# Construction of a genetically improved *Streptomyces clavuligerus* strain that might have the potential for a larger clavulanic acid production

Luisa Fernanda Patiño Cervantes

A Thesis submitted in partial fulfillment of the requirements for the degree of

**Master of Chemical Engineering**

**Advisor:**

Rigoberto Ríos-Estepa, PhD

**University of Antioquia**

**Department of Engineering**

**Faculty of Chemical Engineering**

**Bioprocess group**

**2023**

Maestría en Ingeniería Química, Cohorte V.

Grupo de Investigación Bioprocesos.

Centro de Investigación Ambientales y de Ingeniería (CIA).


*To my mother and father*

## ACKNOWLEDGMENTS

First, I would like to express my deepest gratitude to my supervisor Dr. Rigoberto Ríos for the tutoring and support given throughout these years. Specially, his patience and the trust gave me despite the hurdles faced during this research. I also want to thank him, the liberty received for explore my own ideas. In an invaluable way, this helping me to become a better person and scientist.

I thank to Dr. Laura Inés Pinilla, and León Felipe Navarro to giving me the opportunity to be part of clavulanic acid research project. For the support and advice, to them, my gratitude. I also want to thank Dr. Mariana Peñuela for given me the opportunity to pave my way as scientist in the University of Antioquia Bioprocess laboratory.

I specially want thank Vanessa Aguirre to her guidance and support during the experimentation, for her invaluable friendship, and the spend time in the laboratory trying to understand *S. clavuligerus*. I want to thank my colleague and friend, Yeison Agudelo for the long conversations about our dreams and shared frustrations, and for cheering up my days in the lab. I also want to thanks to my dear friend, Jaison Cuartas, for his advice on the manuscript and, his help in the construction of the pan genome. I also want to thank him for motivating me to the end to finish this thesis.

To my loves, Stiven and Matias, for filling my life with happiness and love, for being the reason, and my greatest motivation. To Stiven, for encouraging me to be a better person and finishing this research. Finally, this thesis would not have been possible without the love and support of my family. To my mother, for her care and love. To my father, for his generous advice, to being my friend and, for pushed me not give up.

**TABLE OF CONTENTS**

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATION

| | |
|---|---|
| ATCC | American Type Culture Collection |
| NRRL | Agricultural Research Service Collection |
| DSM | German Collection of Microorganisms and cell cultures GmbH |
| bp | Base pair |
| dH2O | Distilled water |
| $OD_{600}$ | Optical density at 600 nm |
| $OD_{590}$ | Optical density at 590 nm |
| dNTP | Deoxynucleotide Triphosphate |
| RE | Restriction endonuclease |
| ORF | Open Reading Frame |
| rpm | Revolution per minute |
| kb | Kilo base |
| DMSO | Dimethyl Sulfoxide |
| SDS | Sodium dodecyl sulfate |
| PCR | Polymerase chain reaction |
| RT-qPCR | Quantitative Reverse Transcription PCR |
| HPLC | High performance liquid chromatography |
| DWC | Dry Weight of Cells |
| CFU | Colony-forming unit |

## ABSTRACT

*Streptomyces clavuligerus* (*S. clavuligerus*) is a Gram-positive soil bacterium capable of producing several β-lactam metabolites, including the β-lactamase inhibitor clavulanic acid (CA). CA effectively inhibits the activity of β-lactamase enzymes, qualifying it as responsible for most of the resistance developed towards β-lactam antibiotics. Despite the significant number of studies for submerged cultivations related to CA production using *S. clavuligerus*, low titers (~1 $g.L^{-1}$) rendered low yields obtained when using a wild-type strain. In contrast, more than 7 $g.L^{-1}$ are produced using genetically modified strains. Furthermore, regulatory mechanisms and pathways governing CA biosynthesis are not fully depicted and characterized for *S. clavuligerus*. Therefore, the implementation of novel strategies for understanding and improving CA biosynthesis in *S. clavuligerus* is required. In this study the effect of overexpression of two regulatory genes involved in CA biosynthesis pathway were separately assessed. The *claR* gene was overexpressed under the control of the constitutive promoter *ermE** in the wild type *S. clavuligerus* strain. Transformant *S. clavuligerus*/pICLAR obtained via chromosomal integration, increased CA production approximately 1.4-fold using ISP and GSPG media. No morphological differences were observed between *S. clavuligerus*/pICLAR and control strain, though, we did not evaluate possible mRNA expression differences. Furthermore, the effect of *orf21* overexpression on CA production was analyzed for wild type *S. clavuligerus* strain. The role of *orf21* on CA biosynthesis has been previously studied; nevertheless, its performance as CA production regulator is not fully understood. The findings obtained in this study suggest that the role of *orf21* on CA biosynthesis is strongly determined by the environmental conditions of the fermentation. As for GSPG, a defined medium, *S. clavuligerus*/pIORF21 increased CA production 2.6-fold. Likewise, the *orf21* overexpression in GSPG stimulated expression of the late CA pathway genes *gcas* and the genetic regulator *adpA*, as determined by real-time PCR. In contrast, for ISP, a soy protein based medium, the transformant *S. clavuligerus*/ pIORF21 decreased CA production by 1.8-fold compared to the control strain. The *orf21* overexpression using ISP medium negatively modulates CA biosynthesis, possibly repressing indirectly the expression of the regulatory gene, *claR*.

The results obtained during this research allowed to improve CA production throughout the construction of two recombinant strains exhibiting higher CA yields. Specifically, *S. clavuligerus*/pICLAR can be used as a promising tool for studies aimed at enhancing CA biotechnological production. Conversely, *S. clavuligerus*/pIORF21 is not a recommended transformation alternative for CA production improvement in rich media based on soy protein, as it is strongly affected by environmental conditions; the study of *orf21* highlights the importance of understanding the role of ORF21 and other regulatory proteins behind the CA biosynthesis regulatory network in *S. clavuligerus*.

# 1. INTRODUCTION

## *1.1.* General characteristics of the genus *Streptomyces*

In terms of number and variety of identified species, the phylum Actinobacteria represents one of the largest taxonomic units among the 18 major lineages currently recognized within the domain Bacteria[1]. Actinobacteria are a group of Gram-positive bacteria with high guanine and cytosine content in their DNA, which are widely distributed in both terrestrial and aquatic ecosystems, especially in soil, where they play a crucial role in the recycling of biomaterials by decomposition and humus formation [1,2]. The phylum Actinobacteria exhibits a wide variety of morphologies: coccoid (*Micrococcus* sp.) or rod-coccoid (*Arthrobacter* sp.), fragmented hyphal forms (*Nocardia* sp.) or permanent and highly differentiated branched mycelium (*Streptomyces* sp.).

The Actinobacteria produce a variety of secondary metabolites exhibiting high pharmacological and commercial interest [2]. Approximately, 10,000 bioactive metabolites are produced by members of this phylum, which is close to 45% of all bioactive microbial metabolites discovered so far [2]. The use of Actinobacteria to produce valuable molecules in pharmaceutical industry, started since the streptomycin discovery.

In 1943, Albert Schatz (member of Waksman's group), found the antibiotic streptomycin from *Streptomyces griseus* [3]. Streptomycin was used effectively as the first treatment against tuberculosis and paved the way for the search of new drugs against human infections caused by pathogenic microorganisms[3]. After streptomycin implementation success, Waksman's group, American companies, and other countries, continued isolating and screening actinomycetes. Notably, many of the new antibiotics identified from the late 1940s to the late 1960s belonged to *Streptomyces* species. The genus *Streptomyces* belongs to the phylum XXVI of Gram-positive, GC-rich Actinobacteria, order XIV Streptomycetales, family I Streptomycetaceae [4]. The genus *Streptomyces* was described in 1875 by Ferdinand Cohn and was initially named, *Streptothrix* ("twisted hair") [3]. Nonetheless, it was later that the genus adopted the name of *Streptomyces* meaning "twisted fungus" (Waksman & Henrici 1943) [3].

The genus *Streptomyces* spp. has a complex life cycle that includes the formation of widely branched mycelium and aerial mycelium leading to spores formation [5]. The aerial morphogenesis and formation of reproductive spores provide *Streptomycetes* with a mechanism to disperse the species to new habitats. Under suitable nutritional and environmental conditions, the *Streptomyces* spp. growth in solid medium begins with the germination of the spore. The emerging germ tube, referred to as "substrate hyphae", are non-unicellular filamentous hyphae, which grows apically and branches forming a tangled network of filamentous cells, referred to as a "substrate mycelium" [6]. The substrate mycelium divides into, early substrate mycelium (MI) and late substrate mycelium (MII); both lack a hydrophobic layer which is associated with the development cycle. As the substrate mycelium ages and/or goes through environmental conditions changes, the cell undergoes a differentiation process; in a next stage a second type of filamentous cell is produced, which grows up into the air away from the substrate hyphae, usually occurring around 24 and 48 hours for *Streptomyces coelicolor*, under laboratory conditions.

During this process, specific enzymes are released allowing the substrate mycelium to undergo a programmed cell death process (PCD). Cellular changes associated with cell wall and cell membrane deterioration, DNA degradation, and cytoplasm release to extracellular matrix, are programmed; this PCD allows the nascent aerial mycelium to feed and grow [7]. The aerial mycelium does not branch but undergoes a process of septation (pre-spore compartments); later, each compartment, with a single chromosome, will form a new spore [7,8] . The aerial hypha formation coincides with the biosynthesis of a diverse array of secondary metabolites, many of which possess a significant medical use as antibiotics [7]. The antibiotics or other secondary metabolites released by substrate mycelium are used to inhibit the growth of other microorganisms that compete for nutrients, as well as to support aerial mycelium growth and sporulation.

The genus *Streptomyces* has been one of the most studied groups of microorganisms among actinomycetes, not just for its complex life cycle, but also for its ability to produce a plethora of secondary metabolites with application in human medicine, agriculture, and animal health (see Table 1). Many of these secondary metabolites have been widely studied for its industrial importance (e.g., clavulanic acid), the regulatory mechanisms that control their biosynthesis, as well as the biology of its production. *Streptomyces* spp. remains as a valuable source of bioactive natural products since in its genome harbors

silent secondary metabolites biosynthetic gene clusters (smBGC) that are cryptic under laboratory conditions, and yet remain to be discovered (Table 1).

**Table 1.** *Streptomyces* and its potential as a producer of secondary metabolites.

| Organism | Genome size (Mb) | Importance | smBGC* | Reference |
|---|---|---|---|---|
| *S. milbemycinicus* NRRL 5739 | 11.62 | Antiparasitic (Milbemycins) | 81 | [9,10] |
| *S. venezuelae* ATCC10712 | 8.22 | Antibiotic (Chloramphenicol and Jadomycin) | 30 | [11] |
| *S. albus* J1074 | 6.9 | Cytotoxic and antifungal (Antimycin) | 60 | [12] |
| *S. antibioticus* DSM 41481 | 8.47 | Siderophore (2′-Deoxycoformycin). Antineoplastic drug (Desferrithiocin) | ~27 | [13,14] |
| *S. avermitilis* | 9.1 | Antiparasitic and Insecticidal (Avermectin) | 38 | [15] |
| *S. coelicolor* A3(2) | 8.7 | Model microorganism. Antibiotics (Actinorhodin, Prodigines, and Coelimycin P1) | 23 | [16,12] |
| *S. clavuligerus* ATCC 27064 | 8.70 | Antibacterial (Clavulanic acid, Cephamycin C, Holomycin and Tunicamycin) | 58 | [17] |
| *S. rimosus* | 9.36 | Antibiotics (Oxytetracycline) | 45 | [18,19] |
| *S. spectabilis* | 9.81 | Antibiotics (Kanamycin, Neomycin B, Spectinomycin) | 44 | [18,20] |

* Silent secondary metabolites biosynthetic gene clusters (smBGC).

### *1.2. Streptomyces clavuligerus* and clavulanic acid

*Streptomyces clavuligerus* (*S. clavuligerus*) was isolated from a South American soil sample in 1971 [21]. The genome of *S. clavuligerus* is composed of a linear chromosome (6.75-Mbp), three linear short plasmids, pSCL1(10.5 kb), pSCL2 (149.4 kb), and pSCL3 (442.2 kb), and a linear mega plasmid named pSCL4 (1,8 Mbp) [17]. Genome sequencing has revealed that it contains many secondary metabolite biosynthetic gene clusters (see Table 1) [17]. Among them, 30 and 28 smBGCs were found from the linear chromosome and the plasmid pSCL4, respectively [17].

*S. clavuligerus* is recognized for producing β-lactam compounds, including cephamycin C antibiotic and the β-lactamase inhibitor clavulanic acid (CA). Cephamycin C and CA are synthesized simultaneously, though via different metabolic pathways in *S. clavuligerus* [22]. The antibiotic cephamycin C is synthesized from three precursors, the aminoadipic acid that comes from L-lysine produced from aspartate in a branched pathway and the amino acids L-Cysteine and L-Valine [23]. Cephamycin C proceeds through the formation of the tripeptide, α-aminoadipyl-cysteinyl-valine, which cyclizes to form isopenicillin N, the first compound with antibiotic activity. In the three middle steps of the pathway, isopenicillin N is converted to deacetoxycephalosporin C, which in turn forms cephamycin C in the final steps of the pathway [23]. Although there are no biosynthetic enzymes shared by the cephamycin C and CA pathways, the genes encoding CA biosynthetic enzymes are located in a cluster adjacent to the cephamycin C gene cluster (Figure 3) and, the two biosynthetic pathways are co-regulated by the same transcription activation protein, CcaR [24].

Originally, cephamycin C was highly resistant to the β-lactamase enzymes that cleave the β-lactam ring. However, due to bacterial resistance to β -lactam antibiotics, cephamycin C is currently used as an intermediate for important semisynthetic antibiotics such as cefoxitin, cefmetazole, and cefotetan [25]. β-lactamases are enzymes which open the β-lactam ring of penicillins and cephalosporins (β-lactam antibiotics), whereupon their anti-bacterial effectiveness is lost. β-lactamases are synthetized by many bacteria, therefore their resistance towards penicillins and cephalosporins (resistance acquisition)[26]. Clavams belong to a class of antibiotics seemly to β-lactam penicillin antibiotic, even

though in its chemical structure an oxygen atom is substituted by a sulfur. Regarding its chemical structure, CA has 3R,5R stereochemistry as opposed to the 3S,5S stereochemistry of the other clavams, which are commonly referred to as the 5S clavams (Figure 1). CA inhibits class A and D β-lactamases by binding irreversibly to the active site serine hydroxyl group of the β-lactamase producing a stable acylated intermediate, which results in the inactivation of the enzyme [27].

Although, CA is active against a broad spectrum of Gram-positive and Gram-negative bacteria, in relation to other broad-spectrum antibiotics, its antibiotic activity is relatively low [23]. Therefore, it has been proven advantageous to use β-lactam antibiotics in a mixture with CA or its pharmaceutically acceptable salts, maintaining the effectiveness of the β-lactam even in the presence of β-lactamase-producing bacteria. One antibiotic mixture example is the commercially available combination of amoxicillin and clavulanic acid potassium salt, which is broadly used in the control of infectious diseases [26] .

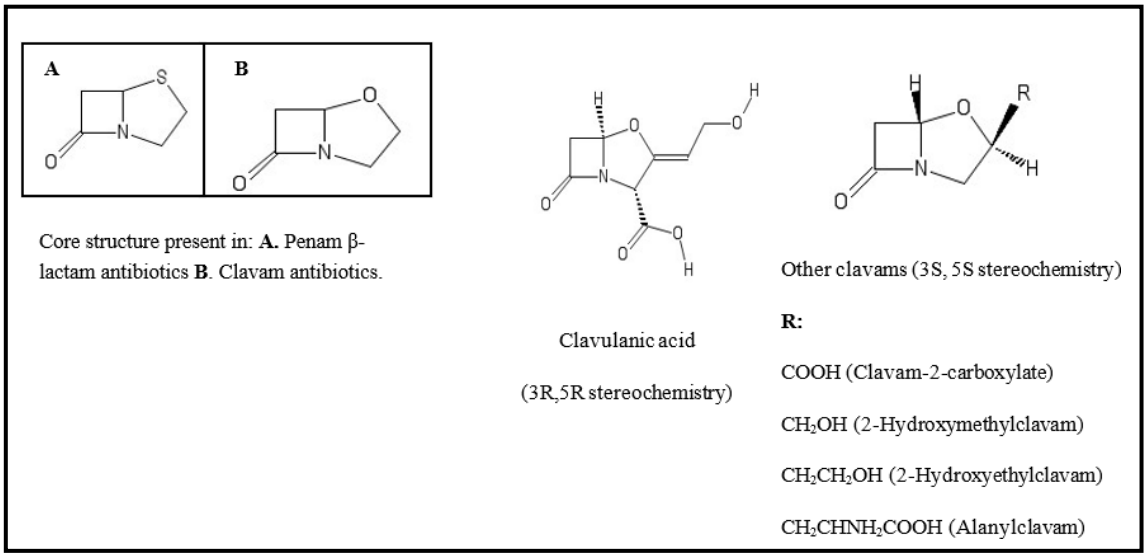


**Figure 1.** Primary skeleton structure that defines the β-lactam antibiotics and clavams. β-lactamase inhibitor clavulanic acid and the related 5S clavams; the core structure present in 5S clavams is shown and the corresponding side groups (designated by R) present in the different members.

CA is obtained by fermentation using *S. clavuligerus*. Depending on the culture medium composition and fermentation operating conditions, it has been possible to obtain yields around 1000 mg $L^{-1}$ of CA using the wild-type strain. *S. clavuligerus* preferably uses

glycerol as carbon source; the inclusion of glycerol enhances the production of CA as compared to other carbohydrates (e.g., starch), as glycerol provides a higher energy content on a weight by weight basis [28]. Furthermore, culture media formulations include a nitrogen source to provide adequate amino acid concentrations. The choice of soybean flour or isolate soy protein in complex media, significantly improves CA production [29], reporting an increased CA production for media containing soy protein isolate (698 mg / L) rather than soybean flour (338 mg / L) [28]. However, the influence of the primary nitrogen source (amino acids), as well as salt composition and other metabolites present in the culture media, can negatively affect CA biosynthesis. For instance, the presence of glutamate in GSPG medium negatively influence CA production. In the biosynthetic pathway in *S. clavuligerus*, glutamate can be converted to 2-oxoglutarate, which enters the citric acid cycle [30]. The conversion is catalyzed by a glutamate dehydrogenase and releases ammonium, which inhibits CA biosynthesis, hence explaining the negative influence of glutamate on CA production [30]. In general, the production of ammonium is an important issue since ammonium ion depresses the production of some antibiotics in *Streptomyces* spp. Thus, in CA production, ammonium ions can inhibit some enzymes related to CA biosynthesis as well as degrade the CA molecule present in the culture medium [31,32]. An increase in ammonium concentration above 15 mM (0,26 g/L) gives rise to the CA degradation rate constant, induced by CA degrading enzymes, as mentioned by Roubos et al., (2002) [33,32]. Based on the foregoing, culture medium optimization is an important milestone, which has made it possible to improve the low levels of CA production obtained using wild-type strains [28]. Notwithstanding, the maximum CA production reported so far has been achieved by the genetic modification of the industrial strain *S. clavuligerus* (see Table 2).

**Table 2.** Genetic modifications to increase CA production and to study regulatory aspects of its biosynthesis.

| Strain | Genetic modification | Relevant characteristics | Reference |
|---|---|---|---|
| *S. clavuligerus* Gap15-7-30 | Disruption of *gap1* | CA increment in 110%. | [34] |
| *S. clavuligerus* SB1918 *lat*::*apr-cvm1*::*apr* | Disruption of *lat* and *cvm1* | CA increment in 10% | [35] |
| *S. clavuligerus* IDG3 | Overexpression of *ccaR* | 25.9-fold CA increment | [36] |

| | | | |
|---|---|---|---|
| *S. clavuligerus* GV61 | Overexpression of *cas2* | 3.8-fold CA increment | [36] |
| *lat gene disrupted S. clavuligerus.* | Disruption of *lat* | CA increment in 10% | [37] |
| *S. clavuligerus ΔoppA2::aac* | Disruption of *oppa2* | No effect on CA production. Genes for holomycin are strongly upregulated. Production of *N*-acetylglycyl-clavaminic acid. | [38] |
| *S. clavuligerus ΔadpA* | Deletion of *adpA* | CA production decreases to about 10%. Lack of sporulation. | [39] |
| *S. clavuligerus* pIJadpA | Overexpression of *adpA* | CA increment in 218% | [39] |
| *S. clavuligerus* SMF5713 | Disruption of *orf23* | Effect on cell growth and sporulation. CA production was reduced to 40%. o*rf23* affects transcription of *claR* gene. | [40] |
| *S. clavuligerus/pICA2* | Overexpression of *orf21* | Hyperproduction of spores and 1.43-fold increased CA production by | [41] |
| *S. clavuligerus* M13-19 | Random mutagenesis | CA increment in 33%. Reporter gene *neo* fused downstream of *claR* | [42] |

Other CA and cephamycin C producing bacteria, besides *S. clavuligerus*, are *Streptomyces jumonjinensis* and *Streptomyces katsurahamanus*. Their genome sequences revealed that both contain CA BGCs in the same order as *S. clavuligerus*, except that the genes *orf18* (*pbpA*), *orf20*, *orf21*, *orf22*, and *orf23* are absent downstream of the CA gene cluster [43]. Additionally, *S. jumonjinensis* and *S. katsurahamanus* also contain CephC BGCs but lack the clavam and paralog gene cluster, which explain why they do not produce the 5*S* clavams as compared to *S. clavuligerus* [43]. Even though the three species altogether have the most important genes for CA biosynthesis, *S. clavuligerus* wild type strain has the highest production levels compared to the others. It has been previously suggested that higher CA yields in *S. clavuligerus* might be explained in part by increased precursors supply for biosynthesis due to the presence of the paralog and clavam gene clusters in this species [43].

The following full genome dendrogram was constructed to display the genetic relationship among the wild-type *S. clavuligerus*, the industrial strains (*S. clavuligerus* F613-1), and other CA-producing *Streptomyces* strains using the pangenome analysis software Roary [44]. It also serves to illustrate how genetic diversity (accessory genome) among the evaluated strains is one of the reasons for such variable CA titers.

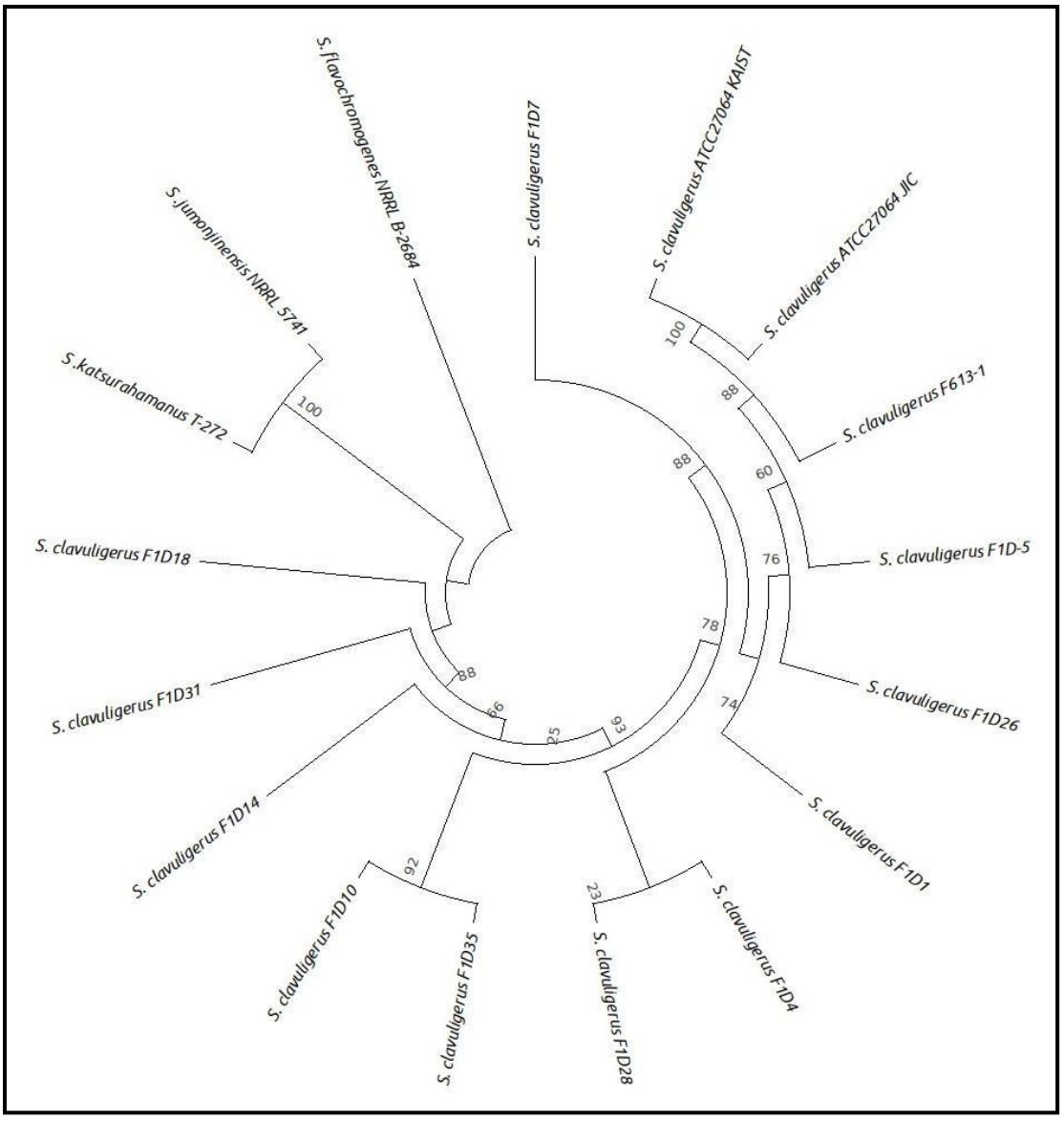


**Figure 2.** Genetic relationships deduced from CA producing *Streptomyces* strains pan-genome. Reference genomes reported in GenBank were used for all *Streptomyces* strains except for *S. jumonjinensis* and *S. katsurahamanus*. The genome of *S. flavochromogenes* NRRL B-2684 was used to root the three.

### *1.3.* Clavulanic acid gene cluster organization

In *S. clavuligerus*, the genes involved in the biosynthetic pathways for CA and 5S clavams reside in three distinct gene clusters that are not physically linked (see Figure 3), even though these metabolites share the early steps of the CA biosynthetic pathway (Figure 4) [45]. The CA gene cluster is located immediately downstream the cephamycin gene cluster, thus forming a larger gene cluster often referred to as the β-lactam supercluster (ceph-clav) (Figure 3A), situated in the 4.86-Mba *S. clavuligerus* chromosome region [45,46,47]. This CA gene cluster contains genes encoding enzymes involved in the early shared stages of the CA and 5S clavam pathway (early genes), as well as genes encoding proteins involved only in the later stages of CA biosynthesis (late genes). Table 3 shows the core genes of CA pathway and their function.

The second group of genes, referred to as the clavam cluster (Figure 3B) is also located in the chromosome in the 3.44-Mba region. This cluster includes one gene (*cas1*) involved in the biosynthesis of both CA and 5S clavams, as well as other genes involved exclusively in the 5S clavams biosynthesis [46,48,47]. Finally, the third group located in the plasmid pSCL4, is referred to as the paralogue gene cluster (Figure 3C) , and carries secondary copies for all genes, except *cas1*, found in both, CA and 5S clavam biosynthetic early stages in the CA gene cluster [48].

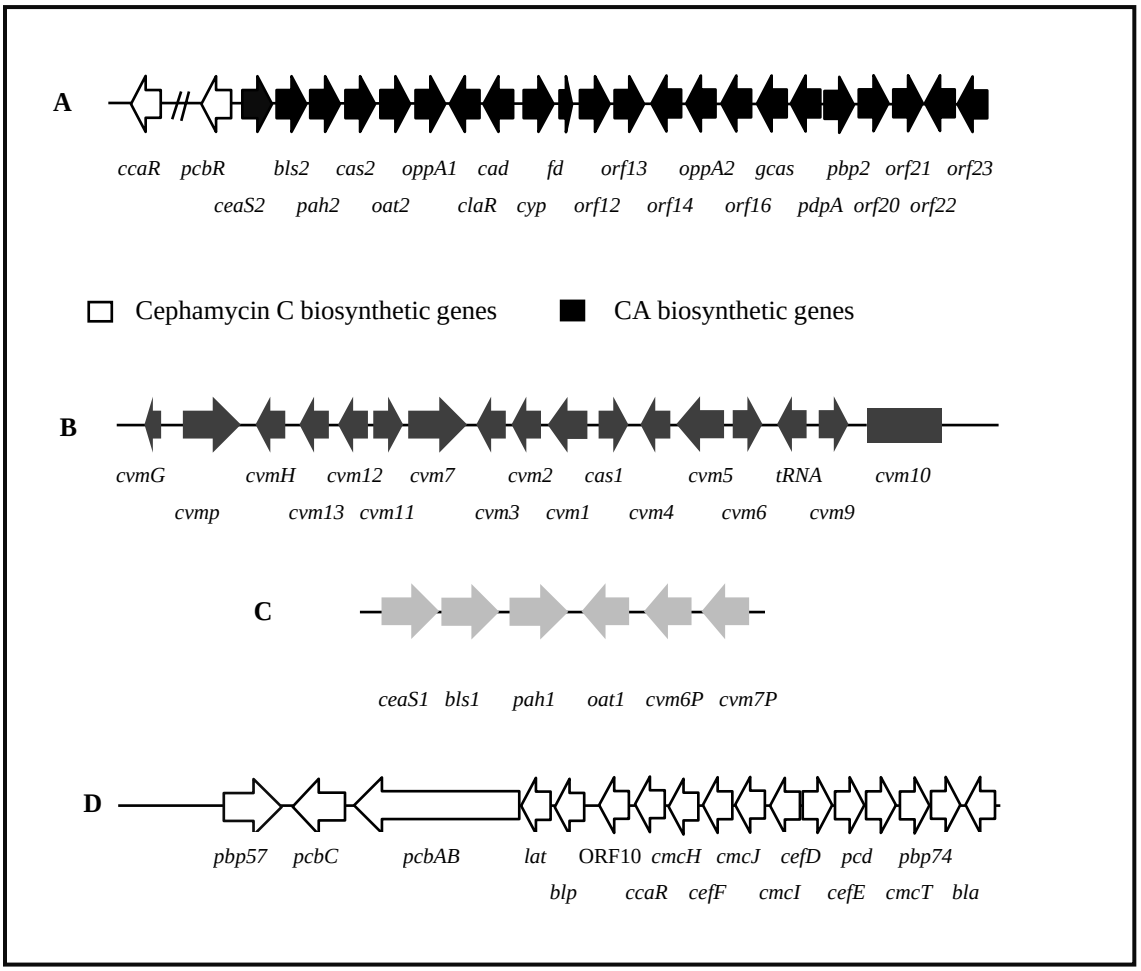


**Figure 3.** Gene cluster organization for CA and 5S clavams biosynthesis in *S. clavuligerus.* The arrows represent each gene with the arrowheads indicating their orientation (A) β-lactam supercluster (ceph-clav). (B) Clavam gene cluster. (C) Paralogue gene cluster. (D) Cephamycin C gene cluster.

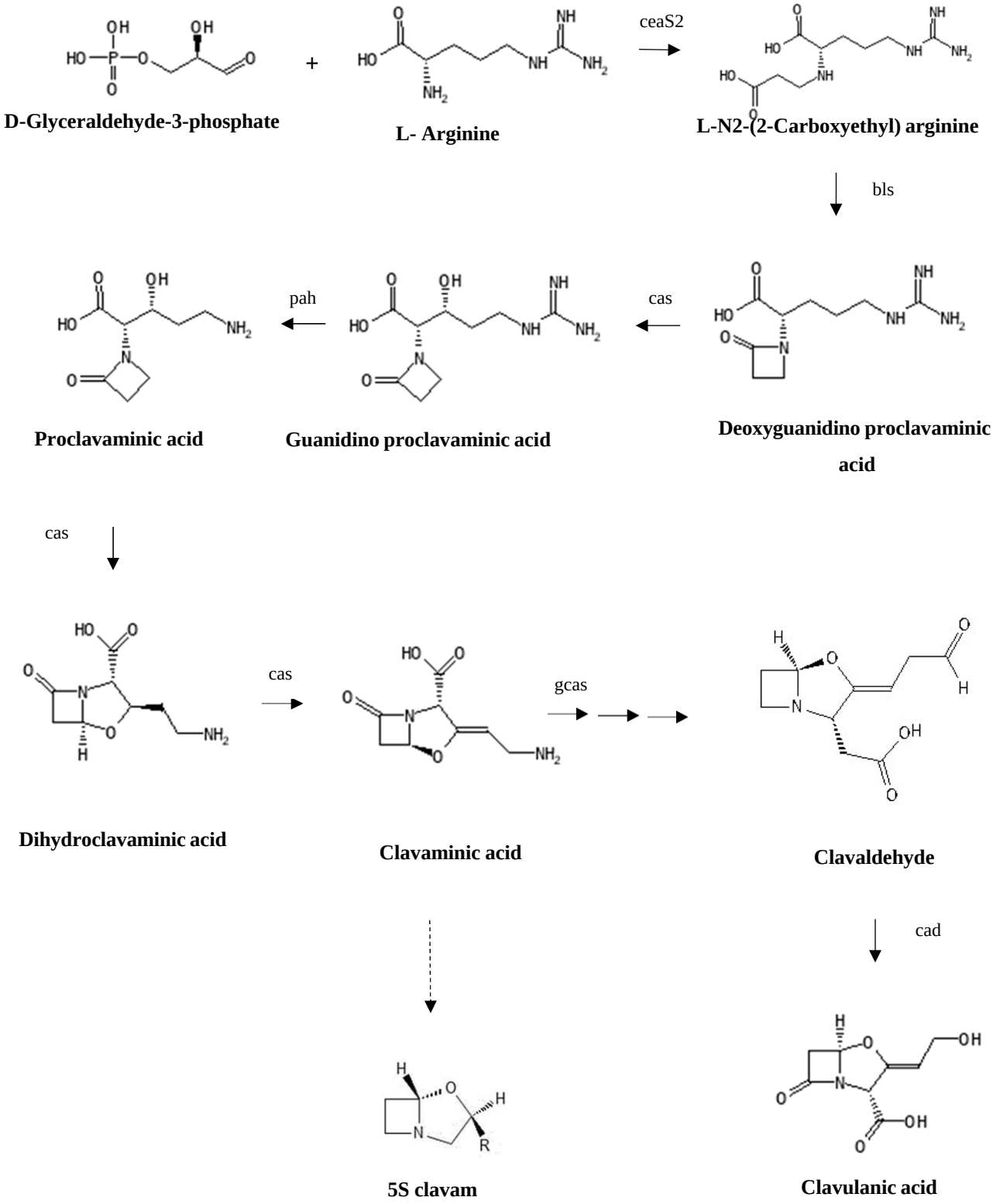


**Figure 4.** Biosynthetic pathway leading to CA and 5*S* clavam production in *S. clavuligerus*. Enzymes involved are indicated above the solid arrows. The dotted arrow represents unknown multiple reactions leading to different 5S clavam metabolites.

## *1.4.* Clavulanic acid biosynthesis

### 1.4.1. The early CA biosynthetic pathway (conversion of early precursors to clavaminic acid)

The biosynthesis of both CA and 5S clavams initiates with the condensation of glyceraldehyde-3-phosphate (G3P) and L-arginine to form $N^2$-(2-carboxyethyl) arginine, catalyzed by carboxyethylarginine synthase (CeaS2), which is encoded by the *ceaS2* gene [49]. The second step generates the β-lactam ring found in clavams; this reaction is catalyzed by β-lactam synthase (Bls2), which is encoded by two separate genes (*bls1* and *bls2*), to give the deoxyguanidino proclavaminate molecule [45]. Then, clavaminate synthase (Cas) and proclavaminate amidinohydrolase (Pah) alternate back and forth to catalyze reactions that form the bicyclic intermediate clavaminic acid [49,45] (see Figure 4). The formation of clavaminic acid is an important branch point in the CA pathway since it can either be converted to (3R, 5R)- clavulanic acid, or to the 3S, 5S clavams, by the late reactions of the pathway [50].

### 1.4.2. The late CA biosynthetic pathway (conversion of clavaminic acid to CA)

Up to this point, the clavaminic acid biosynthesis steps (including 3S, 5S stereoisomers) have been shared for CA and 5S clavams biosynthetic pathways; a next step leads to a divergence metabolic node [49]. The *cvm6P* gene directs part of the clavaminic acid metabolic pathway flux leading to 5S clavams. For CA biosynthesis, different compounds result from the intermediate steps involved in the transition of the clavaminic acid into CA. The first late reaction is the conversion of clavaminc acid to N-glycyl-clavaminic acid, catalyzed by N-glycyl-clavaminic acid synthase (Gcas); then a possible acetylation of N-glycyl-clavaminic acid leads to N-acetylglycyl-clavaminic acid molecule, which binds to the OppA2 protein to be transported to the place in the cell where the next enzymatic reactions occur [38]. Subsequently, through some enzymatic reactions that have not been well characterized, the N-acetylglycyl group from N-acetylglycyl-clavaminic acid is released by an oxidative deamination to produce clavaldehyde. The final step of the pathway is the conversion of clavaldehyde to CA by clavaldehyde dehydrogenase (Cad) [51] (Figure 4).

**Table 3.** Genes involved in the clavulanic acid pathway.

| Gene/ORF | Function/Putative role | Reference |
|---|---|---|
| **Genes in the early steps of the CA pathway** | | |
| *ccaR* | *ccaR* encodes the CcaR protein. CcaR, positively regulates both cephamycin C and CA production. | [49] |
| *ceaS2, ceaS1* | Condensation of L-arginine and glyceraldehyde-3-phosphate to produce N2-(2-carboxy-ethyl) arginine. | [52] |
| *pcbR* | Penicillin-binding protein involved in β-lactam resistance and part of the cephamycin C gene cluster. | [45] |
| *bls2, bls1* | β-lactam synthase that catalyzes the Cyclization of N2-(2-carboxy-ethyl) arginine to form deoxyguanidino proclavaminate. | [53] |
| *pah2, pah1* | Proclavaminate amidinohydrolase. Hydrolysis of guanidine-proclavaminate to form proclavaminate. | [49] |
| *oat2, oat1* | A predicted ornithine acetyltransferase. Somehow regulates CA production in response to arginine levels. | [54] |
| **Genes in the late steps of the CA pathway** | | |
| *oppA1, oppA2* | Oligopeptide permease binding proteins. Unknown function. | [52] |
| *claR* | Transcriptional activator. Positively regulates late steps of the CA pathway. | [55] |
| *cad* | Clavaldehyde dehydrogenase. Reduction of clavaldehyde to CA | [52] |
| *cyp* | Cytochrome P-450 protein. Involve in oxidative reactions. | [49] |
| *fd* | Ferredoxin protein. | [49] |
| *orf12* | ORF12 is, functionally, like the esterase VIII superfamily. ORF12 is involved in the requisite epimerization step during the conversion of clavaminic acid to CA. | [56] |
| *orf13* | Predicted to encode a membrane transport protein. Possible role in transporting the clavam metabolites out of the cell. | [52] |
| *orf14* | Acetyltransferase | [52] |
| *orf16* | Hypothetical protein. Role in conversion of *N*-acetylglycylclavaminic acid to *N*-acetylclavaminic acid. | [49] |
| *gcas (orf17)* | *N*-glycyl-clavaminic acid synthetase. Clavaminate to N-glycylclavaminate | [57] |
| *orf20* | Cytochrome P-450. Unknown function. | [43] |
| *orf21* | Putative sigL sigma factor | [41] |
| *orf22, orf23* | Two-component regulatory system. Sensor kinase/response regulator. | [58] |

## *1.5.* Regulation of clavulanic acid in *Streptomyces clavuligerus*

### 1.5.1. The pathway specific regulators of CA biosynthesis

Usually, genes responsible for the biosynthesis of a specific antibiotic are arranged in clusters. Within the cluster there are structural genes that encode enzymes responsible for catalyzing the chemical reactions involved in biosynthesis of the metabolite of interest. In addition, these clusters contain regulatory functions encoded by these gene arrangements [49]. Among them, the expression of pathway-specific regulators located in each gene cluster, in a few cases outside of the regulated cluster, are controlled primarily by genes present in the same cluster [59]. However, increasing evidence suggests that some of these genes control distant clusters or genes outside their cluster. For this reason, the name cluster-situated regulator (CSR) has been proposed instead of pathway-specific genes[59]. Likewise, there are many other regulators that exert their functions at higher levels. Global or pleiotropic regulators, for instance, are often located far side their cluster and control the expression of genes involved in morphological development and/or CSR genes, indirectly affecting the production of antibiotics [60]. Furthermore, there are global regulators directly affecting structural antibiotic biosynthesis genes [60].

The CA gene cluster in *S. clavuligerus* is controlled by a network of regulators that connect secondary metabolism with the cell environment. Global regulators in response to changes in nutritional or environmental factors bind to the promoters of CSR genes, which in turns controls the expression of the biosynthetic genes in the respective clusters [59]. The CSRs belong to different protein families; one of these classes comprises the SARP (*Streptomyces* antibiotic regulatory proteins), typified by the CcaR protein of *S. clavuligerus* [24] and AfsR in different *Streptomyces* spp [60]. CcaR is encoded by the *ccaR* gene located in the cephamycin C gene cluster and is required for the biosynthesis of CA and cephamycin C. CcaR is an autoregulatory protein that controls its own expression, cephamycin C gene cluster expression, and the expression of other genes in the CA gene cluster. Thus, the CcaR protein binds to a heptameric sequences located upstream the *lat* gene and the intergenic *cmcI-cefD* region in the cephamycin C gene cluster, and in the promoter region upstream of *ceasS2* and *claR*, activating the transcription of the early genes and, indirectly, the late genes of CA pathway (see Figure 3). In the case of

cephamycin C, the *lat* gene encodes for lysine-amino transferase, the first enzyme of the cephamycin C pathway that form the α-aminoadipic acid precursor from lysine in a two-step reaction [61]. The low or no production of lysine-amino transferase produces no or very low levels of Cephamycin C in the wild type *S. clavuligerus,* and raised specific production of CA (see Table 2) [37].

In the CA gene cluster, the *claR* gene encodes the ClaR protein of the LysR family, an important pathway-specific regulator located at the bottom of the regulatory network that positively regulates only the late biosynthetic genes involved in the conversion of clavaminic acid to CA. Both CcaR and ClaR in *S. clavuligerus,* are controlled by global regulatory mechanisms, including, (a) transcriptional regulators as *bld* genes, (b) production of autoregulators, and (c) stringent response.

### 1.5.2. Global regulators associated with CA biosynthesis.

The expression of transcriptional regulators (e.g., *ccaR*) is controlled by global regulatory mechanisms, bound to the production of antibiotics, depending on environmental conditions and bacteria physiological state, e.g., via AdpA protein. AdpA, which is ubiquitously distributed in streptomycetes as a member of the AraC/XylS family regulators, affects transcription of hundreds of genes involved in morphological differentiation and antibiotic biosynthesis [62]. Thus, AdpA affects aerial mycelium formation in *S. clavuligerus* and positively modulates the production of CA in transformants carrying multiple *adpA* copies [63]. Similarly, nutrient starvation environmental signals, stress generated by adverse growth conditions, concentration of chitin or chemical signals, trigger this cellular response. Pinilla et, al. (2019), established that CA synthesis in *S. clavuligerus* occurred under stress conditions induced by exhaustion of amino-acid content [64].Under this depletion, the overexpression of metalloprotease SCLAV_4359 at the onset of cell lysis was observed, suggesting that the synthesis of this protease may be directly involved in morphological differentiation and indirectly towards CA production [64].

As for *Streptomyces coelicolor* (*S. coelicolor*), several genes have been characterized exerting functions on the global regulation of morphological differentiation. Such genes are referred to as *bld* (bald) genes since mutants lack the characteristic fuzzy coating of aerial hyphae noticed in the wild-type strain. Some of the *bld* mutants that have been isolated are also defective in the production of secondary metabolites, such as antibiotics,

suggesting that the corresponding *bld* genes are involved in the regulation of both processes [65].

Many *bld* genes are known to control both sporulation and antibiotic production in *S. coelicolor*, but only *bldA*, *bldH*, *bldN* and *bldG* have been studied in other *Streptomyces* spp [66]. In the case of *S. clavuligerus*, the *bldG* gene was the first *bld* regulatory gene identified that functions upstream of *ccaR* and *claR* in the CA biosynthesis regulatory cascade, affecting CA production and the morphological differentiation of *S. clavuligerus* [65]. The *bldG* gene encodes a putative anti-anti-sigma factor that shows homology to anti-anti-sigma factor proteins of *Bacillus* spp. Generally, these kinds of proteins function along with a cognate anti-sigma factor protein to regulate and target the activity of a sigma factor. For instance, an interesting regulatory mechanism occurs in *Bacillus subtilis*; the SpoIIAB anti-sigma factor regulates the activity of the sporulation-specific sigma factor $\sigma^F$ by binding to it and preventing the association of the sigma factor with the core RNA polymerase. Likewise, the anti-anti-sigma factor is regulated by phosphorylation, since only the unphosphorylated form of the protein is able to bind to the anti-sigma factor [66,65].

Alternatively, the *bldA* gene in *S. clavuligerus* encodes the leucine-specific transfer RNA, which is required to translate the rare UUA codon. Regarding *S. coelicolor, bldA* mutants are defective in sporulation and antibiotic production, suggesting that the *bldA* tRNA is required for translation of genes involved in both processes [67]. The *ccaR* gene in *S. clavuligerus* contains a TTA codon; however, disruption of the *bldA* gene in *S. clavuligerus* had no effect on translation of *ccaR* or on cephamycin C and CA production, since, for the case of *S. clavuligerus,* the *ccaR* gene is mistranslated efficiently [67] .

Sigma factors also control antibiotic production in *Streptomyces* spp. Bacterial sigma (σ) factors are an essential component of RNA polymerase and determine promoter selectivity. The substitution of one σ factor for another can redirect some or all of the RNA polymerase in a cell to activate the transcription of genes that would otherwise be silent [68]. Therefore, sigma factors are the essential nodes in gene regulatory and metabolic network fluxes that govern further interactions and processes in the cell [69].

Sigma factors regulate the transcription of many genes in the cell. The extracytoplasmic function (ECF) σ factors constitute a subfamily of sigma factors, related to the regulation

of gene transcription during various stress responses which leads to morphological development. In general, in the bacterial cell, the ECF σ factors are co-transcribed with one or more negative regulators; often, these include a transmembrane protein functioning as an anti-σ factor that binds, and inhibits, the cognate σ factor. Upon receiving a stimulus from the environment, the σ factor is released and can bind to RNA polymerase to stimulate gene transcription (see Figure 5). In many ways, these anti- σ:σ pairs are analogous to the more familiar two-component regulatory systems consisting of a transmembrane histidine protein kinase and a DNA-binding response regulator. Both are mechanisms for coordinating a cytoplasmic transcriptional response to signals perceived by protein domains external to cell membrane [68].

For *Streptomyces* species, several ECF σ factors are known to be involved in stress response and have indirect roles in regulating secondary metabolite biosynthesis by association with the activation of genes related to morphological development. BldD is a transcriptional regulator essential for morphological development and antibiotic production in *Streptomyces coelicolor*. *S. coelicolor bldD* mutants are blocked at the earliest stage of morphological differentiation and also fail to produce the secondary metabolites, actinorhodin, prodigionines, methylenomycin and CDA [70]. Among the known BldD-regulated targets, there are two genes encoding sigma factors ($\sigma^{BldN}$ and $\sigma^{WhiG}$) that, in turn, play crucial roles in *Streptomyces* development [70]. Conversely, deletion of SigK (SCO6520), a sigma factor in *S. coelicolor*, caused an earlier switch from vegetative mycelia to aerial mycelia and higher expression of *chpE* and *chpH* than the wild type strain. Loss of SigK also resulted in accelerated and enhanced production of antibiotics, actinorhodin, and undecylprodigiosin, and increased expression of *actII-orf4* and *redD* genes. These results suggested that SigK had a negative role in secondary metabolism and morphological transitions [71].

The role of sigma factors in the bacteria morphology and CA production is poorly understood, as it is the case of *S. clavuligerus*. The *orf21* gene encoded a sigma factor that positively regulates CA production, as *orf21*-deleted mutants produce less antibiotic compared to the wild type strain. Conversely, *orf21* overexpression increases CA production [41]. Jensen et, al. (2008) identified the open reading frame denoted as *orf21*, which was annotated as a putative sigma factor, located downstream of the CA gene

cluster in *S. clavuligerus* ATCC27064 [40]. Later, Jnawali et al., (2010) studied the *orf21* role in CA production. The authors showed that, under specific fermentation conditions in SA (Starch-Asparagine) media, both the overexpression and disruption of *orf21* affects CA production in *S. clavuligerus* NRRL3585 [41]. Although the physiological role of *orf21* has been previously studied in *S. clavuligerus*, its relationship with CA biosynthesis remains a matter of discussion [41,40,49].

### 1.5.3. Two component regulatory systems

Another regulatory mechanism present in bacteria is the two-component system that controls many responses of bacteria to changes in their environment, including antibiotic production. In each of these regulatory systems, one protein, called a *histidine kinase sensor*, contains a latent histidine kinase transmitter domain that is regulated in response to environmental changes, detected by a sensor domain (Figure 5). When activated, the transmitter domain transfers the ATP γ-phosphate to a histidine residue located in the transmitter domain. The second protein, called a *response regulator*, contains a *receiver* domain harboring an aspartic acid residue that is phosphorylated [72] (Figure 5A).The response regulator contains a second functional domain that is regulated by the receiver domain phosphorylation. In many cases, this domain is a sequence-specific DNA-binding domain that functions either as a repressor or as an activator, regulating the transcription of specific genes. Further, the effector domain possesses other functions as well, such as controlling the direction in which the bacterium swims in response to nutrient concentration gradients [72].

Response regulators involved in antibiotic biosynthesis are often located within their biosynthetic gene clusters. As for the case of *S. clavuligerus* NRRL3585, a pair of genes (*orf22* and *orf23*), that encode bacterial two component regulatory systems, were found next to the CA gene cluster [73]. Studies conducted by Jnawali (2008), showed that a two-component regulatory system (TCS) participates as a CA biosynthesis positive regulator; therefore, increasing levels of *orf22*/*orf23* gene products can contribute to enhanced CA production in *S. clavuligerus.* Nevertheless, the mechanisms by which *orf22/orf23* affects CA production had not been further depicted. The paired genes *cagS* and *cagR*, which are annotated in *S. clavuligerus* ATCC 27064 as *orf22* and *orf23* respectively, encode a bacterial two-component regulatory system (TCS) located next to the CA biosynthetic gene cluster in *S. clavuligerus* F613-1[58]. The TCS CagRS system mainly regulates genes

involved in primary metabolism, modulating, in a negative manner, the biosynthesis of arginine and, positively, glyceraldehyde 3-phosphate metabolism. In addition, TCS CagRS may affect CA production both directly, through the CA biosynthetic gene cluster (*oat1*, *oat2*, *ceaS1*, and *claR*) and indirectly by affecting arginine and G3P metabolism [58].

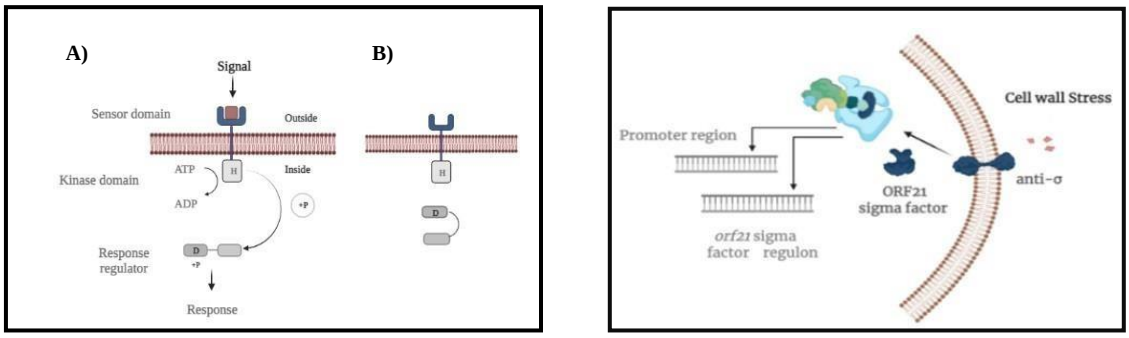


**Figure 5.** Hypothetical schematic representation of some global regulatory mechanisms associated with CA biosynthesis. On the right side, ORF21 trigger mechanism as an example of ECF sigma factor regulation. An extracellular signal induces a conformational change in the anti-sigma -sigma factor complex, releasing the sigma factor (ORF21) and allowing the transcription of its regulon. On the left, the two component systems (Adapted from Sharan et al., 2017 [74]); A) Active, B) Inactive.

### 1.5.4. Alarmone and stringent response

In harsh environments, where there is a lack of amino acids, limitation of minerals, or any stress condition, bacterial metabolism responds producing small signaling molecules (stringent response), known as alarmone; it works as an alarm to the presence of stress and provides a signal to slow down unnecessary processes in most bacteria, directing gene expression to adaptation to stress [59,75]. The first signaling molecules to be identified were the γ -butyrolactones (also known as autoregulatory factor) from *Streptomyces* spp. in the 1960s.

To date, there are several γ -butyrolactones identified in different *Streptomyces* species and, they are mainly classified into three groups based on minor structural differences in the C2 side chain: (1) the virginiae butanolide (VB) type, exemplified by VB-A to VB-E from *Streptomyces virginiae*, which contains a 6-α-hydroxy group; (2) the IM-2 type, exemplified by *Streptomyces lavendulae* IM-2, FRI-5 and SCB1 from *Streptomyces coelicolor*, A3(2), which contains a 6-β-hydroxy group; and (3) the A-factor type,

containing a 6-keto group [76]. As for *Streptomyces* species like *S. virginiae*, *S. lavendulae* and *S. coelicolor*, γ-butyrolactone autoregulators have specific receptors that act as signal transducers, suggesting that receptor-mediated signal transduction is common in *Streptomyces* spp.[76].

The γ -butyrolactones regulate the production of antibiotics in nano molar concentrations and in some cases they regulate differentiation [75]. In general, γ-butyrolactones bind to cytoplasmic receptor proteins (e.g., ArpA in *S. griseus*) and inhibit their binding to specific DNA targets. Thus, most of these receptor proteins act as repressors, so that binding to γ -butyrolactones induces expression of the target genes [75]. For *S. clavuligerus*, the *scaR* gene encodes the ScaR autoregulator receptor showing affinity towards IM-2 type autoregulators containing a long C2 side chain. However, it is unclear whether ScaR participates in the regulation of secondary metabolites as CA [76].

Similarly, *brp* gene encodes the Brp butyrolactone receptor protein [77]. This receptor contains a DNA-binding domain which binds specifically to consensus binding sequences for butyrolactone receptors called $ARE_{brp}$ (located 61 nucleotides upstream the *brp* gene), $ARE_{ccaR}$ (located 815 bp upstream the *ccaR* gene [77]), and to $ARE_{adpA}$, located upstream the *adpA* gene [63]. Regarding CA, Brp can act as a repressor protein of CA and cephamycin C biosynthesis by its binding to the ARE sequences located in the promoter region of *ccaR* and *adpA* genes, inhibiting their expression [77,63]. Despite efforts to understand the relationship between γ -butyrolactones and the biosynthesis of secondary metabolites in *S. clavuligerus*, especially in terms of CA production, the specific ScaR and and Brp receptors autoregulators have not been characterized yet.

Another type of alarmones that modulate the response to stress triggered by stringent response, are several purine derivatives like guanosine pentaphosphate (pppGpp), guanosine tetraphosphate (ppGpp), collectively known as (p)ppGpp. During nutrient rich growth condition, the basal cellular level of (p)ppGpp in *E. coli* is lower than 0.2 mM. Upon stress induction, the level of (p)ppGpp may increase from 10 to 100-fold depending upon the stress type and enzymes involved in the biosynthesis of different biomolecules [78]. As for *E. coli*, the intracellular level of (p)ppGpp is controlled by RelA, a (p)ppGpp synthetase encoded by the *relA* gene, and the (p)ppGpp 39-pyrophosphohydrolase

enzyme encoded by *spoT* gene. RelA is a ribosome associated protein and is activated, presumably by conformational changes, when uncharged tRNAs bind to the ribosome A site [78]. In *S. clavuligerus*, *relA* gene encodes the (p)ppGpp synthetase; furthermore, relA null mutants are unable to form aerial mycelium or to sporulate, and CA production increased three- to fourfold, suggesting a negative regulation of CA biosynthesis by (p)ppGpp in *S. clavuligerus* [79].

### *1.6.* Genetic modifications as a tool to improve clavulanic acid production in strains of *Streptomyces clavuligerus*

Various genetic modifications have been made in *S. clavuligerus* to increase CA production (see Table 2). Methods for increasing the productivity of industrial microorganisms range from the classical random mutagenesis to the use of more rational methods (e.g., CRISPR-related genetic manipulations). One of these rational methods is metabolic engineering where, to maximize product yields, primary metabolic fluxes are redirected toward the target thus reducing byproducts. This is performed by the introduction of genetic modifications through recombinant DNA technology, in a manner that supports high secondary metabolite productivities [80].

A crucial prerequisite for genetic engineering is efficient DNA transfer techniques for recombinant genetic material. However, depending on the exogenous DNA origin, it must be modified before being transferred to *Streptomyces* to avoid its degradation by the potent methyl-specific restriction barrier that some *Streptomyces* such as *S. clavuligerus* possess. To overcome this restriction barrier, the exogenous DNA can be previously isolated from a methylation deficient *E. coli* host. Subsequently, *Streptomyces* can be transformed with the recombinant DNA by electroporation, intergeneric conjugation from *E. coli* into *Streptomyces*, or by protoplasts in the presence of polyethylene glycol (PEG). Electroporation uses the addition of a short electric current to create transient pores in the membrane for the uptake of DNA. However, conditions for electroporation are strain specific, and therefore optimization is required based on the *Streptomyces* species [80]. Conjugation involves the transfer of genetic material through direct cell-to-cell contact. Intergeneric conjugation between *E. coli* and *Streptomyces* is possible with the use of bi-functional vectors, which are species specific [80]. There are several advantages for using

conjugation over transformation as a DNA transfer method: (1) There is no need for protoplast formation and regeneration. (2) Restriction barriers may be surpassed if single stranded plasmid DNA is transferred.
Alternatively, the formation and regeneration of protoplasts is an efficient technique for *S. clavuligerus* transformation described by Okanishi *et al*, (1974) [81]. However, it is a delicate technique as it requires removing the thick cell wall of *Streptomyces* to allow the uptake and incorporation of recombinant DNA in the presence of PEG 1000, that could increase the binding of cells and DNA, being more sensitive and easier to obtain positive transformants.

The cloning of a gene can be carried out in different types of plasmids and under the control of different promoters (constitutive or inducible). Constitutive promoters are widely used in expression studies of actinomycetes, including overexpression of genes, and heterologous expression of genes and gene clusters [82]. For example, the *ermE* constitutive promoter (*ermEp*) of *Saccharopolyspora erythraea* controls the expression of *ermE* gene, which encodes the 23S rRNA methyltransferase and confers resistance to the macrolide and lincosamide antibiotics erythromycin and lincomycin, respectively [82]. The *ermEp* presents a complicated structure, where two separate promoters, *ermEp1* and *ermEp2*, have been identified. The -10 regions of *ermEp1* and *ermEp2* resemble a sequence of the consensus prokaryotic promoter, whereas the -35 region, which is also recognizable, shows much greater variability. Trinucleotide TGG deletion within the -35 region of *ermEp1* has been reported to increase the level of promoter activity by approximately five-fold [82]. This upregulated mutant of *ermEp* is known as *ermEp** and, it is present in the integrative plasmid, pIB139 [83] (Figure 6) which was used for the recombinant studies carried out in the present work.

The activities of certain constitutive promoters, however, are not constant and can vary significantly under cultivation conditions or growth stages. For example, constitutive promoters based on the widely used *ermE* lead to significantly stronger gene expression in the stationary growth phase than in the log phase, whereas the constitutive promoters based on *rpsL* increase the active transcription of genes in the earlier growth phases of actinomycetal cultures [82]. Therefore, promoter actions must be considered when planning an experiment of gene cloning.

In the case of plasmids, these can be integrative or multicopy. For instance, integrating plasmids as pIB139 [83], provides a single copy of the recombinant gene and optimal segregational stability. Integration of pIB139 into the chromosome of *S. clavuligerus* is carried out via homologous recombination between the natural *att*P site from a phage region present in the plasmid and the *att*B sites present in the genome of *S. clavuligerus.*

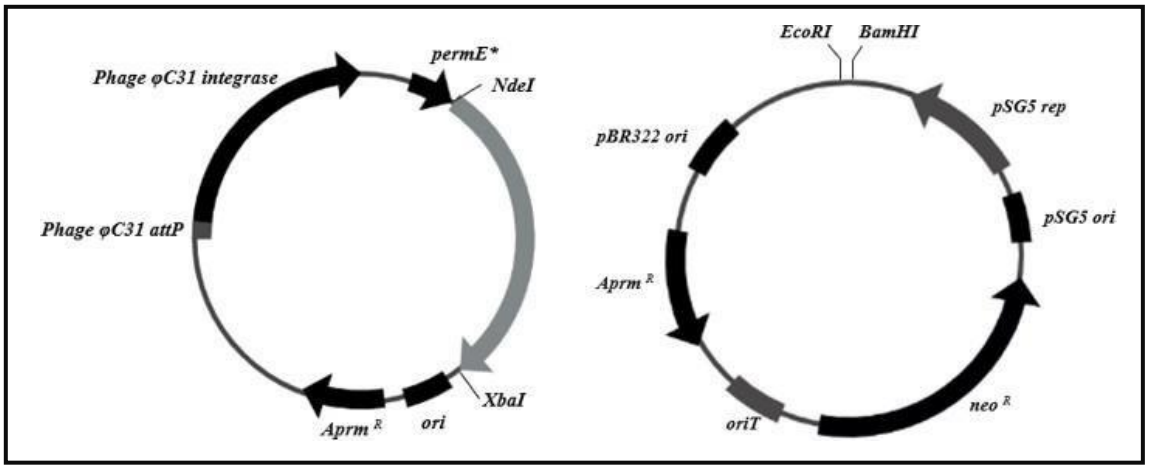


**Figure 6.** Plasmids used for *S. clavuligerus* transformation. On the right, the high- copy number plasmid pSOK201 (Adapted from Zotchev et al., 2000 [84]). On the left, the integrative plasmid pIB139 with a recombinant gene cloning under the control of the constitutive promoter, p*ermE**.

Conversely, high-copy number plasmids allow to obtain many copies per cell of the cloned gene. Nevertheless, these plasmids can be unstable in the cell, as they represent a significant metabolic burden for their hosts. pSOK201 is a self-replicating plasmid with intermediate copy number, which can be transferred to *Streptomyces* spp by protoplasts or conjugation. Some important genetic arrangements in pSOK201 are: the origin of transfer (*oriT*) and pSG5 replicon (from *Streptomyces ghanaensis*) [85]. *oriT* element allow the mobilization of the plasmid construct during conjugation from an appropriate *E. coli* strain into *Streptomyces* spp, while the replicon pSG5, allows plasmid elimination from its host cell. pSG5 replicon is stable at temperatures below 34°C but is lost at incubation temperatures above this. Thus, plasmids constructed using the pSG5 replicon, as pSOK201, are suitable for application of gene disruption techniques by allowing the selection of cells that undergo the mutation [84,85].

As mentioned early, there are different techniques to disrupt gene expression. One of these techniques is based on the homologous recombination (HR) mechanism, which can

be carried out by single crossover or double crossover [86]. HR can be defined as an exchange between two DNA sequences in the region of *shared* homology, where the two DNA molecules are broken and rejoined to each other forming a crossover [87]. During meiosis, in eukaryotic cells, HR allows the exchange of genetic material between homologous chromosomes. Among bacteria, fragments of DNA may be recombined into the chromosome after entering the cell as a result of transformation, transduction, or conjugation. Moreover, HR mostly serves to repair double strand break or single strand gaps in bacteria's DNA and, during mitosis in eukaryotic cells [87]. Figure 7 illustrates the mechanisms of single crossover and double crossover used for gene disruption.

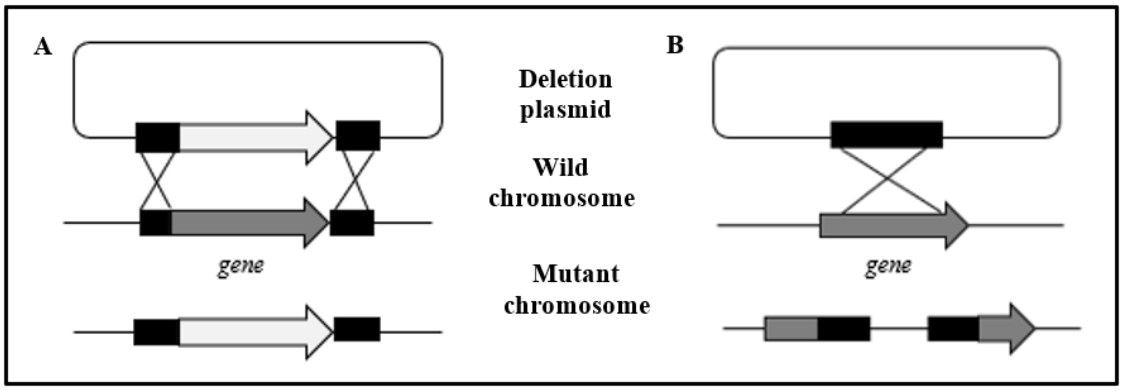


**Figure 7.** Illustrative procedure of recombination between a deletion plasmid and a linear chromosome by the region of shared homology (in black). A) Generation of *gene*-deleted mutant via double- crossover integration. In this case, the gene to be deleted in the wild chromosome is replaced, for example, with an antibiotic resistance gene. B) Generation of *gene*-deleted mutant via single- crossover integration. The deletion plasmid is integrated into the gene for its disruption.

### *1.7.* The present study

As mentioned before, CA is a secondary metabolite of great importance for the health industry due to its pharmacological benefits in the treatment of bacterial resistance to beta-lactam antibiotics. Nevertheless, the biotechnological CA production is hampered, given the molecule chemical instability and the low CA titers attained by the *S. clavuligerus* wild type. Consequently, much of the research on CA production has been focused on deciphering *S. clavuligerus* genetics for strains manipulation and redirection of carbon fluxes towards CA biosynthesis. Despite multiple studies concerning *S. clavuligerus* genetic manipulation, it is worth highlighting that unlike gram-negative bacteria such *E. coli*, the genetic manipulation of Actinobacterias as *Streptomyces*, is

complicated and time consuming. Thus, for *S. clavuligerus*, the lack of compatible molecular tools, limited cloning and DNA transfer methods, DNA degradation, genetic instability, high genome guanine-cytosine content (GC-content), and different codon usage for newly introduced foreign DNA, make genetic manipulation challenging [88]. Hereby, contributing to *S. clavuligerus* genetic enhancement, as well as strengthening the knowledge regarding its genetics, will allow not only to generate strategies to improve CA production, but also to use these principles as a basis for the engineering of other secondary metabolites.

In general, CA production in *S. clavuligerus* is controlled by complex regulatory mechanisms, where pathway-specific regulators and global regulators connect the bacterium metabolism with the environment. Naturally, microorganisms form complex communities where multiple species coexist in the same place. Thus, the constant interactions among bacteria induce the production of secondary metabolites e.g., antibiotics, to allow a better adaption to their surroundings. Furthermore, the biosynthesis and regulation of some secondary metabolites is not an isolated mechanism since the bacterium can regulate expression synchronously to improve its adaptation. A striking example of evolution and natural adaptation to microbial resistance in *S. clavuligerus* is the co-regulation of beta-lactam antibiotic (cephamycin C) and the beta-lactamase inhibitor, CA [49]. Thus, dissociating the biosynthesis of secondary metabolites that share metabolic resources is one of the strategies used to improve CA yields.

Although research efforts aiming to improve CA production yields, have made possible to understand multiple aspects of *S. clavuligerus* biology, many of the molecular mechanisms that activate and control CA production in *S. clavuligerus* are still unknown. The review on the regulatory role of some poorly studied genes located in the CA cluster, will allow to improve the understanding of the relationships governing environmental conditions and CA biosynthesis, and therefore, will further narrow down the biotechnological conditions for its production.

Finally, this research seeks to improve CA production by performing genetic engineering techniques to obtain *S. clavuligerus* recombinant strains with either, reduced capacity for cephamycin C biosynthesis, and/or a CA gene cluster regulatory gene(s) overexpressed. Initially, it was proposed to disrupt the *lat* gene and overexpress the *claR* gene in order to construct a double *S. clavuligerus* mutant strain to redirect some of the metabolites shared between cephamycin C and CA metabolic pathway, towards the late CA biosynthetic

stages. However, given the intrinsic complexity of *S. clavuligerus* and the challenges evolved during its genetic manipulation, it was not possible to generate the double mutant strain. Hence, our research was focused on the construction of two *S. clavuligerus* recombinant strains over-expressing one CA regulatory gene each. This final approach explored the regulatory effect of a sigma factor (ORF21) and a specific CA pathway regulator (CLAR) under different nutritional conditions. Likewise, the effect of gene overexpression was studied on attributes such as: bacterial morphology, primary metabolism (consumption of substrate and ammonium production), gene expression and fermentation kinetic parameters.

# 2. OBJECTIVES

**General objective:**

To investigate and apply genetic engineering techniques aim at overexpressing a regulatory gene, and/ or disrupting an early gene in a competing pathway to enhance clavulanic acid production in *Streptomyces clavuligerus*.

**Specific objectives:**

1. To generate a *S. clavuligerus* mutant strain that might have the potential to promote carbon flux towards CA biosynthesis, by performing either one or both of the following strategies:
   *a)* Standardization of a genetic transformation strategy necessary for disrupting an early gene in the cephamycin C biosynthetic pathway in *S. clavuligerus*.
   *b)* Standardization of a strategy for overexpressing a regulatory gene of the CA metabolic pathway in *S. clavuligerus*

2. To eventually determine high levels of CA production using the mutant strain(s) of *S. clavuligerus.*
   *a)* By performing cell suspension cultures at lab scale using the mutant(s) of *S. clavuligerus*
   *b)* By evaluating both the wild type and mutant cell performance in terms of biomass growth and levels of CA biosynthesis

3. To explore strategies that might allow constructing a double mutant strain of *S. clavuligerus* just upon successful completion of both the gene disruption and overexpression.

# 3. MATERIALS AND METHODS

## 3.1. Materials

### 3.1.1. Bacterial strain

Table 4 shows the list of bacterial strains used in this study along with a brief description.

**Table 4.** List of bacterial strains used in this study.

| Bacteria | Description | Source / reference |
|---|---|---|
| *Escherichia coli* DH5α [(1)] | *F-φ80lacZΔM15 Δ(lacZYA-argF) U169 recA1 endA1 hsdR17 ($r_k$ -, $m_k$ +) phoA supE44 λ- thi-1 gyrA96 relA1* | General culture collection Bioprocess group at University of Antioquia (Colombia) |
| *Escherichia coli* DSM11539 (Migula 1985) [(2)] | GM48Sm$^r$ (*thr leu thi lacY galK galT ara tonA tsx supE44), DeltalacproF1 traD36 proAB lacIQZDeltaM15*, lambda$^-$. Note: the strain has lost its dam and dcm markers. | DSMZ German collection of microorganisms |
| *Escherichia coli* DSM11539 (pUZ8002) [(3)] | *E. coli* DSM11539 carrying pUZ8002 plasmid. | This study |
| *Streptomyces clavuligerus* ATCC 27064 | Wild type strain, clavulanic acid and cephamycin C producer. | Kindly provided by the INBIOTEC institute at Leon University (Spain). |
| *S. clavuligerus*/pIB139 | pIB139 without gene insert integrated into *S. clavuligerus* ATCC27064 (control strain). | This study |
| *S. clavuligerus*/pICLAR | *claR* overexpressed in the wild type *S. clavuligerus*. | This study |
| *S. clavuligerus*/pIORF21 | *orf21* overexpressed in the wild type *S. clavuligerus*. | This study |

(1) Due to its high transformation efficiency this strain has been used in the transformation and amplification of plasmid DNA.

(2) It has a modification in the methylation system, so that it is not able to methylate the exogenous DNA that is being introduced.

*(3)* For intergenic transfer of pSOK201 from *E. coli* DSM11539 (pUZ8002) into *S. clavuligerus.*

### 3.1.2. Cloning vectors

Table 5 introduces the plasmids used in this study and a brief description of their main characteristics.

**Table 5.** List of plasmids used in this study

| Plasmid | Characteristic | Source / reference |
|---|---|---|
| pIB139 | *int* $^{\Phi C31}$, *att* $^{\Phi C31}$, *oriT*, *ermE**, Aprm$^{R}$ | |
| pSOK201 | ColE1, *oriT*, pSG5 rep, and Aprm$^{R}$, Neo$^{R}$, 7.1kb. | Kindly provided by the INBIOTEC institute at Leon University (Spain). |
| pUZ8002 | RK2 derivative with a defective *oriT*. Contains *tra* genes for non-sexual transfer of genetic material. | |
| TOPO-TA | Amp$^{R}$, Neo$^{R}$, pUC ori, *lac*Z | Thermo Fisher |
| pICLAR | pIB139 with *S. clavuligerus claR* gene at its *Xba*I-*Nde*I recognition site. | This study |
| pSLAT | pSOK201 with a 1.2 kb fragment of *S. clavuligerus lat* gene at its *Eco*RI-*Bam*HI | This study |
| pIORF21 | pIB139 with *S. clavuligerus orf21* gene at its *Xba*I-*Nde*I recognition site. | Kindly provided by the Bioprocess group at UdeA (Colombia) |

### 3.1.3. Enzymes

#### 3.1.3.1. Restriction enzymes

**Table 6.** Restriction enzymes used in this study and their cutting site.

| Restriction Enzyme | Cutting site | Suppliers |
|---|---|---|
| *Xba*I | T↓CTAGA | PROMEGA |
| *Nde*I | CA↓TATG | PROMEGA |
| *Eco*RI | G↓AATTC | PROMEGA |
| *Bam*HI | G↓GATCC | PROMEGA |

#### 3.1.3.2. Other enzymes

| | |
|---|---|
| Lysozyme | SIGMA |
| T4 DNA ligase | PROMEGA |

### 3.1.4. Composition and culture media

| **Tryptic Soy Broth (TSB)** | **g/L** |
|---|---|
| Tryptic Soy Broth | 30 |

For solid medium, agar was added at 15 g/L

| **Luria Broth (LB)** | **g/L** |
|---|---|
| NaCl | 10 |
| Peptone | 10 |
| Yeast extract | 5 |

For solid medium, agar was added at 15 g/L.

| **Yeast extract-malt extract (YEMEG)** | **g/L** |
|---|---|
| Yeast extract | 3 |

| | |
|---|---|
| Malt extract | 3 |
| Peptone | 5 |
| Glycerol | 10 |
| Sucrose | 150 |

The pH was adjusted to 7,0. After sterilization for 20 min at 121°C, the following sterile components were added to 100 mL of culture medium.

| | |
|---|---|
| $MgCl_2$ (2,5 M) | 0,2 mL |
| Glycine (20%) | 2,5 mL |

**R2YEG protoplast regeneration medium.**

Modification of R2YE[81,89] medium, where glucose is replaced by glycerol.

| **Component** | **g/L** |
|---|---|
| Sucrose | 103 |
| $K_2SO_4$ | 0,25 |
| $MgCl_2.6H_2O$ | 10,12 |
| Glycerol | 10 |
| Casamino acids | 0,1 |
| Yeast extract | 5 |
| Trace element solution* | 0,002 |
| For solid medium, agar was added at 2,2% | |

After sterilization during 30 min at 121°C, the following sterile components were added to 80 mL of medium:

| **Component** | **mL** |
|---|---|
| $KH_2PO_4$ (0,5%) | 1 |
| $CaCl_2$ (0,25 M) | 8 |
| L-proline (20%) | 1,5 |
| TES buffer (5,73% pH 7,2) | 10 |
| NaOH (1M) | 0,5 |

Each of the components above mentioned were autoclaved separately.

| **Trace element solution *** | **mg/L** |
|---|---|
| $ZnCl_2$ | 40 |
| $FeCl_3$. $6H_2O$ | 200 |
| $CuCl_2 \cdot 2H_2O$ | 10 |

| **Mannitol Soy flour (MS) agar** | **g/L** |
|---|---|
| Soy flour | 20 |
| Agar | 20 |
| Mannitol | 20 |

**Sterilize** twice for 15 min at 121°C.

| **GYM** | **g/L** |
|---|---|
| Glucose | 4,0 |
| Yeast extract | 4,0 |
| Malt extract | 10 |
| $CaCO_3$ | 2,0 |
| Agar | 15 |

Adjust pH to 7.3 prior to agar addition.

| **MYM** | **g/L** |
|---|---|
| Maltose | 4 |
| Yeast extract | 4 |
| Malt extract | 10 |
| Agar | 18 |

| **GSPG medium** | **g/L** |
|---|---|
| Glycerol | 15 |
| Sucrose | 20 |
| NaCl | 5 |
| $K_2HPO_4$ | 2 |

| | |
|---|---|
| MOPS* | 21 |

*(3-(N-morpholino) propanesulfonic acid)

The pH was adjusted to 6,8. After 20 minutes of sterilization at 121 ° C, the following sterile components were added to the culture medium. Each of the components listed below was autoclaved separately. Glutamate was sterilized by filtration.

| | **g/L** |
|---|---|
| $CaCl_2$ | 0,4 |
| $FeCl_3$. $6H_2O$ | 0,1 |
| $MgSO_4$. $7H_2O$ (pH 7.0) | 1 |
| $MnCl_2$. $4H_2O$ | 0,1 |
| $ZnCl_2$ | 0,05 |
| Proline | 2,5 |
| Glutamate | 1,5 |

Note: GSPG was used as pre-culture and culture medium

| **Seed medium** | **g/L** |
|---|---|
| Glycerol | 15 |
| Peptone | 10 |
| Malt extract | 1,0 |
| $MgSO_4$ | 0.75 |
| $MnCl_2$ | 0.0001 |
| $FeSO_4$ | 0.001 |
| $ZnSO_4$ | 0,001 |
| MOPS | 21 |
| $K_2HPO_4$ | 0.8 |

The pH was adjusted to 6,8.

| **Soy protein isolate medium (ISP)** | **g/L** |
|---|---|
| Glycerol | 30 |
| Soy protein isolate | 25 |
| $K_2HPO_4$ | 0.8 |
| $MgSO_4.7H_2O$ | 0.75 |
| $MnCl_2 \cdot 4H_2O$ | 0.0001 |
| $FeSO_4.7H_2O$ | 0.001 |
| $ZnSO_4.7H_2O$ | 0.001 |
| MOPS | 21 |

After completely dissolving the components, the pH was adjusted to 6.8. Subsequently, the medium was sterilized for 20 minutes at 121 ° C.

Note: The pre-culture medium has the same composition as ISP, except for glycerol concentration which was adjusted to 15 g/L.

### 3.1.5. Buffers and solutions

**Plasmid and chromosomal DNA isolation**

| **SET buffer** | |
|---|---|
| NaCl | 0,75M |
| EDTA, pH 8,0 | 25 mM |
| Tris HCl, pH 7,5 | 20 mM |

| **TE buffer** | |
|---|---|
| Tris –HCl, pH 7.5 | 10mM |
| EDTA, pH 8.0 | 1mM |

**Agarose gel electrophoresis**

| **Tris-Acetate-EDTA (TAE) buffer 10X** | |
|---|---|
| Tris base | 24.5 gr |

| | |
|---|---|
| Glacial acetic acid | 5.71mL |
| EDTA 0.5 M, pH 8.0 | 10 mL |
| $dH_2O$ added to | 500 mL |

**Loading buffer (6X)**

| | |
|---|---|
| Bromophenol blue | 0.25 % |
| Xylene cyanol FF | 0.25 % |
| Sucrose in water | 40 % |

**SYBR Safe**

| | |
|---|---|
| Syber safe (10,000X) | Thermofisher |

**DNA size markers**

| | |
|---|---|
| Marker 2-log DNA ladder | Norgen |
| O'GeneRuler 1 kb Plus DNA Ladder | Thermofisher |
| 100-3000 pb DNA ladder | Solis Biodyne |

**Competent *E. coli* preparation**

**Solution $CaCl_2$**

| | |
|---|---|
| $CaCl_2$ | 75 mM |
| Glycerol | 15% |
| KCl | 10mM |
| MOPS | 10mM |

Adjust pH to 6.5 with 0.2M NaOH. Complete the solution volume with $dH_2O$ and sterilize.

**Antibiotic stocks**

Table 7 specifies the antibiotics used in this study. Antibiotics were sterilized by filtration and dissolved in MilliQ water.

**Table 7.** Antibiotics and their concentrations

| Antibiotic | Stock concentration | Final concentration for *S. clavuligerus* | Final concentration for *E. coli* |
|---|---|---|---|
| Apramycin (aprm) | 50 mg/ mL $dH_2O$ | 40 μg/ mL | 50 μg/ mL |
| Kanamycin (Neo) | 50 mg/ mL $dH_2O$ | - | 50 μg/ mL |
| Nalidixic acid (NA) | 25 mg/ mL NaOH 1N | - | 50 μg/ mL |
| Ampicillin (Amp) | 100 mg/mL $dH_2O$ | - | 100 μg/ mL |
| Chloramphenicol (Cm) | 25 mg/ mL ethanol | - | 25 μg/ mL |

### 3.1.6. Kits

- E.Z.N.A Plasmid DNA Mini kit (OMEGA): Plasmid purification
- E.Z.N.A Cycle pure kit (OMEGA): PCR products purification
- TOPO-TA Cloning kit (THERMO): cloning PCR products with A-overhang
- 5X HOT FIREPol® Blend Master Mix (SOLIS BIODYNE): PCR amplifications
- TaqMan 2X PCR Master Mix (NORGEN): PCR amplifications

### 3.1.7. Oligonucleotides (primers)

**Table 8.** Primers used during this study with the restriction sites underlined.

| Primers | Sequence (5'→3') | Restriction site | Description |
|---|---|---|---|
| *claR*-1_Fw | ccc**catatg**gtgggacactgcggagacctcatggaa | *Nde*I | PCR amplification |
| *claR*-1_Rv | ccc**tctaga**tcagccggacatccgggcccc | *Xba*I | |
| *lat*_Fw | ccc**ggatcc**cgatttctcggacgtgggaa | *Bam*HI | PCR amplification |
| *lat*_Rv | ccc**gaattc**ctccgtgtacatgagccgca | *Eco*RI | |
| aac (3) IV_Fw | cagcttctcaaccttggggt | - | To confirm insertion of pIB139 in *S. clavuligerus.* |
| aac (3) IV_Rv | atgatctgctctgcctgtgg | - | |

| | | | |
|---|---|---|---|
| *ermE**_Fw | atgctagtcgcggttgatcg | | To confirm by sequencing, *claR* insertion in pIB139 |
| *orf21*-2_Fw | gccggatcgtcatcaac | - | Amplification of *orf21* by RT-qPCR |
| *orf21*-2_Rv | cgacaggatctgactggt | - | |
| *hrdB*_Fw | cgcggcatgctcttcct | - | Amplification of *hrdB* by RT-qPCR |
| *hrdB*_Rv | aggtggcgtacgtggagaac | - | |
| *gcas* _Fw | ggtcaactggagcctgtgta | - | Amplification of *gcas* by RT-qPCR |
| *gcas*_Rv | ccgcgaacttggcatagtc | - | |
| *adpA*_Fw | caggtctttaccagaggaaatc | | Amplification of *adpA* by RT-qPCR |
| *adpA*_Rv | cgaactgctcgtgaagatg | - | |
| *ccaR*_Fw | tcgcggactccatcgacctctt | - | Amplification of *ccaR* by RT-qPCR |
| *ccaR*_Rv | ggcgggccccttccacag | - | |
| *claR*-2_Fw | cgggcggcggttctt | - | Amplification of *claR* by RT-qPCR |
| *claR*-2_Rv | tcgtcgagcaggggttcc | - | |

## 3.2. Methods

### 3.2.1. Media and culture conditions

Luria broth liquid medium (LB) and agar plates (LA) were used to grow *Escherichia coli* cells at 37°C and 250 rpm for suspension cultures during 16 h. *E. coli* cells were preserved on agar plates, whereas for long term storage 20% glycerol stocks were prepared and kept at -80°C. For spore production, *S. clavuligerus* was grow in GYM-Agar or MYM[90] (with apramycin when necessary) at 28ºC for 10 to 15 days. The spores were scraped off the agar surface using a sterile loop and resuspended in water. After filtering through cotton wool, the spore suspension was serially diluted to $10^9$ CFU $mL^{-1}$ and stored in 20% glycerol at 80°C. Spores of *Streptomyces clavuligerus* wild type and related mutant strain(s) were grown for 36 h ($OD_{600}$ of 7) in TSB at 28°C and 220 rpm for suspension cultures in a rotary shaker. Mycelium cultures were stored as 20% glycerol stock at -80ºC.

The seed medium was inoculated with 1mL of *S. clavuligerus* spores (for 24 h, at 28 ° C and 220 rpm). Afterwards, 5 mL of seed medium were transferred to 45 mL of pre-culture medium, and it was incubated under the same operating conditions. For CA production, either ISP or GSPG culture medium was inoculated with pre-culture medium at 10% v/v. All *S. clavuligerus* cultures were performed in 250-baffled Erlenmeyer flasks containing 50 mL of medium. Cultures for CA production were incubated for 144 h, at 220 rpm and 28 ºC. All experiments were performed in triplicate.

### 3.2.2. Preparation of competent *E. coli* cells

Competent *E. coli* cells were obtained using calcium chloride [91,92]. Single *E. coli* DH5α and *E. coli* DSM11539 colony from fresh LA culture were inoculated into 3 mL of LB and incubated by shaking at 37°C overnight. 500 µL of this culture were transferred to 100 mL of LB in 200 mL Erlenmeyer flask and incubated in an orbital shaker at 37°C and 250 rpm until reaching an $OD_{590}$ of 0,375. The culture was then 50 mL aliquoted in Falcon tubes and incubated on ice for 10 min. Then, the cells were centrifuged at 3500 rpm, and 4°C for 7 min. After decanting the supernatant, the pellets were dissolved in 10 mL of ice cold $CaCl_2$ solution. The resuspended cells were centrifuged (at 2500 rpm and 4 °C for 5 min) and the supernatant was discarded. The cells were resuspended in 10 mL of ice cold $CaCl_2$ solution, and incubated on ice for 30 min. Then, the cells were again centrifuged as above. The supernatant was decanted, and the cells were resuspended in 2 mL of ice cold $CaCl_2$ solution. The cell suspension was 100 µL aliquoted in 1.5 mL Eppendorf tubes. The cells were kept on ice throughout the procedure. Finally, the tubes were stored at -80 °C for 24h.

### 3.2.3. Transformation of competent *E. coli* cells

Competent *E. coli* cells were transformed using the "heat-shock" method[91,92]. -80°C competent cell stocks were thawed on ice and 1-50 ng plasmid DNA or ligation product was added to 50 µL of the cell suspension. After 20 min incubation on ice, tubes were placed in a 42°C water-bath for 45 secs to apply a heat-shock. Immediately after the heat-shock, tubes were placed on ice for 10 min. 800 µL sterile LB was added to each tube and incubated at 37°C for 90 min and 100 rpm. Subsequently, cells were centrifuged at 3500 rpm for 15 min and most of the supernatant was discarded, leaving only 200 µL of LB to

re-suspend the pellet. Finally, cells were spread onto selective LA plates and left for overnight incubation at 37°C. The next day, single colonies were picked and screened for possible recombinants.

### *3.2.4.* Plasmid isolation from *E. coli*

The E.Z.N.A Plasmid DNA Mini kit (Omega Bio-tek) was used for isolation of plasmids from *E. coli* according to the manufacturer's instructions. Manual isolation of plasmids from *E. coli* was performed according to the plasmid extraction procedure described by Birnboim *et al*., 1979 [93] with some modifications. According to this, 3 mL of LB medium with antibiotic was inoculated with a transformant colony and incubated at 37ºC (250 rpm) for 16 h. 3 mL of the culture medium was precipitated by centrifugation at 13000 rpm for 1 min and the supernatant was decanted and the cells were re-suspended in 100µL of TE buffer and treated with RNasa A (5 mg/mL). Immediately, 450 µL of lysis solution was added and mixed by inversion. Next, 225 µL of potassium acetate 5M (pH 5.2) was added and tubes were kept in ice for 10 min. Then, cells were centrifuged at 13000 rpm and the supernatant was transferred to a new tube of 1.5 mL. Next, 700 µL of 100% ice-cold ethanol were added and the tubes were centrifuged again at 13000 rpm, 4ºC for 5 min. The supernatant was discarded, and the precipitate washed with 1 mL of 70% ethanol followed by a further centrifugation at 14000 rpm for 2 min. The supernatant was discarded, and the excess ethanol was left for air dry. The precipitate was re-suspended in 50 µL TE buffer. The DNA was stored at -20 °C until use.

### 3.2.5. *S. clavuligerus* genomic DNA isolation

*S. clavuligerus* genomic DNA was isolated by the salting out method described by Kieser (2000)[94], with some modifications. 50 mL of TSB were inoculated with 500 µL of mycelium stock and incubated at 28°C (220 rpm) for 48-60h. 30 mL of this suspension were precipitated by centrifugation at 4500 rpm for 15 min and the supernatant was decanted. Cells were re-suspended in 5 mL of SET buffer and treated with 100 µL of lysozyme (50 mg/mL) for 30 min at 37°C. Next, 2 mL of 5M NaCl were added and thoroughly mixed by inversions. Next, 5 mL of phenol/chloroform (25:24 V/V) solution was added and mixed by inversions. The solution was centrifuged at 4500 rpm for 15 min and the supernatant was transferred to a new tube (a new extraction with only chloroform

was made as above). Following the transfer, 3 mL of isopropanol were added and mixed again by inverting the tube several times. The solution was put on ice for 10 min and centrifuged for 15 min at 4500 rpm. The DNA pellet was washed with 5 mL of 70% ethanol; the ethanol was removed, and the DNA pellet left for air dry. Finally, the DNA material was dissolved in 1 mL of TE at 55ºC.

### 3.2.6. Transformation of *S. clavuligerus* by protoplast

Plasmids were introduced into *S. clavuligerus* by a modified polyethylene glycol (PEG)-mediated protoplast transformation procedure [95,96]. Spores of *S. clavuligerus* were inoculated in 250 mL baffled flasks with 50 mL of YEMEG medium at 30°C, 220 rpm for 60h ($OD_{600}$ from 7 to 9). The culture was aliquoted into 50 ml Falcon tubes using 10 mL of culture medium; cells were centrifuged and washed twice with 10.3% sucrose solution. Next, the pellet was suspended in 2 mL of P buffer, containing 50 µL of lysozyme (20 mg/mL) and incubated at 30°C for 20 min. After dilution of the lysis solution with 8 mL of P buffer, the protoplast suspension was centrifuged twice and suspended in 0,8 mL of P buffer. The protoplast suspension was incubated at 42°C for 10 min. 1 µg of plasmid DNA was added to 100 µL of protoplast and mixed with 0.2 mL of PEG 1000 (40%). Subsequently, after about 1 min at room temperature, a dilution was made by adding 1mL of P buffer, and the protoplasts were collected by centrifugation at 4000 rpm and suspended in 1 mL of P medium. The protoplasts were plated (0.1 mL per plate) in R2YEG medium (previously dehydrated by drying them in a laminar flow cabinet for 2 h) and incubated at 28°C. After 36 h, 2 mL of sterile $dH_2O$ containing 40 µg/mL of apramycin was spread onto the plates and incubation continued for 4 further days. Subsequently, protoplasts of *S. clavuligerus* were grown in selective TSA medium for 3 days.

### 3.2.7. Intergeneric conjugation between *E. coli* DSM11539 (pUZ8002) and *S. clavuligerus*

The method used by Flett *et al*. (1997) for intergenic conjugation using non-methylating *E. coli* as donor was slightly modified [97]. pUZ8002 of competent *E. coli* DSM11539/pUZ8002 cells was maintained by chloramphenicol (25 µg/mL). Apramycin (50 µg/mL) was also added to maintain the *ori*T-containing vectors. Recombinant *E. coli*

DSM11539/pUZ8002 with the desired plasmid was inoculated into 10 mL of LB media containing chloramphenicol and apramycin and grown overnight at 37°C at 200 rpm. The next day, cells were diluted 1:100 in fresh LB containing antibiotics and grown at 37°C to an $OD_{600}$ of 0.4 – 0.6. After incubation, cells were washed twice with an equal volume of fresh LB (10 mL) and re-suspended in 1 mL of LB. *Streptomyces* mycelia were harvested from 3-4 days old culture growing on TSA using 3-4 mL of 20% glycerol, vortexed thoroughly, and 0.5 mL of it was mixed with 0.5 mL of *E. coli* culture. The cell mixture was mixed and spun down briefly. Supernatant was mostly poured off and the pellet was re-suspended in residual fluid. 200 μL of cell suspension were plated out on MS agar containing 10 mM MgCl2 and incubated at 30°C for 16-20 hours. After incubation, the plates were overlaid with 1 mL of distilled water containing 0.5 mg of nalidixic acid and 1mg of apramycin. Nalidixic acid kills *E. coli* but has no effect on *Streptomyces* viability due to its natural resistance [8]. Following, incubation at 30°C for 3-4 days was carried out until potential exconjugants colonies were observed

### 3.2.8. Agarose gel electrophoresis

The most frequently used concentration of agarose was 1 %; however, depending on the size of fragments to be separated, the concentration was modified accordingly. 1X TAE was used as electrophoresis buffer and Syber safe 1X was used for DNA staining. The bands were visualized on Safe Smartblue transilluminator (Accuris, USA). The molecular weights of DNA bands were determined using 2-log DNA ladder.

### 3.2.9. Purification and elution of DNA fragments

The DNA fragments eluted from agarose gel were purified with the E.Z.N.A cycle pure kit (OMEGA) according to the manufacturer's instructions. The extraction yield was measured with a NanoDrop® ND-2000 (Thermo Scientific) spectrophotometer.

### 3.2.10. DNA manipulation for *claR* and *lat* gene

#### 3.2.10.1. Polymerase chain reaction

*S. clavuligerus* ATCC 27064 DNA was used as template to amplify the *claR* gene and a fragment of the *lat* gene by PCR. Primers designed for PCR amplification were synthesized by Macrogen, Korea (Table 8). For one reaction of 20 µL, PCR reaction was prepared as follow: 5 X PCR Master Mix, 4 µL; DMSO (5%), 1 µL; each forward and reverse primer (10µM), 1 µL; template DNA (50 ng), 2 µL; and 11µL of $dH_2O$.

PCR started with an initial denaturation step (12 min at 95ºC), followed by 30 cycles of amplification (30s at 95 °C for denaturation, 30s at 55°C for annealing, 2 min at 72 °C for extension), and ended with a final extension step of 72°C for 5 or 20 min (20 min for poly(A) tail).

#### 3.2.10.2. Colony PCR

Colony PCR was performed as described by Azevedo et al., 2017 [98] with slight modifications to adapt the procedure to *S. clavuligerus*. Samples taken from the colonies on agar plate were re-suspended in 50 µL $dH_2O$ and lysed at 95 °C for 30 min. Then, 5 µL cell lysate were used as template in the standard PCR mixture.

### 3.2.11. Digestion with restriction enzymes

The DNA double digestion was performed according to the instructions of the enzyme manufacturer. For *Eco*RI and *Bam*HI digestion, both enzymes were initially added, and the reaction was left for 3 h. Some modifications to the procedure were made to facilitate cutting of the enzymes, *Xba*I and *Nde*I. For this, to each reaction the enzyme *Nde*I was initially added, and the reaction was left for 1:30 h. Subsequently, the *Xba*I enzyme was added, and the reaction was left for an additional 1:30 h.

The digestion reactions were inactivated by temperature. However, to dissociate the *Xba*I enzyme from the substrate, SDS was added to a final concentration of approximately 0.5% to the gel loading dye.

The amount of enzyme was adjusted to the amount of DNA to be digested considering that the enzyme volume should not be greater than 10% of the total reaction volume.

### 3.2.12. Ligations of DNA fragments

Fresh gel extracted PCR products were ligated with TOPO-TA vector via TA cloning using the TOPO-TA Cloning kit (Thermo). The ligation mixture composition was: 1µL of TOPO vector, 4 µL Fresh PCR product and 1 µL of salt solution. Ligation was performed for 20 min at room temperature. Then, the reaction was placed on ice while proceeding with *E. coli* transformation.

Following excision from the TOPO-TA, the *claR* gene and a 1.2 kb *lat* gene fragment were ligated with target vector pIB139 and pSOK201, respectively, as follows: 1 µL of 1-3 u/µL T4 DNA ligase (Promega), 4 µL of 2X reaction buffer, 50 ng vector DNA, 150 ng insert DNA mixed, and the volume of the mixture was completed to 20 µL with $dH_2O$. Overnight incubation at 4 °C was provided for ligation reactions to take place.

### 3.2.13. Construction of a plasmid carrying the *claR* gene.

To overexpress the *claR* gene in *S. clavuligerus*, the PCR product was first cloned into TOPO-TA and then into pIB139 plasmid using the *Nde*I and *Xba*I restriction enzymes. The insertion of *claR* into the plasmid was confirmed by 1% (w/v) agarose electrophoresis gel and sequencing.

### 3.2.14. Construction of a deletion plasmid for *lat* disruption

To disrupt the *lat* gene by single- crossover recombination, the 1.2 kb PCR product was cloned into TOPO-TA and then into pSOK201 plasmid using the *Eco*RI and *Bam*HI restriction enzymes. The insertion of the *lat* gene fragment into the plasmid was confirmed by 1% (w/v) agarose electrophoresis gel and sequencing.

### 3.2.15. Clavulanic acid fermentation and analytical techniques

Cell cultures were performed by growing three biological replicates of *S. clavuligerus* wild-type and its recombinants, including their control. For sampling, aliquots of 2 mL were taken from each flask every 24 h during the fermentation process. For CA quantification, cell culture samples were centrifuged at 14,000× *g* for 10 min at 4°C and

filtered (0.22 µm). CA was determined by HPLC Agilent 1200 (Agilent Technologies, Waldbrom, Germany) equipped with a Diode Array Detector (Agilent Technologies, Palo Alto, CA, USA) at 312 nm, using a reverse phase ZORBAX Eclipse XDB-$C_{18}$ (4.6 × 150 mm, 18µm Agilent Technologies, Palo Alto, CA, USA) column; 94% v/v $KH_2PO_4$ (50 mM, pH 3,2) and a 6% v/v methanol solution was used as mobile phase at 0.7 mL/min. CA was imidazole-derivatized at a ratio of 1:3; the reaction was kept at 28°C for 15 min [99].

Glycerol concentration was determined by HPLC-RID[100]. Biomass growth was determined by using the dry weight techniques; for this, cells were washed twice with deionized water and dried at 80°C for 24 h before measuring (gDWC $l^{-1}$)[41].

### 3.2.16. RNA extractions and RT-qPCR analysis

RNA was extracted from mutant and wild-type *S. clavuligerus* cell cultures at 72 h - 96 h of cultivation, when CA concentration was the highest under the tested conditions. Samples were centrifuged at 10,000× *g* for 15 min; cell pellets were immediately stored at −80 °C for subsequent RNA isolation protocol. Total RNA was isolated using Trizol® (Sigma®), following the manufacturer's instructions [101]. All RNA preparations were treated with RNase-free DNase I (Promega®) to eliminate genomic DNA contamination; RNA purity and concentration was determined using a QIAEXPERT® equipment (QIAGEN®). Gene expression analysis and quantification of some CA synthesis related genes, were measured from two samples coming out from two different culture media (GSPG and ISP), using RT-qPCR as described by R. Álvarez-Álvarez et al. [102] according to the $2^{-\Delta\Delta Ct}$ method [103] and the constitutive housekeeping *hrdB* gene as a control for replication threshold [104]. Data for relative gene expression were normalized using the log 2 $[2^{(-\Delta\Delta Ct)}]$ method. The cDNAs were synthesized as described by R. Álvarez-Álvarez et al [39]. Negative controls were used to determine DNA contamination. As for the transcriptional analysis using RT-qPCR, the following genes were tested: *ccaR*, *adpA*, *claR*, *gcas*, *orf21* and SCLAV_4359. All primers, used in the experiment, are presented in Table 7. Equal RNA quantity (1µg) was used for all RT-qPCR experiments.

### 3.2.17. Phylogenetic analysis and prediction of DNA motifs bound.

Phylogenetic inferences were performed using ORF21 homologous proteins and different extracytoplasmic function (ECF) sigma factors. ORF21 homologous were mined using the NCBI protein BLAST algorithm (www.ncbi.nlm.nih.gov/BLAST/) with a PAM 70 as substitution matrix. RefSeq protein sequences were chosen according to its identity percentage (≥ 50%) and adequate *E*-value confidence criterion (≤ 1e_37). Following, a Bayesian tree was constructed using MrBayes (MB) V3.2 [105] setting the Jones Gamma distributed amino acid substitution model, with 1.000.000 generations sample every 1,000 generations and the other analysis parameters as default value. The convergence of the Markov Monte Carlo interactions was assessed with the potential scale reduction factor (PSFR=1) [106] and the standard deviation of split frequencies (0.008).
The ECF sigma factor sequences from different Streptomyces species were selected from GenBank. The phylogenetic tree for ECF sigma factors were inferred using the Maximum Likelihood method [107] and JTT matrix based-model [108]. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analysis. The rates among sites were treated as a Gamma distribution using 4 Gamma Categories (Gamma Distribution option). This analysis involved 30 amino acid sequences. Evolutionary analyses were conducted in MEGA11 [109].

Six putative regulatory sequences were compiled from the 250 nucleotide-long sequences located upstream of the *claR* gene of *S. clavuligerus, S. jumonjinensis* and *S. katsurahamanus,* and *adpA, ccaR, gcas* genes of *S. clavuligerus*. The sequences were analyzed in MEME SUITE 5.1.1 [110], while motifs found were analyzed using TOMTOM[111] for further comparisons. The highest-scoring motif was selected from all generated motifs.

### 3.2.18. Statistical analysis

The experimental results were statistically analyzed using the Statgraphics® software (version 18). For biomass, glycerol and CA results, data were evaluated by the Shapiro-Wilk normality test. Parametric data were subjected to analysis of variance (ANOVA) and Tukey's HSD test. Significance level was set as $p < 0.05$.

# 4. RESULTS

## 4.1. Disruption of the *lat* gene

The simple homologous recombination mechanism was the proposed strategy for *lat* gene disruption (see Figure 7). For this, primers were designed for amplifying a *lat* gene fragment of 1.2 kb from *S. clavuligerus* ATCC27064 genomic DNA (see Table 8), according to the sequence obtained from GenBank accession number AY742798.1. Figure 8 shows a band of approximately 1.2 kb obtained by PCR amplification, which is consistent with the expected size. The 1.2 kb PCR product was subcloned into the TOPO-TA vector and transferred into *E. coli* DH5α cells via transformation. After selection in the LB medium containing kanamycin (50 μg/mL), colony PCR using the lat_Fw and lat_Rv primers was performed for some putative transformed colonies; several colonies amplified for the 1.2 kb fragment. Then, the recombinant plasmid was extracted, and the gene insertion was verified by sequencing. The recombinant plasmid was designated as TOPO-LAT. The resulting plasmid TOPO-LAT was linearized at the *Bam*HI and *Eco*RI, sites located upstream and downstream of the *lat* gene, respectively. The 1.2 kb *lat* gene was then subcloned into pSOK201 as an *Bam*HI −*Eco*RI fragment to create the plasmid pSLAT with the *lat* gene truncated at both ends. The insertion of 1.2 kb *lat* gene into pSOK201 was verified with restriction enzyme (RE) digestion (Figure 9).

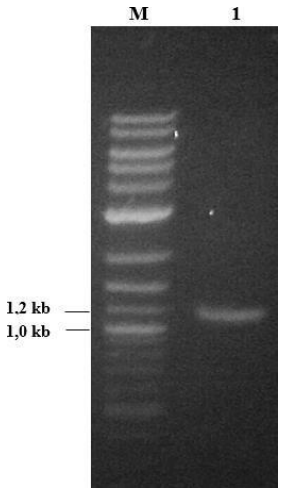


**Figure 8.** PCR amplification of 1.2 kb *lat* gene fragment. Agarose gel Electrophoresis (1%) stained with SYBR Safe. M: high Ranger 1Kb DNA ladder (NORGEN), 1: 1.2 kb *lat* gene fragment obtained from *S. clavuligerus* genomic DNA template.

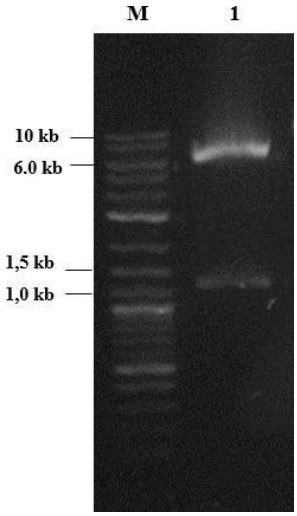


**Figure 9.** Recombinant plasmid verification by *Bam*HI-*Eco*RI. Agarose gel Electrophoresis (1%). M: 2-log DNA ladder, Lane 1: Linearized pSOK201 vector and the released 1.2 kb *lat* gene fragment.

### 4.1.1. *Streptomyces clavuligerus* ATCC 27064 transformation with pSLAT

*E. coli* DSM11539 was transformed separately with the recombinant plasmid pSLAT, to bypass the potent methyl-specific restriction barrier of *S. clavuligerus* cells. After transformation, the plasmid was extracted, and its concentration determined by Nanodrop 2000. Next, *S. clavuligerus* wild-type strain was transformed by protoplasts with pSLAT and pSOK201, separately. After 3 days of incubation with apramycin, regenerated protoplasts, picked from R2YEG agar plates, were grown on TSA plates containing apramycin (40 μg/mL). However, in TSA the regenerated protoplasts were not able to replicate to give separate colonies of *S. clavuligerus.*

Conjugational transfer with *E. coli* is a more common technique that can be applied to nearly all *Streptomyces* spp. Therefore, pSLAT and pSOK201 recombinant plasmids were introduced, separately, into *E. coli* DSM11539 (pUZ8002) cells prior to conjugation with wild-type *S. clavuligerus*. Nonetheless, after conjugation, it was not possible to detect *S. clavuligerus* exconjugants colonies.

## 4.2. Cloning of the *claR* gene

Specific primers were designed for amplifying *claR* from *S. clavuligerus* ATCC27064 genomic DNA, according to the sequence obtained from GenBank accession number CP027858.1. Figure 10 shows the 1320 bp nucleotide sequence fragment amplified by

PCR using claR_Fw and claR_Rv primers pair (Table 8). Figure 11 shows a band of approximately 1.5 kb obtained by PCR amplification, which is consistent with the expected size. Gel eluted PCR product, exhibiting the predicted size, was then ligated to TOPO-TA cloning vector, and transferred into *E. coli* DH5α cells via transformation. After selection in LB medium containing kanamycin (50 μg/mL), colony PCR using the claR_Fw and claR_Rv was performed for 2 putative transformed clones; only one of the colonies amplified for the *claR* gene. Afterward, the recombinant plasmid was extracted, and the gene insertion was verified with restriction enzyme (RE) digestion (Figure 12) and sequencing. The recombinant plasmid was designated as TOPO-CLAR.

ACCGACGTCTGCTGGTGTCAGCCGATGCGATCTGTCTTTACTGGACGCGGTGGGACACTGCGGAGACCTC
ATGGAAGTGGCCCGGCGGACGGGTGTCCGGCACGGCACCGTGGAGCGCAGACTCGACCGGCTGGACCGGA
TCGTCGGCCTCCCGCTGACCCTGCGCAGCCGCCACACGGCCCGGCTGACCACCGCGGGCTCCCGCATCCT
CGTCGCCGGGCGGCGGTTCTTCCACCAGGTCGACCTGGCCGCGCGGACGCATATCTTCGGCCATGGCTCC
GAGGCCGTGGACGCCCCGGAGGTGCTGTCGCTGGTCTCCACGGAACCCCTGCTCGACGAGGTGGTGGAGG
ACGCGGCGGCCTCGCTGGACCTGCTGCTGTCGGTCCGGCACGAGGCCCCGCACCAGGTCGCCGCCCAACT
GGCGGGCTACCAGGTGGACGCGGCCTACACCTGGAGCCTCCAGTCCCCCCGGCACAGCCTGGAGCGGTCC
GTGCGCACCTGTGAGGTGCTGGACGACCCGCTGTGGGTGATCCTGCCCCGGGACCATCCGCTGGCCGCCC
GGCGGGAGGTCTCGCTCGCCGATCTGCGGGACGAGACCTGGGTGTCCGAGACGGGACCCGGCTCGGAGAT
CCTGGTGACACGTGTCTTCCAGCTGGCCGGGCTCACCGCGCCCACCCGGCTCCACATCACGGGGGCATCG
GTTGCCCGGGGCATCCTGCGCCGCGGGGACGCGATAGGTCTCGGCTCGCCCACCCACCCGGCGGTGCAGG
ACCCCTCGCTGGTGCGCCGCTCCCTGGCGGAGCGCCCGCGCCGCACCACGAGTCTGCTCGTCGACCCCAC
CATCGTGCCCCGGGCGCTGGCGGGACGGCTGGCCGCGCTGATCGCCGAGGTCCAGCTCCGGCGCTTCGCC
GAACACCACCGCGACCTGCTGGACGAGCCCTGGTGGGCGCAGTGGTACGCGGAGCGCACCGGCGCGGACG
CCCGCCGCTTCGGGGCGGGACCCGACCAGGGCTCCGTGCCCGGCCAGGCCGAGGGCCGCAAACTGGATGT
GGACGATCTCCATCTGCTCCAGGCCGTGGCCCGGCACGGCAGCATCAACCGGGCCGCGGCGGTGCTGTCG
ATCAGCCAGTCGGCGCTCACCCGCCGGATTCACCGGCTGGAGCAGTCCCTCGGCGCCCGGCTGCTGCTGC
GCAGCCCGCGGGGGACCAGCCTGACCGGCCCGACCCGGCAGTTCCTGCGCCAGCTCGCGCTGTACGAGGC
GGAGTTCCGCGAGGCCGCTCTCGCCTGCCGCAGCGTGGAACGGCCCCTGGCGCAGGGCCACTGGCCGATC
CGGCGCGGGGTCGCGGCCGGGGCCCGGATGTCCGGCTGAGCGGGCCGCGACCGGGTCCGGACCGGACCGG

**Figure 10.** Nucleotide sequence of *claR* (GenBank accession number U87786.2). In blue, the gene coding region (1299 bp); in underlined, the location of claR_Fw and claR_Rv primers.

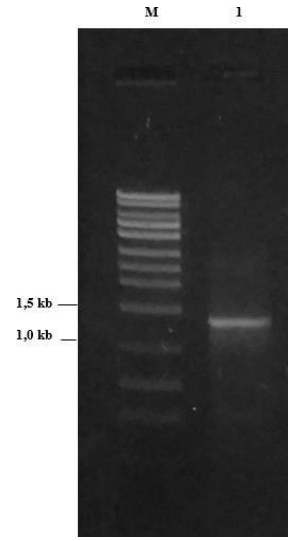


**Figure 11.** PCR amplification of *claR.* Agarose gel Electrophoresis (1%) stained with SYBR Safe. **M**: high Ranger 1Kb DNA ladder (NORGEN), **1**: *claR* gene obtained from *S. clavuligerus* genomic DNA template.

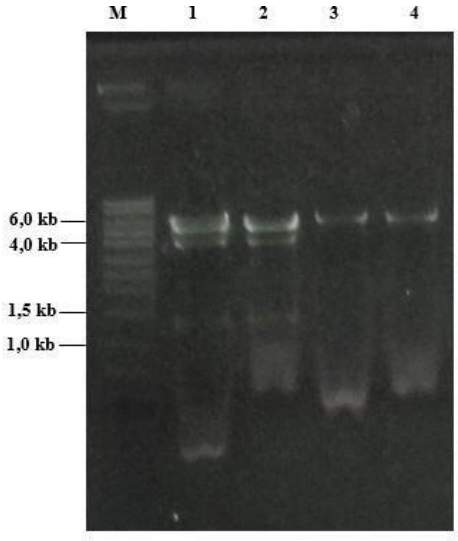


**Figure 12.** *Nde*I-*Xba*I digestion of TOPO-CLAR and pIB139. Agarose gel Electrophoresis (1%) stained with SYBR Safe. **M**: high Ranger 1Kb DNA ladder (NORGEN), **1-2**: Linearized TOPO-CLAR vector and the released *claR* gene (approximately 1.5 kb), **3-4**: Linearized pIB139 vector (5924bp).

Double digestion with *Nde*I and *Xba*I, released the *claR* gene from TOPO-CLAR and linearized the pIB139 vector (Figure 12). Both fragments were gel purified and ligated. Afterwards, the ligation was transformed into *E. coli* DH5α competent cells. Twenty

possible recombinant cells were selected from LA plates containing 50 μg/mL of apramycin. Verification of *claR* gene cloning on pIB139 was performed by colony PCR using claR_Fw and claR_Rv primers. Only one of the colonies analyzed amplified for the *claR* gene (Figure 13). The recombinant plasmid was designated as pICLAR and its integrity was verified by sequencing using the primers ermE*_Fw and claR_Rv (Table 8).

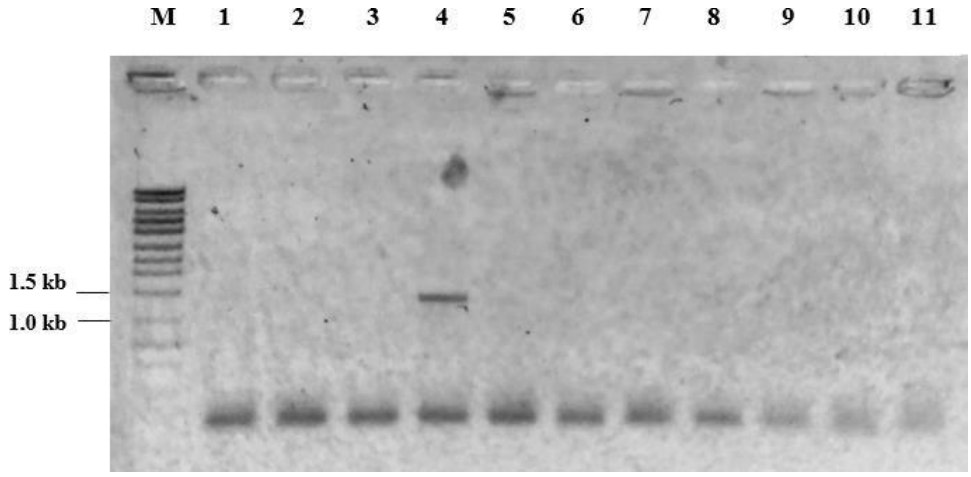


**Figure 13.** Verification of recombinant pICLAR via PCR colony**.** Electrophoresis on 1% agarose gel stained with SYBR Safe. **M:** high Ranger 1Kb DNA ladder (NORGEN), **1**: Negative control (No template), **2-3**: No PCR amplicon was obtained for *claR* gene. **4**: PCR amplicon obtained for *claR* gene (~ 1.5 kb), **5-11**: No PCR amplicon was obtained for *claR* gene.

### 4.3. *Streptomyces clavuligerus* ATCC 27064 protoplast transformation with pICLAR and pIORF21

*E. coli* DSM11539 was transformed separately with the recombinant plasmid pICLAR and pIORF21. After transformation, the plasmid was extracted, and its concentration determined by Nanodrop 2000. *S. clavuligerus* wild-type strain, was transformed by protoplasts with pICLAR and pIORF21, separately. After 3 days of incubation with apramycin, regenerated protoplasts (Figure 14A and Figure 15A), picked from R2YEG agar plates, were grown on TSA plates containing apramycin (40 μg/mL), to obtain isolated clones as observed in Figure 11B and Figure 12B.

**A** **B**

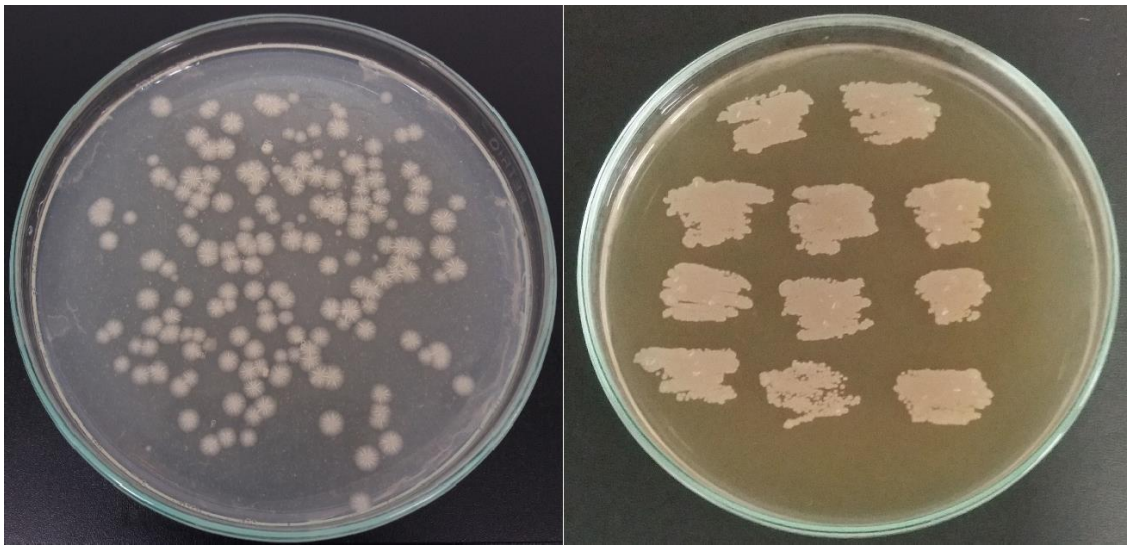

**Figure 14.** Result of *S. clavuligerus* transformed with pICLAR. **A)** Regenerated protoplasts of *S. clavuligerus* after 4 days of growth in R2YEG medium with apramycin (40 μg/mL). **B)** Transformed clones of *S. clavuligerus* with pICLAR in TSA medium with apramycin (40 μg / mL).

**A** **B**

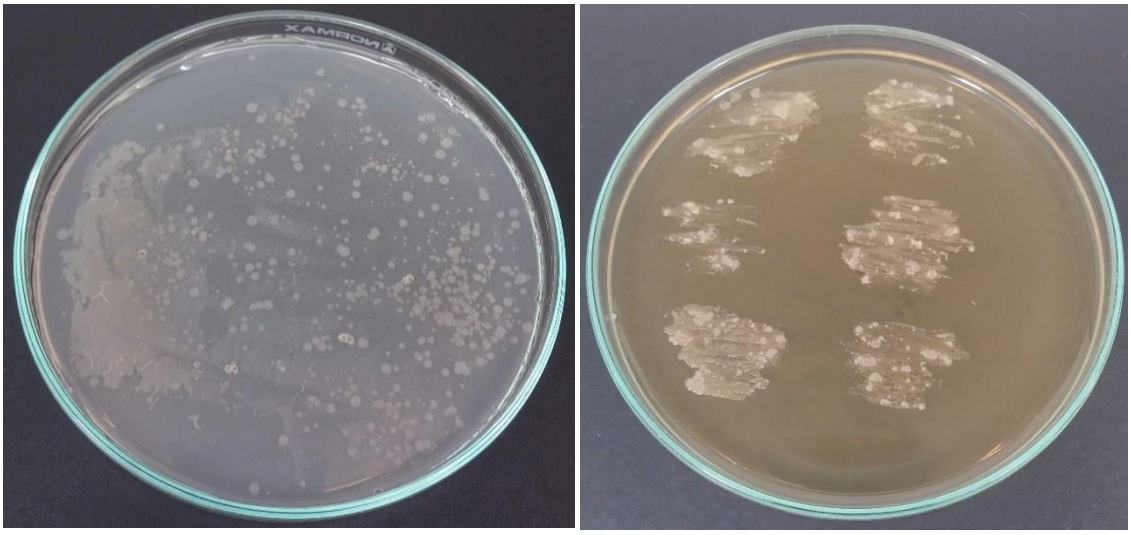

**Figure 15.** Result of *S. clavuligerus* transformed with pIORF21. **A)** Regenerated protoplasts of *S. clavuligerus* after 3 days of growth in R2YEG medium with apramycin (40 μg/mL), **B)** Transformed clones of *S. clavuligerus* with pIORF21 in TSA medium with apramycin (40 μg / mL).

Regenerated protoplasts of *S. clavuligerus* were screened for the presence of apramycin resistance gene via colony PCR by using the primer pairs aac-IV_Fw and aac-IV_Rv (Figure 16A and 16B). Only the colonies that contain the vector were able to grow in

TSA with apramycin. The recombinant strains with pICLAR and pIORF21 were named *S. clavuligerus*/pICLAR and *S. clavuligerus*/pIORF21, respectively. The recombinant strain named *S. clavuligerus*/pIB139 (*S. clavuligerus* / pIB139 having a plasmid without any inserted gene) was generated to determine the potential effect of pIB139 plasmid on CA yield in the recombinant strains. *S. clavuligerus /* pIB139 will be used as control strain to evaluate the effect of the cloned gene (*claR* or *orf21*) on CA production without consider effects caused by pIB139 presence on CA.

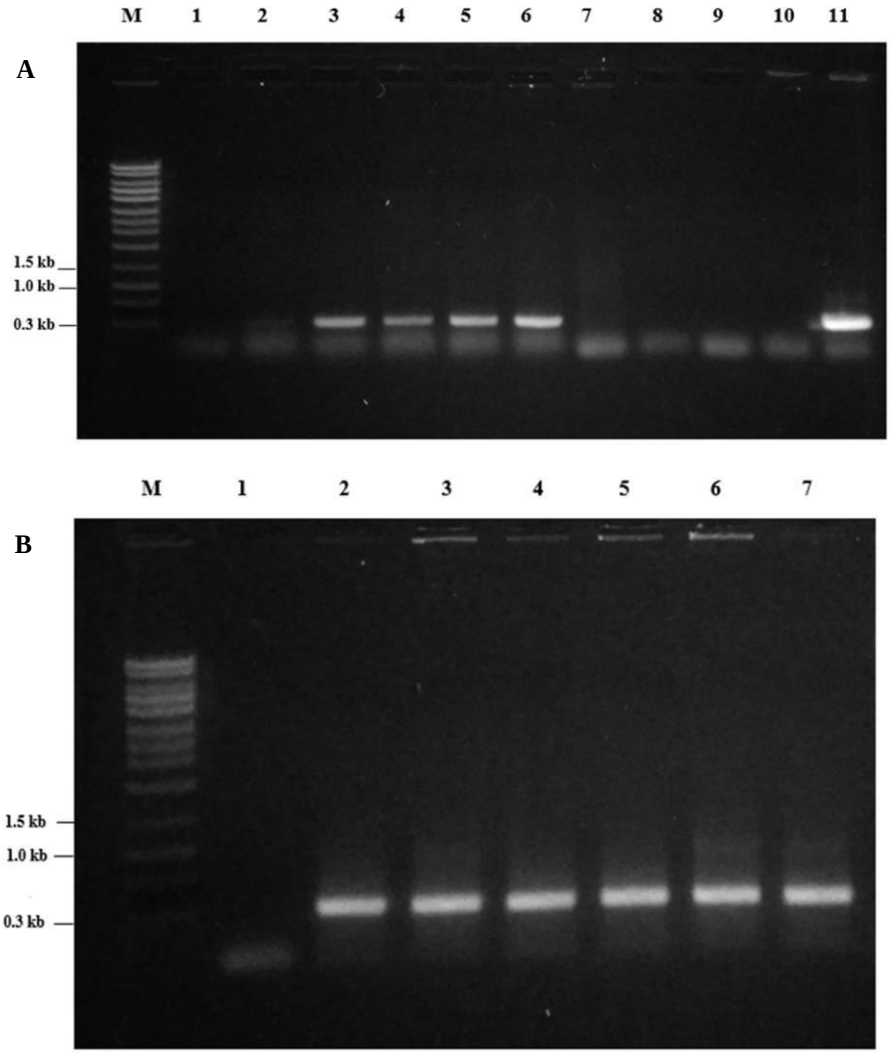


**Figure 16.** Verification of apramycin resistance gene (aac (3)-IV) into *S. clavuligerus* by colony PCR. Agarose gel Electrophoresis (1%) stained with SYBR Safe. A) *S. clavuligerus* transformed with pICLAR. M: high Ranger 1Kb DNA ladder (NORGEN), 1: Negative control with the wild type *S. clavuligerus* genomic DNA used as template, 2-6: PCR amplicon (310 bp) obtained by using regenerated protoplasts of *S. clavuligerus* as template, 11: Positive control with pICLAR used as template DNA, yielding 310 bp

amplicon. B) *S. clavuligerus* transformed with pIORF21. M: high Ranger 1Kb DNA ladder (NORGEN), 1: Negative control with the wild type *S. clavuligerus* genomic DNA used as template, 2-7: PCR amplicon (310 bp) obtained by using regenerated protoplasts of *S. clavuligerus* as template.

### 4.4. Ammonium, biomass growth and morphology of the phenotype

#### *S. clavuligerus/*pICLAR

There was no significant difference ($p > 0.05$) in the maximum biomass production reached from both strains (Figure 17). Nevertheless, the growth of *S. clavuligerus/*pICLAR was slightly faster than that of *S. clavuligerus* /pIB139, in both culture media (Figure 18). For instance, at 72 hours of fermentation in ISP, both strains produced almost the same amount of biomass. However, from this time on the growth of *S. clavuligerus/*pICLAR decreases significantly ($p > 0.05$) up to 144 hours (Figure 17). Moreover, no morphological differences were observed between *S. clavuligerus/*pIB139 and *S. clavuligerus/*pICLAR, in terms of aerial mycelium formation (Figure 18).

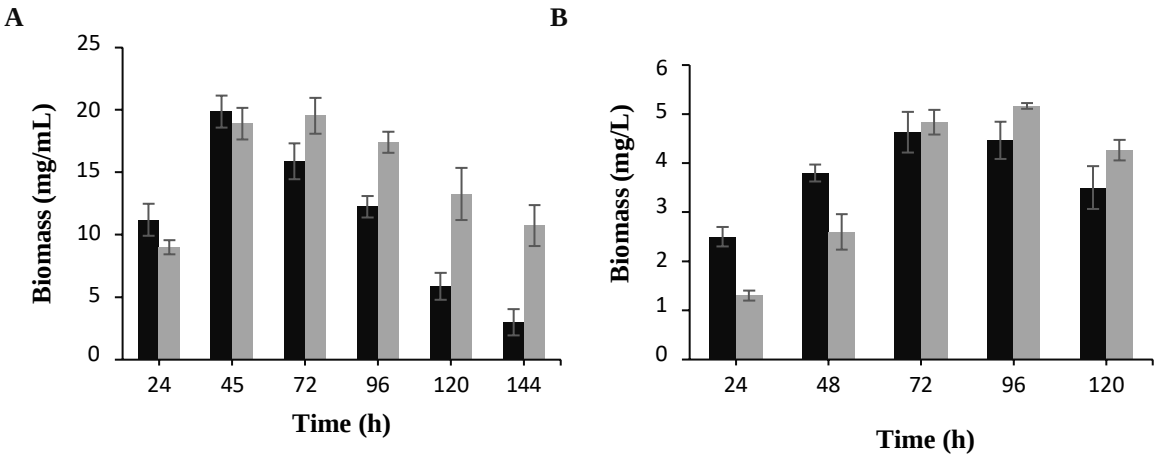


**Figure 17.** Biomass production in GSPG and ISP. In black, *S. clavuligerus*/ pICLAR. In gray, *S. clavuligerus*/pIB139. A) ISP. B) GSPG.

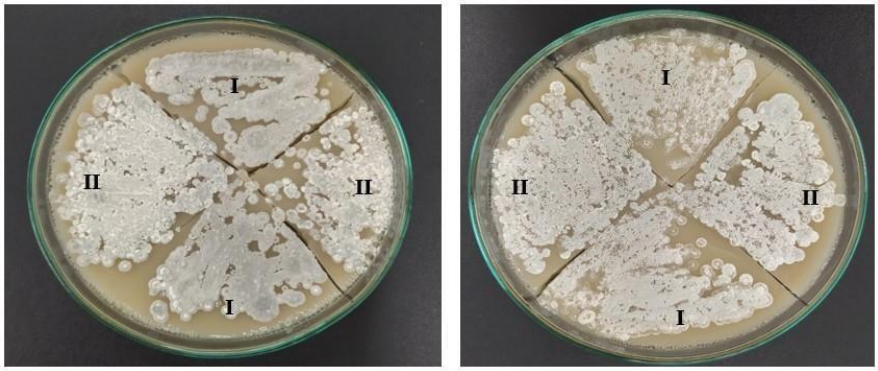


**Figure 18.** Morphology of (I) *S. clavuligerus*/pICLAR and (II) *S. clavuligerus* /pIB139, growing in GYM agar plates (After 8 days of growth).

### *S. clavuligerus*/pIORF21

In GSPG, there was a significant ($p<0.05$) decrease in biomass growth for *S. clavuligerus*/pIORF21 compared to the control strain, during the exponential growth phase (48 h and 72 h), as shown in Figure 19A and 19B. This reduction, as well as a longer stationary phase, might be related to the kinetics of CA production in *S. clavuligerus*/pIORF21 (Figure 19A), using GSPG. Concerning biomass growth in ISP, a similar behavior ($p > 0.05$) was observed for *S. clavuligerus*/pIORF21 and *S. clavuligerus*/pIB139 (see Figure 19C and 19D). Moreover, a lower ammonium production was observed for ISP compared to GSPG. For each medium, ammonium production did not vary significantly ($p >0.05$) between strains (Figure 20).

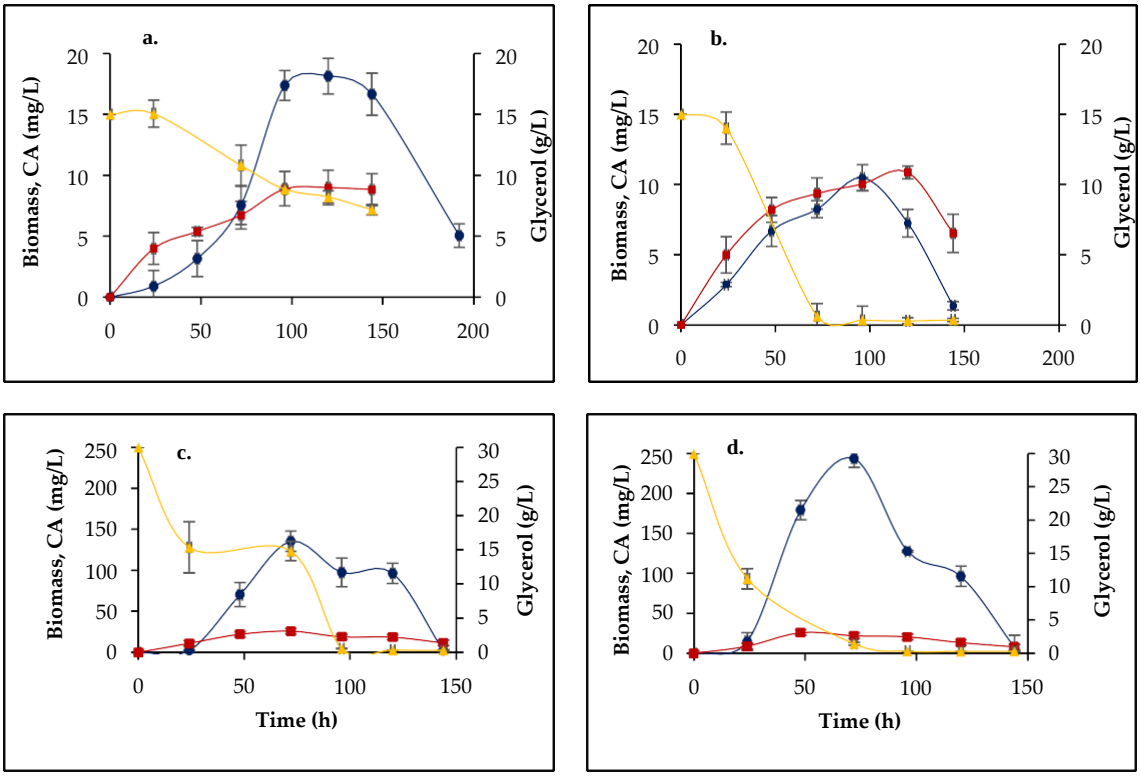


**Figure 19.** Dynamics of biomass, glycerol, and CA production. **a)** *S. clavuligerus*/ pIORF21 in GSPG. **b)** *S. clavuligerus*/pIB139 in GSPG. **c)** *S. clavuligerus*/ pIORF21 in ISP. **d)** *S. clavuligerus*/pIB139 in ISP. (Blue) CA; (Red) Biomass and (Yellow) Glycerol. Each value corresponds to the mean of three flask replicates (n=3). Error bars indicate standard deviation values.

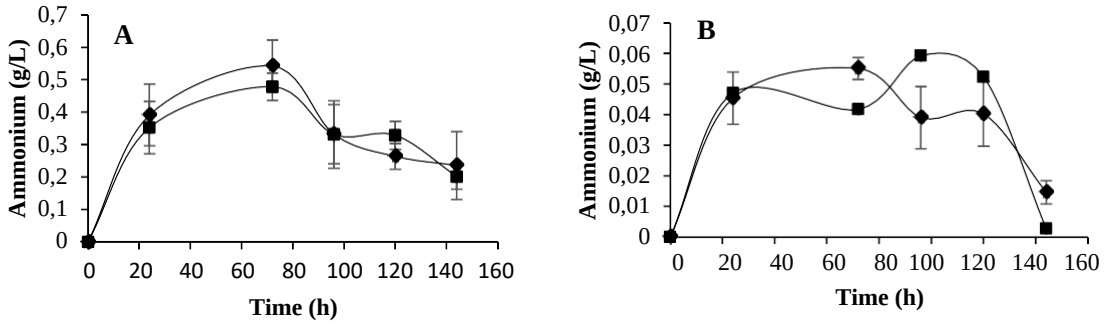


**Figure 20.** Ammonium production during *S. clavuligerus* batch culture. A) GSPG medium and B) ISP medium. *S. clavuligerus*/ pIORF21(♦) and *S. clavuligerus*/pIB139 (■).

Figure 21 shows the results for aerial mycelium formation of *S. clavuligerus*/pIORF21, in the GYM and MYM media. As observed, *S. clavuligerus* /pIORF21 favored aerial

mycelium after culturing for 5-8 days; only sparse mycelium was observed in control cultures (Figure 21). Likewise, a delay in spore formation was noted in *S. clavuligerus*/pIORF21 compared to the control strain (Figure 21B).

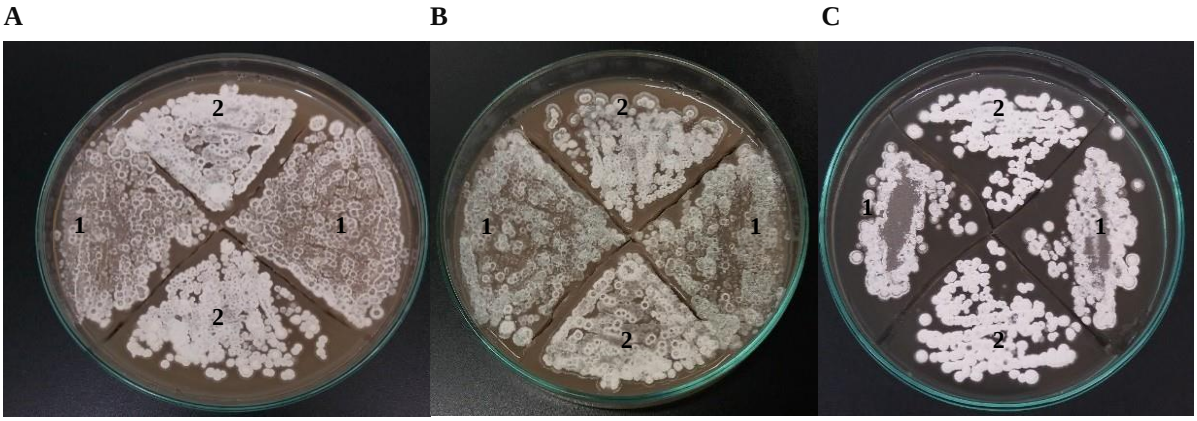


**Figure 21.** Aerial mycelium formation in solid medium with apramycin 40µg/mL. (1), *S. clavuligerus* / pIB139 (2), *S. clavuligerus* / pIORF21. A) After 5 days of growth in GYM. B) After 15 days of growth in GYM. C) After 8 days of growth in MYM.

### 4.5. Clavulanic acid and glycerol production

#### *S. clavuligerus/*pICLAR

Different fermentation studies were performed for *S. clavuligerus/*pIB139 and *S. clavuligerus/*pICLAR for the purpose of evaluating the effect of culture medium composition on the regulatory role exerted by CLAR on CA biosynthesis. In ISP, *S. clavuligerus/*pICLAR reached its maximum CA production (363,6 mg/L) peak at 72 h of fermentation (see Figure 22A), representing an increase in CA production of 1.3-fold ($p<0.05$) compared to *S. clavuligerus/*pIB139 (286 mg / L). Regarding the GSPG medium, *S. clavuligerus/*pICLAR enhanced CA production by 1.4-fold ($p<0.05$) compared to the control strain; the larger peak of CA production was observed at 45 h (see Figure 22B). Conversely, for the control strain, the maximum CA production was reached at 96 h of fermentation (see Figure 22B).

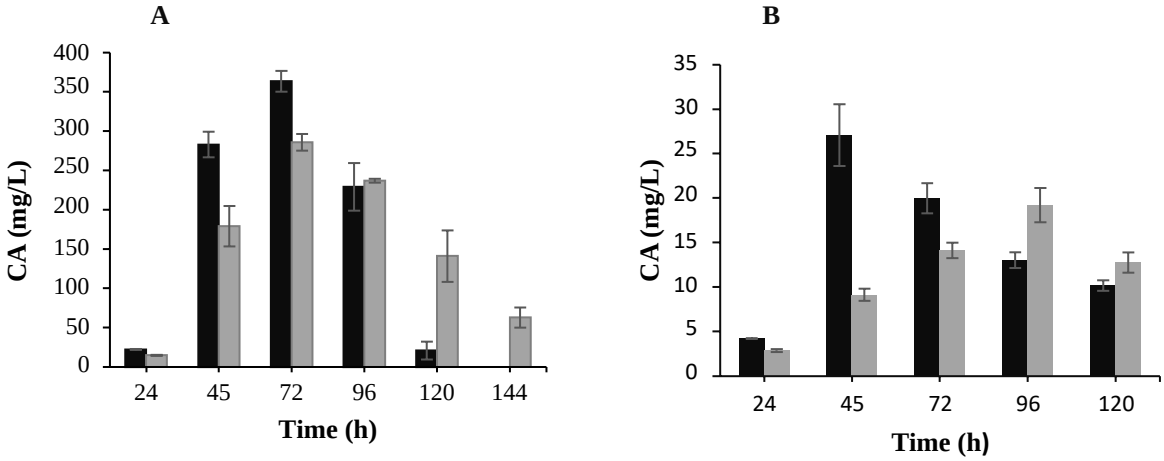


**Figure 22.** Kinetic of CA production in ISP and GSPG. In black, *S. clavuligerus*/ pICLAR. In gray, *S. clavuligerus*/pIB139. A) ISP. B) GSPG.

In both culture media, *S. clavuligerus/*pICLAR obtained a higher CA yield per gram of glycerol consumed and biomass produced (see Table 9). There were no differences in the rate of substrate consumption ($Q_s$) between the strains evaluated, for each medium (see Table 8). Therefore, there is a noticeable difference in the rate of substrate consumption between both media, which reflects the contrast in the kinetics of nutrient consumption and growth of *S. clavuligerus*.

In addition, after 48 h of culture in ISP, *S. clavuligerus* /pICLAR significantly ($p< 0.05$) decreased glycerol consumption, leaving 5.4 g/L of glycerol remaining at the end of fermentation (see Figure 23A). Notwithstanding, in GSPG, there was no significant difference ($p>0.05$) in glycerol consumption between *S. clavuligerus/*pICLAR and the control strain (Figure 23B).

**Table 9.** Kinetic parameters for CA biosynthesis by *S. clavuligerus* in GSPG and ISP media

| **Culture medium** | **Strain** | **$Y_{P/S}$ ($g.g^{-1}$)** | **$Y_{P/X}$ ($g.g^{-1}$)** | **$Q_s$ ($g.g^{-1}.h^{-1}$)** |
|---|---|---|---|---|
| GSPG | *S. clavuligerus* /pICLAR | 0,0054 | 17,61 | 0,035 |
| | *S. clavuligerus* /pIB139 | 0,0013 | 4,17 | 0,029 |
| ISP | *S. clavuligerus* /pICLAR | 0,057 | 44,31 | 0,0098 |
| | *S. clavuligerus* /pIB139 | 0,027 | 15,75 | 0,0071 |

$Y_{P/S}$: substrate to product yield; $Y_{P/X}$: biomass to product yield.

$Q_s$: Specific rate of substrate consumption (gram of substrate per gram of biomass per hour)

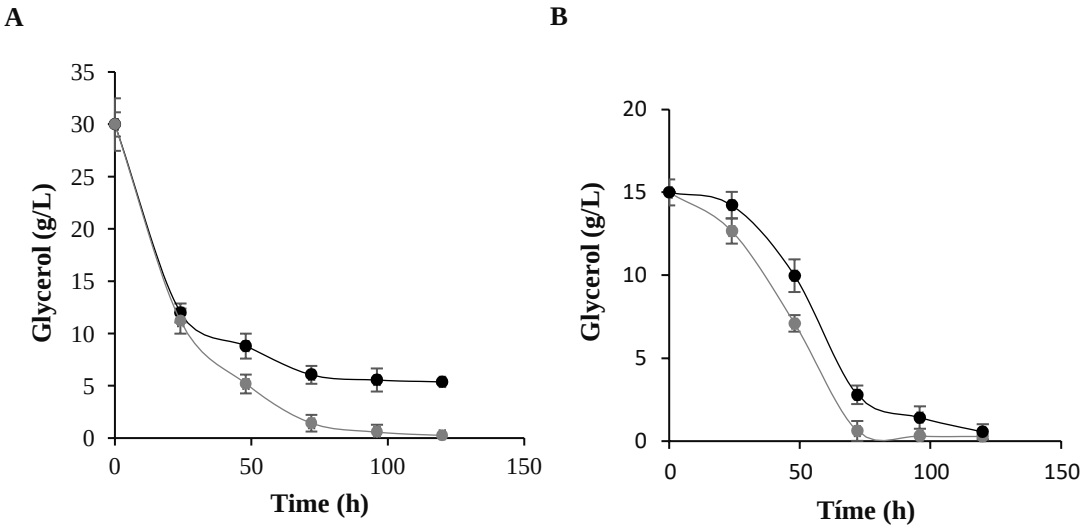


**Figure 23.** Glycerol consumption in GSPG and ISP. In black, *S. clavuligerus*/ pICLAR. In gray, *S. clavuligerus*/pIB139. A) ISP. B) GSPG.

### *S. clavuligerus*/pIORF21

CA production from *S. clavuligerus*/pIORF21 and the control strain was evaluated in two different culture media. For GSPG, the most significant increment in CA concentration was observed in *S. clavuligerus*/pIORF21 after 96 h (17,4 mg/mL) and 120 h (18,1 mg/mL) of fermentation, representing 1.7 and 2.6-fold ($p < 0.05$) increase, respectively; CA production for the control strain, was 10 mg/mL and 7 mg/mL, at 96 h and 120 h, respectively (see Figure 24A).

Unexpectedly, CA production was maintained for *S. clavuligerus*/pIORF21 until 144 h of cultivation, in contrast to the control strain, for which a continuous reduction in product accumulation was observed, from 96 to 144 h (Figure 24A). As for ISP, *S. clavuligerus*/pIORF21 performance, in terms of CA production, was quite different; interestingly, *orf21* overexpression, under the evaluated conditions, showed a low CA yield, compared to control (Figure 24B).

Furthermore, in GSPG, the wild type and control strains showed a similar CA production (Figure 24A). Nevertheless, using the ISP medium, for the control strain (*S. clavuligerus*/pIB139) CA production was approximately 4.1-fold larger ($p < 0.05$)

compared to wild-type (Figure 24B). This substantial difference may be related to a stress response induced by plasmid integration on host chromosome and/or by host physiology[112].

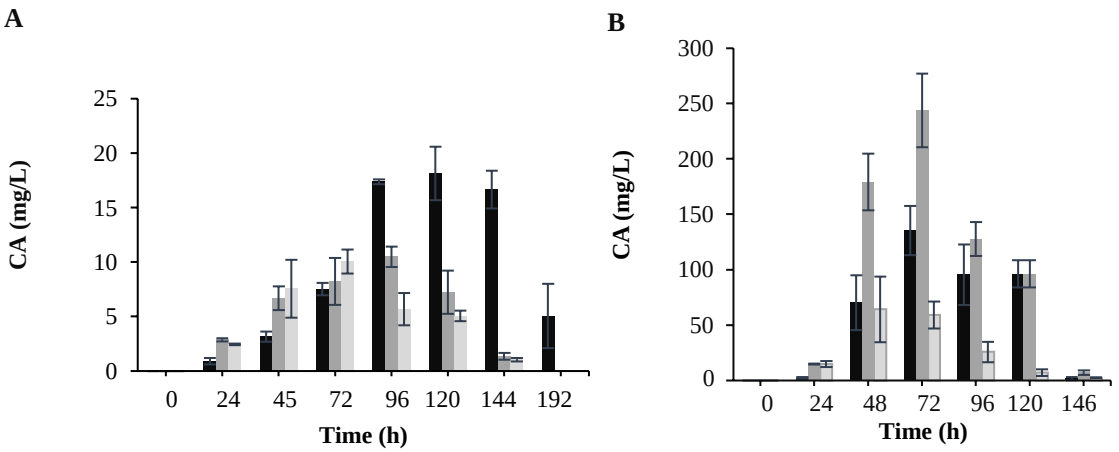


**Figure 24.** Evaluation of CA production in GSPG and ISP medium. In black, *S. clavuligerus*/ pIORF21. In dark gray, *S. clavuligerus*/pIB139. In light gray, *S. clavuligerus* wild-type strain. A) GSPG. B) ISP.

Concerning the substrate consumption, the uptake rate of glycerol showed by *S. clavuligerus* / pIORF21, and the control strain were significantly different ($p < 0.05$) in both culture media (Figure 19). In *S. clavuligerus*/pIORF21 (Figure 19A and Figure 19C), glycerol concentration slightly decreased, leading to 7.2 g/L of residual glycerol at 144h fermentation time, in the GSPG medium. In contrast, for *S. clavuligerus*/pIB139 (Figure 19B and Figure 19D), glycerol concentration rapidly decreased until 72 h of cultivation, when glycerol was completely depleted. For ISP, the specific consumption of substrate in *S. clavuligerus* / pIORF21 decreased notably (see Table 10), where at 72 h there were 14.8 g / L of glycerol compared to the control (72h, 1.43 g / L). Thus, for each strain, glycerol uptake rate and its relationship with CA production varied depending on the culture medium composition (see Table 10). Furthermore, the specific rate of substrate consumption was higher for both strains in GSPG than that in ISP, which shows that ISP, due to its high nutritional content, favors biomass production and allows obtaining higher CA yields per grams of glycerol consumed.

**Table 10.** Kinetic parameters for CA biosynthesis by *S. clavuligerus* in GSPG and ISP media

| Culture medium | Strain | $Y_{P/S}$ ($g.g^{-1}$) | $Y_{P/X}$ ($g.g^{-1}$) | $Q_s$ ($g.g^{-1}.h^{-1}$) |
|---|---|---|---|---|
| GSPG | *S. clavuligerus* /pIORF21 | 0,0025 | 3,44 | 0,0063 |
| | *S. clavuligerus* /pIB139 | 0,0005 | 1,40 | 0,0142 |
| ISP | *S. clavuligerus* /pIORF21 | 0,2480 | 8,90 | 0,0003 |
| | *S. clavuligerus* /pIB139 | 0,0234 | 17,60 | 0,0062 |

$Y_{P/S}$: substrate to product yield; $Y_{P/X}$: biomass to product yield.

$Q_s$: Specific rate of substrate consumption (gram of substrate per gram of biomass per hour)

### 4.6. Transcriptional analysis of clavulanic acid biosynthetic related genes in *S. clavuligerus*/pIORF21

To assess whether some genes associated with CA biosynthesis (*ccaR*, *claR*, *adpA*, SCLAV_4359, *orf21* and *gcas*) are under *orf21* control, a transcription analysis was carried out, using RT-qPCR for *S. clavuligerus*/pIORF21 and *S. clavuligerus*/pIB139 used as control, grown in two different culture media, (see Figure 25). Transcriptional analysis was performed using, as template, RNA isolated from strains cultured in GSPG medium during 96 h; at this time CA production started to increase, compared to control (Figure 24A). The gene *gcas* was partly expressed compared to control, where transcription rate increased by 1,7-fold (Figure 25). Only *adpA*, a pleiotropic regulator related to the biosynthesis of antibiotics and morphological differentiation in *S. clavuligerus* [63], showed a slight expression increment (1,9-fold). No differences in amplification for *ccaR*, *orf21*, *claR* and SCLAV_4359 were observed between the studied strains (data are not presented).

The RNA samples for the transcriptional studies in ISP, were harvested at the peak of maximum CA production (72 h) (see Figure 24B). No differences in gene expression profiles were observed, except for *claR*, for which, gene expression was greatly reduced (200-fold) (Figure 25).

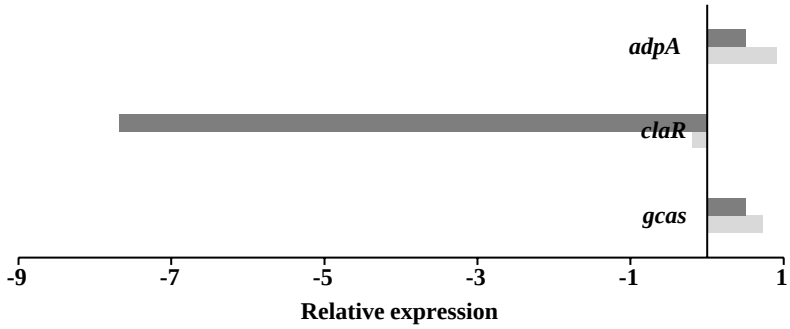


**Figure 25.** Effects of *orf21* overexpression on the transcription of genes related to CA biosynthesis. RT-qPCR results for *S. clavuligerus*/pIORF21 compared to *S. clavuligerus*/pIB139. In light gray the results for GSPG. In dark gray the results for ISP.

### 4.7. Phylogenetic inferences and DNA-binding motifs prediction

To infer phylogenetic relationships and interplays among ORF21 and other regulators of CA biosynthesis, sixty homologous protein sequences for ORF21 were mined. These proteins belong to microorganisms of the order Actinomycetales, e.g., the genus Nonomuraea, Actinomadura and Streptomyces. It was found that ORF21 sequence shares homology to SigL RNA polymerase sigma factor (Genbank access numbers) WP_153520658.1 and WP_153484481.1 from *S. jumonjinensis* and *S. katsurahamanus*, displaying 80.57% and 81.14% identity values, respectively; these sequences are more related to the ORF21 than any of the other ORF21 homologous proteins (Figure 26). Additionally, the clustering of ORF21 and related homologous proteins from CA producers is robust with a branch support of 100% (Figure 26). Also, the ORF21 protein was phylogenetically grouped with SigL proteins from the genus Streptomyces, as shown in Figure 27.

Figure 28 shows that, out of the six promoter regions analyzed using MEME, the promoter region corresponding to the *claR* gene from *S. clavuligerus*, *S. jumonjinensis* and *S. katsurahamanus*, as well as the *adpA*, *ccaR* and *gcas* genes from *S. clavuligerus*, harbored the putative motif observed in Figure 28B. The motif, analyzed using TOMTOM, was found to share homology to the transcription factor BldD, a known regulatory motif from *S. coelicolor* (Figure 28A).

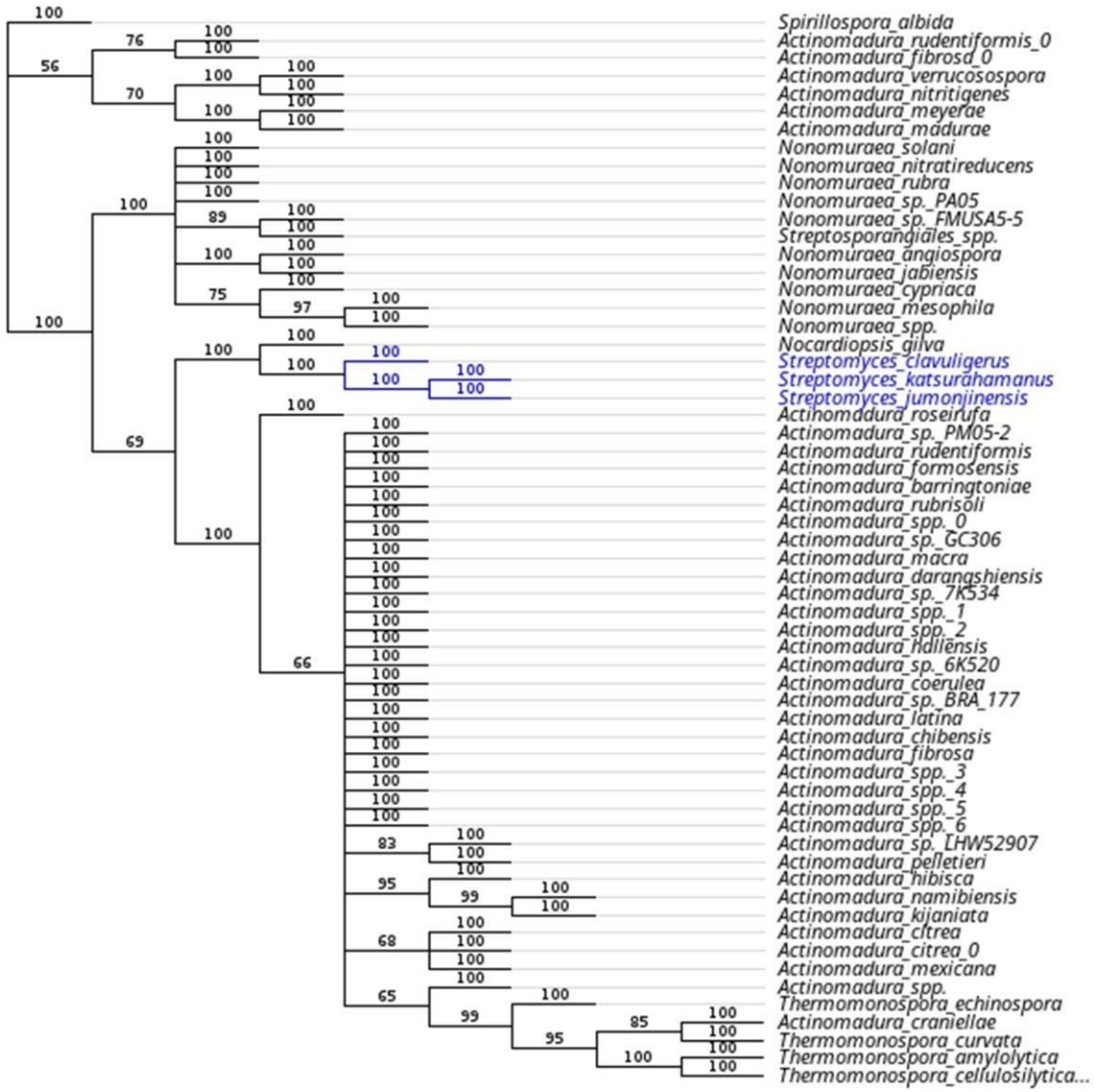


**Figure 26.** Phylogenetic tree of homologous proteins to *Streptomyces clavuligerus* ORF21 protein. Homologous proteins were taken from NCBI BLAST (http:// www.ncbi.nlm.nih.gov/BLAST/) search. In blue, the clustering for CA producer. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1.000.000 replicates) are shown next to the branches.

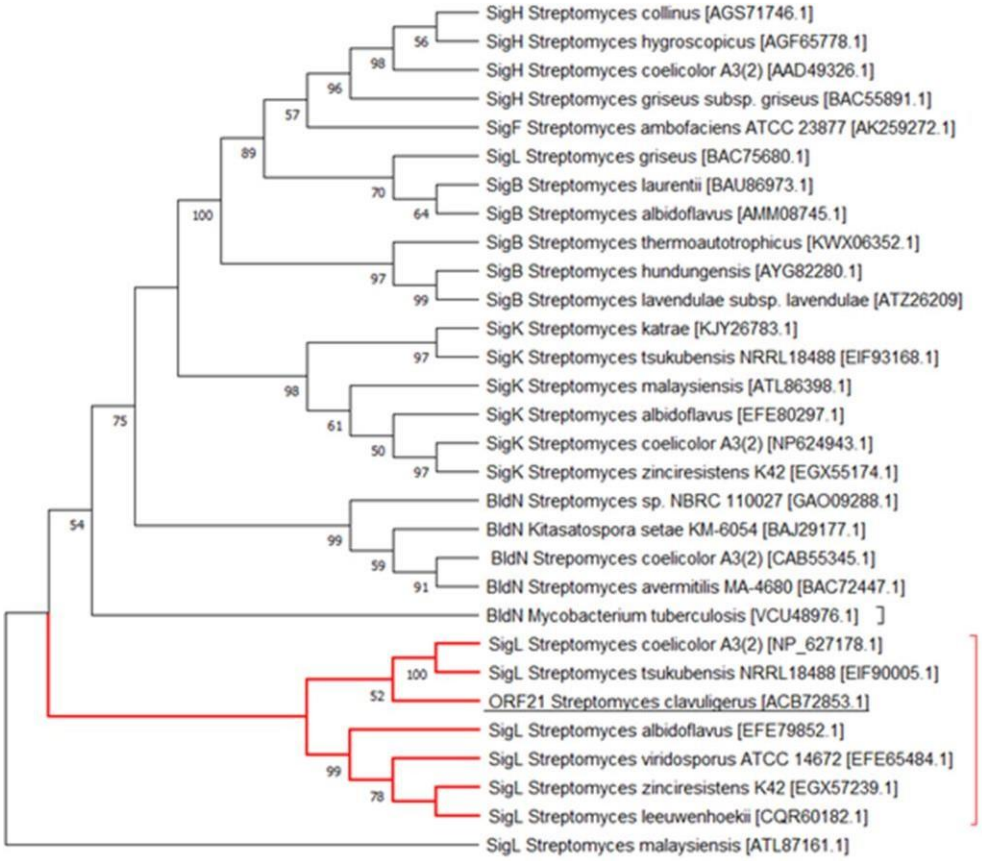


**Figure 27.** Phylogenetic tree of some ECF sigma factors and ORF21. The GenBank accession numbers for the protein sequences used in the analysis are included with the names of each species. The clustering for SigL sigma factors is show in red. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches.

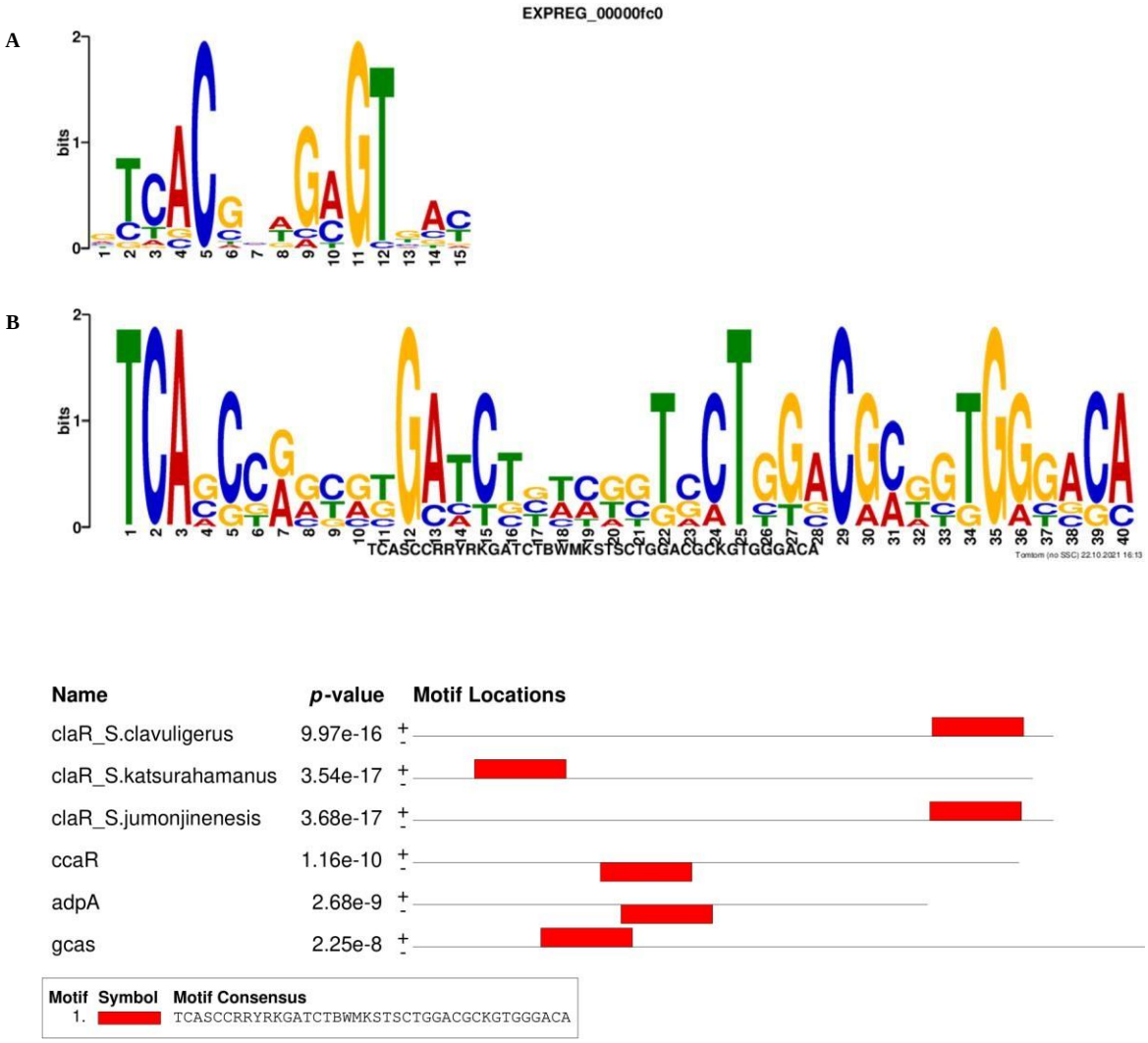


**Figure 28.** A). Putative motif found by TOMTOM (*p*-Value: $5.4e^{-0.4}$) against a collection of bacterial transcription factors. TOMTOM shows that the query motif resembles the binding motif for the transcription factor BldD. B). Putative binding motif logo identified by MEME suite (E-value: $2.8e^{-0.06}$) in the promoter region of the *claR* gene showing affected transcription in *S. clavuligerus*/ pIORF21.

# 5. DISCUSSION

*S. clavuligerus* produces low CA titers, considering the complexity of the production process and the potential hydrolysis the product experiences once it is synthesized. CA production is controlled by multiple interactions of pleiotropic and cluster situated transcriptional factors working on different metabolic networks and cascades [113]. Taking into account, the low CA production obtained from *S. clavuligerus* wild-type strain [114,41,29], and its clinical and economic importance, diverse strategies have been developed to improve CA production yields. Strategies to enhance CA production include control of bioreactor conditions [115], optimization of culture medium [115], as well as genetic engineering [36,41,116,117]. *S. clavuligerus* has been genetically modified in multiple studies to increase CA yields as well as to study the regulatory mechanisms that control CA biosynthesis (Table 2). However, the genetic improvement of *S. clavuligerus* as one of the best tools to enhance CA production, is still a field of intense research. Therefore, the primary goal of this study was to construct a wild-type *S. clavuligerus* recombinant strain to improve CA production. Different molecular biology strategies were proposed; among them, the overexpression of regulatory genes located in the CA gene cluster and the interruption of the expression of a gene involved in the biosynthesis of a metabolite that generally compete with CA biosynthesis (cephamycin C). Nevertheless, due to the interaction of multiple factors, not all the proposed strategies were achieved. The potential factors that limited gene disruption in *S. clavuligerus*, as well as the discussion of the other results obtained are considered in the following sections.

## 5.1. Construction of a mutant strain of *S. clavuligerus* with the *lat* gene disrupted

To disrupt the *lat* gene expression on the wild-type *S. clavuligerus* chromosome, the shuttle plasmid pSLAT was constructed. However, attempts to transfer pSLAT or the control plasmid pSOK201 to *S. clavuligerus* via both protoplast transformation and conjugation, failed to give any transformants, even though the plasmid pSOK201 contained the genetic elements to be introduced into *S. clavuligerus* and for its replication (see Figure 6). Although a method for efficient *S. clavuligerus* transformation via protoplast was established (see Methodology), after running the protoplasts transformation protocol, a relatively high number of small, regenerated protoplasts

appeared on R2YEG plates. Regardless of that, these putative transformed protoplasts failed to replicate and growth on fresh selective TSA media with apramycin.

pSOK201 is derivative from pGM11 plasmid; it contains the pSG5 replicon, frequently used for gene cloning in *Streptomyces* spp. Roth *et al*., (1994), analyzed whether the pGM-vectors vary in their segregational stability, and whether there were relations between plasmid structure and segregational stability in *Streptomyces lividans* [118]. They found that the pGM11 plasmid presented low stability under the evaluated conditions compared to other plasmids of the pGM-family. The reduction in segregational stability was related to a lower number of plasmid copies available for partitioning at cell growth [118]. Thus, the probability for the appearance of plasmid-free cells during growth of a population increases with decreasing copy number [118].

On the other hand, Zotcher *et al*., (2000), constructed the pSOK201 plasmid to extend the choice of vectors for gene disruption experiments in *Streptomyces noursei* (*S. noursei*). However, transformation of *S. noursei* protoplast with pSOK201 failed [84]. In the case of conjugation experiments using pSOK201, Zotcher *et al*., obtained putative transconjugants, even so the colonies stopped growing after 3±4 d incubation and did not develop into normal colonies. Any attempt to transfer these small colonies to fresh media resulted in slow-growing and poorly sporulating colonies from which no plasmid DNA could be isolated [84]. Conversely other research works have efficiently transformed *Streptomyces natalensis* by conjugation using the plasmid pSOK201, although with lower frequencies of transformation than obtained with other plasmid [119].

Considering the difficulties mentioned above, it can be inferring that pSOK201 is an unstable plasmid strongly influenced by the plasmid gene products and plasmid structure as well as by host functions. Thus, all these variables could limit the transformation of *S. clavuligerus* with pSOK201 in the present work.

### 5.2. Effect of overexpressing *claR* on clavulanic acid production and other features of *S. clavuligerus* metabolism

The *claR* gene, located downstream from *oppA1(orf*-7) in the CA gene cluster (Figure 3), encodes a LysR-type pathway-specific transcriptional regulators (LTTRs)[120], a critical regulator for CA biosynthesis [55]. The LTTRs are the most abundant in prokaryotes,

characterized by having the conserved DNA binding HTH (helix-turn-helix) motif at the N-terminus. They also have a regulatory domain for substrate or inducer binding, at the C-terminus [121]. Thus, they have been found regulating the expression of genes coding for proteins involved in very diverse functions like β-lactamase, transporter, amino acids biosynthesis, metabolic signaling, secretion, oxidative-stress response, cell division, quorum sensing, virulence, motility, detoxification, and attachment [121,122].

ClaR controls the expression of genes involved in the late steps of the CA biosynthesis pathway, therefore, an increase in *claR* expression may favor CA production yield. To check the effect of *claR* overexpression in *S. clavuligerus* ATCC27064, recombinant and control strain (Table 4) were grown in GSPG and ISP media, exhibiting a CA production increase of 1,3 and 1,4-fold, respectively, compared to control strain. These results are consistent with previous studies, where *claR* overexpression effectively induced a positive effect on CA production [36,42,114,116] .

Hung Viet et al (2006), constructed the recombinant multicopy plasmid pIBRHL2, containing *claR* [114]. The plasmid was expressed in *S. clavuligerus* NRRL3585 along with the *ermE** promoter. The transformants carrying multiple copies of *claR* enhanced CA production (1.5-fold) over their parental wild type strain [114]. Next, Kurt Kizildoğan et al (2017), overexpressed *claR* in the industrial strain of *S. clavuligerus*, namely DEPA. They used the constitutive promoter *ermE** (*permE**) and the *claR* own promoter to add its additional copy into the chromosome, as well as the *glpF* (strong glycerol inducible promoter) promoter for its multicopy expression in the cell. *claR* expression controlled with its own promoter or *ermE** and *glpF*-mediated amplification in an industrial strain brought about merely 1.2-fold increase in the volumetric CA titers [36]. Likewise, Soo Hang et al (2017), evaluated the *claR* overexpression into the industrial *S. clavuligerus* OR strain, finding only a 10% increase in CA production [116].

Although the *claR* gene overexpression in different *Streptomyces* strains does increase CA production, this enhancement can be considered as low or limited, when compared to *ccaR* gene overexpression or other regulatory genes associated with CA biosynthesis (see Table 3). These results might be attributable to the fact that ClaR main regulatory effects are involved in late steps of CA metabolic pathway, though Martínez-Burgo et al (2015)

reported that ClaR could also partially regulate early steps genes [55]. CcaR regulates CA production by binding to heptameric target sequences located upstream of *ceaS2* and *claR*, exerting transcriptional control over the entire CA genes cluster [49]. Therefore, even if the *claR* gene is overexpressed, either in a multicopy or integrative plasmid, the expression of the early genes will not be strongly affected. In consequence, carbon flux from early pathway steps to clavaminic acid synthesis would restrict CA production, due to a synergy effect on *ccaR* and *claR* genes during CA production; co-expression of *ccaR* and *claR* is required to obtain higher CA yields in genetically modified strains [116].

Secondary metabolite production and morphological differentiation in bacteria are controlled by external signal molecules that bind to cellular receptors and trigger intracellular regulatory cascades [68]. In this process, genes are activated and repressed to adapt bacterial metabolism to different environmental conditions e.g., by producing antibiotics to compete with other microorganisms under auxotrophic conditions. This research wanted to evaluate whether *claR* effect on CA biosynthesis was affected by the fermentation medium, considering that *claR* gene expression product is a transcriptional regulator located at the bottom of the CA regulatory network. GSPG has been determined as low CA production medium compared to other defined media [29,30] ; this medium uses L-glutamate and L-proline as nitrogen source. In contrast, ISP is a complex medium; when used, high CA yields are obtained [123]. The high CA production using ISP is attributed to complex nitrogen source (soybean flour) which provides high free amino acids amounts [28,64,124]. Despite the nutritional differences in each culture medium, CA production was not affected (see Figure 22).

### 5.3. The culture medium composition influences the ORF21 role on clavulanic acid production and some characteristics of primary metabolism

Several transcription regulators are involved in the regulation of CA production, some of them specific to the metabolic pathway (e.g., ClaR) and others involved in different cellular processes, such as sigma factors. In the bacterial cell, RNA polymerase (RNAP) binds the gene and/or operon promoter region by sigma factors recognition, directing specific genes expression [121]. Sigma factors are a typical mode of global regulation that control antibiotic production in *Streptomyces* [121,49]. One of the most important and diverse

groups of sigma factors are the extracytoplasmic function σ factors (ECFs) [59]. ECFs coordinate transcriptional responses to extracellular signals; yet, its role and mechanisms of regulation are largely unknown [68]. According to the ECF hub, an open access data repository for ECF classification, the putative SigL sigma factor (Figure 27) encoded by *orf21* belongs to the actinobacterial ECF17 group [125]. The proteins in the ECF17 group, likely participate in enhanced protein secretion, cell envelope lipid synthesis and secreted protein modification, as evidenced by studies of *S. coelicolor* SigU and *Mycobacterium tuberculosis* SigL [126]. In addition, the proteins present in the ECF17 group might be regulated by anti-σ factors [125]. Generally, these proteins are co-transcribed with a transmembrane anti-σ factor, with an extracytoplasmic sensory domain and intracellular domain [68]. Upon interaction with an extracytoplasmic signal (e.g., a protein or small molecule), a typical bacterial ECF factor is released and is free to bind to core RNAP [68,127] (Figure 5).

The sigma factor *orf21* has a role in CA production; however, its effect is variable across different laboratories and seems to be linked closely with operational differences in the fermentation e.g., composition of the culture medium [49,41,40]. Jnawali et, al (2011) found that *orf21* overexpression in a multicopy plasmid increased CA production 1.43 fold during cultivation in SA medium [41]. Likewise, *orf21* was disrupted and its mutation in *S. clavuligerus* resulted in a 10-15% decrease in CA production in the same cultivation medium [41]. In contrast, Jensen et, al (2009) found that *orf21* disruption had no effect on CA production in a glycerol-arginine (GA) medium [40].

Since *orf21* transcription rate is activated by unknown regulatory mechanisms involving diverse environmental cues, in this contribution we assessed the effect of overexpressing *orf21* under *erm*E* promoter to increases its transcription rate and, connect ORF21 to CA pathway. In addition, the medium GSPG was chosen to evaluate the regulatory effect of ORF21 on CA biosynthesis due to its high salt concentration, limited nutritional content, as well as its known tendency to induce ammonium accumulation [32,64]. All these nutritional stress conditions and environmental imbalance in GSPG could affect the molecular mechanisms by which ORF21 step on CA regulation, substrate consumption and morphological differentiation, compared to its regulatory role under nutrient rich conditions.

In GSPG, the overexpression of *orf21* in *S. clavuligerus* ATCC27064 increased the production of CA 2.6-fold with a higher CA production per gram of substrate (Table 10). During cultivation, glycerol is used as the primary metabolic precursor of D-glyceraldehyde-3-phosphate (G3P) and as carbon source for maintaining cell metabolism [117]. In this manner, *S. clavuligerus*/pIORF21 is more efficient than the control strain for redirecting metabolic fluxes towards secondary metabolism rather than to growth ($Y_{P/X}$: 3,44 $g \cdot g^{-1}$) (Table 10). To gain further insights about ORF21 incidence on primary metabolism, we analyzed the rate of glycerol uptake as an indirect measure of G3P direct primary metabolic precursor of CA. We found that, overexpression of *orf21* reduced glycerol consumption rate in both, the GSPG and ISP media (see Table 10). In *S. clavuligerus*, glycerol metabolism is controlled by the *gylR-glpF1K1D1* gene cluster, encoding a regulator of the IclR family, a protein for glycerol transport, a glycerol kinase, and a glyceraldehyde-3-phosphate dehydrogenase (GAPDH), respectively [128]. Different physiological conditions would require different glycerol kinase/G3P dehydrogenase ratios (G3P generation versus oxidation) and therefore different regulatory mechanisms [129,130]. In the wild-type strain of *S. clavuligerus*, about 20% of G3P is used as a precursor for CA biosynthesis and gluconeogenesis. The remaining 80% enters the glycolytic pathway through the GAPDH encoded by the *gap1* gene [131]. The induction and regulation of *gylR-glpF1K1D1* operon in *S. clavuligerus* influences CA production by managing G3P supply [128]. We suspect that *orf21* overexpression would play a role in the transcription rate of genes related to the uptake and/or transport of glycerol into the cell, given the decrease in glycerol consumption by *S. clavuligerus*/pIORF21 (see Figure 19). In enteric bacteria, for example, the inhibition of glycerol uptake could be exerted through modulation of the activity of glycerol kinase [129,132]. However, in *Streptomyces*, the control of the GAPDH enzyme would provide a mechanism for regulating the glycerol catabolism [51]. Although we did not evaluate the expression of genes related to glycerol metabolism in *S. clavuligerus*, transcriptional studies, previously obtained in our laboratory, suggest that *gap1* gene expression in ISP increased around 12-fold in *S. clavuligerus*/pIORF21 compared to the wild-type strain. However, this presumption about the regulatory role of *orf21* on *gap1*, must be evaluated in experiments in which *S. clavuligerus*/pIB139 can be used as control in both, ISP and GSPG [64,133].

*orf21* is located upstream of the *orf22* and *orf23* genes in the CA gene cluster of *S. clavuligerus* ATCC 27064. In *S. clavuligerus* F613-1, the paired genes *cagS* and *cagR*, which are annotated as *orf22* and *orf23* in *S. clavuligerus* ATCC 27064, respectively, encode a bacterial two-component regulatory system (TCS) known as CagRS TCS. Here, CagRS could bind the intergenic region of *claR* and *oat2* genes affecting CA production. Likewise TCS Cags, TCS CagRS mainly regulates genes involved in fatty acid degradation, G3P and arginine metabolism [58]. Two-component regulatory systems constitute a family of proteins that mediate adaptation to changing environments by modifying the phosphorylated state of a pair of proteins: a sensor histidine kinase and a response regulator [58]. Similarly, *orf21* may be part of an alternative mechanism that regulates carbon metabolism with CA biosynthesis in response to an environmental stimulus.

Furthermore, we evaluated ammonium production to infer certain characteristics that would help to understand the differences in the regulatory nature of ORF21 on CA production in both culture media. The low CA production in GSPG can be conferred by the low availability of free amino acids [64] and glutamate metabolism. As mentioned earlier, in the CA biosynthetic pathway in *S. clavuligerus*, glutamate can be converted to 2-oxoglutarate by glutamate dehydrogenase, which causes the release of ammonium [30]. Ammonium production has a negative effect on CA biosynthetic pathway at concentrations above 0.26 g/L. In GSPG, the maximum ammonium production was 0.5 g /L, while in ISP it was merely 0.06 g/L (Figure 20). ORF21 is an ECFs that can be released and activated in response to an external inducing signal. Therefore, high ammonium concentration could indirectly modulate the expression of genes regulated by ORF21 in *S. clavuligerus/* pIORF21.

Different regulatory mechanisms relate stress signals with secondary metabolism in *Streptomyces*. The pleiotropic orphan nitrogen regulator GlnR is the main regulator involved in response to changes in nitrogen availability, in *S. coelicolor* and other *Streptomyces* species e.g., *Streptomyces venezuelae*, controlling the expression of genes at the transcriptional level [59]. Under nitrogen limitation, GlnR activates the transcription of genes related to ammonium metabolism and transport, secondary metabolism, and amino acid biosynthesis [59]. Another type of control over nitrogen metabolism genes is

exerted by the AfsQ1 − AfsQ2 two-component system. The effect of *afsQ1*/*Q2* over thesecondary metabolism has been known to occur in *S. coelicolor* in minimal medium supplemented with L-glutamate as the sole nitrogen source [59]. Therefore, the expression of ORF21 regulon and its incidence on CA biosynthesis is the result of regulatory mechanisms activated by the metabolism of the culture medium components in *S. clavuligerus.*

### 5.4. The regulatory characteristics of ORF21 on genes associated to clavulanic acid biosynthesis.

The *orf21* overexpression in a constitutive plasmid may well be regulating the production of CA in response to some nutrients e.g., high amino acid concentrations, peptides, and carbohydrates, present in the ISP culture medium. Therefore, the effect of *orf21* overexpression on the transcription rate of some genes related to CA biosynthesis was evaluated. Our results suggest that *orf21* overexpression indirectly decreases the transcriptional level of *claR* (Relative fold change: 0.005; Figure 25). Martinez-Burgo *et al* (2015), showed that a Δ*claR* mutant did not produce CA and exhibited a poor expression of genes for the late steps of the CA pathway [55]. Consequently, the decrease in the transcriptional level of *claR*, might explain the decrease in CA titers, obtained in ISP medium by *S. clavuligerus*/pIORF21 (see Fig. 24B).

Apart from *S. clavuligerus*, *S. jumonjinensis* and *S. katsurahamanus* are the only species known to produce CA along with cephamycin C [43]. AbuSara et al., reported that *orf21* gene is not present in *S. jumonjinensis* and *S. katsurahamanus* genomes, and therefore, it is not part of the core genes required for CA biosynthesis [43]. Although, the genes that encodes for WP_153520658.1 and WP_153484481.1 in *S. jumonjinensis* and *S. katsurahamanus* are located far from CA gene cluster and have a different synteny respect to *S. clavuligerus*, these proteins could be involved in the expression of genes related to CA biosynthesis through trans-regulatory mechanisms. Thus, the promoter regions of the *claR* gene of *S. jumonjinensis* and *S. katsurahamanus* were used as a support for In- silico prediction of the theoretical DNA motif reported in Figure 28B. The theoretically predicted DNA motif closely resembles the binding motif for the BldD (UnitProtKB:

Q7AKQ8) transcription factor (Regulatory mode: 28% repression) of *S. coelicolor* (Figure 28B). In *S. coelicolor*, BldD mainly acts as a repressor of genes involved in morphological differentiation and /or secondary metabolism [70]. Therefore, we speculate the expression of SCLAV_0719 (*bldD*), for *S. clavuligerus* grown in ISP, may well be affected by *orf21* overexpression and, as a result, exert a negative regulatory influence on the transcriptional level of some genes related to CA production i.e., *claR*, *gcas* and *adpA* (see Figure 25). On this subject, it might be reasonable to propose that ORF21 indirectly affects the transcription rate of *claR* by BldD. Nevertheless, this inference must be further studied in works where the transcriptional level of the *bldD* gene in *S. clavuligerus* / pIORF21 would be evaluated.

In contrast, in Figure 2, *S. jumonjinensis* and *S. katsurahamanus* were grouped in a distant cluster from *S. clavuligerus*. This genomic distance show that, although *S. jumonjinensis* and *S. katsurahamanus* have the core genes for CA production, the genes present in the clavam, and paralogue gene cluster are important to produce CA. In addition, regulatory elements present in *S. clavuligerus* could be essential for enhancing CA production.

To evaluate the effect of *orf21* overexpression on genes associated with CA production in GSPG, the transcriptional levels of well-known CA related genes were evaluated. We found no differences in amplification of *ccaR* transcript (Figure 25) between the evaluated strains. Nonetheless, Jnawali (2010) reported that ORF21 induces CA production by activating some genes (*ccaR*, *cas2* and *ceas2*) involved in the early steps of the CA metabolic pathway, at 60 h of culture in SA medium [41]. In this study, the RNA samples were taken in the stationary phase (96 h), associated with morphological differentiation and nutritional stress [133]. During this time, glycerol was completely depleted (Figure 19a), which could induce the transcription of genes of ORF21 regulon related to CA late steps. Hence, ORF21 would not only affect the early steps of CA formation but also the late steps.

To assess the relationship between *orf21* and CA late steps, we studied the transcriptional levels of *adpA* gene for the conditions assessed in this work. *adpA* encodes for the AdpA protein, a pleiotropic regulator that acts positively on the CA regulatory cascade and, is involved in the biochemical and morphological differentiation of *S. clavuligerus* [39]. We

found that *adpA* was slightly overexpressed (see Figure 25) in GSPG. AdpA mainly affects the expression levels of early biosynthetic genes e.g., *ceas2, bls2, pah2* and *cas2*, and late biosynthetic genes e.g., *claR* and *ccaR,* of the CA metabolic pathway [63]. Similarly, Pinilla et al., performed a comparative analysis of the transcriptome of *S. clavuligerus* ATCC 27064 wild-type strain, and found an increase in the expression of the *adpA* gene which concurs with a high CA concentration [64].

Moreover, we evaluated the transcriptional levels of the synthetic *gcas* gene; we observed the *gcas* gene was partially overexpressed (see Figure 25). The enzyme N-glycyl-clavaminic acid synthase controls the branch point between CA and the (3S, 5S)-clavams. An increment in the *gcas* overexpression would promote carbon flux towards CA biosynthesis. In the *S. clavuligerus* Δ*adpA* mutant reported by Lopez Garcia et al., (2017) genes co-transcribed by the *ceas2* promoter and genes in the *oppA2-gcas* operon were down-regulated. Thus, there was a large decrease in the expression of genes encoding biosynthetic enzymes for the first steps of CA formation and a minor decrease in gene expression for the late steps [63]. Although, in the present study, an increase in the expression of the genes *ccaR* and *claR* was not obtained (see Figure 25), the slight improvement in the relative expression of *gcas* and *adpA*, might contribute to improve CA production in GSPG (see Figure 24A). Therefore, *orf21* overexpression may explain a possible relationship between ORF21 sigma factor and *adpA* regulatory gene, which in turns, as a pivotal effect, positively affects the expression of *gcas*, perhaps due to the AdpA binding to the regulatory region of the *oppA2-gcas* operon.

### 5.5. ORF21 and its potential relationship with *S. clavuligerus* morphological differentiation

One of the most important characteristics of *Streptomyces* is its complex life cycle, which is closely related to secondary metabolite production [134]. In solid sporulation cultures, development starts with spore germination and the rapid development of compartmentalized hyphae into the medium (early substrate mycelium or MI). After that, programmed cell death (PCD) occurs which triggers the differentiation of the multinucleated (MII) antibiotic-producing hyphae (late substrate mycelium, aerial MII) [134,135]. Modifying the developmental conditions of *S. clavuligerus* to enhance PCD and

MII differentiation, leads to an improvement in secondary metabolite production [134]. In Figure 21A and 21C, it was observed that *S. clavuligerus*/pIORF21 has a greater production of aerial mycelium (after 5-8 days) relative to the control strain. A greater production of aerial mycelium by *S. clavuligerus*/pIORF21, as well as a delay in sporulation in the GYM and MYM media, might be related to the overexpression of *orf21* and, therefore, an increase in CA production.

It is known that mycelial morphology correlates with the production of secondary metabolites in some *Streptomyces* spp [39]. In *S. coelicolor*, for example, morphological differentiation in liquid cultures is comparable to that occurring during pre-sporulation stages in solid cultures: an initial compartmentalized mycelium suffers a programmed cell death, and remaining viable segments then differentiate to a second multinucleated antibiotic-producing mycelium [39]. For *S. clavuligerus*, the pellet morphology in liquid cultures presents a primary mycelium formation within the pellet and secondary mycelium in the periphery, similarly to *S. coelicolor* [133]. Furthermore, there is a relationship between CA accumulation and morphology of disperse hyphae (secondary mycelium) [133]. All these observations together, and the results obtained in the present research, support the relationship between aerial mycelium and CA production in solid medium. Moreover, it is known that the *adpA* gene acts as a switch for aerial mycelium formation; it may also be involved in hydrolytic enzyme expression (lipase, nucleases and proteases) whose action allows cellular differentiation by degradation of cytoplasmic contents causing hyphal dispersion, followed by secondary mycelium formation [136]. Thus, we argue that a higher production of aerial mycelium in *S. clavuligerus*/pIORF21, may well be associated to an increase in the expression levels of the *adpA* gene. In this regard, the sigma factor ORF21 regulates the expression of *adpA* directly or through the expression of other genes belonging to its regulon, related to *adpA* expression.

Lastly, the findings obtained from this contribution suggest that *orf21* overexpression mediates the regulation of the *S. clavuligerus* metabolic response to the associated environmental conditions i.e., media composition. Therefore, the regulatory effect of ORF21 on CA production will depend exclusively on the conditions within which the cultivation is carried out. ORF21 can affect the secondary metabolism of *S. clavuligerus* and improve CA production through the transcription initiation of genes involved in the

late steps of CA pathway, such as *adpA* and *gcas*. Nevertheless, under certain environmental conditions, *orf21* overexpression negatively affect CA production, indirectly repressing *claR* expression. Likewise, this contribution revealed that somehow *orf21* affects morphological differentiation and primary metabolism of *S. clavuligerus* by regulating glycerol consumption. Consequently, ORF21 would be affecting the expression level of multiple genes involved in CA biosynthetic pathway, morphological differentiation, and primary metabolism in wild-type *S. clavuligerus.*

### 5.5.1. Insights about the inductive effect of pIB139 on clavulanic acid production under the influence of host cell physiology

We analyzed whether the insertion of pIB139 into the *S. clavuligerus* ATCC27064 chromosome would significantly affect CA production by the host cell. In GSPG, pIB139 did not have an inducing effect on CA production and its behavior was similar to the wild-type strain (Figure 24). Concerning ISP, *S. clavuligerus*/pIORF21 increased ($p < 0.05$) CA production by 212% compared to the wild-type strain (see Figure 24B). As mentioned above, the plasmid pIB139 in ISP elicits the production of CA; in this respect, *S. clavuligerus*/pIB139 surprisingly boosts ($p < 0.05$) CA production by approximately 381%, compared to the wild-type strain (see Figure 24B). In GSPG and ISP, the impact of pIB139 on CA production by the wild-type strain of *S. clavuligerus* was markedly different. This behavior is not necessarily unusual since plasmid integrations affect host cell metabolism[112]. In general, bacteria respond to the presence of a new plasmid by altering their metabolism to compensate for the new physiological requirements. These alterations could be sufficient to eliminate the plasmid metabolic burden [137]. However, abiotic factors that affect host physiology should be considered as elements that influence the plasmid regulatory role on the expression (over-expression or under-expression) of global transcriptional regulators [127,117].

In addition, it is unexpected that the inducing effect of pIB139 on CA production occurs only in ISP medium, and not in GSPG. It is known that the production of some secondary metabolites is finely controlled by the Quorum sensing (QS) mechanism [138]. QS is the mechanism by which bacteria join forces with other members of the population to increase their survival capacities and act as a big multicellular organism.

Notwithstanding, QS is unproductive when carried out by a single cell and become effective when carried out by a group [138]. Microorganisms as *Streptomyces* species, produce signaling molecules which work as important regulatory factors for coordinating the production of secondary metabolites and sporulation. Antibiotics, besides their role as antimicrobials, may have an additional function as signal molecules at subinhibitory concentrations. When they are present in sub-lethal amounts, they provoke changes in gene expression that affect different processes and behaviors, for example, β-lactam antibiotics such as imipenem or penicillin, and aminoglycoside antibiotics such as tobramycin, stimulate the formation of biofilms in *Pseudomonas aeruginosa* or *Streptococcus pneumonia.* In fact, it has been argued that the primary function of antibiotics compounds in nature is as signaling molecules, and that they only act as actual antibiotics when they reach a certain concentration in the medium [138].

Although additional experiments are required to establish why pIB139 induces CA production in ISP, it can be hypothesized that the high cell density achieved in ISP favors QS mechanism in *S. clavuligerus* cells. Thus, the addition of apramycin to the culture medium acts as a signal molecule that activates the expression of the *aac*IV gene and, at the same time, increases the expression of genes related to CA biosynthesis. Conversely, in GSPG, biomass production or cell density is lower compared to the biomass levels reached in ISP (see Figure 19). Likewise, GSPG is a medium with high salt concentration, which could be inducing a phenomenon of osmotic stress, limiting cell proliferation and, therefore, affecting the QS among bacterial cells. Under those environmental conditions, the presence of pIB139 would not affect CA production.

Moreover, pIB139 is pSET152 derivative plasmid, which could be integrated at several loci of *S. clavuligerus* chromosome via independent recombination events and somehow affects CA production in ISP [139]. pIB139 is a pSET152 derivative plasmid that utilizes the ΦC31 system for its integration at several loci of *S. clavuligerus* chromosome via independent recombination events [139]. ΦC31-based vectors induce metabolic perturbation by inactivating Pirin-encoding gene, whose encoding product is involved in different cell biological processes [112]. Although the effect of pirin like gene disruption has not been studied in *S. clavuligerus*, our results indicate that pIB139 integration at this site of the chromosome could alter the primary and secondary metabolism of *S. clavuligerus* in ISP.

Therefore, by some means these interconnections among environmental conditions and plasmid integration would explain the alteration in CA production by *S. clavuligerus* / pIB139. However, further gene expression analysis on this topic would provide interesting insights into CA regulation and the effect of pIB139 on S. clavuligerus physiology.

# 6. CONCLUSION AND PERSPECTIVES

This work was aim at improving CA production by performing genetic engineering techniques to obtain *S. clavuligerus* recombinant strains with either, reduced capacity for cephamycin C biosynthesis, and/or a CA gene cluster regulatory gene(s) overexpressed. Initially, disruption of the *lat* gene present in cephamycin C gene cluster was proposed as an objective for this dissertation to improve CA production in wild type *S. clavuligerus*. Nevertheless, it was not possible to transform wild type *S. clavuligerus* using pSOK201 plasmid, a plasmid designed to disrupt *Streptomyces* genes by homologous recombination. Conjugation with *Escherichia coli* and protoplast were used as *S. clavuligerus* transformation protocols. However, no recombinant colonies were obtained using either these methods. Apparently, *S. clavuligerus* was unable to replicate pSOK201 plasmid and / or segregate correctly to the daughter cells. All these issues hampered the possibility of acquiring a *S. clavuligerus*-*lat*-gene-disrupted-mutant strain thus impeding to achieve the third objective. Instead, an additional regulatory gene whose effect on CA production is poorly understood, was explored.

To improve CA production, the wild type *S. clavuligerus* ATCC 27064 was genetically engineered either by *claR* or by *orf21* overexpression, via chromosomal integration using the pIB139$^{ermE^*}$ vector. The CA production capabilities of recombinants *S. clavuligerus*/pICLAR and *S. clavuligerus*/pIORF21 improved, compared to the control strains. Still the effect of *orf21* on CA production was conditioned by the medium composition.

Despite the importance of *claR* in CA metabolic pathway, different studies have shown that regardless of the genetic strategy used to overexpress *claR* in a multicopy or integrative plasmid, the yields for *S. clavuligerus* CA production does not exceed 50%. Thus, the co-expression of *claR* with other regulatory genes such as *ccaR* or genes encoding global regulators related to CA biosynthesis, is a primary strategy that would allow the construction of genetically modified strains with higher CA yields by channeling carbon flux towards the late stages of CA metabolic pathway.

Conversely, in this contribution it was possible to identify the role ORF21 might play on CA biosynthesis; as previously described, the effect is settled by the cultivation

environmental conditions. Just as the two-component system composed by *orf22* and *orf23* are fundamental on CA regulation, the sigma factor encoded by *orf21* is another environmental sensing system that *S. clavuligerus* must trigger for CA biosynthesis and cell adaptation. Hence, future RNA-seq studies will allow a general exploration for the cellular metabolism of *S. clavuligerus*/pIORF21 to understand its CA biosynthesis and regulation. Though further research will be needed to elucidate how pIB139 affects cellular metabolism, including transcriptional and metabolic flux analysis, it is important to note that *S. clavuligerus* /pIB139 significantly enhanced CA fermentation processes in ISP medium, increasing CA production 4-fold compared to yields obtained from the wild type strain. Therefore, future studies may find exciting candidate genes for overexpression studies or other genetically engineering techniques.

Finally, the genetic relationships obtained from the pan genome, as well as other biological inferences deduced throughout this research, allow highlighting the importance for new studies on gene expression of global regulators (e.g., BldD) and 5S clavam, as genetic elements poorly studied in CA biosynthesis.

# 7. BIBLIOGRAPHY


1. Ventura, M. *et al.* Genomics of Actinobacteria : Tracing the Evolutionary History of an Ancient Phylum †. **71**, 495–548 (2007).
2. Jiang, Y. *Actinobacteria: Basics and Biotechnological Applications*. (2016).
3. Hopwood, D. A. *Streptomyces in nature and medicine*. (OXFORD, 2007).
4. Goodfellow, M., Kämpfer, P. & Busse, H.-J. *Bergey's Manual of Systematic Bacteriology*. (Springer).
5. Burgo, Y. M. Análisis transcriptómico de cepas derivadas de Streptomyces clavuligerus ATCC 27064: expresión heteróloga de genes y activación de compuestos bioactivos. (Universidad de León, 2016).
6. Elliot, M. A., Buttner, M. J. & Nodwell, J. R. Multicellular Development in Streptomyces. in *Myxobacteria: Multicellularity and Differentiation* (ed. Whitworth, D. E.) 419–439 (2008).
7. Roy, S. Growth , Differentiation and Secondary Metabolism in Streptomyces. *ACTA Sci. Microbiol.* **3**, 56–59 (2020).
8. Tobias E. Kieser; Mervyn J. Bibb, M. J. B. *Practical Streptomyces Genetics*. (2000).
9. Kumar, Y. & Goodfellow, M. Reclassification of Streptomyces hygroscopicus strains as Streptomyces aldersoniae sp . nov ., Streptomyces angustmyceticus sp . nov ., comb . nov ., Streptomyces ascomycinicus sp . nov. *Int. J. Syst. Evol. Microbiol. (2010)*, 769–775 (2010). doi:10.1099/ijs.0.012161-0
10. Zhou, Z. *et al.* Genome mining-directed activation of a silent angucycline biosynthetic gene cluster in Streptomyces chattanoogensis. *ChemBioChem* **16**, 496–502 (2015).
11. Kim, W. *et al.* Comparative Genomics Determines Strain-Dependent Secondary Metabolite Production in Streptomyces venezuelae Strains. *Biomolecules* **10**, 1–18 (2020).
12. Ward, A. C. & Allenby, N. E. . Genome mining for the search and discovery of

bioactive compounds : the Streptomyces paradigm. *FEMS Microbiol. Lett.* **365**, 1–20 (2018).

13. David, M. *et al.* Negative Correlation between Lipid Content and Antibiotic Activity in Streptomyces : General Rule and Exceptions. *Antibiot. MDPI* 280 (2020). doi:10.3390/antibiotics9060280

14. Blin, K. *et al.* AntiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **47**, 81–87 (2019).

15. Ikeda, H. & Satoshi, K. S. Genome mining of the Streptomyces avermitilis genome and development of genome-minimized hosts for heterologous expression of biosynthetic gene clusters. *J Ind Microbiol Biotechnol* 5–9 (2013). doi:10.1007/s10295-013-1327-x

16. Thomson, N. R. *et al.* Complete genome sequence of the model actinomycete Streptomyces. **3**, (2002).

17. Hwang, S. *et al.* Primary transcriptome and translatome analysis determines transcriptional and translational regulatory elements encoded in the Streptomyces clavuligerus genome. *Nucleic Acids Res.* **47**, 6114–6129 (2019).

18. Lee, N. *et al.* Thirty complete Streptomyces genome sequences for mining novel secondary metabolite biosynthetic gene clusters. *Sci. Data* **7**, 1–9 (2020).

19. Cullum, J. *et al.* Genetics of Streptomyces rimosus , the Oxytetracycline Producer. *Microbiol. Mol. Biol. Rev.* **70**, 704–728 (2006).

20. Kim, K. R., Kim, T. J. & Suh, J. W. The Gene Cluster for Spectinomycin Biosynthesis and the Aminoglycoside-Resistance Function of spcM in Streptomyces spectabilis. *Curre Microbiol Microbiol* **57**, 371–374 (2008).

21. Higgens, C. E. & Kastner, R. E. Streptomyces clavuligerus sp. nov., a beta-lactam Antibiotic producer. *Int. J. Syst. Bacteriol.* **21**, 326–331 (1971).

22. Martínez-burgo, Y., Álvarez-álvarez, R., Pérez-redondo, R. & Liras, P. Heterologous expression of Streptomyces clavuligerus ATCC 27064 cephamycin C gene cluster. *J. Biotechnol.* **186**, 21–29 (2014).

23. Ozcengiz, G. & Demain, A. L. Recent advances in the biosynthesis of penicillins,

cephalosporins and clavams and its regulation. *Biotechnol. Adv.* **31**, 287–311 (2013).

24. Santamarta, I. *et al.* Characterization of DNA-binding sequences for CcaR in the cephamycin-clavulanic acid supercluster of Streptomyces clavuligerus. *Mol. Microbiol.* **81**, 968–981 (2011).

25. Bellão, C., Antonio, T., Araujo, M. L. G. . & Badino, A. . Production of clavulanic acid and cephamycin C by Streptomyces clavuligerus under different fed-Batch conditions. **30**, 257–266 (2013).

26. Krenmüller, F. & Summer, H. Process for the production of clavulanic acid salts. 3–6 (2006).

27. Bush, K. & Bradford, P. A. Beta -Lactams and Beta -Lactamase Inhibitors: An Overview. *Cold Spring Harb. Perspect. Med.* 1–22 (2016).

28. Ser, H. *et al.* Fermentation Conditions that Affect Clavulanic Acid Production in Streptomyces clavuligerus : A Systematic Review. *Front. Microbiol.* **7**, (2016).

29. Pinilla, L., Toro, L. F., Avignone-Rossa, C., Peñuela, M. & Ríos-Estepa, R. Streptomyces clavuligerus strain selection for clavulanic acid biosynthesis: A study based on culture composition effects and statistical analysis. *DYNA* **85**, 111–118 (2018).

30. Da Silva Rodriguez, K. C., Biscaro, D. & Otavio, M. Screening of medium constituents for clavulanic acid production by Streptomyces clavuligerus. *Brazilian J. Microbiol.* (2018). doi:10.1016/j.bjm.2018.01.006

31. Rius Núria & Demain, A. L. Lysine Ɛ-Aminotransferase, the Initial Enzyme of Cephalosporin Biosynthesis in Actinomycetes. *J. Microbiol. Biotechnol.* **7**, 95–100 (1997).

32. Gómez-Ríos, D., Ramírez-Malule, H., Neubauer, P., Junne, S. & Ríos-Estepa, R. Degradation Kinetics of Clavulanic Acid in Fermentation Broths at Low Temperatures. *Antibiotics* 12–16 (2019). doi:10.3390/antibiotics8010006

33. Roubos, J. A., Krabben, P., De Laat, W. T. A. M., Babuška, R. & Heijnen, J. J. Clavulanic acid degradation in Streptomyces clavuligerus fed-batch cultivations. *Biotechnol. Prog.* **18**, 451–457 (2002).

34. Rongfeng, L. & Townsend, C. A. Rational strain improvement for enhanced clavulanic acid production by genetic engineering of the glycolytic pathway in Streptomyces clavuligerus. *Metab. Eng.* **8**, 240–252 (2006).

*35.* Paradkar, A. S. *et al.* Applications of Gene Replacement Technology to. *Society* **67**, 2292–2297 (2001).

36. Kurt Kizildoğan, A., Vanli Jaccard, G., Mutlu, A., Sertdemir, İ. & Özcengiz, G. Genetic engineering of an industrial strain of Streptomyces clavuligerus for further enhancement of clavulanic acid production. *Turkish J. Biol.* **41**, 342–353 (2017).

37. Zhihan, Z. & Yanping, W. Application of lat gene disruption to increase the clavulanic acid production of Streptomyces clavuligerus. *J. Mol. Catal. B Enzym.* **43**, 102–107 (2006).

38. Álvarez-álvarez, R., Rodríguez-García, A., Martínez-Burgo, Y., Martín, J. F. & Liras, P. Transcriptional studies on a Streptomyces clavuligerus oppA2 deletion mutant: Nacetylglycyl- clavaminic acid is an intermediate of clavulanic acid biosynthesis. *Appl. Environ. Microbiol.* **84**, 1–17 (2018).

39. López-García, M. T., Santamarta, I. & Liras, P. Morphological differentiation and clavulanic acid formation are affected in a Streptomyces clavuligerus adpA - deleted mutant. *Microbiology* 2354–2365 (2010). doi:10.1099/mic.0.035956-0

40. Song, J. Y., Kim, E. S., Kim, D. W., Jensen, S. E. & Lee, K. J. A gene located downstream of the clavulanic acid gene cluster in Streptomyces clavuligerus ATCC 27064 encodes a putative response regulator that affects clavulanic acid production. *J. Ind. Microbiol. Biotechnol.* **36**, 301–311 (2009).

41. Jnawali, H. N., Liou, K. & Sohng, J. K. Role of σ-factor (orf21) in clavulanic acid production in Streptomyces clavuligerus NRRL3585. *Microbiol. Res.* **166**, 369–379 (2011).

42. Qin, R. *et al.* Improvement of clavulanic acid production in Streptomyces clavuligerus F613-1 by using a claR-neo reporter strategy. *Electron. J. Biotechnol.* **1**, 0–5 (2017).

43. Abusara, N. F. *et al.* Comparative Genomics and Metabolomics Analyses of Clavulanic Acid-Producing Streptomyces Species Provides Insight Into

Specialized Metabolism. *Front. Microbiol.* **10**, 1–17 (2019).

44. Page, A. J. *et al.* Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691–3693 (2015).

45. Tahlan, K., Anders, C. & Jensen, S. E. The paralogous pairs of genes involved in clavulanic acid and clavam metabolite biosynthesis are differently regulated in Streptomyces clavuligerus. *J. Bacteriol.* **186**, 6286–6297 (2004).

46. Tahlan, K., Park, H. U. & Jensen, S. E. Three unlinked gene clusters are involved in clavam metabolite biosynthesis in Streptomyces clavuligerus . *Can. J. Microbiol.* **50**, 803–810 (2005).

47. Kwong, T. *et al.* 5S clavam biosynthesis is controlled by an atypical two-component regulatory system in Streptomyces clavuligerus. *Antimicrob. Agents Chemother.* **56**, 4845–4855 (2012).

48. Tahlan, K. *et al.* 5S Clavam Biosynthetic Genes Are Located in Both the Clavam and Paralog Gene Clusters in Streptomyces clavuligerus. *Chem. Biol.* **14**, 131–142 (2007).

49. Paradkar, A. Clavulanic acid production by Streptomyces clavuligerus : biogenesis , regulation and strain improvement. *J. antibiot* 1–10 (2013). doi:10.1038/ja.2013.26

50. López-Agudelo, V. A., Gómez-Ríos, D. & Ramirez-Malule, H. Clavulanic acid production by streptomyces clavuligerus: Insights from systems biology, strain engineering, and downstream processing. *Antibiotics* **10**, 1–26 (2021).

51. Romero-Rodríguez, A. *et al.* Transcriptomic analysis of a classical model of carbon catabolite regulation in Streptomyces coelicolor. *BMC Microbiol.* 1–16 (2016). doi:10.1186/s12866-016-0690-y

52. Jensen, S. E. *et al.* Five Additional Genes Are Involved in Clavulanic Acid Biosynthesis in Streptomyces clavuligerus. *Antimicrob. Agents Chemother.* **48**, 192–202 (2004).

53. Ramirez-malule, H., Restrepo, A., Cardona, W. & Junne, S. Inversion of the stereochemical configuration ( 3S , 5S ) -clavaminic acid into ( 3R , 5R ) -clavulanic acid : A computationally-assisted approach based on experimental

evidence. *J. Theor. Biol.* 1–11 (2016). doi:10.1016/j.jtbi.2016.01.028

54. de la Fuente, A., Martín, J. F., Rodríguez-García, A. & Liras, P. Two Proteins with Ornithine Acetyltransferase Activity Show Different Functions in Streptomyces clavuligerus : Oat2 Modulates Clavulanic Acid Biosynthesis in Response to Arginine. *J. bacte* **186**, 6501–6507 (2004).

55. Martínez-burgo, Y., Álvarez-álvarez, R., Rodríguez-garcía, A. & Liras, P. The Pathway-Specific Regulator ClaR of Streptomyces clavuligerus Has a Global Effect on the Expression of Genes for Secondary Metabolism and Differentiation. *Appl. Environ. Microbiol.* **81**, 6637–6648 (2015).

56. Kershaw, N. J. *et al.* Structural and mechanistic studies of the orf12 gene product from the clavulanic acid biosynthesis pathway research papers. *Biol. Crystallogr.* 1567–1579 (2013). doi:10.1107/S0907444913011013

57. Arulanantham, H. *et al.* ORF17 from the Clavulanic Acid Biosynthesis Gene Cluster Catalyzes the ATP- dependent Formation of N -Glycyl-clavaminic Acid *. *J. Biol. Chem.* (2016). doi:10.1074/jbc.M507711200

58. Fu, J. *et al.* The CagRS Two-Component System Regulates Clavulanic Acid Metabolism via Multiple Pathways in Streptomyces clavuligerus F613-1. *Front. Microbiol.* **10**, 1–17 (2019).

59. De Bruijn, F. J. *Stress and Environmental Regulation of Gene Expression and Adaptation in Bacteria*. *Stress and Environmental Regulation of Gene Expression and Adaptation in Bacteria* **1**, (2016).

60. Rabyk, M. V, Ostash, B. O. & Fedorenko, V. O. Gene Networks Regulating Secondary Metabolism in Actinomycetes : Pleiotropic Regulators. *Cytol. Genet.* **48**, 55–67 (2014).

61. Chary, V. K., De la Fuente, J. L., Leitão, A. L., Liras, P. & Martín, J. F. Overexpression of the lat gene in Nocardia lactamdurans from strong heterologous promoters results in very high levels of lysine-6-aminotransferase and up to two-fold increase in cephamycin C production. **53**, 282–288 (2000).

62. Xia, H. *et al.* The Application of Regulatory Cascades in Streptomyces: Yield Enhancement and Metabolite Mining. *Front. Microbiol.* **11**, 1–14 (2020).

63. López-García, M. T., Santamarta, I. & Liras, P. Morphological differentiation and clavulanic acid formation are affected in a Streptomyces clavuligerus adpA-deleted mutant. *Microbiology* **156**, 2354–2365 (2010).

64. Pinilla, L., Toro, L. F., Laing, E., Alzate, J. F. & Ríos-Estepa, R. Comparative transcriptome analysis of streptomyces clavuligerus in response to favorable and restrictive nutritional conditions. *Antibiotics* **8**, (2019).

65. Bignell, D. R. D., Tahlan, K., Colvin, K. R., Jensen, S. E. & Leskiw, B. K. Expression of ccaR , Encoding the Positive Activator of Cephamycin C and Clavulanic Acid Production in Streptomyces clavuligerus , Is Dependent on bldG. *Antimicrob. Agents Chemother.* **49**, 1529–1541 (2005).

66. Liras, P., Gomez-Escribano, J. P. & Santamarta, I. Regulatory mechanisms controlling antibiotic production in Streptomyces clavuligerus. *J. Ind. Microbiol. Biotechnol.* **35**, 667–76 (2008).

67. Trepanier, N. K., Jensen, S. E., Alexander, D. C. & Leskiw, B. K. The positive activator of cephamycin C and clavulanic acid production in Streptomyces clavuligerus is mistranslated in a bldA mutant. *Microbiology* **148**, 643–656 (2002).

68. Helmann, J. D. The Extracytoplasmic Function ( ECF ) Sigma Factors. *Adv. Microb. Physiol.* **46**, (2002).

69. Strakova, E., Zikova, A. & Vohradsky, J. Inference of sigma factor controlled networks by using numerical modeling applied to microarray time series data of the germinating prokaryote. *Nucleic Acids Res.* **42**, 748–763 (2014).

70. Den Hengst, C. D. *et al.* Genes essential for morphological development and antibiotic production in Streptomyces coelicolor are targets of BldD during vegetative growth. *Mol. Microbiol.* **78**, 361–379 (2010).

71. Mao, X., Zhou, Z., Hou, X., Guan, W. & Li, Y. Reciprocal Regulation between SigK and Differentiation Programs in Streptomyces coelicolor. *J. Bacteriol.* **191**, 6473–6481 (2009).

72. Lodish, H., Arnold, B., Kaiser, C. A., Monty, K. & Anthony, B. *Molecular cell biology*. (2016).

73. Jnawali, H. N., Oh, T. J., Liou, K., Park, B. C. & Sohng, J. K. A two-component

regulatory system involved in clavulanic acid production. *J. Antibiot. (Tokyo).* **61**, 651–659 (2008).

74. Sharan, A., Soni, P., Nongpiur, R. C., Singla-Pareek, S. L. & Pareek, A. Mapping the 'Two-component system' network in rice. *Sci. Rep.* **7**, 1–13 (2017).

75. Takano, E. γ-Butyrolactones: Streptomyces signalling molecules regulating antibiotic production and differentiation. *Curr. Opin. Microbiol.* **9**, 287–294 (2006).

76. Hyun, S. K. *et al.* Cloning and characterization of a gene encoding the γ-butyrolactone autoregulator receptor from Streptomyces clavuligerus. *Arch. Microbiol.* **182**, 44–50 (2004).

77. Santamarta, I., Pérez-Redondo, R., Lorenzana, L. M., Martín, J. F. & Liras, P. Different proteins bind to the butyrolactone receptor protein ARE sequence located upstream of the regulatory ccaR gene of Streptomyces clavuligerus. *Mol. Microbiol.* **56**, 824–835 (2005).

78. Das, B. & Bhadra, R. K. (p)ppGpp Metabolism and Antimicrobial Resistance in Bacterial Pathogens. *Front. Microbiol.* **11**, (2020).

79. Gomez-Escribano, J. P., Martín, J. F., Hesketh, A., Bibb, M. J. & Liras, P. Streptomyces clavuligerus relA-null mutants overproduce clavulanic acid and cephamycin C: Negative regulation of secondary metabolism by (p)ppGpp. *Microbiology* **154**, 744–755 (2008).

80. Bekker, V., Dodd, A., Brady, D. & Rumbold, K. Tools for metabolic engineering in Streptomyces. *Bioeng. Bugs* **5**, 293–299 (2014).

81. Britain, G. Formation and Reversion of Streptomycete Protoplasts : Cultural Condition and Morphological Study. 389–400 (1974).

82. Myronovskyi, M. & Luzhetskyy, A. Native and engineered promoters in natural product discovery. *Nat. Prod. Rep.* **00**, 1–14 (2016).

83. Wilkinson, C. J. *et al.* Increasing the Efficiency of Heterologous Promoters in Actinomycetes. *Microb. Biotechnol.* **4**, 417–426 (2002).

84. Zotchev, S. *et al.* Identification of a gene cluster for antibacterial polyketide-

derived antibiotic biosynthesis in the nystatin producer Streptomyces noursei ATCC 11455. *Microbiology* 611–619 (2000). doi:10.1099/00221287-146-3-611

85. Muth, G., Nuflbaumer, B., Wohlleben, W. & Piihler, A. A vector system with temperature-sensitive replication for gene disruption and mutational cloning in streptomycetes. 341–348 (1989).

86. Kuzminov, A. Homologous Recombination—Experimental Systems, Analysis and Significance. *HHS Public Access* **4**, 765–770 (2011).

87. Clark, D. P., Pazdernik, N. J. & McGehee, M. R. Recombination. in *Molecular Biology* (eds. Clark, D. P., Pazdernik, N. J. & McGehee, M. R.) 880–895 (Elsevier, 2019). doi:10.1016/B978-0-12-813288-3.00027-6

88. Tong, Y. *et al.* CRISPR–Cas9, CRISPRi and CRISPR-BEST-mediated genetic manipulation in streptomycetes. *Nat. Protoc.* **15**, 2470–2502 (2020).

89. Illing, G. T., Normansell, I. D. & Peberdy, J. F. Protoplast Isolation and Regeneration in Streptomyces clavuligerus. *J. Gen. Microbiol.* 2289–2291 (1989).

90. Shepherd, M. D., Kharel, M. K., Bosserman, M. A. & Rohr, J. Laboratory maintenance of streptomyces species. *Curr. Protoc. Microbiol.* 1–10 (2010). doi:10.1002/9780471729259.mc10e01s18

91. Sambrook, J., E. F. Fritsch, and T. M. *Molecular cloning: A laboratory Manual*. *Molecular Cloning A Laboratory Manual [3rd ed.]* (Cold Spring Harbor Laboratory Press, 1989).

92. Sambrook, J. & Green, M. R. *Molecular cloning: A laboratory Manual*. (John Inglis, 2012).

93. Laboratoire, D., Moleculaire, B. & Received, F. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. **7**, 1513–1523 (1979).

94. Tobias E. Kieser; Mervyn J. Bibb, M. J. B. Practical Streptomyces Genetics.pdf. (2000).

95. Garcia-dominguez, M. & Juan, F. Efficient Plasmid Transformation of the beta-lactam producer Streptomyces clavuligerus. *Appl. Environ. Microbiol.* **53**, 1376–1381 (1987).

96. Ávila, N. N. Efecto de reguladores específicos y reguladores globales en la biosíntesis de antibioticos de Streptomyces clavuligerus. (Universidad de León (España), 2015). doi:10.18002/10612/5176

97. Flett, F., Mersinias, V. & Smith, C. P. High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting Streptomycetes. *FEMS Microbiol. Lett.* **155**, 223–229 (1997).

98. Flávio Azevedo, Humberto Pereira, and B. J. Colony PCR. **1620**, 129–139 (2017).

99. Gomez-Rios, D., Ramírez-malule, H., Neubauer, P., Junne, S. & Ríos-estepa, R. Data of clavulanic acid and clavulanate- imidazole stability at low temperatures. **23**, (2019).

100. Hadjikinova, R., Petkova, N., Hadjikinov, D., Denev, P. & Hrusavov, D. Development and Validation of HPLC-RID method for Determination of Sugars and Polyols. *J. Pharm. Sci. Res.* **9**, 1263–1269 (2017).

101. Merck. TRI Reagent® (Guanidine Thiocyanate & Phenol) Protocol. (2021).

102. Álvarez-Álvarez, R. *et al.* Transcriptomic analysis of Streptomyces clavuligerus ΔccaR:: Tsr: Effects of the cephamycin C-clavulanic acid cluster regulator CcaR on global regulation. *Microb. Biotechnol.* **7**, 221–231 (2014).

103. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real- Time Quantitative PCR and the 2 -ΔΔCt method. *METHODS* **25**, 402–408 (2001).

104. Schmittgen, T. D. & Zakrajsek, B. A. Effect of experimental treatment on housekeeping gene expression : validation by real-time , quantitative RT-PCR. *Biochem. Biophys. methods* **46**, 69–81 (2000).

105. Ronquist, F. *et al.* Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542 (2012).

106. Gelman, A. & Rubin, D. Inference from iterative simulation using multiple sequences. *Stat. Sci.* **10**, 354–363 (1992).

107. Nei, M. & Kumar, S. *Molecular Evolution and Phylogenetics.* (2000).

108. Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation

data matrices. *Bioinformatics* **8**, 275–282 (1992).

109. Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).

110. Bailey, T. L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Second Int. Conf. Intell. Syst. Mol. Biol.* 28–36 (1994).

111. Gupta, S., Stamatoyannopoulos, J. A., Bailey, T. L. & Noble, W. S. Quantifying similarity between motifs. *Geneome Biol.* **8**, (2007).

112. Talà, A. *et al.* Pirin: A novel redox-sensitive modulator of primary and secondary metabolism in Streptomyces. *Metab. Eng.* **48**, 254–268 (2018).

113. Martín, J. F. & Liras, P. The Balance Metabolism Safety Net : Integration of Stress Signals by Interacting Transcriptional Factors in Streptomyces and Related Actinobacteria. **10**, 1–19 (2020).

114. Trinh Viet Hung, Kenji Ishida, Niranjan Parajuli, Kwangkyoung Liou, Hei Chan Lee, and J. K. S. Enhanced Clavulanic acid production in Streptomyces clavuligerus NRRL3585 by overexpression of regulatory genes. *Biotechnol. Bioprocess Eng.* 2–6 (2006).

115. Ortiz, S. C. A., Hokka, C. O. & Badino, A. C. Utilization of soybean derivatives on clavulanic acid production by Streptomyces clavuligerus. *Enzyme Microb. Technol.* **40**, 1071–1077 (2007).

116. Soo, H. *et al.* Improved production of clavulanic acid by reverse engineering and overexpression of the regulatory genes in an industrial Streptomyces clavuligerus strain. *J. Ind. Microbiol. Biotechnol.* **46**, 1205–1215 (2019).

117. Jnawali, H. N., Hei, L. C. & Sohng, J. K. Enhancement of Clavulanic Acid Production by Expressing Regulatory Genes in gap Gene Deletion Mutant of Streptomyces clavuligerus NRRL3585. *Microb. Biotechnol.* **20**, 146–152 (2010).

118. Roth, M., Hoffmeier, C., Geuther, R., Muth, G. & Wohlleben, W. SEGREGATIONAL STABILITY OF pSGS-DERIVED VECTOR PLASMIDS. *Biotechnol. Lett.* **16**, 1225–1230 (1994).

119. Tunca, S. *et al.* An efficient gene transfer system for the pimaricin producer Streptomyces natalensis. *FEMS Microbiol. Lett.* **257**, 312–318 (2006).

120. Paradkar, A. S., Aidoo, K. A. & Jensen, S. E. A pathway-specific transcriptional activator regulates late steps of clavulanic acid biosynthesis in Streptomyces clavuligerus. **27**, 831–843 (1998).

121. Romero-Rodríguez, A., Robledo-Casados, I. & Sánchez, S. An overview on transcriptional regulators in Streptomyces. *Biochim. Biophys. Acta* **1849**, 1017–1039 (2015).

122. Issa, K. H. B., Phan, G. & Broutin, I. Functional Mechanism of the Efflux Pumps Transcription Regulators From Pseudomonas aeruginosa Based on 3D Structures. (2019). doi:10.3389/fmolb.2018.00057

123. Sanches Henao, C. P., Gomez Grimaldos, N. A. & Quintero Diaz, J. C. Production of clavulanic acid by fermentation of Streptomyces clavuligerus : Evaluation of different culture media and mathematical modeling. *DYNA* (2012).

124. Toro, L., Pinilla, L., Avignone-Rossa, C. & Ríos-Estepa, R. An enhanced genome-scale metabolic reconstruction of Streptomyces clavuligerus identifies novel strain improvement strategies. *Bioprocess Biosyst. Eng.* **41**, 657–669 (2018).

125. Casas-pastor, D. *et al.* Expansion and re-classification of the extracytoplasmic function ( ECF ) sigma factor family. *Nucleic Acids Res.* **49**, 986–1005 (2021).

126. Staron, A. *et al.* The third pillar of bacterial signal transduction : classification of the extracytoplasmic function ( ECF ) sigma factor protein family. *Mol. Microbiol.* **74**, 557–581 (2009).

127. Guzina, J. & Djordjevic, M. Promoter Recognition by Extracytoplasmic Function sigma Factors : Analyzing DNA and Protein Interaction Motifs. *J. Bacteriol.* **198**, 1927–1938 (2016).

128. Dekun, G., Youbao, Z. & KeQian, Y. Coordination of glycerol utilization and clavulanic acid biosynthesis to improve clavulanic acid production in Streptomyces clavuligerus. *Sci. CHINA Life Sci. Sci.* **56**, 591–600 (2013).

129. Hindle, Z. & Smith, C. P. Substrate induction and catabolite repression of the Streptomyces coelicoior glycerol operon are mediated through the GylR protein.

*Mol. Microbiol.* **12**, 737–745 (1994).

130. Sonia, B., Pérez-Redondo, R., Koekman, B. & Liras, P. Glycerol Utilization Gene Cluster in Streptomyces clavuligerus. *Appl. Environ. Microbiol.* **75**, 2991–2995 (2009).

131. Kirk, S., Avignone-rossa, C. A. & Bushell, M. E. Growth limiting substrate affects antibiotic production and associated metabolic fluxes in Streptomyces clavuligerus. *Biotechnol. Lett.* **22**, 1803–1809 (2000).

132. Sanchez, S. *et al.* Carbon source regulation of antibiotic production. *J. Antibiot. (Tokyo).* **63**, 442–459 (2010).

133. Yepes-García, J., Caicedo-Montoya, C., Pinilla, L., Toro, L. F. & Ríos-Estepa, R. Morphological differentiation of streptomyces clavuligerus exposed to diverse environmental conditions and its relationship with clavulanic acid biosynthesis. *Processes* **8**, (2020).

134. Manteca, Á. & Yagüe, P. Streptomyces Differentiation in Liquid Cultures as a Trigger of Secondary Metabolism. *Antibiotics* **7**, 41 (2018).

135. Manteca, A., Alvarez, R., Salazar, N. & Yagu, P. Mycelium Differentiation and Antibiotic Production in Submerged Cultures of Streptomyces coelicolor. *Appl. Environ. Microbiol.* **74**, 3877–3886 (2008).

136. Kato, J., Suzuki, A., Yamazaki, H., Ohnishi, Y. & Horinouchi, S. Control by A-Factor of a Metalloendopeptidase Gene Involved in Aerial Mycelium Formation in Streptomyces griseus. *J. Bacteriol.* **184**, 6016–6025 (2002).

137. San Millan, A. *et al.* Integrative analysis of fitness and metabolic effects of plasmids in Pseudomonas aeruginosa PAO1. *ISME J.* **12**, 3014–3024 (2018).

138. Sanchez, M. Quorum sensing in Streptomyces coelicolor: Regulation of the SCB signalling system that controls the synthesis of antibiotics. (University of Groningen, 2016).

139. Combes, P., Till, R., Bee, S. & Smith, M. C. M. The Streptomyces genome contains multiple pseudo-attB sites for the φC31-encoded site-specific recombination system. *J. Bacteriol.* **184**, 5746–5752 (2002).