## RESUMO

ORTIZ-CASTRO, R.E. **Caracterização de genes putativos biossintéticos do antitumoral cosmomicina D.** 2019. Tese (Doutorado em Microbiologia) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2018.

*Streptomyces olindensis* DAUFPE 5622 foi isolado na década de 1960 a partir de uma amostra de solo brasileira. Produz a antraciclina cosmomicina D, que tem atividade antitumoral e atraiu interesse por causa de seu padrão peculiar de glicosilação. Após anos de estudos sobre a molécula da cosmomicina e sua via biossintética, o estudo do cluster genético permitiu identificar os produtos codificados em cada ORF e agrupá-los, de acordo com sua funcionalidade, em genes envolvidos na biossíntese de agliconas, modificação da aglicona, biossíntese de açúcares, glicosiltransferases, genes de resistência e genes com função desconhecida. Entretanto o tamanho exato do cluster e o número de genes permaneceram desconhecidos.

Neste estudo concluímos que o cluster biossintético da cosmomicina D em *Streptomyces olindensis* é constituído por 38 ORFs, tendo como extremos definidos ORF17 *cosU* e ORF54 *cosV* G1Pdt (glicose-1-fosfato-timidiltransferase); e que o cluster da cosmomicina tem um tamanho de aproximadamente 40 kb (39,972 pb).

Estudamos também os genes *cosY* e *cosM*. Quando *cosY* foi inativado, a produção de cosmomicina aumentou, indicando uma potencial atividade reguladora. Para *cosK,* a inativação produziu um desvio na via biossintética levando à síntese de substâncias possivelmente inéditas. Estudos adicionais são necessários para estabelecer a real função desses dois genes.



**Palavras-chave:** *Streptomyces olindensis*, antraciclinas, cosmomicina, cluster biossintético, gene regulador.

## ABSTRACT

ORTIZ-CASTRO, R.E. **Characterization of putative biosynthetic genes of the antitumor cosmomycin D**. 2019. Ph.D. thesis (Microbiology) – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2019

*Streptomyces olindensis* DAUFPE 5622 was isolated in the 1960s from a Brazilian soil sample. It produces the anthracycline cosmomycin D, which has antitumor activity and has attracted interest because of its distinctive glycosylation pattern.

After many years and studies on the cosmomycin molecule and its biosynthetic pathway, the study of the gene cluster has allowed to identify the products encoded by each ORF and grouped them according to their functionality in genes involved in aglycone biosynthesis, aglycone modification, sugar biosynthesis, glycosyltransferases, resistance genes and genes with unknown function. However, the exact size of the cluster and number of genes remained unknown.

In this study we can conclude that the cosmomycin D cluster in *Streptomyces olindensis* is comprised of 38 ORFs, having as defined ends ORF17 *cosU* and ORF54 *cosV* G1Pdt (glucose-1-phosphate thymidyltransferase); and that the cosmomycin cluster has a size of approximately 40Kb (39.972bp).

We also studied the *cosY* and *cosM* genes. When *cosY* was inactivated the production of cosmomycin increased, indicating a potential regulatory activity. The *cosK* inactivation lead to a deviation in the biosynthetic pathway leading to the synthesis of probably new substances. Additional studies are necessary to establish the real functions of these two genes.



**Key words:** *Streptomyces olindensis*, anthracyclines, cosmomycin, biosynthetic cluster, regulatory gene.

## INTRODUCTION

### Actinobacteria: Streptomyces genus

Microorganisms of the genus Streptomyces spp. are members of the Actinobacteria phylum; are Gram-positive, non-acid-alcohol resistant bacteria, which form an extensively branched substrate and aerial mycelium. They are chemoorganotrophs and have an oxidative metabolism. The vegetative hyphae rarely fragment. Aerial mycelium forms chains of three to many spores when it ripens. Some species show short chains of spores in the mycelium of the substrate. Spores are not motile. The colonies are discrete and lichenoid, like leather. Often the colonies show a smooth surface at first, but later they develop a net of aerial mycelium that may appear flaky, granular, dusty, or velvety. They can produce a wide variety of pigments responsible for the color of the vegetative and aerial mycelium. They may also form diffusible colored pigments. Many strains can produce one or more antibiotic substances. They are catalase positive. They usually reduce nitrates to nitrites and degrade polymeric substances, as well as adenine and l-tyrosine. Most species use a wide variety of organic compounds as unique sources of carbon for energy and growth. The optimum temperature for most species is in the range of 25-35 °C; some species, can grow at temperatures between the psychrophilic and thermophilic interval; the optimum pH range is 6.5 - 8.0. Widely distributed and abundant in soil, marine environment and associated with animals and plants. Some species are pathogenic to animals and humans, while some are phytopathogenic. Content G + C (mol%): 66-78 (Tm). Species type: *Streptomyces albus* (Bergey's Manual. BOONE, 2012).

### Life cycle

The life cycle of *Streptomyces* resembles in many respects the cycle of filamentous fungi: both grow into branched hyphae that constitute the vegetative mycelium and disperse spores formed in specialized reproductive structures known as aerial hyphae that emerge from the surface of the colonies. When a typical spore of *Streptomyces* finds the appropriate conditions and nutrients, the germination process begins, and one or two germ tubes emerge to form the hypha. The tubes grow apically branching and forming the vegetative mycelium; as a response to nutrient depletion, and other signs both antibiotic production and morphological differentiation are initiated leading to the formation of air hyphae which rupture the surface tension escaping from the

aqueous environment of the vegetative mycelium and growing in the presence of air. Aerial hyphae are divided by a process of controlled cell division into long chains of pre-spore compartments, which later develop the thick wall of the spores, synthesize the gray pigment and confer other characteristics typical of mature spores (KIESER, 2000; FLARDH, et al., 2009).

**Genome characteristics**

The genome of *Streptomyces* strains is arranged on a linear chromosome spanning ~ 8Mb size (LIN, *et al*., 1994, OMURA, *et al.,* 2001, BENTLEY, *et al*., 2002), its size becomes extraordinarily large when compared to other bacterial genomes described as *Escherichia coli* (4.6 Mb) and *Bacillus subtilis* (4.2 Mb). The chromosome has a high content of guanines and cytosines (~ 70%) with inverted repeats (TIRs) and proteins covalently attached at the 5 'ends.

Replication of the chromosome occurs bidirectionally from the typically central (oriC) origin of replication, where the continuous synthesis strand is replicated to the right of the chromosomal terminal end, and the discontinuous strand aces a problem associated with removal of the RNA primer from the last Okasaki´s fragment that leaves a space at the 3'-end in the form of a single strand of DNA. This terminal end is replicated through the action of terminal proteins (TP) which act as primers to complete the synthesis of the complementary strand (HOPWOOD, 2006).

Different gene elements are distributed variably between the center and the distal regions of the chromosome and are expressed at different stages of the life cycle. The central region of approximately 6.5 Mb contains mostly essential metabolic genes, such as those responsible for cell division, DNA replication, transition, and translation. Among different species of *Streptomyces* this region seems to be highly conserved as to genetic content as well as in gene order. On the other hand, the telomeric region of the arms of the chromosome shows substantial and polymorphic variability from one species to another and contains genes associated with secondary metabolism as a product of events that are adaptive to stress conditions (KALLIO, 2008).

The high frequency of genetic changes in the arms of the *Streptomyces* chromosome was defined as "genetic instability". The variability is not related to linearity, but to those gene elements that generate large DNA rearrangements, deletions and duplications;

these regions contain long sequences of inverted repeats (TIRs) that range from hundreds to thousands of kilobases and participate in events that result in intra-chromosomal homologous recombination, substitutions, and rearrangements (FISCHER, et al., 1998).

**Secondary metabolites and gene clusters**

The production of secondary metabolites and the diversity of bioactive compounds produced by microorganisms of the genus *Streptomyces* is very large, it is estimated that around 70% of those used in human and animal health as well as in the environment are synthesized by them (BERDY, 2005). This potential has aroused great interest because of the multiple clinical and pharmacological applications of many of these molecules, including antibiotics and anticancer compounds. Antibiotics are molecules synthesized as a product of secondary metabolism, that are low in molecular weight, and whose biological activity consists basically in inhibiting the growth of other organisms without affecting the producer; are produced at low concentrations conferring a selective advantage.

Unlike genes involved in primary metabolism (dispersed in the genome), those participating in pathways of secondary metabolites are continuously arranged in groups between 10-200 kb in loci referred to as clusters (AHLERT, et al., 2002). This organization allows the detection and isolation of functional biosynthetic pathways for biochemical study and manipulation. Clusters typically consist of genes required for the synthesis of the specific metabolite including those associated with carbon chain assembly, modification and regulation (KALLIO, 2008).

The first cloned cluster of a polyketide compound was from *Streptomyces coelicolor* in 1985 by HOPWOOD, *et al*., to produce actinorhodin. In the last 30 years, clusters involved in the anthracycline biosynthesis of different species of *Streptomyces* were isolated and many of their genes were characterized. Some of the best known are: the biosynthetic cluster of daunorubicin (dau) from *Streptomyces* sp. C5 (DICKENS, et al., 1995; STROHL 1997); (rdm) in *S. purpurascens* whose similarity with the dau/dnr genes is studied in combinatorial biosynthesis (NIEMI, *et al*., 1994; NIEMI, *et al*., 1995); aclacinomycin (akn) from *S. galilaeus* (ATCC 31133) (OKI, et al., 1975; FUJII, *et al*., 1997; RATY, *et al*. 2000).

With the development of molecular biology techniques, sequencing and genomics, it was possible to sequence the first genome of an Actinobacteria, in this case the genome of *S. coelicolor* A3 (2) (8.6 Mbp) in 2002 (BENTLEY, et al, 2002). Within this genome about 8000 coding sequences of proteins and around 20 clusters (45% of their total genome), encoding biosynthetic pathways for natural products were identified (BENTLEY, et al., 2002). In the second genome sequenced, that of *S. avermitilis*, 25 gene clusters involved in the biosynthesis of secondary metabolites were described (IKEDA, et al., 2003).

**Polyketide compounds**

The polyketides are divided into three groups (AUSTIN, et al., 2003):

- PKS type I: They are complex molecules that can be large or small, are synthesized by type I synthases; statins and erythromycin are present in this group.
- PKS type II: These are mainly aromatic molecules, synthesized by the synthases II, and are multifunctional enzymes, in this group are the anthracyclines (antitumoral).
- PKS type III: These polyketides are mainly produced by higher plants, in this group are the flavonoids.

PKS II enzymes are a group of multifunctional proteins that catalyze the formation of cyclic aromatic compounds and do not require the reduction or dehydration (HERTWECK, *et al*., 2007; HUTCHINSON; *et al*., 1995). The complex enzymatic activity is called minimal PKS, and has three subunits KSα, KSβ and ACP. The KSα subunit is involved in condensation, the KSβ subunit determines the length of the chain and the ACP unit is responsible for transporting the groups acyl. However, the action of the enzymes encoded by the minimal PKS does not determine the formation of the final compound, so that enzymes such as cyclases, aromatases and reductases give the final formation of type II polyketides, in this group are the anthracyclines that have antitumor activity (KEATINGE-CLAY *et al*., 2004; METSÄ-KETELÄ, *et al*., 1999).

**Anthracyclines**

One of the most studied natural products in the last quarter of a century is the anthracyclines produced by *Streptomyces* spp. The first anthracycline, rhodomycin

was isolated in 1939 from *S. purpurascens* on forest soil in Götingen (BROCKMANN, *et al*., 1950). From that moment on, the search for anticancer compounds began as an initiative of the Farmitalia research laboratories in Milano, Italy (CASSINELLI, 2016). Since the antimicrobial properties of anthracyclines were already known, the chemistry of these metabolites was only investigated in the mid 1960s. After the isolation of similar novel compounds, BROCKMANN and BROCKMANN Jr. named as the anthracyclines, the group of glucosides derived from quinines 7,8,9,10-tetrahydro-5,12-naphthacene. In 1958, daunorubicin belonging to the new class of anthracyclines, called daunomycin was reported by Farmitalia (GREIN, et al., 1963) and rubidomycin by Rhône-Poulenc (DUBOST et al., 1963); this molecule demonstrated activity against bacteria, fungi and cytotoxic properties against some types of tumors. In 1967 the doxorubicin of the mutant strain *S. peucetius* subsp. *caesius* as a result of a daunorubicin mutation was approved by the FDA, for commercial use as an anticancer agent in 1974 (CASSINELLI, 2016). In addition, doxorubicin exhibits an activity against a wide range of tumors (HORTOBAGYI, 1997; MINOTTI et al., 2004), as well as being less toxic (ARCAMONE, 1981). Doxorubicin has been the most successful anticancer agent and the most widely used chemotherapeutic agent in clinical use. The most widespread clinical anthracyclines are daunorubicin (daunomycin, rubomycin), doxorubicin (adrinamycin), idarubicin, epirubicin, zorubicin and aclacinomycin A (aclarubicin).

Although active against a wide variety of solid tumours and haematological malignancies, the clinical use of anthracyclines is hindered by tumour resistance and toxicity to healthy tissue ((HORTOBAGYI, 1997).

**Antitumor activity of anthracyclines**

Despite extensive clinical use, the mechanisms of action of anthracyclines in cancer cells remain a matter of controversy, the following mechanisms have been considered: 1) DNA intercalation, leading to macromolecules synthesis inhibition, 2) Generation of free radicals, which leads to DNA damage or lipid peroxidation, 3) DNA binding and alkylation, 4) interference with DNA in the unwinding or separation of the DNA strands and the activity of the helicase, 5) Direct effects on the membrane, 6) initiation of DNA damage through inhibition of topoisomerase II, and 7) Induction of apoptosis in response to inhibition topoisomerase II (GEWIRTZ, 1999).

**Cosmomycin biosynthesis by *Streptomyces olindensis* DAUPFE 5622**

In the 1960s, several *Streptomyces* strains producing anthracyclines were isolated by the Department of Antibiotics at the Federal University of Pernambuco (DAUFPE), in order to identify those with the capacity to produce antitumor compounds. Among the strains, *Streptomyces olindensis* DAUFPE 5622 was selected, which produces a purple color polyketide complex belonging to the anthracycline group (GONÇALVES DE LIMA, *et al*., 1969). At that time, this polyketide complex was called retamycin, a molecule that had 3 sugars bound to carbon 10 of the aglycone, an event unheard in anthracyclines. Further studies determined that the molecule had three additional sugars bound to carbon 7 (FURLAN *et al.*, 2004). The structure of this new molecule was named cosmomycin D (Fig. 1).

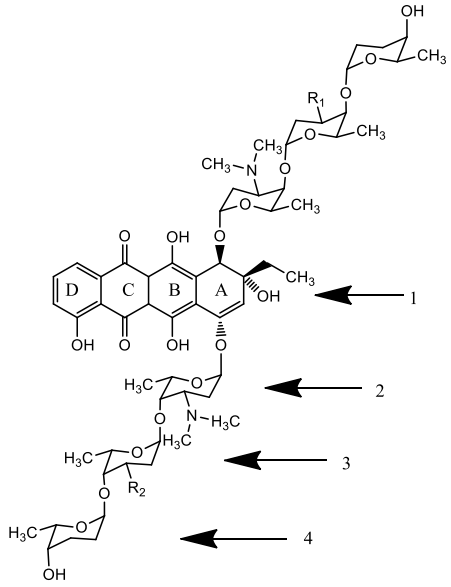


**Figure 1 -** Chemical structure of cosmomycins produced by *S. olindensis*. Cosmomycin D: R 1 = OH and R 2 = OH, cosmomycin C: R 1 = H, R 2 = OH and Citorhodyn N: R 1 = OH, R 2 = H. cosmomycin. 1) β-rhodomycin 2) rhodamine 3) L-rhodinose or L-deoxy-fucose 4) L-rhodinose. (GARRIDO, 2005).

The cosmomycins belong to a very interesting group of molecules that present one of the most complex glycosylation patterns present among anthracyclines, with two trisaccharide chains located on carbon 7 and on carbon 10 of aglycone γ-rhodomicinone (FURLAN et al., 2004). Differences between the trisaccharide fractions of these molecules were determined from mass spectroscopy studies, varying in the position of sugars: L-rhodosamine, 2-deoxy-L-fucose and L-rhodinose in the aglycone (Table 1).

**Table 1 -** Characteristics of cosmomycin A, B, C and D. (HIRAYAMA, 1987)

| **Cosmomycin** | **A** | **B** | **C** | **D** |
|---|---|---|---|---|
| **Molecular weight (g mol$^{-1}$)** | 756 | 772 | 1173 | 1189 |
| **Molecular formula** | $C_{40}H_{53}O_{13}N$ | $C_{40}H_{53}O_{14}N$ | $C_{60}H_{55}O_{21}N_2$ | $C_{60}H_{59}O_{22}N_2$ |

Sequencing studies by GARRIDO (2005) found part of a region (14kb) responsible for cosmomycin biosynthesis. BORDA (2007) studied some of the genes with unknown function as *cosS* and *cosY*, assigning to them functions of transcriptional regulator and ornithine cyclodesaminase, respectively.

With the sequence of the *S. olindensis* genome (ROJAS *et al*, 2014), it was possible to identify the total biosynthetic cluster of cosmomycin D, through the alignment of the sequence described by GARRIDO (2006) with the contigs generated. In order to identify the cosmomycin D cluster, CONTRERAS (2013) analyzed 40 ORFs in several categories: aglycone synthesis, regulators, glycosyltransferases, deoxy sugars, resistance and with unknown functions. CONTRERAS (2013) established the main gene functions of the cosmomycin cluster (Table 2), additional studies on the resistance genes within this cluster were made by ARTEAGA (2015), that presented a resistance model that is expressed in conjunction with antibiotic biosynthesis. FERREIRA-TORRES (2015) made genomic analyzes of *S. olindensis*, focusing attention on the cryptic pathways aiming at obtaining new secondary metabolites of biotechnological interest.

**Table 2 -** Inferred function from the cosmomycin D cluster genes (CONTRERAS, 2013; ARTEAGA, 2015).

| **Gene** | **Hypothetical Function** | **Higher similarities** Identity % (*Blast P) |
|---|---|---|
| ORF 15 | Hypothetical protein | Serine protease – *Streptomyces afghaniensis*, 100% |
| ORF 16 | Dehydrogenase / reductase | Dehydrogenase / reductase – *Streptomyces* sp. NRRL WC-3641, 99% |
| ORF 17 *cosU* | Antibiotic resistance protein "UvrA-like"type | Daunorubicin resistance protein *drrC* - *Streptomyces* sp. NRRL WC-3641, 100% |
| ORF 18 *cosP* | Glutathione peroxidase | Glutathione peroxidase - *Streptomyces* sp. NRRL WC-3641, 99% |
| ORF 19 *cosJ* | ABC-Transporter Membrane Protein | Permease ABC transporter – *Streptomyces* sp. NRRL WC-3641, 99% |
| ORF 20 *cosI* | ABC Transporter ATP Binding Protein | ABC transporter ATPase – *Streptomyces* sp. NRRL WC-3641, 100% |
| ORF 21 *cosH* | dTDP-glucose 4,6 dehydratase | dTDP-glucose 4,6 dehydratase – *S. violaceus*, 95% |
| ORF 22 *cosG* | Glycosyltransferase | *RhoG* Glycosyltransferase – *S. violaceus*, 96% |
| ORF 23 *cosT* | Cytochrome P450 Glucosyltransferase auxiliar protein | *RhoF – S. violaceus*, 74% |

| | | |
|---|---|---|
| ORF 24 *cosF* | Acyltransferase for starter unit | *AknF* – *S. galilaeus*, 68% / Acyltransferase – *S. zinciresistens*, 74% |
| ORF 25 *cosE* | Ketoacylsynthase for starter unit | 3-oxoacyl-ACP synthase – *S. zinciresistens*, 74% |
| ORF 26 *cosC* | Ketosynthase - beta subunit | *AknC* – *S. galilaeus*, 66% / beta-ketoacyl synthase *S. zinciresistens*, 83% |
| ORF 27 *cosB* | Ketosynthase - alpha subunit | *AknB* – *S. galilaeus*, 73% / alpha-ketoacyl synthase – *S. zinciresistens*, 96% |
| ORF 28 *cosX* | Monoxygenase | *AknX* – *S. galilaeus*, 57% / PadR transcriptional regulator – *S. zinciresistens*, 84% |
| ORF 29 *cosY* | Unknown function | Ornithine cyclodeaminase – *Burkholderia* sp., 40% / *S. zinciresistens*, 56% |
| ORF 30 *cosS* | Possible PadR transcriptional regulator | *AclS* – *S. galilaeus*, 54% / PadR transcriptional regulator – *S. zinciresistens*, 84% |
| ORF 31 *cosK* | Glycosyltransferase | *AknK* – *S. galilaeus*, 65% / Glycosyltransferase – *S. zinciresistens*, 78% |
| ORF 32 *cosL* | dTDP-4-hexose 3,5 epimerase | dTDP-4-dehydro-rhamnose-3,5-epimerase – *S. zinciresistens*, 78% |
| ORF 33 *cosM* | NAD dependent epimerase / dehydratase | NAD dependent epimerase / dehydratase – *S. zinciresistens*, 74% / *SnoG* – *S. nogalater*, 54% |
| ORF 34 | NDP-hexose 2,3- dehydratase | NDP-hexose 2,3- dehydratase – *S. zinciresistens*, 74% / *SnogH* – *S. nogalater*, 61% |
| ORF 35 | NDP-deoxyhexose 3-aminotransferase | Glutamine transaminase – *S. zinciresistens*, 87% / *AclZ* – *S. galilaeus*, 72% |
| ORF 36 | Methyltransferase | Methyltransferase - *S. zinciresistens*, 88% |
| ORF 37 | Cyclase | Cyclase - *S. zinciresistens*, 88% / *S. steffisburgensis*, 73% / *S. nogalater*, 72% |
| ORF 38 | Cyclase | Nuclear transport factor 2 - *S. zinciresistens*, 68% / *AknV* – *S. galilaeus*, 58% |
| ORF 39 | Keto redutase | Dehydrogenase / reductase - *S. zinciresistens*, 78% / *SnoaF* C-7 keto-redutase – *S. nogalater*, 67% |
| ORF 40 | Aromatase | *RdmK* – *S. purpurascens*, 90% |
| ORF 41 | Keto-redutase | *RdmJ* - *S. purpurascens*, 94% / *AknA* – *S. galilaeus*, 77% |
| ORF 42 | NDP-hexose-3,4-dehydratase | *RdmI* – *S. purpurascens*, 95% |
| ORF 43 | Glycosyltransferase | *RdmH* - *S. purpurascens*, 92% |
| ORF 44 | Cytochrome P-450 | *RdmG* - *S. purpurascens*, 83% |
| ORF 45 | Keto-redutase | *RdmF* - *S. purpurascens*, 94% |
| ORF 46 | Aclavinone-11 hydroxylase | *RdmE* - *S. purpurascens*, 91% |
| ORF 47 | Methyltransferase | *RdmD* - *S. purpurascens*, 87% |
| ORF 48 | 10-carbomethoxy-13-deoxycarminomycin esterase | *RdmC* - *S. purpurascens*, 95% |
| ORF 49 | Aglycone 10-Hydroxylase | *RdmB* - *S. purpurascens*, 90% |
| ORF 50 | Methylaklanonic acid cyclase | *RdmA* – *S. purpurascens*, 98% |
| ORF 51 | Aklanonic acid methyltransferase | *AknG* – *S. galilaeus*, 66% / *SnoaC* – *S. nogalater*, 69% |
| ORF 52 *cosX* | SARP family pathway specific regulatory protein | SARP Regulator – *S. zinciresistens* / *AknI* – *S. galilaeus*, 65% |
| ORF 53 *cosA* | Acyl Charger Protein | *AknD* Acyl Charger Unit (ACP) – *S. galilaeus*, 54% |
| ORF 54 *cosV* | Glucose-1-phosphate thymidylyltransferase | Glucose-1-phosphate thymidylyltransferase – *S. zinciresistens*, 85% / *AclY* – *S. galilaeus*, 73% |
| ORF 55 | Putative Oxidoreductase | Aldo/keto reductase – *S. zinciresistens*, 85% |
| ORF 56 | Aclacinomycin oxidoreductase | Berberine domain– *S. zinciresistens* / *AclO* – *S. galilaeus*, 63% |
| ORF 57 | Nitroreductase | Nitroreductase – Multispecies *Streptomyces* sp. 100% |
| ORF 58 | Hypothetical protein | Hypothetical protein - *Streptomyces* sp. NRRL WC-3641, 99% |

### *CosY*

According to GARRIDO (2006), the cosY gene corresponds to an open reading phase of 978 bp that shows no protein homologies from the databases. The proteins with the highest degree of homology to the product of this gene are ornithines

cyclodeaminases, which are involved in ornithine to proline conversion. Despite the low homologies and having no apparent function in cosmomycin biosynthesis, *cosY* has a codon distribution with the other genes described for *S. olindesis* and a possible ribosome binding site located at the -12 gene start position (AGGACG), which indicates that it can be translated normally (GARRIDO, 2006).

***CosM***

*CosM* is a 936 bp gene, was reported by GARRIDO (2006) having the greatest homology to the *aclM* gene of *Streptomyces galileus*. The probable function of the protein encoded by this gene is that of a 4-ketoreductase responsible for the final step in dTDP-L-2-deoxyfucose biosynthesis. A possible ribosome binding site (RBS) of this gene starts at position -9 and the hexamer sequence is AGCACG which overlaps the *cosL* stop codon, so a coupled translation between the two genes could occur (GARRIDO, 2006).

## JUSTIFICATION

According to Globocan (FERLAY, 2018) during 2018 there were 18,1 million new cases of cancer worldwide and 9,6 million deaths related to this kind of diseases. Of these new cases, 60% occur in Africa, Asia and Central and South America; and 70% of the world's cancer deaths are accounted by these regions. It is estimated that in the next 20 years there will be an increase of 70% in the number of new cases of cancer. Currently, the treatment against this kind of pathologies includes chemotherapy, radiotherapy, surgery and immunotherapy.

Considering the importance of antitumor antibiotics in chemotherapeutic treatments, knowing that the *Streptomyces* genus is the producer of most of the identified bioactive compounds, and taking into account all the work, previously done in our laboratory in the study of the cosmomycin biosynthetic pathway in *Streptomyces olindensis*, the main objective of this study was to establish the exact number of genes that comprises the cosmomycin biosynthetic cluster and characterize the genes *cosY* and *cosM*, whose function remain unknown.

## CONCLUSIONS

A) The transcriptomic analysis let us to conclude that the cosmomycin D cluster in *Streptomyces olindensis* is comprised of 38 ORFs, having as defined ends ORF17 *cosU* and ORF54 *cosV* G1Pdt (glucose-1-phosphate thymidyltransferase); and that the cosmomycin cluster has a size of aprox. 40Kb (39.972bp)

B) We achieved the knockout of the two genes and studied them, although with the data available so far, is not possible to define the function of the *cosY* gene and new possibilities are opened to those expected by homology, with respect to the *cosM* gene. Additional studies are necessary to establish the function of these two genes.

## REFERENCES


ALEXEEV, I., *et al*. Aclacinomycin oxidoreductase (*AknOx*) from the biosynthetic pathway of the antibiotic aclacinomycin is an unusual flavoenzyme with a dual active site. **PNAS**, v. 104, n. 15, p. 6170-75, 2007.

ALTSCHUL, S. F., *et al*. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, n. 3, p. 403-10., 1990.

ARCAMONE, F. **Doxorubicin, Anticancer Antibiotics.** Medicinal Chemistry, a series of Monographs. Ed de Stevens, G. Academic Press, New York. vol. 17. 1981.

ARTEAGA, R. D. C. **Caracterização dos mecanismos de resistência ao antibiótico antitumoral cosmomicina D**. 2015. 117 p. Dissertação (Mestrado em Microbiologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2015.

AUSTIN, M. B.; NOEL, J. P. The chalcone synthase superfamily of type III polyketide synthases. **Natural Product Reports**, v. 20, n. 1, p. 79–110, 2003.

BENSON, D. A. *et al*. GenBank. **Nucleic Acids Research**, v. 43, n. D1, p. D30–D35, 2015.

BENTLEY, S. D., *et al*. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). **Nature**, v. 417, n. 6885, p. 141-7, 2002.

BERDY, J. Bioactive microbial metabolites. **Journal of Antibiotics (Tokyo)**. v. 58, n. 1, p. 1-26. 2005.

BOONE, D. R., *et al*. **Bergey's manual of systematic bacteriology**. New York, 2012.

BOOSER, D. J., *et al*. Anthracycline antibiotics in cancer therapy. Focus on drug resistance. **Drugs,** v. 47, n. 2, p. 223-58, 1994.

BORDA, C. C. **Estudo de genes envolvidos na via biossintética do antibiótico antitumoral Cosmomicina**. 2007. 125 p. Tese (Doutorado em Biotecnologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2007.

BROCKMANN, H., et al. Rhodomycin, ein rotes Antibioticum aus Actinomyceten. **Naturwissenschaften**. v. 37, n. 21, p. 492-493. 1950.

BROCKMANN, H., et al. Rhodomycin, II. **Chemische Berichte.** v. 86, n. 2, p. 261-269. 1953.

CARVALHO H, *et al*. DNA damage induced by the anthracycline cosmomycin D in DNA repair-deficient cells. **Cancer Chemotherapy and Pharmacology**. v. 65, p. 989 –994, 2010.

CASSINELLI, G. The roots of modern oncology: from discovery of new antitumor anthracyclines to their clinical use. **Tumori**, v. 102, n. 3, p. 226-235. 2016.

CHUNG, J. Y., *et al*. (2002). Expression, purification, and characterization of AknX anthrone oxygenase, which is involved in aklavinone biosynthesis in *Streptomyces galilaeus*. **Journal of Bacteriology**, v. 184, n. 22, p. 6115-22, 2002.

CONTRERAS, C.A. **Caracterização de genes biossintéticos do antitumoral cosmomicina D**. 2013. 106p. Dissertação (Mestrado em Microbiologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2013.

DICKENS, M. L., *et al*. Analysis of clustered genes encoding both early and late steps in daunomycin biosynthesis by *Streptomyces* sp. strain C5. **Journal of Bacteriology**, v. 177, n. 3, p. 536-43, 1995.

DUBOST, M.P., *et al*. Un novel antibiotique à proprieties cytostatiques: la rubidomycine. **C.R. Acad. Sci. Paris**, v. 257, p. 1813-1815. 1963.

FERLAY J, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. **International Journal of Cancer**. 144(8):1941–1953. 2019.

FERREIRA-TORRES, M. A. **Análise genômica de *Streptomyces olindensis* DAUFPE 5622 e de suas vias crípticas para a obtenção de novos metabólitos secundários de interesse biotecnológico.** 2015. 104p. Dissertação (Mestrado em Microbiologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo. 2015.

FISCHER, G. Chromosomal arm replacement generates a high level of intraspecific polymorphism in the terminal inverted repeats of the linear chromosomal DNA of Streptomyces ambofaciens. **PNAS**, v. 95, n. 24, 14296-14301. 1998.

FLARDH, K., *et al*. *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. **Nature Reviews Microbiology**, v. 7, n. 1, p. 36-49, 2009.

FLETT, F., *et al*. High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting streptomycetes. **FEMS Microbiology Letters**, v. 155, n. 2, p. 223-229, 1997.

FUJII, I., *et al*. Anthracycline Biosynthesis in *Streptomyces galilaeus*. **Chemical Reviews**, v. 97, n. 7, p. 2511-2524, 1997.

FURLAN, R. L., *et al*. DNA-binding properties of cosmomycin D, an anthracycline with two trisaccharide chains. **The Journal of Antibiotics**, v. 57, p. 647– 654, 2004.

FURLAN, R. L. **O antitumoral retamicina: produção e efeitos biológicos.** 2002. 102p. Tese (Doutorado em Biotecnologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2002.

FURLAN, R. L. **Caracterização de Mutantes de *Streptomyces olindensis*, Produtor do Antibiótico Antitumoral Retamicina.** 1997. 76p. Dissertação (Mestrado em Biotecnologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 1997.

GARRIDO, L. M., *et al*. Insights in the glycosylation steps during biosynthesis of the antitumor anthracycline cosmomycin: characterization of two glycosyltransferase genes. **Applied Microbiology and Biotechnology**, v. 73, n. 1, p. 122-31, 2006.

GARRIDO, L. **Identificação e caracterização de genes envolvidos na biossíntese da antraciclina cosmomycina D produzida por *Streptomyces olindensis* DAUFPE 5622**. 2005. 203p. Tese (Doutorado em Biotecnologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2005.

GARRIDO, L. **Construção de uma biblioteca genômica de *Streptomyces olindensis* e triagem de clones envolvidos na biossíntese do antitumoral retamicina.** 2000. Dissertação (Mestrado em Biotecnologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo. 2000.

GEWIRTZ, D. A. A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. **Biochemical Pharmacology**. v. 57, n. 7, p. 727-741. 1999.

GONÇALVES DE LIMA, O., *et al*. Primeiras observações sobre o complexo antibiótico e antitumoral -retamicina- produzido por *Streptomyces olindensis* nov. sp. IAUPe. **Revista Instituto de Antibioticos,** v. 9, p. 27-38. 1969

GREIN, A., et al. Descrizione e classificazione di un attinomicete (*Streptomyces peucetius* sp. *nova*) produttore diumaa sostanza attività antitumorale-La daunomicina. **Giorn Microbiol**, v. 11, p. 109-118. 1963

GRIMM, A., K., *et al*. Characterization of the *Streptomyces peucetius* ATCC 29050 genes encoding doxorubicin polyketide synthase. **Gene**, v. 151, n. 1-2, p. 1-10, 1994.

HE, Y., *et al*. Two pHZ1358-derivative vectors for efficient gene knockout in *Streptomyces*. **Journal of Microbiology and Biotechnology**, v. 20, n. 4, p. 678-82, 2010.

HERTWECK, C.; *et al*. Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. **Natural Product Reports**, v. 24, n. 1, p. 162–190, 2007.

HIRAYAMA, K., *et al*. Field Desorption Tandem Mass Spectrometry of Anthracyclines Antibiotics, Cosmomycin A, B, A', B', C and D. **Biomedical and Environmental Mass Spectrometry**, v. 14, p. 305-312. 1987.

HOPWOOD, et al. Production of 'hybrid' antibiotics by genetic engineering. **Nature**, v. 314, n. 6012, p. 642-4. 1985.

HOPWOOD, D. A. Soil to genomics: the *Streptomyces* chromosome. **Annual Review of Genetics**, v. 40, p. 1-23, 2006.

HORTOBÁGYI, G.N. Anthracyclines in the treatment of cancer. An overview. **Drugs**. v. 54, suppl. 4. p. 1-7- 1997.

HUTCHINSON, C. R. Biosynthetic Studies of Daunorubicin and Tetracenomycin C. **Chemical Reviews**, v. 97, n. 7, p. 2525-2536, 1997.

HUTCHINSON, C. R.; *et al*. I. Polyketide synthase gene manipulation: a structure-function approach in engineering novel antibiotics. **Annual Review of Microbiology**, v. 49, p. 201–238, 1995.

IKEDA, H., J., *et al*. Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. **Nature Biotechnology**, v. 21, n. 5, p. 526-31, 2003.

ISHIKAWA, J., *et. al*. FramePlot: a new implementation of the frame analysis for predicting protein-coding regions in bacterial DNA with a high G+C content. **FEMS Microbiology Letters**, v. 174, n. 2, p. 251-3, 1999.

JANSSON, A., H., *et al*. Aclacinomycin 10-hydroxylase is a novel substrate-assisted hydroxylase requiring S-adenosyl-L-methionine as cofactor. **The Journal of Biological Chemistry**, v. 280, n. 5, p. 3636-44, 2005.

JANSSON, A., J., *et al*. Crystal structure of aclacinomycin-10-hydroxylase, a S-adenosyl-L-methionine-dependent methyltransferase homolog involved in anthracycline biosynthesis in *Streptomyces purpurascens*. **Journal of Molecular Biology**, v. 334, n.2, p. 269-80, 2003.

JANSSON, A., J., *et al*. Crystal structure of aclacinomycin methylesterase with bound product analogues: implications for anthracycline recognition and mechanism. **The Journal of Biological Chemistry**, v. 278, n. 40, p. 39006-13, 2003.

KALLIO, P. **Type II aromatic polyketide biosynthetic tailoring enzymes: diversity and adaptation in *Streptomyces* secondary metabolism.** 2008. 49p. University of Turku, Turku, Finland, 2008.

KEATINGE-CLAY, A. T.; *et al*. An antibiotic factory caught in action. **Nature Structural & Molecular Biology**, v. 11, n. 9, p. 888–893, 2004.

KELSO, C.; *et al*. Characterisation of Anthracyclines from a Cosmomycin D-Producing Species of *Streptomyces* by Collisionally-Activated Dissociation and Ion Mobility Mass Spectrometry. **European Journal of Mass Spectrometry**, v.15, p. 73-81, 2009.

KIESER, T. **Practical *Streptomyces* Genetics**. Norwich, John Innes Foundation, 2000.

KIM, R. Y., *et al*. mu-crystallin is a mammalian homologue of *Agrobacterium* ornithine cyclodeaminase and is expressed in human retina. **PNAS**, v. 89, n.19, p. 9292-6, 1992.

KORMAN, T. P., *et al*. Structural analysis of actinorhodin polyketide ketoreductase: cofactor binding and substrate specificity. **Biochemistry**, v. 43, n. 6, p. 14529-38, 2004.

LARSON, C. B. *et al*. PCR-Independent Method of Transformation-Associated Recombination Reveals the Cosmomycin Biosynthetic Gene Cluster in an Ocean Streptomycete. , **Journal of Natural Products**, v. 80, n. 4, p. 1200-1204, 2017.

LI, S., *et al*. Genome-wide identification and characterization of reference genes with different transcript abundance for *Streptomyces coelicolor*. **Nature Scientific Reports**, v. 5, p. 15840, 2015.

LIN, Y. S., *et al*. The chromosomal DNA of *Streptomyces lividans* 66 is linear. **Molecular Microbiology**, v. 14, n. 5, p. 1103. 1994

LOMBÓ, F., *et al*. Elucidation of oxygenation steps during oviedomycin biosynthesis and generation of derivatives with increased antitumor activity. **ChemBioChem**, v. 10, n. 2, p. 296-303. 2009

LOMOVSKAYA, N., *et al*. Doxorubicin overproduction in *Streptomyces peucetius*: cloning and characterization of the dnrU ketoreductase and dnrV genes and the doxA cytochrome P-450 hydroxylase gene. **Journal of Bacteriology**, v. 181, n. 1, p. 305-18, 1999.

MEDEMA, M. H., K. Blin, *et al*. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. **Nucleic Acids Research**, v. 39 (Web Server issue), p. W339-46. 2011.

MENÉNDEZ, N., *et al*. Deoxysugar transfer during chromomycin A3 biosynthesis in *Streptomyces griseus* subsp. *griseus*: new derivatives with antitumor activity. **Applied Environmental Microbiology**, v. 72, n. 1, p. 167-77, 2006.

METSÄ-KETELÄ M., *et al.* **Anthracycline Biosynthesis: Genes, Enzymes and Mechanisms.** In: Krohn K. (eds) Anthracycline Chemistry and Biology I. Topics in Current Chemistry, vol 282. Springer, Berlin, Heidelberg, 2007.

METSÄ-KETELÄ, M.; *et al.* An efficient approach for screening minimal PKS genes from *Streptomyces*. **FEMS Microbiology Letters**, v. 180, n. 1, p. 1–6, 1999.

MINOTTI, G., *et al*. Anthracyclines: Molecular Advances and Pharmacologic Developments in Antitumor Activity and Cardiotoxicity. **Pharmacological Reviews.** v. 56, n. 2, p. 185-229. 2004.

MONACHE, D., *et al.* Chemical composition of retamycin, an antibiotic complex with anti-tumor activity produced by *Streptomyces olindensis* nov. sp. (IAUFPe 5622). **Annali dell'Istituto Superiore Di Sanità**, v. 6 (6) 357-64.

MOORE, B.S., *et al*. Discovery of a new bacterial polyketide biosynthetic pathway. **Chembiochem**, v. 2, n. 1, p. 35-8, 2001.

NIEMI, J., *et al*. **Anthracycline biosynthesis: stps, enzymes and genes** (review) in Microbial secondary metabolites: biosynthesis, genetics and regulation. Research Signpost 37/661. Kerala, India, eds. Francisco Fierro & Juan Francisco Martin. 2002

NIEMI, J. *et al.* Nucleotide sequences and expression of genes from *Streptomyces purpurascens* that cause the production of new anthracyclines in *Streptomyces galilaeus*. **Journal of Bacteriology**, v. 177, n. 10, p. 2942-5, 1995.

NIEMI, J., *et al*. Hybrid anthracycline antibiotics: production of new anthracyclines by cloned genes from *Streptomyces purpurascens* in *Streptomyces galilaeus*. **Microbiology**, v. 140, Pt. 6, p. 1351-8, 1994.

OKI, T., *et al*. New antitumor antibiotics aclacinomycins A and B. **The Journal of Antibiotics (Tokyo)**, v. 28, n. 10, p. 830-4, 1975.

OMURA, S., *et al*. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. **PNAS**, v. 98, n. 21, p. 12215-20, 2001.

RATY, K., *et a*l. A gene cluster from *Streptomyces galilaeus* involved in glycosylation of aclarubicin. **Molecular Genetics and Genomics**, v. 264, n. 1-2, p. 164-72, 2000.

RODRIGUES, S. C. **Análise do crescimento e produção do complexo antibiótico sintetizado por *Streptomyces olindensis*.** 2004. 98p. Dissertação (Mestrado em Biotecnologia) - Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo. 2004.

ROJAS, J. D., *et al*. Genome Sequence of *Streptomyces olindensis* DAUFPE 5622, Producer of the Antitumoral Anthracycline Cosmomycin D. **Genome announcements**, v. 2, n. 3, p. 4-5, 2014.

SAMBROOK, J.; *et al*. **Molecular cloning: a laboratory manual. cold spring harbor laboratory press**. New York: Cold Spring Harbor, 1989. 3 v.

STROHL, W. R. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. **Nucleic Acids Research**, v. 20, n. 5, p. 961-74, 1992.

STROHL, W. R. **Biotechnology of antibiotics**. New York, Marcel Dekker, 1997.

STUTZMAN-ENGWALL, K. J., *et al.* Regulation of secondary metabolism in *Streptomyces* spp. and overproduction of daunorubicin in *Streptomyces peucetius*. **Journal of Bacteriology**, v. 174, n. 1, p. 144-54, 1992

SUN, Y., *et al*. Analysis of functions in plasmid pHZ1358 influencing its genetic and structural stability in *Streptomyces lividans* 1326. **Applied Microbiology and Biotechnology**, v. 82, n. 2, p. 303-10, 2009.

TORKKELL, S., *et al*. The entire nogalamycin biosynthetic gene cluster of *Streptomyces nogalater*: characterization of a 20-kb DNA region and generation of hybrid structures. **Molecular Genetics and Genomics**, v. 26, n. 2, p. 276-88, 2001.

TORKKELL, S., *et al*. Characterization of *Streptomyces nogalater* genes encoding enzymes involved in glycosylation steps in nogalamycin biosynthesis. **Molecular Genetics and Genomics**, v. 256, n. 2, p. 203-9, 1997.

VINCZE, T., *et al*. NEBcutter: A program to cleave DNA with restriction enzymes. **Nucleic Acids Research**, v. 31, p. 3688-3691, 2003.

YLIHONKO, K., *et al*. A gene cluster involved in nogalamycin biosynthesis from *Streptomyces nogalater*: sequence analysis and complementation of early-block mutations in the anthracycline pathway. **Molecular Genetics and Genomics**, v. 251, n. 2, p. 113-20, 1996.