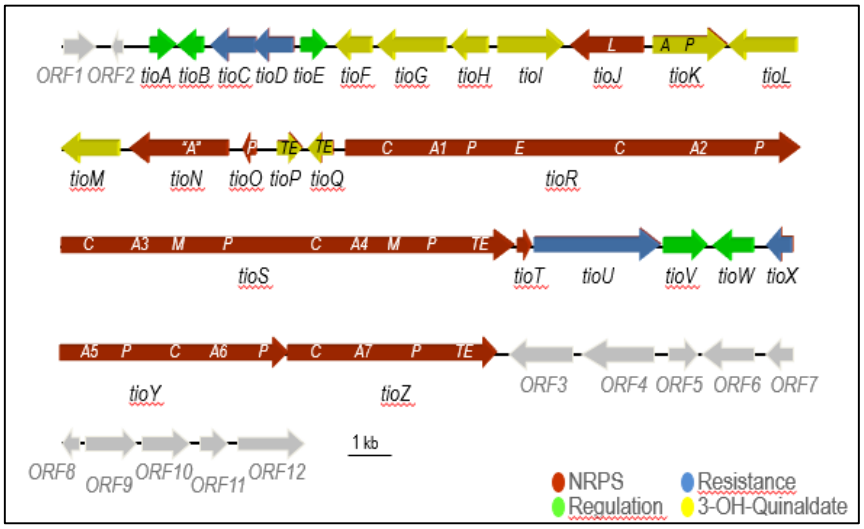


Figure 5: Schematic representation of thiocoraline gene cluster (Lombó et al., 2006).

Experiments of specific insertional inactivation of gene in *Streptomyces* together with a combinatorial study allowed the deduction of the function of each protein involved in thiocoraline biosynthesis.

Therefore, according to Figure 5, it is assumed that:

- The 6 yellow genes in Figure 1 are involved in the synthesis of thiocoraline starter unit, the 3-hydroxy-quinaldic-acid (3HQA): *tioF*, *tioG*, *tioH*, *tioI*, *tioK*, *tioL*, *tioM*, *tioP*, *tioQ*;
- The blue genes are involved in generating a thiocoraline resistance: *tioC* and *tioD* (coding for an ABC transporter), *tioU* (coding for an UV-repair enzyme) and *tioX* (coding for a thiocoraline-binding enzyme to avoid intracellular toxicity) (Lombó et al, 2006; Biswas et al., 2010).
- The green ones are regulation genes that acts at different levels during thiocoraline biosynthesis: *tioA*, *tioB*, *tioE*, *tioV*, *tioW*;
- The red ones are conserved domain in NRPS family: *tioJ*, *tioN*, *tioO*, *tioR*, *tioS*, *tioT*, *tioY*, *tioZ*.

2.4.3 The thiocoraline starter unit 3HQA

The majority of antitumoral compounds produced by terrestrial and marine microorganisms are characterized by a distinct chemical feature that allow them to specifically interact with DNA helix and gain the name of bisintercalators. All these compounds contains a cyclic structure, which protrude at both distant positions with two heteroaromatic chromophores. The presence of these protruding moieties from the peptide scaffold give reason to the DNA intercalation capability. During this interaction, both chromophores are inserted in between the minor groove base pairs, unwinding and extending the DNA helix. Structural similarity between all 21 known natural bisintercalators revealed that all of them are characterized by the evolutionary use of two types of chromophores as starter units such as 3HQA or quinoxaline-2-carboxylic acid (QXC) which are both derived from L-tryptophan (Sheoran et al., 2008).

2.4.3 Biosynthetic pathway of 3HQA

Combinatorial analysis within the cluster genes of natural bisintercalator compounds such as echinomycin, quinomicin and triostin, agree with considering L-tryptophan as the first actor compound in the biosynthesis of the both 3HQA and QXC, as proved by feeding experiments with labeled precursors (Watanabe 2006).

So, considering thiocoraline cluster gene, L-Trp is firstly bound to a small NRPS protein formed by an A-T module that is reperesented by TioK in thiocoraline cluster gene which works together with a small MbtH-like protein TioT to achieve the adenylation step and obtain a hetero-tetramer. The attached L-Trp is then β-hydroxilated by the action of a cytochrome P450 coded by TioI. The resulting β-hydroxy-L-Trp is then released by the action of a free-thioesterase TioQ. Indole ring opening at the β-hydroxy-L-Trp is achieved by the action of a Trp-2,3-dioxigenase TioF generating N-formyl-β-hydroxykynurenine (Fernandez et al., 2011).

Loss of the formyl moiety in order to obtain the 3-OH-kynurenine could be a spontaneous event or other enzymes could be involved, such as TioL (as it has been proposed in literature) and/or TioM both of each roles remains still unclear in thiocoraline biosynthetic pathway (Fernandez et al., 2011).

Combinatorial analysis in quinomycin family biosynthetic clusters have demonstrated that although an arylformamidase is an essential enzyme for kynurenine biosynthesis, there has not been found any corresponding gene with a formamidase function. Initially, *tioL* in the case of thiocoraline and *ecm14* in the case of equinomycin were proposed to catalyze this reaction but low-sequence homology suggested lack of this gene in all quinomycin gene clusters (Hirose et al., 2011).

Once the β-hydroxykynurenine is released, a TioG aminotransferase generates the bicyclic heteroaromatic intermediate 3,4-dihydroxy-quinaldic acid which is further converted into 3HQA by a TioH oxidoreductase (Figure 6).

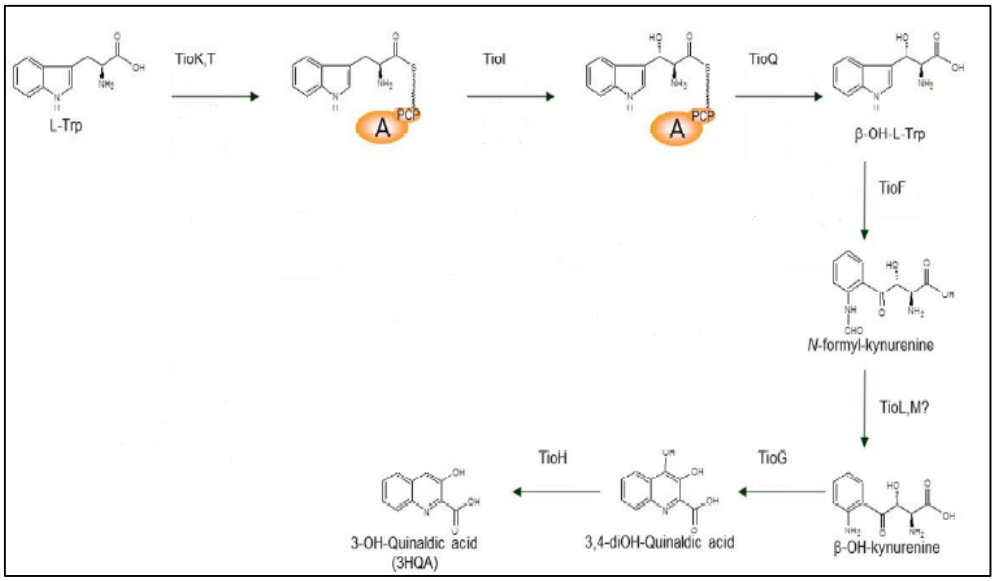


Figure 6: Biosynthetic pathway of thiocoraline starter unit 3 HQA (Fernández et al., 2014).

2.4.4 Elucidating *tioL* and *tioM* roles in 3HQA biosynthesis

As previously observed in the explanation of the proposed biosynthetic pathway of 3HQA, the formyl-kyneurine that is formed by the action of a dioxigenase enzyme (*tioF* coding gene) could be then deformylated by the action of a gene coded by the only two genes that still remain a mystery in thiocoraline starter unit biosynthesis: *tioL* and *tioM*.

Combinatorial studies by means of bioinformatics tools have failed to deduce a clear and reasonable function to neither of these two coding genes, except from revealing a gene and structural similarity between *tioM* and a vanadium peroxidase owning to the so-called PAP family and which is involved in the oxidation of halides by using vanadate and hydrogen peroxide.

Previous experiments of insertional inactivation gene of each and both *tioL* and *tioM* genes in the bacterial strain *S. albus*-pFL1049 have demonstrated their probable involvement in thiocoraline production, due to its reduced production *in vivo*.

To give a better understanding about the mutants that have been studied all along this work, it is worth to mention the procedure by means they have been obtained and that is schematically represented in Figure 7.

In particular, the procedure has been carried out by a series of sequential steps:

1. PCR amplification of a truncated version of the coding sequences of *tioL* and *tioM* and their incorporation into a plasmid by means of a cloning procedure (the plasmids used are shown in Figure 7);

2. Gene disruption procedure in order to obtain:

    a. the single mutants *ΔtioM* and *ΔtioL* thanks to a single crossing-over event between the plasmid vectors containing mutated versions of the coding sequences of *tioL* and *tioM* and their corresponding genes in the bacterial chromosome of *S. albus*-pFL1049 strain;
    b. the double mutant *ΔtioLM* thanks to a double crossing-over event between the plasmid vector containing an interrupted version of the coding gene sequences of *tioL* and *tioM* and their corresponding genes

in the bacterial chromosome of *S. albus*-pFL1049 strain. Therefore, a gene substitution event occurred with the introduction of an erythromycin resistance gene in place of *tioL* and *tioM*;

3. Isolation of positive mutants clones by a selective growth.

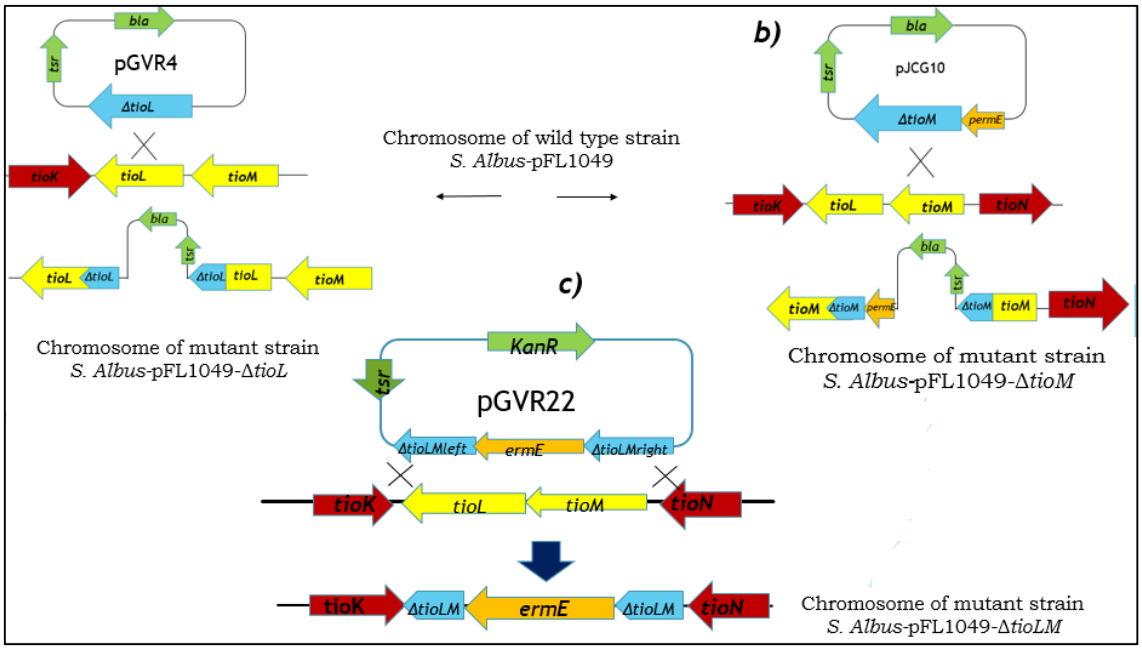


Figure 7: Schematic representation of the genetic construction of mutants.

3. Objectives

This work has been focused on the study and characterization of the three mutant strains of *S. albus*-pFL1049: *ΔtioM*, *ΔtioL* and *ΔtioLM*. These strains had been previously obtained and here they will be compared to *S. albus*-pFL1049 production levels.

In particular, the main objectives that have been pursued all along this study are:

1. Testing of the 3 mutant strains

2. Optimization of the best culture conditions for thiocoraline production in the strain *S.albus*-pFL1049 and in its mutant forms *S. albus-ΔtioL*, *S. albus-ΔtioM* and *S. albus-ΔtioLM*

3. Comparison of thiocoraline production levels in *S. albus*-pFL1049 and his mutant forms *S. albus-ΔtioL*, *S. albus-ΔtioM* and *S. albus-ΔtioLM* in standardized and optimal culture conditions

4. Comparison of 3HQA intermediates production in *S. albus* pFL1049 and its mutant forms *S. albus-ΔtioL*, *S. albus-ΔtioM* and *S. albus-ΔtioLM* in standardized and optimal culture conditions

4. Materials and Methods

4.1 Bacterial strains

The main strain that has been used all along this work is the *S. albus*-pFL1049 as thiocoraline producer platform together with the wild type strain *S. albus* J1074 (Chater et al., 1976).

Three different mutants were also used in this work:

a) *S. albus*-pFL1049-Δ*tioL*: this mutant was generated by gene disruption on *tioL* gene. The plasmid used for this purpose was pJCG10 (Figure 7);

b) *S. albus*-pFL1049-Δ*tioM:* this mutant was generated by gene disruption on *tioL* gene, thanks to the construction plasmid GVR4 (Figure 7);

c) *S. albus*-pFL1049-Δ*tioLM:* this mutant was generated by gene replacement on *tioL* and *tioM* genes. The plasmid used for this was pJCG22 (Figure 7).

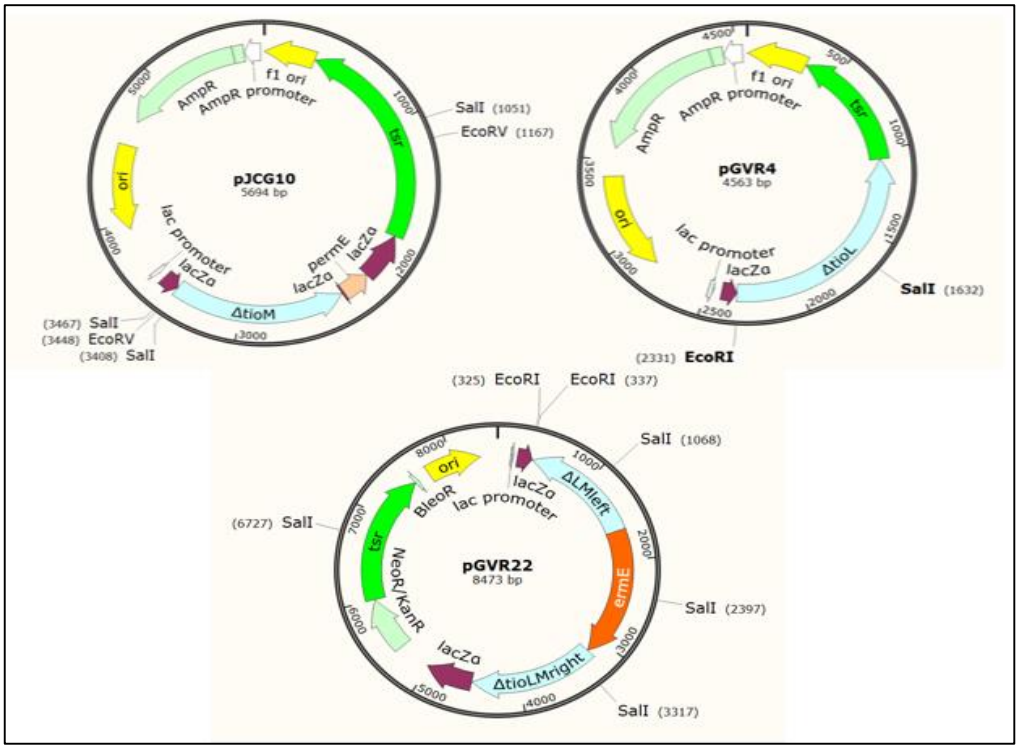


Figure 8: Construction plasmids pJCG10, pGVR4 and pGVR22, used for generation of the three mutant strains.

4.2 Culture conditions

The main culture media that have been used all along this work are schematically represented in Table 1 (Kieser et al., 2000). The ingredients of each media is referred to a total volume of 1 L of distilled water. After the preparation of the media they had to be autoclaved at 121 °C and 2 atm for 20 min, in order to get them sterilized.

Culture Media	Ingredients (1L)	Use
TSB	Pepton of casein → 17 g Pepton of soy→ 3 g NaCl→ 5 g Dipotassium phosphate→ 2,5 g Dextrose or Glucose→ 2,5 g	Growth in flasks for *S. albus*
Bennet	Glucose→ 10 g Pepton of meat→ 1 g Yeast extract→ 1 g Casein hydrolysate→ 2 g	Sporulation on plates for *S. albus*
R5A	Sacarose→ 103 g K_2SO_4→ 0,25 g $MgCl_2 \times 6H_2O$→ 10,12 g Glucose → 10 g Casein Hydrolizate → 0,1 g Casein Hydrolizate → 0,1 g Yeast extract → 5 g MOPS → 21 g Traces →2 mL	Growth in flask for *S. albus*
Medium A	MOPS → 21 g Glucose → 5 g MOPS → 21 g Glucose → 5 g Yeast extract → 0,5 g Meat extract → 0,5 g Casa-aminoacids → 1 g	Growth on Petri dishes for *S. albus*

Medium K	Soybean meal→ 15 g Starch→ 25 g $CaCO_3$→ 4 g Dried Yeast → 2 g	Growth in flasks for *S. albus*
Medium A1	Starch→ 1 g Peptone → 2 g Yeast extract → 4 g	Growth in flasks for *S. albus*
Medium YEME 17%	Glucose→ 10 g Yeast extract→ 3 g Meat peptone→ 5 g Malta extract→ 3 g Saccharose→ 170 g After the autoclave add: $MgCl_2$ 2,5→ 2 mL Glycine 20%→ 37,48 mL	Preparation of *S. albus* protoplasts Extraction of genomic DNA from *S. albus*

Table 1: Schematic representation of used media.

The optimization of the culture of *S. albus* p-FL1049 and of its mutants forms have been carried out with different protocols in order to find the best parameters of growth and production of the desired metabolites. Therefore, depending on the final purpose, it could be distinguished two types of culture:

1. Plating out on solid agar plates of Bennet medium or A medium from a stock culture tube, or directly from a single colony.
2. Liquid culture in different types of liquid media (R5A, YEME 17%, K, A1, TSB) of 25 mL or 250 mL, depending on the final purpose.

Each culture medium had to be added with the relative antibiotic of resistance of each strain on order to obtain their selective growth both in liquid as solid.
The conditions of culture varied depending on the final aim. During this work there have been carried out different procedures of culture:

4.2.1 Sporulation of *S. albus* wild type and mutant forms

After melting a bottle of Bennet medium, the antibiotic of selection was added and the media was poured in plates. The plate was then divided in 4 quadrants in order to obtain a larger amount of spores from a single CFU under strictly sterilized

conditions. The plate was left into the incubator at 30 °C and for 3 days. After the appearing of white spores, 9 Bennet plates could be inoculated from a single quadrant in order to let them grow during 3 days at 30 °C and then proceed with the spores' quantification and with a glycerol stock of them.

4.2.2 Production of thiocoraline

The strain *S. albus* p-FL1049 and its mutant forms were inoculated in a liquid media (like for example R5A, K and A1) in flasks of 25 mL from 1 mL of a pre-inoculum of 10 mL in TSB. The culture was left grow for 3 days at 30 °C and 250 rpm.

4.2.3 Antibiotics

The antibiotics that have been used during this work are schematically represented in the Table 2.

Antibiotic	Solvent	Initial Concentration (mg/mL)	Final Concentration (mg/mL)
Ampicillin	H_2O	100	100
Apramycin	H_2O	200	15
Lyncomycin	H_2O	100	200
Thiostrepton	DMSO	50	100

Table 2: Schematic representation of the antibiotics used.

4.3 Selection procedure for mutant testing

A selective growth for the testing of the double mutant has been carried out by using different culture conditions in Bennet plates, since this method allowed both the differentiation between the original strain (*S.albus*-pFL1049) and the efficacy of the procedure of double crossing which was used to obtain the double mutant of *tioL* and *tioM* genes.

An aliquot of 20 µL from a glycerol stock of *S. albus*-pFL1049-*ΔtioLM* was inoculated in a Bennet plate together with the corresponding antibiotic of selection, the lincomycin (which is analogous to erythromycin, the antibiotic of resistance for this mutant). After 3 days growth of the mutant, a ¼ amount of the plate was recovered in an eppendorf tube with a solution of saccharose 10.3% and X-triton 0.1%. The mix obtained was then filtrated in sterilized conditions and the tube was centrifuged at 12000 rpm for 5 min. After discharging the supernatant, the pellet was then washed

for two times with 1 mL of saccharose 10.3%. Once obtain the washed pellet it was diluted in 1 mL of saccharose 10.3% a series of decimal dilution was carried out, in order to have a lower amount of cells that would be used to obtain a culture plate with isolated colonies in the next step.

So, an amount of 100 µL of the 10^{-1}, 10^{-2} and 10^{-3} cell dilutions was inoculated in 4 Bennet plates for each dilution, in the presence of lyncomycin at 200 mg/mL.

After a 7 day incubation period at 30 ° C, a single colony from the plate of 10^{-3} cell dilution, was replaced in two different labeled Bennet plates with the help of a sterilized loop. The homologous plates differed each other from the antibiotic which in one case was absent and in the other was thiostrepton (the selective antibiotic for the plasmid used in the construction of the double mutant *S. albus*-pFL1049-*ΔtioLM)*.

4.4 Molecular biology procedures

Molecular biology procedures have been carried out in order to test and prove the mutants forms.

4.4.1 Extraction of genomic DNA

The extraction of genomic DNA from the strain *S. albus* was obtained by using a 1.5 ml volume of a one-day culture that had been grown in a 10 ml flask of YEME 17% medium at 30 °C and 250 rpm. The protocol used was described in the kit of Thermo Scientific Genomic DNA Purification Kit. Briefly, after centrifuging the sample for 7 min at 10000 rpm, the supernatant was discharged and a volume of 180 µL Resuspension Solution buffer was added to the eppendorf tube. Since the proper resuspension of the pellet, the tube was put at 37 °C for 30 min. Then, a mix of 200 µL of lysis solution and 20 µL of Proteinase K was added to the sample, which was then incubated at 56 °C for another half an hour, until the solution became visible clear. So, a series of washing passages (firstly with ethanol, then with Wash Buffer I and II) and purifications through a column allowed the discarding of any cellular debris. Finally, the elution of DNA in a new eppendorf tube was achieved. To test the effectiveness of the protocol and at the same time to obtain the DNA concentration, an absorbance measure at 260 nm was performed, by the use of a spectrometer Biophotometer (Eppendorf).

4.5 Analytical chemistry procedures

A series of analytical procedure have been carried out in order to extract and identify the microbial compounds and metabolites.

4.5.1 Thiocoraline extraction

The protocol of thiocoraline extraction was performed in order to optimize its analysis in control and mutant samples. Generally, it was achieved by a first treatment of a defined amount of culture cellular pellet with a 100% acetone solution in a 2:1 proportion in order to effectively disrupt the cells. After an incubation period of about one hour and a half, the sample was centrifuged at 10,000 rpm for 5 min in order to clearly separate the acetone phase from the cellular debris. After the evaporation of the total amount of acetone phase, 500 µL ethyl acetate were added. After a brief vortex of the samples, a centrifuge step of 5 min at 14,000 rpm was carried out, in order to separate the organic phase from the water one. Finally the organic phase of interest was transferred into a new tube and the solvent was then evaporated in vacuum in order to obtain the dried extract which was characterized by a yellow appearance. Extractions from culture supernatants were carried out starting in the ethyl acetate step.

4.5.2 3HQA intermediates extraction

The presence of 3HQA intermediates has been investigated during this work in the mutants and control strains for thiocoraline production, with a particular focus on β-OH-tryptophan, *N*-formyl-L-kyneurine and L-kyneurine.

The polar nature of these aromatic intermediates determined the setting up of an extraction protocol for both the organic and the aquatic phase. The protocol consisted on a first destruction treatment of the cells by using of a 100% acetone solution in a 2:1 proportion. After a waiting period of about 1 hour, the samples were centrifuged at 10,000 rpm for 5 min. Then, the acetone phase was transferred to other new tubes in which a corresponding volume of ethyl acetate was added in a proportion of 1:1. After vortexing the samples, a centrifugation step of 5 min at 14,000 rpm was applied in order to separate the organic phase from the water one. Both phases were let evaporated in vacuum, until obtaining of a dried pellet.

4.5.3 Thiocoraline and 3HQA intermediates analysis by HPLC-MS

The qualitative and quantitative measurement of thiocoraline and 3HQA intermediates was achieved with a liquid chromatography-mass spectrometry (HPLC-MS) in which the physical separation guaranteed by the HPLC is combined with the mass analysis capability of the MS. The knowing of thiocoraline fragmentation pattern and retention time allowed its accurate detection as a mass of 1157.2476 m/z in a retention time of 14.40 min.

All the extracts were first dissolved in a 1:1 DMSO:Methanol solution just prior to HPLC-MS analysis. Depending on the concentration of the samples and the interference of some medium derived compounds, the extracts could have been diluted and centrifuged prior the analysis.

The equipment used for the detection was an HPLC Dionex Ultimate 3000 (Thermo Scientific, Bremen) and the chromatographic column was a Waster Cortecs with a C18 stationary phase, 2.1x150 mm size and a particle diameter of 2.7 μm. The 0.25 mL/min mobile phase used for the analysis was:

- Phase A: 0,1% formic acid in water
- Phase B: 0,1% formic acid in acetonitrile

The program lasted 30 min in which a gradient of Phase B was applied to the column in order to have the following distribution:

- A column stabilization phase of 16 min with a 2% Phase B;
- A first 5 min elution with 100% Phase B
- A final elution of 9 min with 2% Phase B

The injection volume was 0.3 or 1 μL, depending on the sample concentration.

The parallel MS analysis was made with the equipment Bruker Impact II Q-ToF, by using the method ESI and a scanning of a range of masses between 100 and 2000 m/z.

The results obtained with the HPLC-MS analysis were later processed and elaborated by using the analysis software Bruker DataAnalysis 4.3.

5. Results and Discussion

5.1 Genetic testing of the mutants

The first step in this work was to test the three mutant strains of *S. albus*-pFL1049: *s ΔtioL*, *ΔtioM* and *ΔtioLM*.

In particular, two methodologies were developed:

- A Southern Blotting for the single mutants test
- A selection culture for the double mutant test

5.1.1 Southern blot testing of *S. albus*-pFL1049-*ΔtioL* and *S. albus*-pFL1049-*ΔtioM* mutant strains

The Southern blot was previously obtained from other studies and demonstrated the presence of the mutated versions of *tioL* and *tioM* genes in the genomic DNA of the strains studied herein. In particular, the DNA probes used in the hybridization step of this Southern blot were the same plasmid DNA that were used as vectors for the disruption gene procedure in a way that their hybridization with the control strain *S. albus*-pFL1049, the *S. albus* pFL1049-*ΔtioL* and the *S. albus* pFL1049-*ΔtioM* resulted with a band pattern of respectively 2.5 kbp, 1.2 kbp+6.3 kbp and 2.4 kbp +5.9 kbp (Figure 10).

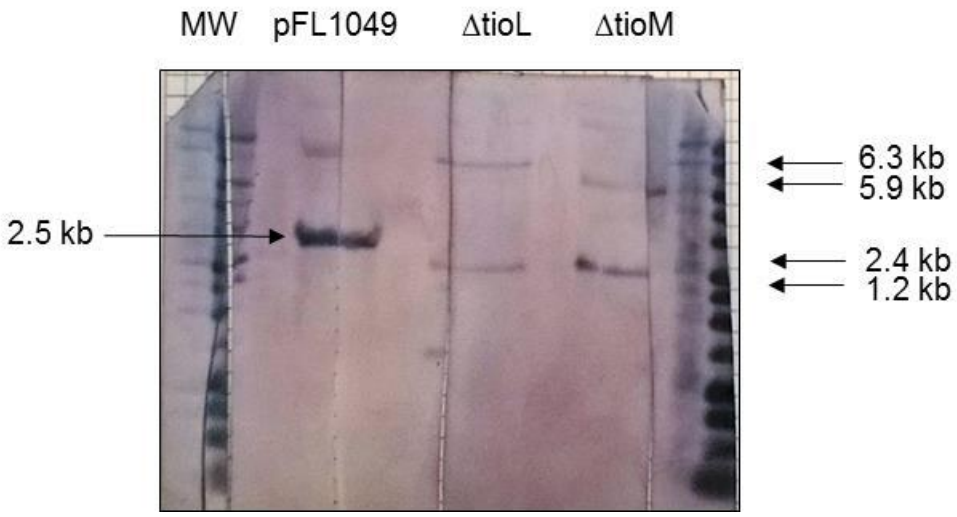


Figure 10: Southern Blotting of *S. albus*-pFL1049, *ΔtioL* and *ΔtioM* strains.

5.1.2 Selection culture of the double mutant

The application of a microbiological experiment to test the double mutant was possible thanks to the use of different antibiotics for the genetic construction of *S. albus*-

pFL1049-*ΔLM* strain. In fact the plasmid that had been used to obtain the genetic construction of the double mutant (the pGVR22 in the Figure 7) contained the antibiotic resistance of thiostreptone which was then lost during the double cross-over event in which the erythromycin resistance gene replaced the coding sequences of both *tioL* and *tioM*. So, the differential growth of the same CFU previously obtained from a Bennet plate with lincomycin.

The impossibility for the *S. albus*-pFL1049-*ΔLM* strain to grow in a thiostreptone media, as it could be seen in Figure 11, make it possible to confirm the occurred insertion event of the erythromycin resistance and the absence of the plasmid DNA used for this purpose which contained the thiostreptone resistance coding gene.

At the same time, a genetic evidence needs to be carried out in order to complement and definitely confirm the correct insertion event that occurred in *S. albus*-pFL1049-*ΔLM* strain even if there was not the time to include these results in this work.

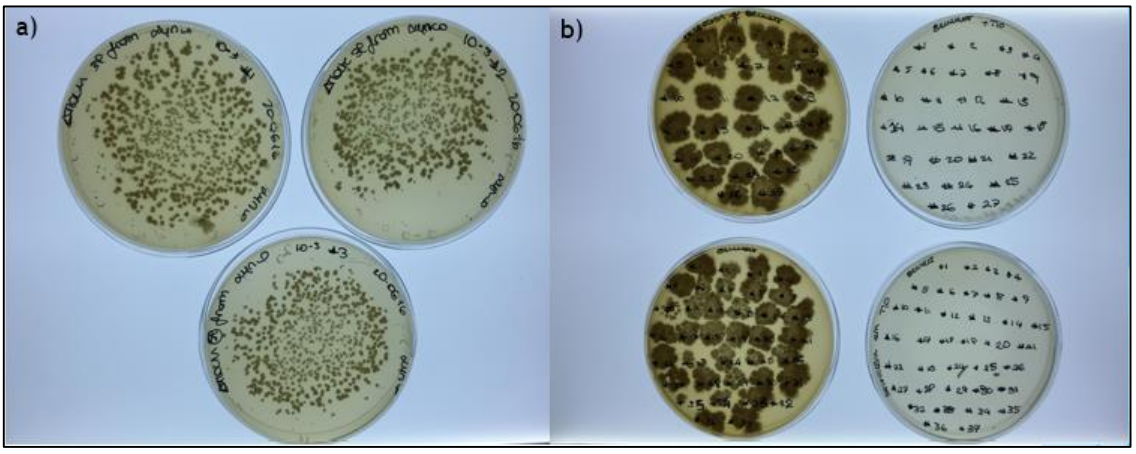


Figure 11: Photos of the procedure to test the *S. albus* pFL1049-*ΔtioLM* strain in a) the step of isolation of single colonies in lincomycin Bennet plates and b) the step of selective growth of the same colony with (on the right) or without thiostreptone (on the left)

5.2 Culture optimization of thiocoraline production

Thiocoraline production has been optimized by using 3 different culture media (media R5A, media A1 and media K) and verifying its production every 24 h during 3 days. In order to do that, 1 mL of a 24 h preculture in TSB medium from each strain (inoculated with ¼ plate of substrate mycelium from Bennet Petri dishes of each strain) was used in each case as inoculum starting material for these 250 mL buffled flasks containing 25 mL of the corresponding medium.

The results obtained from this experiment are shown in Figure 12.

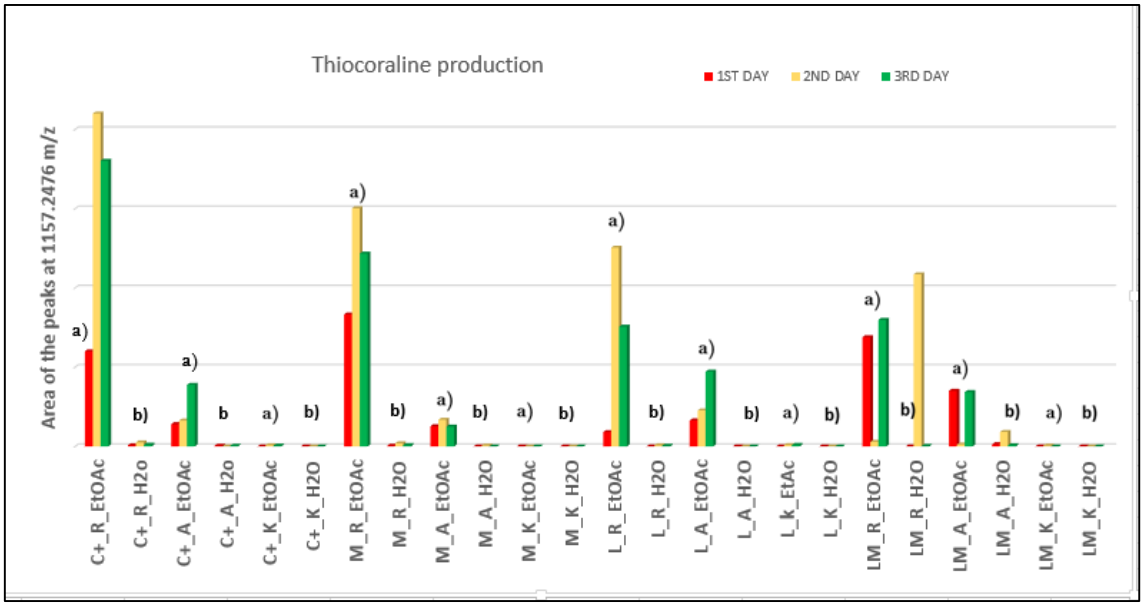


Figure 12: Histogram showing the comparison of thiocoraline production levels in *S. albus*-pFL1049 (C+ in the graph) in contrast with the mutant strains (*S. albus*-pFL1049-*ΔtioM*, *S. albus*-pFL1049-*ΔtioL*, *S. albus*-pFL1049-*ΔtioLM)* at days 1, 2 and 3 and in the 3 tested media (R5A, A1 and K). "a" indicates the extracted thiocoraline from cell pellet, and "b" from culture supernatant.

The results obtained from this study clearly demonstrate the following statements:

1. The R5A media is the best one for thiocoraline production in *S. albus*

2. Thiocorlaine is mainly retained inside the cells, as its extraction is mainly observed in the cell pellet (the bars targeted with an "a")

3. Under these culture conditions, we were unable to compare the four strains with respect to final amounts of thiocoraline, as these experiments were done only to compare culture media and the best extraction protocol

5.3 Comparison of production levels of thiocoraline

Since the mutants were unable to sporulate there was no possibility a priori to know the exact amount of cells in each culture of the previous experiment. So, a different strategy has been set up in order to obtain standardized conditions. A preinocula was made in 10 mL of R5A in order to get a previous grown from each stock of the tested strains. The comparison of thiocoraline production levels was achieved by carefully weighing a fixed amount of culture pellet that was 0.2 g ± 0.1% in 15 mL falcon tubes.

These 25 mL R5A cultures were carried out in 250 mL buffled flasks as described in section 5.2.

In order to obtain reliable data about thiocoraline production of the mutant strains studied, we also included the complemented mutant strains which were previously obtained from other works. Briefly, by including these 3 complemented mutant strains, we could study the effect of overexpression of *tioL* and *tioM* genes on thiocoraline production, as these genes were independently cloned in a multicopy number plasmid, and overexpressed under the control of the constitutive P_{ermE}* promoter.

The extraction of thiocoraline was achieved by using the standard protocol that has been previously described in the section of "Material and Methods". The subsequent analysis of the organic phase extracts by HPLC-MS gave reason to the results shown in the Figure below.

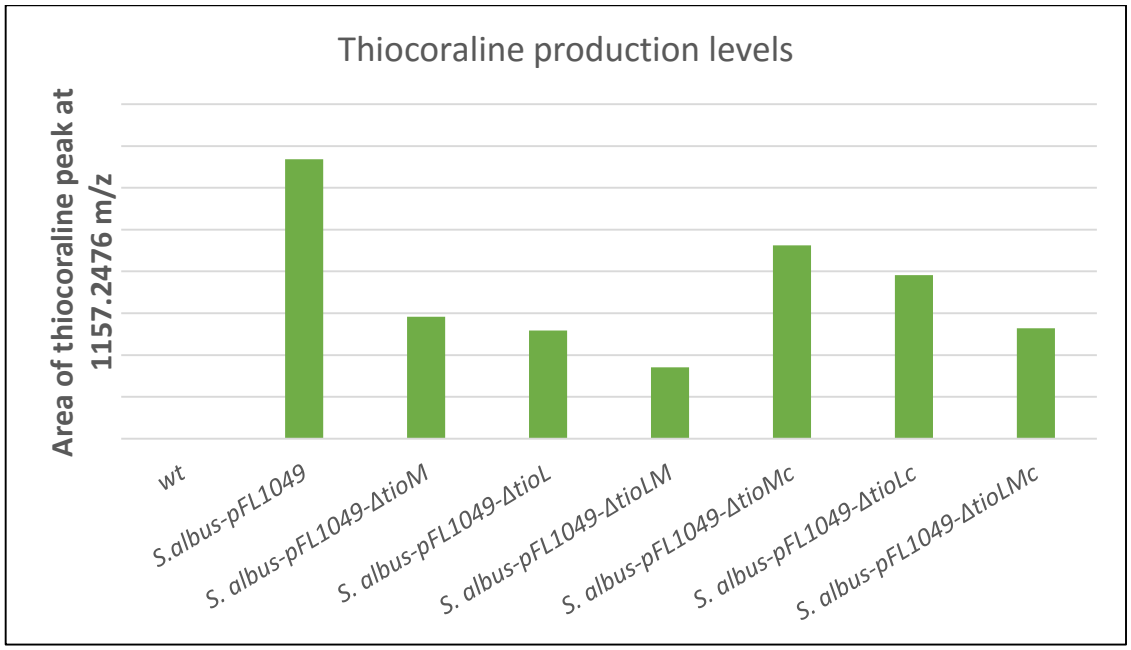


Figure 13: Graphic representation of thiocoraline production levels for wild type strain of *S. albus* (wt), *S. albus*-pFL1049 strain, its mutant forms (*ΔtioM*, *ΔtioL* and *ΔtioLM*) and the corresponding complemented versions (*ΔtioMc*, *ΔtioLc and ΔtioLMc)*.

The results obtained clearly define the following statements:

1. Thiocoraline production levels is clearly diminished in each *ΔtioM*, *ΔtioL* and *ΔtioLM* mutants, with lower production levels in the double mutant strain.

2. The complemented mutant strains *ΔtioLc* and *ΔtioLMc* seem to be more affected by the genetic manipulation compared to *ΔtioMc*.

3. Total thiocoraline production levels are never totally recovered in the complemented strains to the *S. albus*-pFL1049 strain levels.

5.4 Identification of 3 HQA intermediates

In relation to the previously announced hypothesis about the probable formamidase roles of *tioL* and *tioM* genes, we set up in this work an experiment of searching for 3HQA intermediates both in *S. albus*-pFL1049 strain and in the 3 mutant strains.

In particular, referring to the Figure 6, the extracts obtained from the mentioned strains were analyzed for the presence or absence of the probable β-hydroxytryptophan, N-formyl-3-hydroxykynurenine, 3-hydroxykynurenine, 3,4-dihydroxyquinaldic acid and 3-hydroxyquinaldic acid.

Due to the chemical properties of all these compounds, which show a very high polarity, we analyzed in the HPLC-MS only the water phase after acetone extraction.

The absence of neither a specific pattern of separation nor a known retention time in the HPLC for any of the searched compounds results on a hard work of analysis and study of all the chromatograms obtained from the HPLC-MS analysis.

5.4.1 Identification of 3-hydroxytryptophan

The β-3-hydroxytryptophan (Figure 14) is obtained from hydroxylation of β-Trp in the first step of 3HQA biosynthesis (Figure 6) and is characterized by an exact mass of 220.0848 m/z and a very high polarity demonstrated by its high tPSA (Topological Polar Surface Area) that is 95.58.

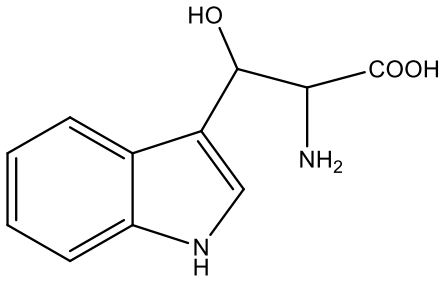


Figure 14: β-3-hydroxytryptophan chemical structure

The HPLC-MS analysis was set up in order to find its protoned mass of 221.092 m/z. In absence of specific information about retention time and fragmentation pattern of this compound, the HPLC-MS analysis identified the searched mass of interest in all

the sample chromatogram giving a huge amount of reasonable peaks at different retention times and with different fragmentation pattern.

First of all, the main screening was to turn the attention only on the lower retention time peaks of each sample because of the polar property of the compound.

Then, the peaks with higher values and a well-defined height were selected in order to being deeply analyzed by using the DataAnalysis software.

A first analysis was performed by comparing the chromatogram profiles of the wild type strain *S. albus* in contrast with the strain *S. albus*-pFL1049 in order to get a reference of an expected peak corresponding to β-3-hydroxytryptophan mass.

As it could be observed from the Figure 15, the analysis revealed the presence of the mass of interest at different time of retention and with a different pattern in both the wild type strain (yellow chromatogram) and in the thiocoraline control strain (red chromatogram). Avoiding the peaks matching with those of the wild type, the higher peak at 4.46 min in the *S. albus*-pFL1049 chromatogram was selected in order to deeply being speculated in its spectrum. As could be seen from Spectrum View in the second image of Figure 15, the corresponding peak of 221.091m/z is not the predominant one in pFL1049 strain, but is 298.1145 m /z.

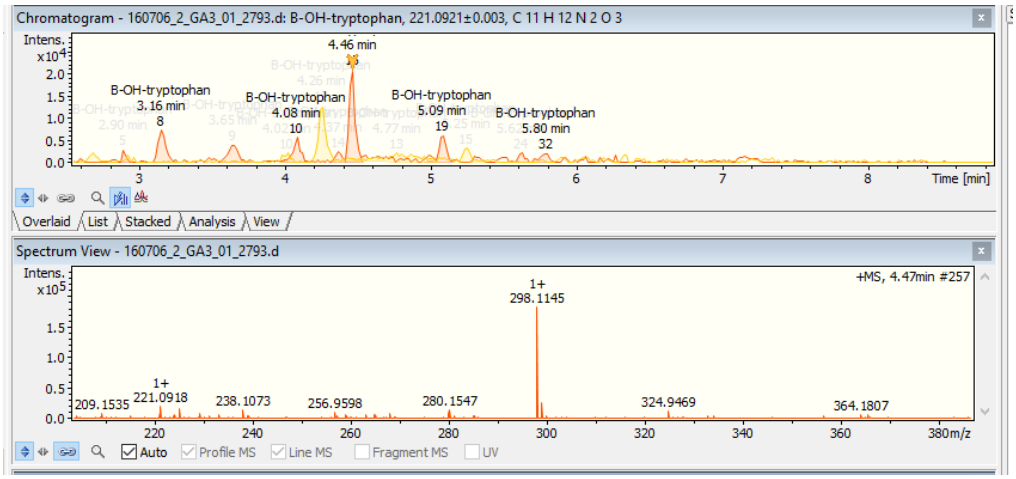
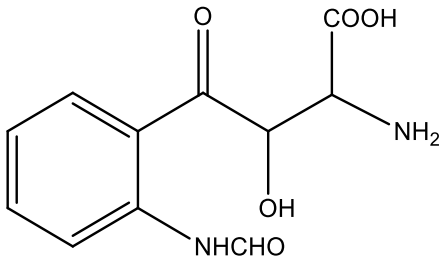


Figure 15: Comparison of wild type (yellow) and positive control pFL1049 (red) chromatograms for β-OH-Trp searching. Analysis of an interesting peak at 4.47 min.

The same ionization and fragmentation data were obtained for the other higher peaks (3.16 min, 4.08 min, 5.09 min and 5.8 min) in the positive control and in the mutant strains. Moreover a similar profile of retention time for the peaks found in positive control and mutant strains was not detected.

5.4.2 Identification of *N*-formyl-3-hydroxykynurenine

The *N*-formyl-3-hydroxykynurenine is obtained from the β-hydroxy-L-Trp thanks to the indole ring opening by a Trp-2,3-dioxygenase TioF, as could be seen in the Figure 6. It is a very polar compound as demonstrated by its high tPSA value of 129.72, so it is thought to have a very low retention time in the HPLC. Its exact mass is 252.0746 m/z and the corresponding protoned mass is 253.0809 m/z.



Exact Mass: 252,07

Figure 16: Chemical structure of *N*-formyl-3-hydroxykynurenine

The same procedure which has been illustrate above was applied for the analysis of the chromatograms obtained in the different strains for the presence of a 253.0809 m/z compound. As the last analysis, in Figure 17 there is shown the overlapping of the chromatograms of the wild type strain (the yellow chromatogram) and the positive control *S. albus*-pFL1049 strain (the red chromatogram).

The searching for *N*-formyl-3-hydroxykynurenine peak revealed an interesting peak in *S. albus*-pFL1049 chromatogram at a retention time of 2.2 min which was absent as in the wild type (yellow chromatogram in Figure 17) as in the mutants (data not shown).

An analysis of the peak at 2.22 min of retention revealed its specific mass spectra in which bigger corresponding peaks were also visible. So it is suggested that the 256.9597 m/z mass found at this retention time only corresponds to a derivative fragment of a bigger compound as it could be seen in Figure 17.

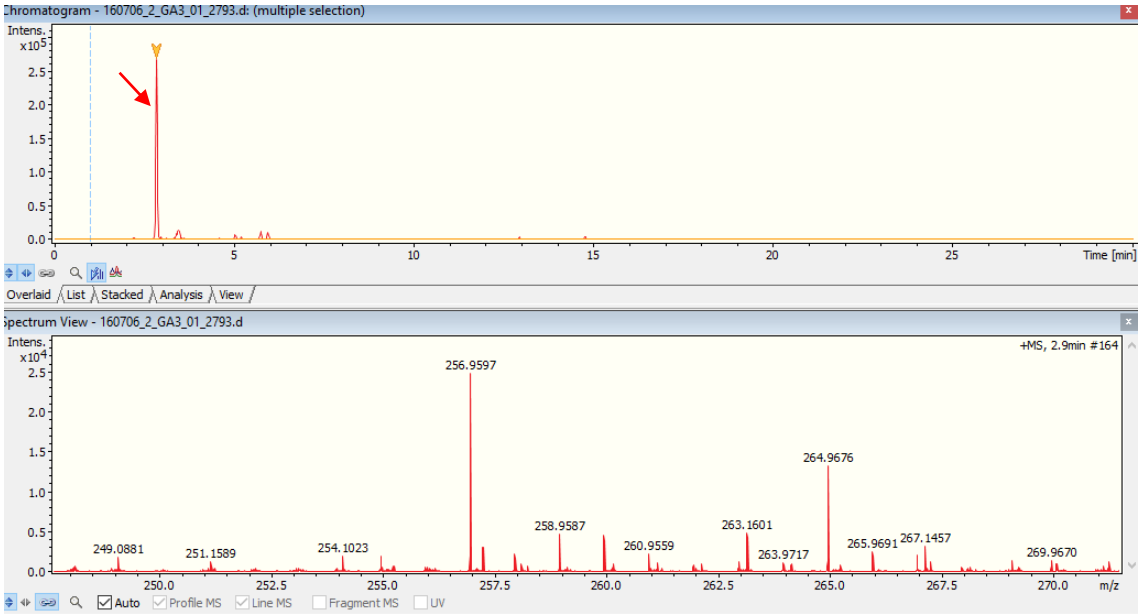


Figure 17: Overlapping chromatograms for *N*-formyl-3-hydroxykynurenine in wild type (yellow) and positive control *S. albus*-pFL1049 (red) strains. Analysis of an interesting peak at a retention time of 2.2 min.

5.4.2 Identification of 3-hydroxykynurenine

The 3-hydroxy-kynurenine is formed by the elimination of a formyl group from the previous compound by the action of an unknown enzyme whose function have been hypothetically attributed to the coding genes which are the object of this study.

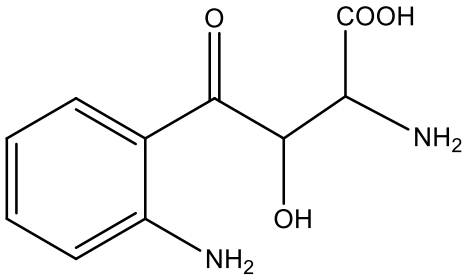


Figure 18: Chemical structure of 3-hydroxykynurenine

The 3-hydroxykynurenine is a polar compound (tPSA value of 126.64) with an exact mass of 224.0797 and a corresponding protoned mass of 225.086 m/z.

The overlapping of each strain chromatogram for the searching of 3-hydroxykynurenine showed the presence of an interesting peak in the positive control strain, *S. albus*-pFL1049 at a retention time of 2.80 min, which is reasonable considering the high polarity of this compound.

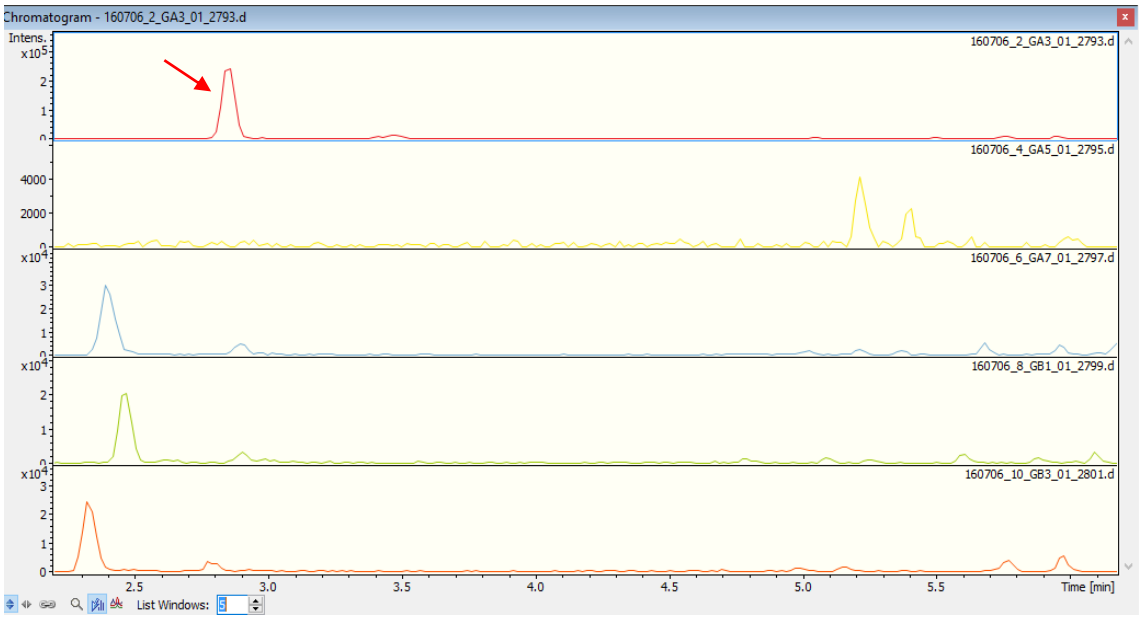


Figure 18: Overlapping chromatographs for 3-hydroxykynurenin research in extracts from wt (yellow), *S. albus*-pFL1049 (red), Δ*tioL* (green), Δ*tioM* (light blue) and Δ*tioLM* (orange) strains.

The analysis of the peak observed at a retention time of 2.8 min was carried out in the control strain *S. albus*-pFL1049 (Figure 20) since it was not present in the mutant strains as it could be seen in Figure 19 where respectively the light blue, green and orange are for *ΔtioM*, *ΔtioL* and *ΔtioLM S. albus* strains.

A deeper analysis of the mass spectra at this retention time show that apparently the 225.0867 m/z compound is the predominant one, which gave reason to likely attribute it to 3-hydroxykynurenine (the images in the bottom of Figure 19).

The absence of a corresponding peak at a retention time of 2.8 min in the other mutant strains could be reasonable explicated by the fact that maybe this compound is worst produced by the mutants, due to the absence of a putative protein (encoded by *tioM* and *tioL* genes) involved in a hypothetical regulation step of *N*-formyl-3hydroxy kynurenine conversion into 3-hydroxykynurenine.

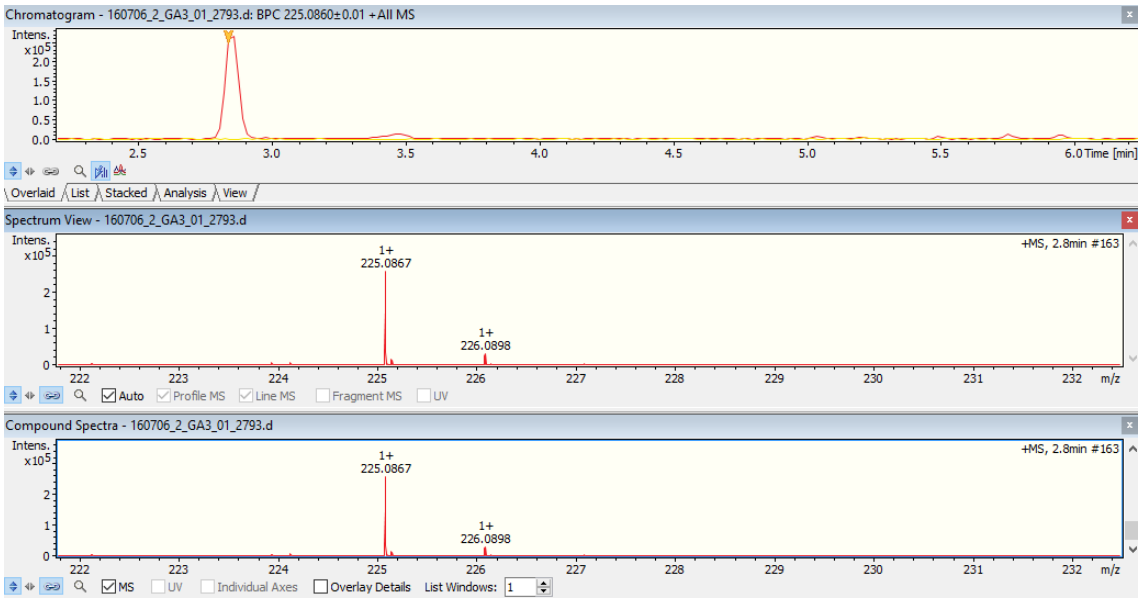


Figure 20: Overlapping chromatograms for 3-hydroxykynurenine in wild type (yellow) and positive control *S. albus*-pFL1049 (red) strains. Analysis of an interesting peak at a retention time of 2.8 min.

5.4.3 Identification of 3,4-DHQA and 3HQA

The 3,4-dihydroxyquinaldic acid is generated from the action of the aminotransferase TioG, and the 3HQA that is finally released by the action of a TioH oxidoreductase. These reactions represent the two final ones for the starter unit biosynthetic pathway.

Both of them are polar molecules that are characterized respectively by an exact mass of 205.04 m/z (tPSA of 90.12 for 3,4-dihydroxyquinaldic acid), and 189.0426 m/z (tPSA of 69.89 for 3HQA).

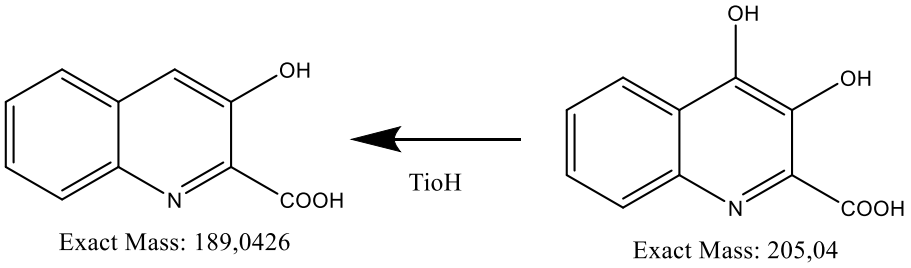


Figure 21: Chemical structure and schematic derivative reaction of 3HQA (on the left) from 3,4DHQA (0n the right)

The two final compounds of 3HQA biosynthesis were supposed to be found together in the HPLC analysis but no reasonable amounts of 3,4-dihydroxy quinaldic acid were identified, as it could be observed in the chromatograms overlapping of the samples in

the Figure 21, where a clear and high peak is not comparing except from one at retention time of 5.27 min which not appear in the wild type strain.

The unspecific peaks observed for 3,4 DHQA could hide the peak of interest.

In fact, the supposed lower production could be correlated with the fact that TioH reaction is so effective that all its substrate (3,4DHQA) is readily transformed into 3HQA.

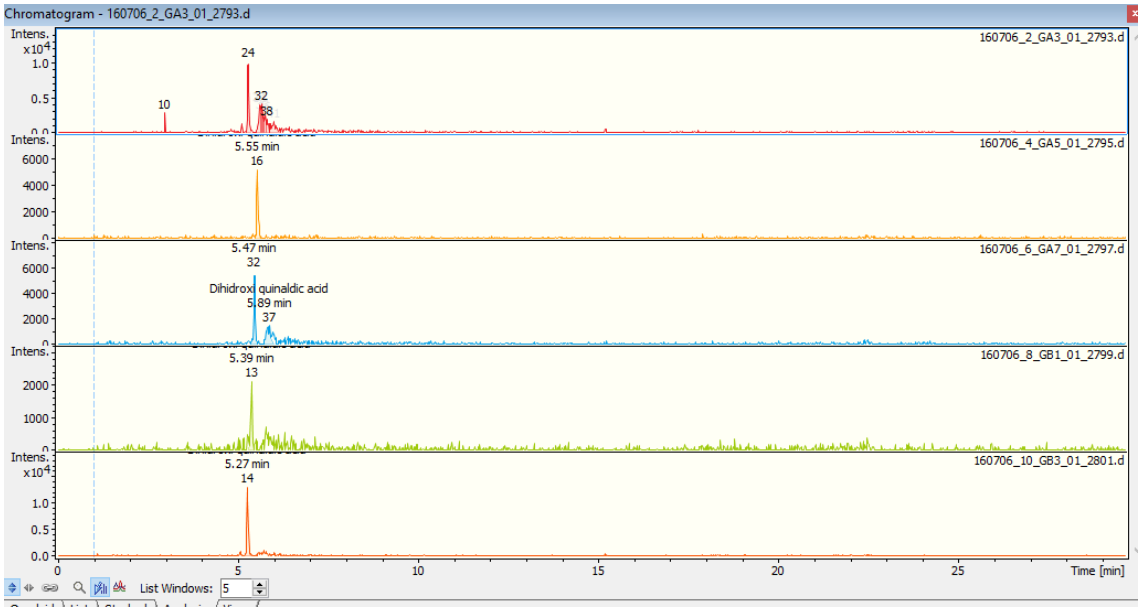


Figure 22: Chromatograms corresponding to the 3,4DHQA m/z in extracts from wt (yellow), *S. albus*-pFL1049 (red), Δ*tioL* (green), Δ*tioM* (light blue) and Δ*tioLM* (orange) strains.

At the same time, the searching for 3HQA (189.0426 m/z and positive ion 190.0499 m/z) was carried out. As it could be observed from the overlapping chromatograms in the Figure 22 there are at least two high interesting peaks which appear in pFL1049 control and in the mutant strains, while in the wild strain the pecks observed at this retention time could be easily related to background noise because of their low values.

So, the peaks which appear at 5.95 min and at 5.75 min are analyzed in Figure 23.

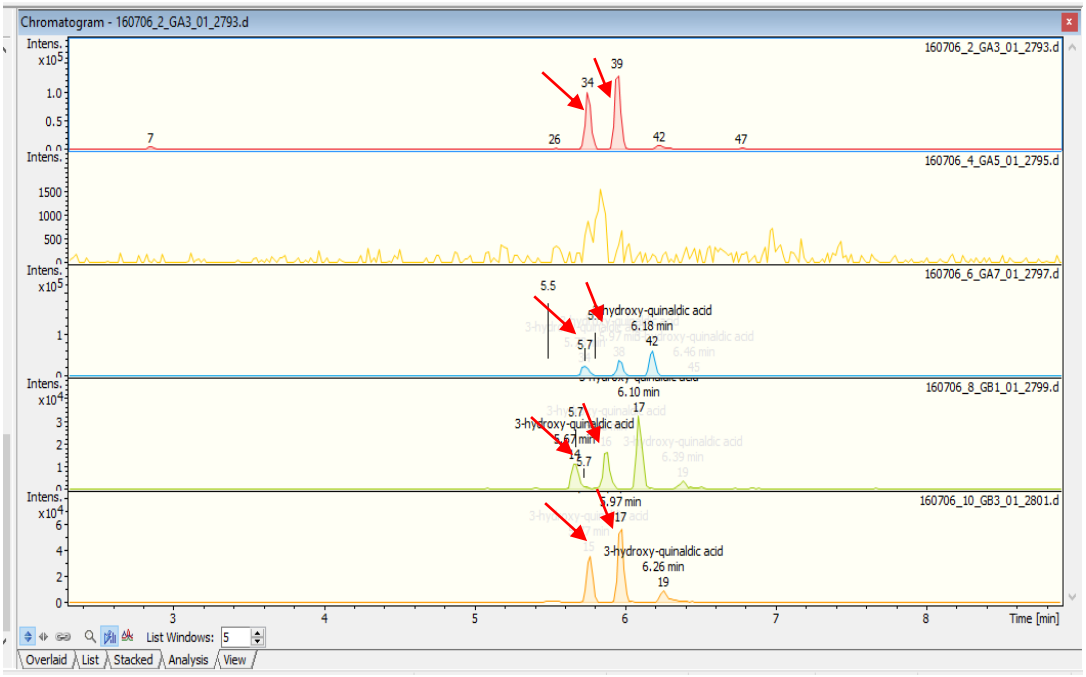


Figure 23: Overlapping chromatograms for 3 HQA peaks in extracts from wt (yellow), *S. albus*-pFL1049 (red), Δ*tioL* (green), Δ*tioM* (light blue) and Δ*tioLM* (orange) strains.

Firstly, it was analyzed the higher peak at the retention time of 5.95 min. The peak analysis in Figure 23 shows that actually the peak at 5.95 min corresponds to a mass spectra in which the compound of 190.0492 m/z could represents a probable adduct or derivative of the bigger one 267.097 m/z compound as it was demonstrated by the same repetitive pattern of fragmentation also in the chromatograms of the mutant strains.

Moreover the availability of a tool for calculating the previsted chemical formula corresponding to the compound of 190.0492 m/z showed 5 different formulas which could be related to that mass in a way that its probability of being the compound of interest decrease even more.

.

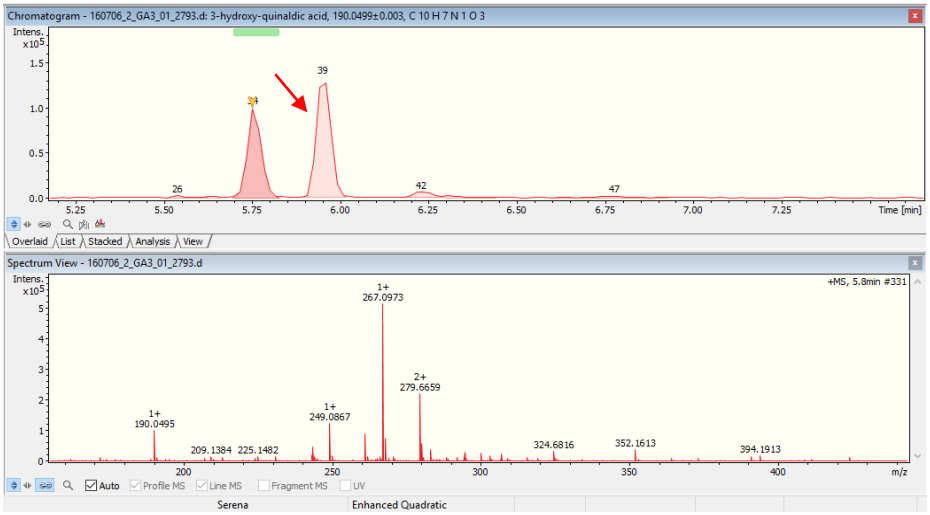


Figure 23: Analysis of the peak at retention time of 5.95 min in the positive control strain, *S.albus*-pFL1049 (red chromatogram of the Figure 22)

The other peak of interest compared at a retention time of 5.76 min and was analyzed by using the same procedure describe above in order to look for its reliability.

In this case, the peak investigation in the control strain demonstrated the presence of a corresponding peak of 190.0495 m/z mass higher than the previous one. Also, it was not accompanied by other bigger ions, since its spectra peak perfectly fit with that of the retention time firstly identified (Figure 24).

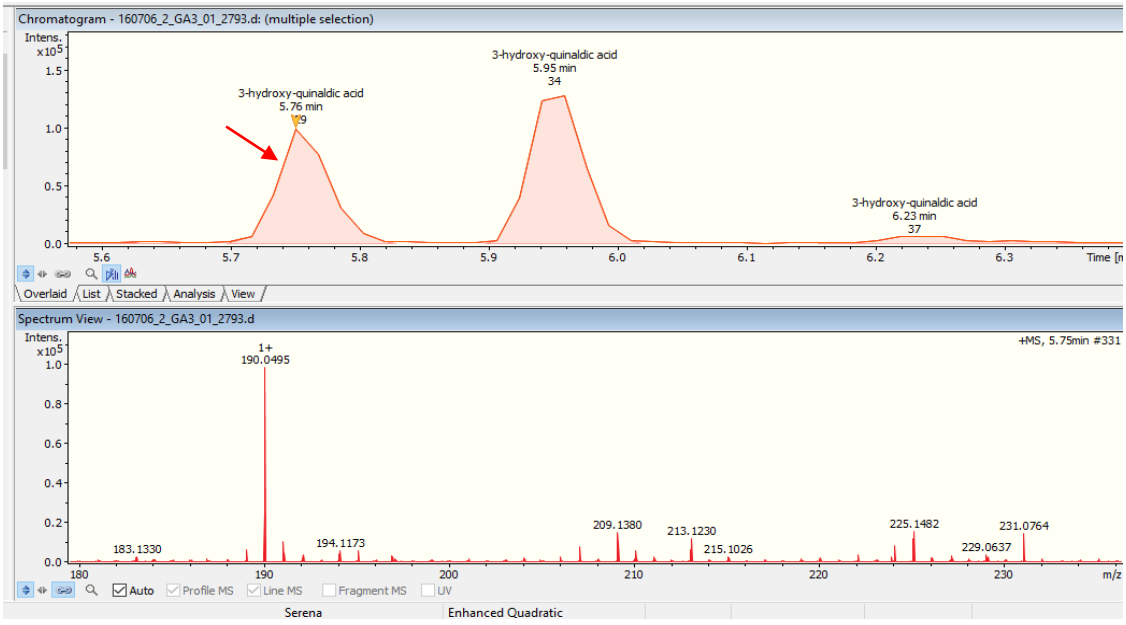


Figure 24: Compound spectra analysis at the retention time of 5.76 min in the positive control, *S.albus*-pFL1049 (red chromatogram of the Figure 22)

The candidate peak for 3HQA was also tested in the *S. albus*-pFL1049-***ΔtioM*** strain (the light blue chromatogram in Figure 25) and a similar spectra pattern was found.

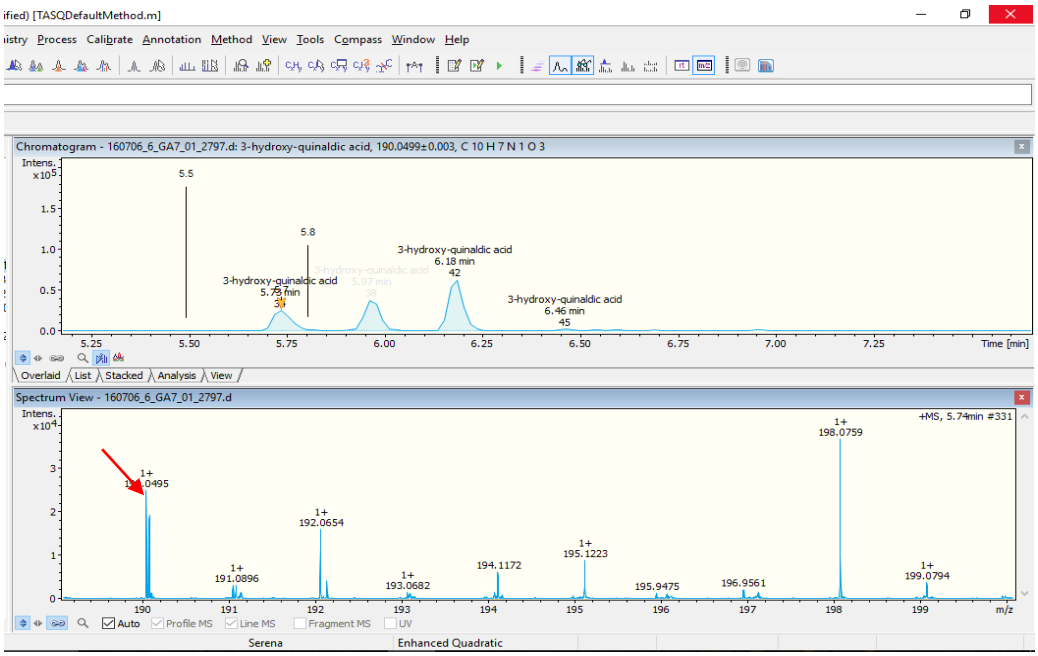


Figure 25: Analysis of 5.74 min peak in *S. albus*-pFL1049-***ΔtioM***

The compound found was then investigated for its specific retention time and its peak exactly corresponds to the retention time firstly observed in the chromatograms of both the control (Figure 24) and the ***ΔtioM*** mutant strain and that was 5.74 min.

Moreover, the prediction of the chemical formula was carried out thanks to an available tool in the software of analysis and it demonstrated that the compound only could correspond to the chemical formula of 3HQA ($C_{10}H_7N_1O_3$).

So, supposing the peak coincidence with the 3HQA, the chromatograms showed in Figure 25 gave reason of its lower production level in mutant strains compared to the control one (*S. albus*-pFL1049) as it was expected from the previous results for thiocoraline production levels.

5.5 Hypothesis for 3HQA production levels

Considering the results that have been previously elucidated about the probable identification of 3HQA in chromatograms of *S. albus*-pFL1049 and in its mutant strains, a schematic graph of these different production levels was obtained (Figure 26).

The area values of the peaks identified at a retention time of 5.74 min in all of these strains were represented in a histogram in order to have an idea about their

availability to synthetize thiocoraline starter unit, since the mutation of *tioL* and/or *tioM* coding genes was hypothesized to have important implications in 3HQA biosynthetic pathway.

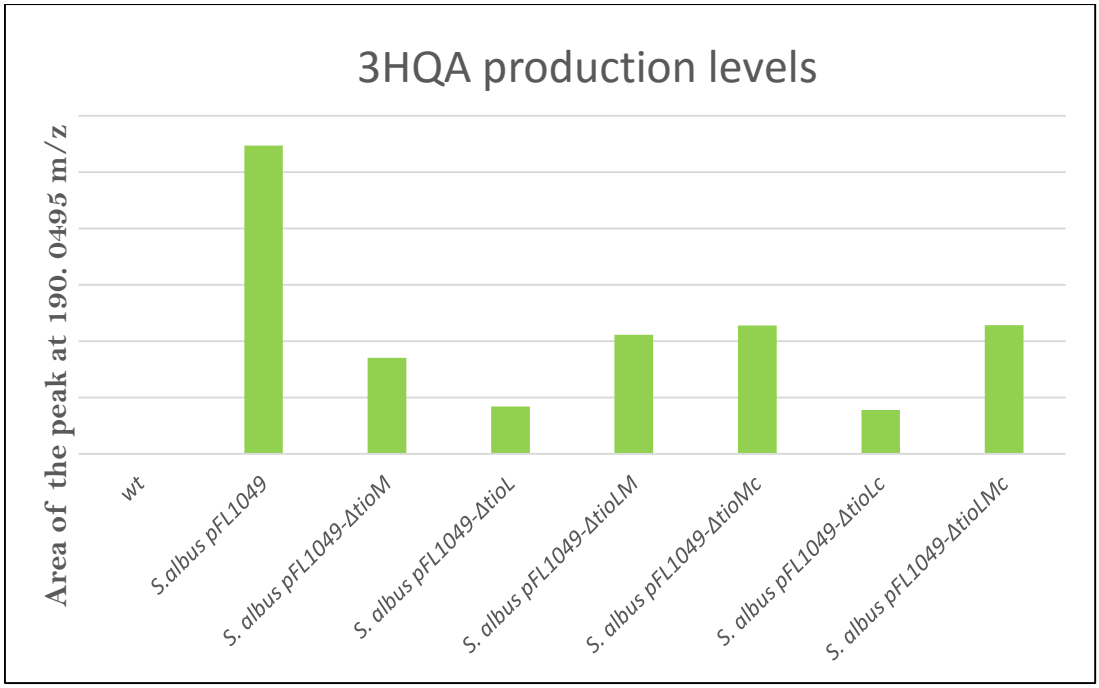


Figure 26: Comparison of the putative 3HQA production in all tested strains.

6. Conclusions

Considering the results explained before and in relation with the objectives that were firstly set up in this work it could be concluded that:

- The mutant strains *S. albus*-pFL1049-Δ*tioL*, *S. albus*-pFL1049-Δ*tioM*, *S. albus*-pFL1049-Δ*tioLM* are definitely mutant strain versions of thiocoraline producer strain *S. albus*-pFL1049;
- Thiocoraline extraction in the original strains as in its mutant strains is optimized by using R5A as culture media and a period of growth of at least 3 days at 30 ºC and 250 rpm.
- The production of the antitumoral compound is greatly diminished in each of the three mutants in a very similar proportion each other, as it could be demonstrated by the corresponding differential production in complemented versions of the mutants. This result suggests that *tioL* and *tioM* genes are maybe involved in such a regulative role of thiocoraline production instead of a critically step point, which could have been true if there would have obtained a totally abolishment of its production.
- The presence of 3HQA intermediates was analyzed in the extracts of the different strains by HPLC-MS. In particular, an interesting peak has been attributed to 3HQA presence both in *S. alb*us-pFL1049 and in the three mutant strains at a retention time of 5.74 min.
- Another intermediate has been putatively detected, 3-hydroxykynurenine, at a retention time of 2.84 min which only appear in the control strain *S. alb*us-pFL1049, which is maybe correlated to the lower capability of the mutant to finalize thiocoraline production form its early stages. As far as the other intermediates could not have been detected, they are maybe correlated with a lower production level.
- The putative comparison of 3HQA production levels has been suggested. This confirms that even in the absence of *tioL* and *tioM* genes, thiocoraline and also its starter unit is till produced, although at a lower level than the original strain *S. albus*-pFL1049. So, their enzymatic attribution to the essential step of *N*-formyl-3-hydroxykynurenine conversion in 3-hydroxykynurenine may be

discharged and redirected to a spontaneous formamidase action. TioL and TioM proteins may act as regulators.

7. Future perspectives

The lessons learned from this study don't allow to elucidate the role of the proteins encoded by *tioL* and *tioM* genes in 3HQA biosynthesis as far as they seem not being directly or uniquely involved in its production. At the same time, a possible regulation role may be attributed to both TioL and TioM which could also act as a complex.

The isolation and purification of the proteins encoded by *tioL* and *tioM* genes could give more information about their physicochemical properties. A prediction of the molecular mechanism of action of the putative proteins could then be proposed, by means of computational studies based on known chemical characteristics.

Moreover an interesting point is to investigate for the production of novel and potential metabolites in mutant strains, which is an issue that could be afforded by the described elaboration procedure of the HPLC-MS chromatograms obtained from the mutant strains extractions. In fact, some new peaks have just been discovered in the available HPLC-MS data which need to be yet validated and tested with new experiments.

6. Acknowledgments

"Sii il cambiamento che vuoi vedere nel mondo" (Mahatma Gandhi).

When I think to the goals achieved during all along this Master, the idea of changing and breaking the molds that is characteristic of my personal attitude is coming to me. During this steeplechase a fundamental rule have taken the great guides I had the pleasure to meet on.

First of all, I would like to thank all the BIONUC research unit to let me take considerably part of their scientific group. That would have not been possible without the leadership of Felipe who has helped and encouraged me all along this work. Also, I would like to thank Nacho, Javi, Saúl, Coral and Kei for their precious advises, the suggestions, the comprehension and of course the atmosphere of happiness that they always generate in the lab.

Moreover I am very grateful to my Spanish an Italian coordinators, Maite and Giovanna, which gave me the opportunity to have a great experience in my professional carrier thanks to the setting up of an international curriculum that gave me the chance to open my mind and be consciousness of my design thinking abilities, professionally talking.

Of course, is never too much my gratitude towards my family that follow me from distance in every obstacle and success that I get encouraging me always in my choices.

At the end, but not the last one, this work could not have been possible without the everyday support of my Spanish family, Maria and Mimi.

6. References

Amador, M. L., Jimeno, J., Paz-Ares, L., Cortes-Funes, H., & Hidalgo, M. (2003). Progress in the development and acquisition of anticancer agents from marine sources. *Annals of Oncology*, *14*(11), 1607–1615.

Brown, D. D., Hitchcock, J. M., & Katz, E. (1980). Evidence for a Constitutive and Inducible Form of Kynurenine Formamidase in an Actinomycin-Producing Strain of Streptomyces parwulus, *Archives of Biochemistry and Biophisics*, *202*(1), 18–22.

Chater, K.F., and Wilde, L.C. (1976) Restriction of a bacteriophage of Streptomyces albus G involving endonuclease SalI. J Bacteriol 128: 644–650).

Das, L., Bhaumik, E., Raychaudhuri, U., & Chakraborty, R. (2012). Role of nutraceuticals in human health. *Journal of Food Science and Technology*, *49*(2), 173–183.

Fernández, J., Marín, L., Álvarez-Alonso, R., Redondo, S., Carvajal, J., Villamizar, G., Lombó, F. (2014). Biosynthetic modularity rules in the bisintercalator family of antitumor compounds. *Marine Drugs*, *12*(5), 2668–2699.

Koketsu, K., Oguri, H., Watanabe, K., & Oikawa, H. (2006). Identification and Stereochemical Assignment of the -Hydroxytryptophan Intermediate in the Echinomycin Biosynthetic Pathway, (8), 9–12.

Romero, F., Espliego, F., Baz, J. P., Quesada, T. G. De, Gravalos, D., Calle, F. De, Leon, A. De. (1997). Thiocoraline, a New Depsipeptide with Antitumor Activity Produced by a Marine Micromonospora I . Taxonomy, Fermentation, Isolation, and Biological Activities, 734–737.

Li, A. N., Li, S., Zhang, Y. J., Xu, X. R., Chen, Y. M., & Li, H. Bin. (2014). Resources and biological activities of natural polyphenols. *Nutrients*, *6*(12), 6020–6047.

Lombó, F., Velasco, A., Castro, A., De La Calle, F., Braña, A. F., Sánchez-Puelles, J. M., Salas, J. A. (2006). Deciphering the biosynthesis pathway of the antitumor thiocoraline from a marine actinomycete and its expression in two Streptomyces species. *ChemBioChem*, *7*(2), 366–376

Mady, A. S. a, Zolova, O. E., Millán, M. Á. S., Villamizar, G., de la Calle, F., Lombó, F., & Garneau-Tsodikova, S. (2011). Characterization of TioQ, a type II thioesterase from the thiocoraline biosynthetic cluster. *Molecular bioSystems*, *7*(6), 1999–2011.

Prasanna, S., & Doerksen, R. J. (2009). Topological Polar Surface Area : A Useful Descriptor in 2D-QSAR, *1848*, 21–41.

Therapy, E. (n.d.). Biotechnology & Experimental Therapy based on Nutraceuticals. Retrieved from http://www.unioviedo.es/IUOPA/index.php?view=article&id=152:biotecnologiayterapiaexperimentalbasadaennutraceuticos&format=pdf&option=com_content&Itemid=158&lang=en

Vidanarachchi, J. K., Kurukulasuriya, M. S., Malshani Samaraweera, A., & Silva, K. F. S. T. (2012). *Applications of Marine Nutraceuticals in Dairy Products. Advances in Food and Nutrition Research* (1st ed., Vol. 65).

Watanabe, K., Hotta, K., Nakaya, M., & Praseuth, A. P. (2009). Escherichia coli Allows Efficient Modular Incorporation of Newly Isolated Quinomycin Biosynthetic Enzyme into Echinomycin Biosynthetic Pathway for Rational Design and, 9347–9353.

Watanabe, K., Hotta, K., Praseuth, A. P., Koketsu, K., Migita, A., Boddy, C. N., Oikawa, H. (2006). Total biosynthesis of antitumor nonribosomal peptides in Escherichia coli, *2*(8), 423–428. Watanabe, K., Oguri, H., & Oikawa, H. (n.d.). Diversification of echinomycin molecular structure by way of chemoenzymatic synthesis and heterologous expression of the engineered echinomycin biosynthetic pathway.