Biochemical and genetic investigations for the production of β-lactam metabolites from *Streptomyces clavuligerus* ATCC27064 and *Streptomyces pratensis* ATCC33331

by

© Nader Fareid AbuSara

A thesis submitted to the School of Graduate Studies in partial fulfillment for the requirements for the degree of

Doctor of Philosophy

Department of Biology

Memorial University of Newfoundland

December 2021

St. John's, Newfoundland and Labrador

Abstract

Streptomyces are recognized for their ability to produce a wide range of antimicrobial agents. In this thesis, two of the *Streptomyces* bacteria, *Streptomyces clavuligerus* and *Streptomyces pratensis* were the subject of my investigations. *Streptomyces clavuligerus* has the ability to produce a diverse set of β-lactams compounds; the β-lactamase inhibitor clavulanic acid (CA), the 5*S* clavams, and cephamycin C (Ceph-C). I first investigated the molecular features of two genes, *cpe (orf12)* and *orf14,* from the CA biosynthetic gene cluster (BGC) of *S. clavuligerus* and their roles in CA and 5*S* clavam biosynthesis. The two genes are essential for the production of CA since deletion of any of them abolishes CA biosynthesis. The two genes are thought to be involved in converting clavaminic acid into clavaldehyde during the late steps of CA biosynthesis. Different protein variants of CPE and ORF14 were prepared by site-directed mutagenesis and used to investigate their effects on CA and 5*S* clavam production. In addition, the regulatory impact of *cpe* on the transcription level of CA and 5*S* clavam biosynthetic genes was tested in *cpe* deleted and overexpressed mutants compared to the wild-type strain.

Next, a comparative genomic study was conducted for CA and CA-like BGCs between the CA producer (*S. clavuligerus*) and non-producers (*Streptomyces pratensis*, *Saccharomonospora. viridis,* and *Streptomyces* sp. M41). One of the main differences is a large gene, *nocE*, resides within the CA-like BGCs of the non-producers species. Deletion and overexpression of *nocE* in *S. clavuligerus* were achieved, and their effects on physiology, general metabolism, and specialized metabolism were investigated.

In the third part of the study, genomic analysis for BGCs was conducted for the *S. pratensis* ATCC33331 genome, which predicted 27 BGCs. None of the predicted products has been previously reported to be produced by this species. Therefore, I followed the one strain many compounds (OSMAC) approach to investigate the ability of *S. pratensis* to produce specialized metabolites. Our results revealed the production of a bioactive substance (SB) that has antimicrobial activity against *Klebsiella pneumoniae*. Biochemical and genetic procedures were performed to discover the identity of this substance.

Acknowledgment

First of all, I would like to acknowledge my supervisor Dr. Kapil Tahlan for accepting me in his lab, for all the support he has given me, for his patience, guidance, and everything he has taught me over the years. Thanks, Dr. Tahlan, you were always there when I needed you. It is my privilege to work with you. I would also like to thank my supervisory committee members, Dr. Dawn Bignell and Dr. Andrew Lang, for the help, guidance, feedback, and editorial input they have provided me for my projects and thesis. Thanks, professors, for your kind support. I am honored to have you as committee members.

Additionally, I would like to acknowledge the School of Graduate Studies and the Department of Biology at Memorial University of Newfoundland for providing me the opportunity to pursue my degrees. The financial support was gratefully provided at Memorial University. My grateful is extended to the Natural Sciences Engineering Research Council of Canada (NSERC) for their support through the Postgraduate scholarship. I also want to thank Dr. Stefani Egli from CCART for their help with the LC/MS analysis portion of this work, and Elena Salogni from Dr. Ted Miller Lab, who helped me with the statistical analysis.

I would like to express my appreciation to Lab mates. Arshad Shaikh for his immense amount of help in GNPS analysis and for being a good friend in the lab. Dr. Santosh Srivastava for his ideas and technical helps that I learned from him. Dr. Lancy Cheng, for your advice and pleasant conversations that always lift my spirits. Dr. Phoebe Li for be a kind and helpful friend in the lab and sharing some conferences together.

Brandon Piercey for the help with his expertise in bioinformatics. And I won't forget Kelcey King, Fraser Davidson, Jody-Ann Clarke, and Alex Byrne for their kind friendship in the lab.

I want to express my eternal gratitude to my parents, without their encouragement and support, I would not have had the opportunities or abilities that I have now. To my brothers and sisters, thanks for your encouragement and lovely words.

To my soul mate, who stands with me during these years, my lovely wife Raseel, this journey would not have been possible without you, and my words won't thank you enough, but as we say in Arabi "جزاك الله خيرا.", always we stay together. To my lovely kids, Yaman, Zeinadeen, Razan, and Bessan. Your smiles, playing around me, and sleeping on my laps while writing this thesis, especially you Bessan, I won't forget it. I won't forget your funny questions about bacteria and covid-19. It will stay in my memories until I die. You, darlings, are another four PhD degrees that I have to accomplish.

Co-Authorship Statement

The initial study concept of chapter III "Molecular investigations into the role of *orf12* (*cpe*) and *orf14* in clavam metabolite biosynthesis in *Streptomyces clavuligerus*" was designed by K. Tahlan, and the experimental methodology was designed by K. Tahlan, and N. AbuSara. N. AbuSara conducted all the described work except the constructions of *S. clavuligerus* Δ*cpe* mutant and the complementation stains *S. clavuligerus*/Δ*cpe*/pHM11a-*cpe* and *S. clavuligerus*/Δ*cpe*/pSET152-*cpe*, and the CPE variants Δ*cpe*/pSET-*cpe*-S173A, Δ*cpe*/pSET-*cpe*-S234A, Δ*cpe*/pSET-*cpe*-S27A, Δ*cpe*/pSET-*cpe*-L89A, and Δ*cpe*/pSET-*cpe*-S206A which were previously prepared in K. Tahlan Lab. N. AbuSara conducted the data analysis and interpretation. The results figures 3.5 and 3.10 were published in PloS ONE [Srivastava S, King K, AbuSara N, Malayny C, Piercey B, Wilson J, & Tahlan K. (2019). *In vivo* functional analysis of a class A β-lactamase-related protein essential for clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *PloS ONE*, 14(4), e0215960. https://doi.org/10.1371/journal.pone.0215960]. Chapter III was drafted and prepared by N. AbuSara and K. Tahlan.

The initial study concept of chapter IV "The investigation of *nocE* and its impact on the physiology and metabolism of *Streptomyces clavuligerus*" was designed by K. Tahlan. The experimental methodology was conceived and designed by K. Tahlan and N. AbuSara. N. AbuSara conducted all the described work except the LC-MS/MS analyses for the metabolomics analysis which was performed by K. Tahlan at University of California, San Diego, USA. The Molecular networks were generated in the Global

Natural Products Social Molecular Networking (GNPS) by the direction of K. Tahlan and great help from Arshad A. Sheikh. N. AbuSara conducted the data analysis and interpretation The main results from this chapter were published in frontiers in Microbiology [AbuSara N, Piercey B, Moore M, Shaikh A, Nothias L, Srivastava S, Cruz-Morales P, Dorrestein P, Barona-Gómez F, & Tahlan K. (2019). Comparative Genomics and Metabolomics Analyses of Clavulanic Acid-Producing Streptomyces Species Provides Insight Into Specialized Metabolism. *Frontiers in Microbiology*, 10, 2550. https://doi.org/10.3389/fmicb.2019.02550]. Chapter IV was drafted and prepared by N. AbuSara and K. Tahlan.

The initial study concept of chapter V "The production of a bioactive substance by *Streptomyces pratensis* ATCC 33331" was designed by K. Tahlan. The experimental methodology was designed by K. Tahlan and N. AbuSara. N. AbuSara conducted all the described work except the LC-MS/MS for the metabolomics analyses which was performed by K. Tahlan in the University of California, San Diego, USA. The Molecular networks were generated in the GNPS by the direction of K. Tahlan and great help from Arshad A. Sheikh. N. AbuSara conducted the data analysis and interpretation. Chapter V was drafted and prepared by N. AbuSara and K. Tahlan.

Table of Contents

Abstract ..... **i**
Acknowledgment ..... **iii**
Co-Authorship Statement ..... **v**
Table of Contents ..... **vii**
List of Figures ..... **xii**
List of Tables ..... **xvii**
List of Symbols and Abbreviations ..... **xx**
CHAPTER I **1**
Introduction **1**
1.1. The *Streptomyces* 1
1.1.1. Growth and physiology 2
1.1.2. Specialized metabolite (SM) production 4
1.1.3. *Streptomyces clavuligerus* 7
1.2. β-lactams, β-lactamases, and β-lactamase inhibitors 9
1.2.1. β -lactam antibiotics and mechanism of action 9
1.2.2. β-lactamases as enzyme-mediated resistance 13
1.2.3. The β-lactamases inhibitors (βLIs) 18
1.3. Clavams and *Streptomyces clavuligerus* 21
1.3.1. Clavams biosynthesis, the "Early Steps" 23
1.3.2. The "Late Steps" of clavulanic acid biosynthesis 25
1.3.3. The "Late Steps" of 5*S* clavams biosynthesis 30
1.3.4. Regulation of clavam production 32
1.4. Thesis objectives and goals. 39
1.5. Figures and tables 41
1.5.1. Figures 41
1.5.2. Tables 53
CHAPTER II **59**
Materials and Methods **59**
2.1. Bacterial culture conditions and general procedures 59

2.1.1. Bacterial strains, cultivation, and maintenance 59
2.1.2. Preparation and transformation of chemically competent cells 59
2.2. General *Streptomyces* procedures 60
2.2.1. *Streptomyces* strains, cultivation, and maintenance 60
2.2.2. Seeding cultures and fermentation media 60
2.2.3. The *Streptomyces* genomic DNA preparation 61
2.2.4. Spore conjugations 62
2.2.5. *Streptomyces* colony PCR method 63
2.3. Nucleic acids extraction, manipulation, and general procedures 64
2.3.1. Plasmids and DNA manipulation 64
2.3.2. Primers and polymerase chain reaction (PCR) 64
2.3.3 RNA isolation and RT-PCR 65
2.4. Cloning, gene deletion, and preparation of *Streptomyces* strains 66
2.4.1. Preparation of *S. clavuligerus*/pHM11a-*cpe*-His 66
2.4.2. Preparation of *S. clavuligerus*/Δ*orf14*/pHM11a-*orf14* and *S. clavuligerus*/Δ*orf14*/ pSET152-*orf14* 67
2.4.3. Preparation of *S. clavuligerus*/pHM11a-*orf14* variants by site-directed mutagenesis 68
2.4.4. Preparation of the *S. clavuligerus* Δ*nocE* and *ermE*p∗-*nocE* Strains 68
2.4.5. Preparation of *S. pratensis*/Sc-*cpe*, *S. pratensis*/Sc-*orf14*, and *S. pratensis*/Sc-*cpe*-*orf14* 70
2.4.6. Preparation of *S. pratensis* Δ*cas2* by insertional inactivation 70
2.5. *Streptomyces* growth measurements 72
2.5.1. Growth curve measurements of *S. clavuligerus*/Δ*nocE* and *S. clavuligerus*/*ermE*p*-*nocE* in liquid media 72
2.5.2. Growth assessment of *S. clavuligerus*/Δ*nocE* and *S. clavuligerus*/*ermE*p*-*nocE* strains on solid media. 73
2.6. Metabolite detection by bioassays 73
2.6.1 Disc-diffusion bioassays 73
2.6.2. Agar plug diffusion bioassay 75
2.7. Protein crosslinking, extraction, and detection 76
2.7.1. Protein crosslinking 76
2.7.2. Protein extraction and purification using nickel affinity resins 77

2.7.3. Protein analysis by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot. 78
2.8. Specific experiments and analysis for *S. pratensis* project 79
2.8.1. Time-course production of the bioactive substances from *S. pratensis* 79
2.8.2. Fermentation of *S. pratensis* strains in different types of broth media 79
2.8.3. Growing of *K. pneumoniae* on SM agar with *S. pratensis* 80
2.8.4 The bacteriostatic effect assay for the bioactive substances of *S. pratensis* 80
2.8.5. Testing the bacteriostatic and bactericidal effects of some antibiotics 81
2.8.6. RNA isolation from *Streptomyces pratensis* cultured on solid agar media 81
2.9. Liquid chromatography-mass spectrometry (LC-MS and LC-MS/MS) analysis 82
2.10. Metabolomics and molecular networking 83
2.11. MS data annotation and analysis 85
2.12. Bioinformatics analysis 86
2.13. Tables 88
2.14. Supplementary Materials 94
CHAPTER III 100
Molecular investigations into the role of *orf12* (*cpe*) and *orf14* in clavam metabolite biosynthesis in *Streptomyces clavuligerus* 100
3.1 Abstract 100
3.2 Introduction 102
3.3. Objectives 105
3.4. Results and discussion 105
3.4.1. Fermentation of the *S. clavuligerus* Δ*cpe* mutant and complemented strains 105
3.4.2. LC-MS analysis for assessing the production of clavulanic acid and 5*S* clavams in the *S. clavuligerus* Δ*cpe* mutant and complemented strains. 106
3.4.3. Bioassays for detecting 2HMC production in the *S. clavuligerus* Δ*cpe* complementation strains 108
3.4.4. Examining the influence of CPE on the expression of clavulanic acid biosynthesis genes. 109
3.4.5. The influence of CPE on the expression of select genes from the 5*S* clavam and paralogue gene clusters 111
3.4.6. Protein crosslinking and detection of CPE by western blot 115
3.4.7. The effect of different variants of CPE on the production of 2HMC 116

3.4.8. Investigations into the involvement of *orf14* from the CA BGC in clavam metabolite production in *S. clavuligerus*. 121
3.4.9. Examining the effect of different variants of ORF14 on the production of clavulanic acid 122
3.5. Conclusion 125
3.6. Figures and tables 126
3.6.1. Figures 126
3.6.2. Tables 148
3.7. Supplementary materials 149
3.7.1. Supplementary Figures 149
3.7.2. Supplementary Tables 151
CHAPTER IV 153
The investigation of *nocE* and its impact on the physiology and metabolism of *Streptomyces clavuligerus* 153
4.1. Abstract 153
4.2. Introduction 154
4.3. Objectives 155
4.4. Results and Discussion: 157
4.4.1. Comparative study of clavulanic acid BGCs between producers and non-producers. 157
4.4.2. NocE characterization and comparative study 159
4.4.3. *nocE* transcription, deletion, and constitutive expression in *S. clavuligerus* 163
4.4.4. The impact of *nocE* on β-lactam metabolite (CA, Ceph-C, and 5*S* clavams) productions in *S. clavuligerus* 165
4.4.5. The role of *nocE* in *S. clavuligerus* growth 166
4.4.6. The impact of *nocE* on the metabolome of *S. clavuligerus* 168
4.5. Conclusions 171
4.6. Figures and tables 173
4.6.1. Figures 173
4.6.2. Tables 191
4.7. Supplementary Materials 198
4.7.1. Supplementary Figures 198
4.7.2. Supplementary tables 199

CHAPTER V 202

The production of a bioactive substance by *Streptomyces pratensis* ATCC 33331 ..202

5.1. Abstract 202

5.2. Introduction 203

5.3. Objectives 205

5.4. Results and discussion 206

5.4.1. Testing *S. pratensis* for the production of bioactive substances using different types of media 206

5.4.2. Testing the activity range of the bioactive substance 208

5.4.3. Examination for the interspecies interactions to produce the bioactive substance. 210

5.4.4. The bacteri(cidal/ostatic) effect of the bioactive substance 211

5.4.5. Genomic analysis of specialized metabolite BGCs in *S. pratensis* 213

5.4.6. Transcriptional analysis of the CA-like and Carb4550-like gene clusters. 215

5.4.7. Heterologous expression of *S. clavuligerus cpe* (*orf12*) and *orf14* in *S. pratensis* 216

5.4.8. The effect of Sp-*cas2* insertional inactivation on the production of the bioactive substance. 217

5.4.9. Metabolomics analysis of *S. pratensis* (GNPS molecular networking) 219

5.5. Conclusion 220

5.6. Figures and tables 223

5.6.1. Figures 223

5.6.2. Tables 241

5.7. Supplementary materials 243

5.7.1. Supplementary Figures 243

CHAPTER VI 245

6.1. Summary and Perspectives 245

References 249

Appendixes 298

List of Figures

Figure 1.1. Chemical structure of β-lactams……. 42

Figure 1.2. Chemical structures of clinically approved β-lactamase inhibitors…… 42

Figure 1.3. Mode of action of clavulanic acid……. 43

Figure 1.4. The three biosynthesis gene clusters in *S. clavuligerus* related to the production of clavams……. 44

Figure 1.5. Clavulanic acid and *5S* clavams biosynthetic pathways in *S. clavuligerus*. ……..... 46

Figure 1.6. The proposed late steps of clavulanic acid biosynthesis pathway in *S. clavuligerus*. …… 47

Figure 1.7. The proposed late steps of *5S* clavams biosynthetic pathway in *S. clavuligerus.* ……. 49

Figure 1.8 A proposed scheme for the regulation mechanisms in clavulanic acid BGC. ……..…….. 50

Figure 1.9. A proposed scheme for the regulation mechanisms in *5S* clavams BGC. 51

Figure 1.10. A proposed scheme for the regulation mechanisms in paralogue gene cluster. ……..…… 52

Figure 3.1. Clavulanic acid bioassay for *cpe* (*orf12*) complemented *S. clavuligerus* strains. ……... 126

Figure 3.2. Detection of 2-hydroxymethyl clavam (2HMC) in *cpe* overexpression strains. ……..……. 129

Figure 3.3. Detection of clavulanic acid and 2-hydroxymethyl clavam (2HMC) in soy media. ……..…….. 130

Figure 3.4. 5*S* clavam bioactivity assay. ……..……. 131

Figure 3.5. Transcriptional analysis of genes from the clavulanic acid biosynthetic gene cluster (BGC) in wt *S. clavuligerus* and the Δ*cpe* mutant…….... 132

Figure 3.6. Transcriptional analysis of genes from clavulanic acid gene cluster in wt *S. clavuligerus*, Δ*cpe* mutant, and the two complemented strains (Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe*). ……..……... 133

Figure 3.7. Transcriptional analysis of essential genes from 5*S* clavam gene cluster in wt *S. clavuligerus,* Δ*cpe* mutant, and the two complemented strains Δ*cpe*/pSET-*cpe* and Δ*cpe/*pHM-*cpe*. ……..…… 134

Figure 3.8. Transcriptional analysis of essential genes from paralogue gene cluster in wt *S. clavuligerus,* Δ*cpe* mutant, and the two complemented strains Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe*. ……..……. 136

Figure 3.9. Schematic diagram for the proposed mechanism of 2HMC production in starch asparagine media. ……..…… 137

Figure 3.10. CPE detection and crosslinking. …….. 138

Figure 3.11. Clavulanic acid bioassay results for wt, Δ*cpe*, Δ*cpe*/pSET152-*cpe*, and CPE variant strains. ……..…… 139

Figure 3.12. Detection of 2-hydroxymethyl clavam (2HMC) in CPE variant strains. 142

Figure 3.13: Detection of clavulanic acid in *orf14* deletion mutant and complementation strains. ……..……. 145

Figure 3.14. Site-directed mutagenesis on *orf14* of *S. clavuligerus*. ……. 146

Figure 3.15. Clavulanic acid bioassay results for *S. clavuligerus* strains carrying ORF14 variants. ……..…… 147

Figure S3.1. CPE crosslinking optimization. ……..…….. 150

Figure 4.1. Genetic comparison between the CA producing/non-producing species. 174

Figure 4.2. NocE features. …… 176

Figure 4.3. Deletion of *nocE* in *S. clavuligerus*. …… 178

Figure 4.4. Clavulanic acid and cephamycin C bioassays in *nocE* deleted and overexpressed strains. …….. 179

Figure 4.5. Metabolites detection for Clavulanic acid and 2-Hydroxymethylclavam (2HMC). …… 181

Figure 4.6. Cellular growth curves of Δ*nocE* (green squares) and *ermE*p-nocE* (orange triangles) mutants in compared to wt *S. clavuligerus* (blue diamonds)…… 182

Figure 4.7: Growth characteristics of the wt *S. clavuligerus*, Δ*nocE,* and *ermE*p∗-*nocE* strains on different solid media. ……. 185

Figure 4.8. Comparative metabolomics of the wt *S. clavuligerus*, Δ*nocE,* and *ermE*p∗-*nocE* strains grown on two different media SA and TSA-S1%. …… 186

Figure 4.9. Metabolomics analysis for the main groups of molecules in wt *S. clavuligerus*, Δ*nocE,* and *ermE*p*-*nocE*. …….. 188

Figure 4.10. Metabolic network constructed using *S. clavuligerus* wt, Δ*nocE*, and *ermE*p∗-*nocE* strains culture extracts. …….. 190

Figure S4.1. Plasmid map of pIJ8668-*ermE*p*-*nocE*. …….. 198

Figure 5.1. The production of bioactive substances by *S. pratensis* grown on different types of media. …… 224

Figure 5.2. The activity range of *S. pratensis* bioactive substance. …….. 225

Figure 5.3. Physical interaction with *K. pneumoniae* is not necessary for bioactive substance production by *S. pratensis.* …….. 226

Figure 5.4. The bacteriostatic effect of *S. pratensis* bioactive substance. ……. 227

Figure 5.5. Antimicrobial susceptibility test for bacteriostatic and bactericidal antibiotics. ……. 228

Figure 5.6. Organization and comparative gene clusters of clavulanic acid and carbapenem MM4550 BGCs in *Streptomyces*. ……. 230

Figure 5.7. Transcriptional analysis of genes from CA-like and Carb4550-like BGCs. ……..…….. 232

Figure 5.8. The production of a bioactive substance for *S. pratensis* strains with heterologous expression for two genes (*cpe* and *orf14*) from *S. clavuligerus*. …… 233

Figure 5.9. *Sp-cas2* cloning into pIJ773 for insertional inactivation. …….. 235

Figure 5.10. *S. pratensis cas2* insertional inactivation. …….. 237

Figure 5.11. Metabolomics analysis for the main groups of molecules in wt *S. pratensis*. ……. 239

Figure 5.12. Metabolic network constructed using wt *S. pratensis* strains culture extracts. ……..…… 240

Figure S5.1. Sp-*carE* cloning into pIJ773 for insertional inactivation. ……. 244

List of Tables

Table 1.1. β-Lactamase Inhibitors' characteristics and their Combinations. …….. 53

Table 1.2. Clavulanic acid (CA) and cephamycin C (Ceph-C) production in *S. clavuligerus* mutants with defects in genes from the clavulanic acid biosynthetic gene cluster. ……. 55

Table 1.3. 5*S* clavams production in *S. clavuligerus* mutants with defects in genes from clavam BGC and paralogue BGC. …… 57

Table 2.1. Stock and experimental concentrations of antibiotics used throughout this study. ……. 88

Table 2.2. Bacterial strains were used in this study. …… 89

Table 2.3. Plasmids and DNA constructs were used in this study…… 92

Table S2.1. Sequences of oligonucleotide primers used in the current study and their details. ……..... 94

Table 3.1. The LC/MS assessment of clavulanic acid and 2-hydroxymethylclavam production in *S. clavuligerus* and Δ*cpe* mutants expressing different variants of CPE grown in SA. ……... 148

Table 3.2. Clavulanic acid bioassay for *orf14* deleted mutant and complemented strains grown in SA and SM media. ……..... 148

Table S3.1. Clavulanic acid bioassay for two *cpe* (*orf12)* complemented strains grew in SA media at two time points.…… 151

Table S3.2. The ion abundance values for the corresponding extracted ion chromatogram for imidazole-derivatized CA $[M+H]^+$ (m/z = 224), the fragmented product $[M\text{-imidazole}]^+$ (m/z = 156) and the 2HMC fragmented product $[M\text{-imidazole}]^+$ (m/z = 144) for SA supernatant samples. ……..... 151

Table S3.3. The assessment of clavulanic acid production in *S. clavuligerus* and Δ*cpe* mutants expressing different variants of CPE grown in SA and SM……. 152

Table S3.4. Clavulanic acid bioassays measurements for *orf14* variants grown in SM media at two-time points.…… 152

Table 4.1. Comparative analysis of NocE in *S. clavuligerus* with other orthologues in Actinomycetes species that predicted to have β-lactam antibiotics biosynthesis gene clusters. ……. 191

Table 4.2. The secretory signal peptide prediction for NocE homologues from Actinomycetes species. …… 193

Table 4.3. Specialized metabolites (SMs) detected with high confidence in wt *S. clavuligerus*, Δ*nocE,* and *ermE*p∗*-nocE* strains using MS-based metabolomics and GNPS analysis. …… 194

Table S4.1. The predicted protein products of the genes surrounding *nocE* 199

(upstream and downstream) in CA producers species *S. clavuligerus, S. jumonjinensis,* and *S. katsurahamanus,* and CA-like BGC holding strain *Streptomyces* sp. M41. ……

Table S4.2. The diameters (in mm) of zones of growth inhibition of *K. pneumoniae* by CA containing supernatants from triplicates cultures of Δ*nocE*, *ermE*p*-*nocE* and wt *S. clavuligerus* strains in two types of media SA and SM…. 200

Table S4.3. The diameter (in mm) of zones of growth inhibition of *E. coli* ESS from cephamycin C containing supernatants from triplicates cultures of Δ*nocE*, *ermE*p*-*nocE* and Wt *S. clavuligerus* strains in two types of media SM and TSB-S…….. 200

Table S4.4. The superclasses of molecules detected by in silico annotation in GNPS, for the extracts of wt *S. clavuligerus*, Δ*nocE*, and *ermE*p**nocE* strains….. 201

Table 5.1: Specialized metabolite (SM) biosynthetic gene clusters predicted in the genome of *S. pratensis* ATCC33331 using antiSMASH 5.0…… 241

List of Symbols and Abbreviations

Δ	deletion
AGCA	*N*-acetyl-glycyl-clavaminic acid
Ala	L-alanine
Amp	ampicillin
ANOVA	analysis of variance
Apr	apramycin
ATCC	American type culture collection
ATP	adenosine triphosphate
βLIs	β-Lactamase inhibitors
BGC	biosynthesis gene cluster
BES	beef extract-Starch media
BLASTP	protein basic local alignment search
BS	bioactive substance
CA	clavulanic acid
CarMM	carbapenem MM4550
cDNA	complementary DNA
Ceph-C	cephamycin C
CoA	coenzyme A
Cpe	cephalosporin esterase gene
DAD	diode array detector
DEPC	diethyl pyrocarbonate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EIC	extracted ion chromatogram
FDA	Food and Drug Administration
2FMC	2-Formyloxymethylclavam
GNPS	Global natural products social molecular networking
His	histidine
2HMC	2-hydroxymethylclavam
HPLC	high pressure liquid chromatography
HRP	horseradish peroxidase
Hyg	hygromycin

ISP-4	International *Streptomyces* project medium 4
KDa	Kilo dalton
LC-MS	liquid chromatography-mass spectrometry
MEGA	molecular evolutionary genetics analysis
MS/MS	tandem mass spectrometry
MSF	Mannitol-Soy flour media
NAP	Network annotation propagation tool
NCBI	National Center for Biotechnology Information
NGCA	*N*-glycyl-clavaminic acid
Ni-NTA	nickel charged affinity resin
NPs	Natural products
NRP	non-ribosomal peptide
OD_{600}	optical density at 600 nm
ORFs	open reading frames
*ori*T	origin of transfer
OSMAC	one strain many compounds
PBP	penicillin binding protein
PBS	phosphate buffer saline
PCR	polymerase chain reaction
PenG	penicillin G
PKS	polyketide synthase
RBS	Ribosome binding site
RNA	Ribonucleic acid
rpm	revolutions per minute
RT-PCR	reverse transcription-PCR
SA	Starch asparagine media
SARP	*Streptomyces* antibiotic regulatory protein
Sc	*Streptomyces clavuligerus*
Scat	*Streptomyces cattleya*
SDM	Site directed mutagenesis
SM	Soy media
SMs	Specialized metabolites
Sp	*Streptomyces pratensis*
sp.	species (singular)
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis

TBE	Tris Borate EDTA buffer
TBO	Tomato Paste-Baby Oatmeal media
TCS	Two components system
Thn	Thienamycin
TOF	time-of-flight
TSA	Tryptic soy agar
TSB	Tryptic soy broth
TSB-S	Tryptic soy broth-starch 1%
UV	ultraviolet
v/v	volume/volume
wt	wild type
w/v	weight/volume
ZOI	Zone of inhibition

CHAPTER I

Introduction

1.1. The *Streptomyces*

The genus *Streptomyces* is Gram-positive aerobic bacteria from the phylum Actinobacteria with over 600 species described so far (Labeda et al., 2012; van der Meij et al., 2017). The *Streptomyces* are found in soil environments as well as in marine and freshwater ecosystems. They undergo filamentous growth and sporulation and are renowned for their capacity to produce a vast array of natural products (Chater, 2016; Hopwood, 1999; van der Meij et al., 2017).

Genetically, *Streptomyces* have some striking features compared to other prokaryotes. They harbour a single large linear chromosome that is approximately 8 - 12 Mb in size and has telomere-like structures (Nett et al., 2009; Wang et al., 2010). The chromosome consists of a central conserved 'core' region of 5 – 7 Mb that contains most essential function genes, and the core is surrounded by 'arms' that vary widely among *Streptomyces* in their gene content (Baltz, 2019; Nett et al., 2009). On average *Streptomyces* chromosomes contain more open reading frames than the genome of the eukaryote *Saccharomyces cerevisiae*, and they harbour double the number of genes found on the *Escherichia coli* (Paradkar et al., 2003). In addition, some *Streptomyces* contain circular and/or linear plasmids, some of which can be very large in size (Hopwood, 1999; Paradkar et al., 2003). *Streptomyces* DNA has a G/C content of approximately 69 - 73 %, which is higher than that of other bacteria such as *E. coli* (which is ~ 50%) (de Lima Procópio et al., 2012).

1.1.1. Growth and physiology

Streptomyces are unusual among prokaryotes in that they follow a complex life cycle that classically encompasses three morphologically distinct growth stages. When nutrients and conditions are favorable, the *Streptomyces* life cycle begins when a uni-genomic dormant spore gives rise to one or two germ tubes, which grow by tip extension to form hyphae (Chater, 2006; Flärdh & Buttner, 2009). The growing of more hyphae and emerging additional branches from the lateral walls ultimately results in a dense network of filaments termed the vegetative (or substrate) mycelium (Flärdh & Buttner, 2009; Kieser et al., 2000). This branching mycelium helps the *Streptomyces* colony to anchor in its substrate and absorb nutrients. During vegetative growth, the chromosomes replicate resulting in multi-genomic hyphae that are delimited by occasional cross-walls (Jones & Elliot, 2018; Kois-Ostrowska et al., 2016). When the nutrients have been used up, the vegetative mycelium cells start to excrete enzymes for the degradation of insoluble nutrients such as chitin or cellulose for use as nutrient sources in the second growth stage (Chater, 2006). In addition to the secretion of extracellular enzymes, three more critical events happen, the onset of secondary (or specialized) metabolism (a metabolism that produces compounds that confer a selective advantage to the bacteria such as antimicrobial compounds in order to protect their nutrients against competitors), initiation of the break down of a proportion of the vegetative mycelium, and initiation of aerial hyphal formation (Chater, 2006; Kieser et al., 2000). In the second stage, the formation of nonbranching aerial hyphae occurs, which rise into the air away from the vegetative mycelium and give a white and fuzzy appearance to the colony surface (Flärdh & Buttner, 2009). The nonbranching aerial growth is under the regulation of several *bld* genes,

named for the ‘bald’ phenotype of the mutants lacking the fuzzy aerial hyphae (Chater, 2016; Plaskitt & Chater, 1995). The aerial hyphae become divided by a developmentally controlled form of cell division into long chains of pre-spore compartments (Flärdh & Buttner, 2009), which then mature to form dormant, grey pigmented spores with each spore containing a single chromosome (Flärdh & Buttner, 2009; Jakimowicz & Van Wezel, 2012). The formation of spores from aerial hyphae is regulated by a number of *whi* (for 'white') genes, named for the fact that *whi* mutants fail to sporulate and produce the grey pigment (Chater, 1993).

Recently, in some *Streptomyces* species, a non-classical stage was noticed in the life cycle. The 'exploration" stage is a different mode of growth in which explorer cells grow as nonbranching vegetative hyphae, in contrast to the branching vegetative mycelium seen in the classic *Streptomyces* life cycle (Jones et al., 2017). This exploratory growth happens in response to fungi and other metabolic triggers (a condition that resembles the biological communities in soil). In this growth mode, the colonies grow rapidly at a rate more than ten times that of classical vegetative growth and spread over the agar and other solid obstacles (Jones & Elliot, 2018). The known *bld* and *whi* genes are not directly involved in the exploration process, however, the regulators and the genes required for exploration are still under investigation (Jones & Elliot, 2018; Jones et al., 2017).

Each stage of the life cycle has its characteristics where different extracellular enzymes (such as chitinases, cellulases, proteases etc.) and specialized metabolites are produced to help the bacteria complete its developmental program. However, these by-

products would eventually have significant industrial and medical applications for human welfare today.

1.1.2. Specialized metabolite (SM) production

One of the most important features of *Streptomyces* bacteria is their metabolic versatility in the production of chemically diverse and biologically active natural products (NPs), which are termed specialized metabolites (SMs), or in many reports as "secondary metabolites" (Baltz, 2019; Traxler & Kolter, 2015). The structural chemistry of SMs is varied and based on a number of different backbone structures, e.g., polyketides, peptides, lipids, steroids, β-lactams, and pyrroles (van Wezel & McDowall, 2011). However, many SMs and their derivatives have important applications in human and animal health, agriculture, and biotechnology. For example, some bacterial SMs are used as antibiotics (e.g., vancomycin, tetracycline), antifungals (amphotericin B), antivirals (virantmycin), anticancer (mitomycin C), antiparasitic (ivermectin), immunosuppressive (rapamycin), or crop protection agents (e.g., ziracin, dalbavancin, spinosyns) (Baltz, 2008; Chater, 2016; Quinn et al., 2020; Traxler & Kolter, 2015).

The production of most SMs is species-specific and generally is influenced and regulated by many factors, such as the external environment, nutrition (including carbon and nitrogen sources), growth rate, and regulatory networks (Bibb, 2005; van Wezel & McDowall, 2011). The function of SMs in the natural environment is still not clear in many instances. However, production of SMs (especially antimicrobial agents) can enhance *Streptomyces* competition with other microorganisms in their habitats (van der Meij et al., 2017). In addition, SMs have important roles in cell signaling and

communication with other microorganisms (Traxler & Kolter, 2015). The exchanging of metabolites in microbial communities significantly impacts growth, development, and SMs production by community members (van der Meij et al., 2017; van Wezel & McDowall, 2011).

Besides interactions with other microorganisms, *Streptomyces* species interact extensively with higher organisms, and they can forge a mutualistic relationship with their host and protect them from infectious pathogens (Chater, 2016). For instance, some *Streptomyces* species form a symbiotic association with plants, where they can use the exudates produced by the plants as a nutrient source, while the antimicrobial metabolites produced by the bacteria inhibit many phytopathogens and suppress plant diseases (Seipke et al., 2012; Taechowisan et al., 2003). Furthermore, some *Streptomyces* may enhance the growth of their plant host by producing the growth hormone auxin, which is vital for the root's growth and development (Seipke et al., 2012). In a similar way, *Streptomyces* can also form a mutualistic symbiosis with insects, such as Southern pine beetles (*Dendroctonus frontalis*), beewolf wasps (*Philanthus* spp.), and leaf-cutter ants (*Acromyrmex octospinosus*) where *Streptomyces* have been found defending against fungal infections (Kaltenpoth, 2009; Seipke et al., 2011, 2012). Recently, in the marine ecosystem, *Streptomyces* were found abundant within the microbiota of marine organisms, such as sponges, seaweed, and marine cone snails (Khan et al., 2011; Peraud et al., 2009). Several *Streptomyces* strains were successfully isolated from these marine organisms, and they were found producing SMs with neurological, antibacterial, and antifungal bioactivities. However, the *in vivo* functions are still unknown (Peraud et al., 2009; Seipke et al., 2012; van der Meij et al., 2017). Seemingly, *Streptomyces* are

desirable guests to many other organisms or environments, and this is often related to their ability to produce a variety of SMs to fight off pathogenic bacteria or fungi, or extracellular enzymes to degrade organic compounds (Seipke et al., 2012; van der Meij et al., 2017). On the other hand, it is conceivable that the chemical diversity of these SMs produced by *Streptomyces* has evolved as a result of interactions with other organisms in highly diverse environments (Chater, 2016; Traxler & Kolter, 2015).

In *Streptomyces*, the genes responsible for the biosynthesis of individual specialized metabolites are usually arranged in clusters (biosynthesis gene clusters, BGCs) that vary in size from a few to over 100 kb and include several operons. Most of these BGCs are located on chromosomal DNA and infrequently on plasmid DNA (Bibb, 2005; Nett et al., 2009). Such clusters typically consist of genes encoding the biosynthetic enzymes as well as genes that regulate their transcription, the metabolite exportation, and resistance genes in the case of antibacterial metabolites (Liu et al., 2013). On average, each *Streptomyces* species harbours ~30 distinct BGCs for producing SMs, but only a few are expressed under typical laboratory conditions, while many of them are silent or poorly expressed (Xia et al., 2020). By using advanced high-throughput genome sequencing methods and various *in silico* genome mining, *Streptomyces* have been found to possess an even more significant number of uncharacterized SMs than previously estimated (Watve et al., 2001; Xia et al., 2020). The activation of these cryptic or silent clusters has been the target of research in recent years, and several successful approaches were conducted for that purpose, such as the use of metabolic signals, manipulation of pathway-specific and pleiotropic regulators, heterologous expression, microbial co-

culturing, CRISPR-Cas9 strategy and others (Baltz, 2016; Onaka, 2017; Ren et al., 2017; Zhang et al., 2017; Kong et al., 2019; Xu & Wright, 2019).

1.1.3. *Streptomyces clavuligerus*

One of the most important industrial streptomycete species is *Streptomyces clavuligerus*, which produces a wide variety of β-lactam compounds. The species was first isolated in 1971 by the American pharmaceutical company Eli Lilly and Co. from a South American soil sample (Higgens & Kastner, 1971). During the first screening test, *S. clavuligerus* was found to produce penicillin N, cephamycin C and deacetoxycephalosporin C (Higgens et al., 1974; Nagarajan et al., 1971). A few years later, *S. clavuligerus* was screened for production of β-lactamase inhibitors and found to produce clavulanic acid (CA), the first naturally occurring β-lactamase inhibitor to be fully characterized (Reading & Cole, 1977). Currently, both cephamycin C and CA are used in the clinic to treat many infectious diseases. In addition, *S. clavuligerus* produces a variety of 5*S* clavams, which are β-lactam compounds belonging to the clavam group and containing a 5*S* configuration structure. The 5*S* clavams include clavam-2-carboxylate (C2C), 2-hydroxymethylclavam (2HMC), 2-formyloxymethylclavam (2FMC), and alanylclavam, none of which exhibit β-lactamase inhibition activity, but instead they have some antifungal and bacteriostatic effects (Brown et al., 1979; Pruess & Kellett, 1983; Jensen, 2012). Also, at least three additional SMs with bioactivity have been reported to be produced by *S. clavuligerus*: the tunicamycin-related complex of antibiotics (Kenig & Reading, 1979), the tacrolimus macrolide immunosuppressant (Kim & Park, 2008), and the dithiolopyrrolone class antibiotic holomycin (Kenig & Reading, 1979; Li & Walsh,

2010). Recently, another dithiolopyrrolone compound predicted to be N-propionylholothin was detected in extracts from *S. clavuligerus* strains lacking the giant linear plasmid pSCL4 (Álvarez-Álvarez et al., 2017).

Metabolomics analysis also revealed the presence of certain plant-associated SMs in *S. clavuligerus* extracts. The citrus flavonoid naringenin and the genes involved in the production of this metabolite were identified (Álvarez-Álvarez et al., 2015). Naringenin exhibits antibacterial, antifungal, and anticancer activities (Kanno et al., 2005; Rauha et al., 2000), and its production by a bacterium was unexpected since it was previously isolated from plants only (Álvarez-Álvarez et al., 2015). However, *S. clavuligerus* possess many terpene-like BGCs of unknown function, potentially involved in the biosynthesis of more plant-associated metabolites (Hwang et al., 2019; Medema et al., 2010).

S. clavuligerus contains a 6.75-Mbp linear chromosome and a 1.8-Mbp linear mega-plasmid (pSCL4), which is one of the largest plasmids ever identified (Hwang et al., 2019; Medema et al., 2010). Also, *S. clavuligerus* contains three plasmids (pSCL1 – 3) with sizes 11.7, 120 and 430 kb, respectively (Netolitzky et al., 1995). *S. clavuligerus* was predicted to have 43 BGCs, 26 on the chromosome and 17 on the giant plasmid pSCL4 (AbuSara et al., 2019). However, with a recent high-quality genome sequence and annotation, *S. clavuligerus* is now predicted to have a total of 58 BGCs; among them, 30 and 28 BGCs were found on the chromosome and the pSCL4 giant plasmid, respectively (Hwang et al., 2019). The relatively high number of BGCs reflects this species' ability to produce a wide variety of specialized metabolites.

1.2. β-lactams, β-lactamases, and β-lactamase inhibitors

1.2.1. β -lactam antibiotics and mechanism of action

Penicillin, a β-lactam, was the first antibiotic discovered by Alexander Fleming in 1928, and the ninety years of steady progress following this discovery makes the β-lactams one of the most successful groups of natural products in medicinal application and chemotherapy. This family of SMs is critically important as pharmaceutical agents to combat bacterial infections, and it makes up 65% of the total antibiotics in the market (Hamed et al., 2013). In addition to fungi, the *Streptomyces* group of bacteria is a key natural source of β-lactam compounds (Demain & Elander, 1999). The β-lactam family of antibiotics includes penicillins, cephalosporins, carbapenems, monocyclic β-lactams, and clavams. From a biochemical point of view, these compounds have a common feature, which is the 3-carbon and 1-nitrogen ring (β-lactam ring) that is highly reactive (Figure 1.1) (Tahlan & Jensen, 2013).

The β-lactam antibiotics exhibit their bactericidal effects by inhibiting enzymes involved in cell wall formation in both Gram-positive and Gram-negative bacteria. Peptidoglycan, which is primarily responsible for the strength of the cell wall, is a layer composed of two alternating sugar derivatives, *N*-acetylmuramic acid (NAM) and *N*-acetylglucosamine (NAG), and short oligopeptides consisting of L-alanine, D-alanine, D-glutamic acid, and either lysine or diaminopimelic acid (Sauvage et al., 2008). At the final stage of cell wall synthesis, the adjacent glycan chains are cross-linked using oligopeptides. The oligopeptide attaches to each NAM unit, and the cross-linking of two D-alanine–D-alanine NAM oligopeptides is catalyzed by transpeptidase enzymes in a

reaction called transpeptidation. These peptidoglycan cross-linkages confer the rigidity of the cell wall (Bush & Bradford, 2016).

The β-lactam ring of the antibiotics is sterically similar to the D-alanine–D-alanine of the NAM dipeptide, and the transpeptidases "mistakenly" binds penicillin or other β-lactam antibiotics during cell wall synthesis. Thus, these transpeptidases are also known as penicillin-binding proteins (PBPs). The binding results in irreversible acylation of the active site serine in PBPs, rendering the enzyme incapable of catalyzing further transpeptidation reactions. As a consequence, the cell wall synthesis ceases, and the cells lyse and die due to the internal osmotic pressure (Bush & Bradford, 2016; Tooke et al., 2019).

Penicillin was the first β-lactam antibiotic discovered and was isolated from a rare variant of *Penicillium chrysogenum* or *P. notatum* (formerly) in 1928 (Demain and Elander, 1999). The penicillin core contains a 6-animopenicillanic acid nucleus (β-lactam plus thiazolidine) linked to other side groups (Figure 1.1). The class of molecules includes natural penicillins (i.e., penicillin G, K, N), β-lactamase-resistant agents (i.e., methicillin and nafcillin), aminopenicillins (i.e., ampicillin and amoxicillin), carboxypenicillins (i.e., carbenicillin and ticarcillin), and ureidopenicillins (i.e., piperacillin and azlocillin) (Bush & Bradford, 2016).

The cephalosporins were first characterized in 1955 (Newton & Abraham, 1955) and are known to be produced by both fungi (e.g., *Acremonium chrysogenum*) and bacteria. Cephalosporins contain a 7-aminocephalosporanic acid nucleus and side chain containing 3,6–dihydro- 2 H-1,3- thiazane rings (Figure 1.1A). Cephalosporins are traditionally divided into five classes. For example, cephalexin (1st generation), cefoxitin

and cefotetan (2nd generation), cefotaxime (3rd generation), ceftazidime (4th generation), and ceftaroline (5th generation) (Demain and Elander, 1999; Bush and Bradford, 2016).

The carbapenems have a typical β-lactam ring fused to a five-membered ring (Figure 1.1A). They differ in that the five-membered ring is unsaturated and contains a carbon atom replacing the sulfur. However, the sulfur is present at the C-2 side chain in most of the carbapenems (Papp-Wallace et al., 2011). Thienamycin was the first carbapenem identified in the mid-1970s (Kahan et al., 1979), and it serves as the parent or model compound for other carbapenems discovered since then. A number of other carbapenems have been identified such as imipenem, meropenem, doripenem and ertapenem (Bush & Bradford, 2016). The carbapenems antibiotics exhibit potent and broad antibacterial activity, and some show an *in-vitro* β-lactamase inhibitory activity (Breilh et al., 2013; Papp-Wallace et al., 2011).

The monocyclic β-lactams are compounds which do not contain a ring fused to the β-lactam ring. The two naturally occurring subfamilies of monocyclic β-lactams are the nocardicins and the monobactams (Hamed et al., 2013). Nocardicin A was the first to be discovered by strains of *Nocardia* and *Streptomyces* (Aoki et al., 1976), and it has weak to moderate activity against Gram-negative bacteria with low activity against Gram-positive organisms (Demain and Elander 1999). Monobactams are produced by a large number of unicellular bacteria such as *Chromobacterium* and *Pseudomonas*. Aztreonam, a monobactam derivative antibiotic, is the only monobactam to gain regulatory approval for clinical use against *Pseudomonas aeruginosa* and aerobic enteric bacteria (Bush & Bradford, 2016).

Clavams are compounds that have a bicyclic nucleus with a β-lactam ring fused to an oxazolidine ring (Figure 1.1) (Demain and Elander, 1999). Clavulanic acid is a known model of clavams and the only one among them that exhibits β-lactamase inhibitory activity and is used in clinical treatments (see section 1.2.3) (Rolinson, 1991). Other examples of clavams, such as clavamycin, valclavam, clavaminic acid, alanylclavam, and 2-hydroxymethylclavam (2HMC), are naturally produced by several bacteria species and show some antibacterial or antifungal properties (Pruess and Kellett 1983; Jensen and Paradkar, 1999).

1.2.1.1. Mechanisms of resistance to β-lactams antibiotics

The extensive abuse of antibiotics has led to the emergence and spreading of β-lactam antibiotic-resistant pathogens, which are a growing public health concern (Geddes et al., 2007). Resistance can occur by multiple molecular mechanisms in both Gram-positive and Gram-negative bacteria, some of which include the following mechanisms:

(i) Modification of the target (mutation or expression of alternative PBPs). This can lower the ability of β-lactam antibiotics to bind to PBPs in the bacterial cell wall, and subsequently increase resistance to them, such as those seen in PBPs of Methicillin-resistant *Staphylococcus aureus* (MRSA), *Streptococcus pneumoniae* and *Neisseria* spp. which through natural transformation and recombination with DNA from other organisms, have acquired highly resistant and low affinity PBPs (Bowler et al., 1994; Drawz & Bonomo, 2010; Page, 2007).

(ii) Reduction in cell permeability through downregulation of porin channels required for β-lactam entry. In order to enter the cell, β-lactams must diffuse through

porin channels in the outer membrane of the Gram-negative bacterial cell walls. Some Enterobacteriaceae (e.g., *Enterobacter* spp., *Klebsiella pneumoniae*, and *Escherichia coli*) exhibit resistance to carbapenem antibiotics based on loss of these porin proteins, leading to reduced antibiotic permeability or their complete exclusion in some cases (Doumith et al., 2009; Hopkins & Towner, 1990; Jacoby et al., 2004). Also, the loss of porin channels in clinical isolates of multidrug-resistant *Acinetobacter baumannii* confer this bacteria resistance against imipenem and meropenem (Mussi et al., 2005). It is notable that sometimes the loss of porin proteins alone is not enough for the resistance phenotype, and typically it is due to a combination with other resistance mechanisms (Drawz & Bonomo, 2010).

(iii) Over-expression of efflux systems to expel drugs from the periplasm to outside the cell. These pumps are responsible for multidrug resistance in many Gram-negative pathogens. For example, in *Pseudomonas aeruginosa* and *Acinetobacter* spp. the upregulation of efflux pumps, in concert with low membrane permeability, confers resistance to penicillins and cephalosporins as well as other antibiotics (Zhang & Poole, 2000; Poole, 2004).

(iv) Antibiotic modification or inactivation due to the production of β-lactamase enzymes is the most common and important mechanism of resistance in bacteria and will be the focus in the next section.

1.2.2. β-lactamases as enzyme-mediated resistance

The primary mechanism of resistance against β-lactams in Gram-negative and many Gram-positive pathogens is the production of β-lactamases. The enzymes hydrolyze

the β-lactam ring by breaking the amide bond, resulting in the inability of antibiotics to bind to the PBPs in the cell wall and kill bacteria (Tyers & Wright, 2019).

The first β-lactamase enzyme was identified in 1940 from *E. coli* before the first administration of penicillin in clinical treatment (Abraham & Chain, 1940). Prior to the end of the 1950s, the enzyme, which was called "penicillinase", was isolated from many other bacterial samples such as *Staphylococcus aureus* (Rolinson, 1991). To date over 4300 β-lactamase enzymes have been identified and characterized at various levels and databases have been established for their curation (www.bldb.eu) (Naas et al., 2017; Tooke et al., 2019). Two major schemes exist for classifying β-lactamases. In the first, β-lactamases are classified into four molecular classes (A to D) based on amino acid sequence homology (Ambler, 1980), whereas the second method employs a functional approach and β-lactamases are classified into three groups (1 to 3) based on functionality, substrates, and inhibition profiles (Bush et al., 1995; Bush and Jacoby 2010). According to the Ambler (or structural) classification, classes A, C and D utilize a serine moiety to hydrolyze the β-lactams, while class B have either a single Zn^{2+} ion or a pair of ions in the active site that facilitate the reaction (Drawz & Bonomo, 2010; Tooke et al., 2019). The following is a brief description of the four classes of β-lactamases enzymes according to the Ambler classification scheme.

1.2.2.1. Class A serine β-lactamases

Class A β-lactamases are the most commonly detected group among clinical Gram-positive and Gram-negative bacterial isolates. They show broad substrate specificity, including penicillins, cephalosporins, and carbapenems (Eiamphungporn et

al., 2018; Toussaint & Gallagher, 2015). TEM-1 β-lactamase (named after the patient Temoneira from which the isolate was taken) and SHV-1 (Sulfhydryl variable) are the most prevalent plasmid-encoded class A β-lactamases in Gram-negative pathogens (e.g., *E. coli* and *K. pneumoniae*), and both TEM-1 and SHV-1 share 68% sequence homology (Drawz and Bonomo 2010; Tooke et al. 2019). Other class A enzymes are encoded on the chromosome or integrons, such as PC1 (penicillinase type 1 from *Burkholderia pseudomallei*) and VEB-1 (from *Pseudomonas aeruginosa* and *Acinetobacter baumannii*), respectively (Bush & Jacoby, 2010; Naas et al., 2008).

Mutations in the parent TEM-1 and SHV-1 genes led to single amino acid changes in the enzymes, enabling the new variants to hydrolyze many of the extended-spectrum cephalosporins antibiotics. This group of β-lactamases is known as class A extended-spectrum β-lactamases (ESBLs) and are responsible for resistance to antibiotics such as cefotaxime, ceftobiprole (to treat methicillin-resistant *S. aureus* MRSA), ceftazidime, and the monobactam aztreonam. (Drawz and Bonomo 2010; Bush and Jacoby 2010). ESBLs cannot efficiently hydrolyze cephamycins, carbapenems, and β-lactamase inhibitors. CTX-M (cefotaximase), another ESBL, emerged by plasmid transfer from pre-existing chromosomal ESBL genes from *Kluyvera* spp., which are Gram-negative non-pathogenic organisms (Bonnet, 2004; Humeniuk et al., 2002). Most of ESBLs belong to SHV, TEM, and CTX-M families; less frequently, they are derived from BES (Brazilian ESBLs), GES-1 (Guyana ESBLs), VEB (Vietnam ESBLs), and PER (*Pseudomonas* extended resistance) enzymes (Bonnet, 2004; Eiamphungporn et al., 2018; Naas et al., 2008).

Carbapenemases, another class A β-lactamases, include NMC-A (non-Metallo-carbapenemases of class A), SME (Serratia marcescens enzyme), and KPC (*K.*

pneumoniae carbapenemase) (Bush and Jacoby 2010; Drawz and Bonomo 2010). Members of this group can hydrolyze carbapenems as well as cephalosporins, penicillins, and the monobactam aztreonam (Queenan & Bush, 2007). These carbapenemases have been identified primarily in *Enterobacter cloacae*, *Serratia marcescens*, and *K. pneumoniae* (Drawz & Bonomo, 2010; Queenan & Bush, 2007). In general, class A enzymes are susceptible to commercially available β-lactamase inhibitors (clavulanic acid, tazobactam, and sulbactam). However, the *K. pneumoniae* carbapenemases KPC are an exception as they confer resistance to all β-lactams and are not efficiently inhibited by typical Class A inhibitors, making them of great concern (Drawz & Bonomo, 2010; Eiamphungporn et al., 2018; Queenan & Bush, 2007).

1.2.2.2. Class B Metallo-β-Lactamases (MBLs)

Class B metallo-β-lactamases use zinc (Zn^{2+}) atom in the active site for inactivating β-lactams antibiotics (Palzkill, 2013; Tooke et al., 2019). They can inactivate a broad range of antibiotics such as penicillins, cephalosporins, carbapenems, and also the clinically used β-lactamase inhibitors (clavulanate, sulbactam, and tazobactam), but not the monobactam aztreonam (Bush and Jacoby 2010; Palzkill 2013). However, Class B MBLs are inhibited by metal chelators, such as ethylenediaminetetraacetic acid (EDTA) (Palzkill 2013). There are several types of class B MBLs, including VIM (Verona integron-encoded metallo-β-lactamase), IMP (imipenemase), and NDM-1 (New Delhi metallo-β-lactamase) (Drawz & Bonomo, 2010; Palzkill, 2013). The MBLs are encoded by genes situated on the chromosome or plasmid where *P. aeruginosa*, *K. pneumoniae*, and *A. baumannii* are producers of class B enzymes (Palzkill, 2013; Walsh et al., 2005).

1.2.2.3. Class C serine β-lactamases

The Class C enzymes or AmpC type β-lactamases hydrolyze diverse β-lactam antibiotics, including penicillins, cephamycins (e.g., cefoxitin and cefotetan), oxyimino cephalosporins (e.g., ceftazidime, cefotaxime, and ceftriaxone) and β-lactamase inhibitors (Thomson 2010; Drawz and Bonomo 2010). The enzymes are produced by Gram-negative bacteria such as *P. aeruginosa, Enterobacter* spp., *Acinetobacter* spp., *Aeromonas* spp., *Citrobacter freundii, E. coli,* and *S. marcescens* (Parveen et al., 2010), and they are usually encoded on the chromosome and less frequent on plasmids (Drawz & Bonomo, 2010; Thomson, 2010). In most genera of Enterobacteriaceae, the *ampC* genes (encoding for AmpC enzymes) are inducible with certain β-lactams, unlike plasmid-encoded *ampC,* where the genes are always expressed constitutively (Parveen et al., 2010; Bush and Jacoby 2010). The most commonly encountered AmpC β-lactamases belong to the ACT-1 (AmpC β-lactamase), P99 (from *Enterobacter cloacae* P99), CMY (cephamycinase), FOX (cefoxitinase), and DHA (Dhahran Hospital in Saudi Arabia) families (Thomson, 2010; Drawz and Bonomo, 2010; Bush and Jacoby, 2010). However, the enzymes are inactivated explicitly by boronic acid and avibactam (Thomson, 2010).

1.2.2.4. Class D serine β-lactamases

Class D β-lactamases or OXA-type β-lactamases (oxacillinases) were initially categorized as "oxacillinases" because of their ability to hydrolyze oxacillin much faster than classical penicillins (Tooke et al., 2019). The OXA enzymes are encoded by genes on both chromosomes and plasmids in a wide diverse species of Gram-negative bacteria (such as *Acinetobacter* spp., *Pseudomonas* spp., *Burkholderia* spp., *Shewanella spp.* and

Enterobacteriaceae) (Antunes & Fisher, 2014), and Gram-positive bacteria (such as *Bacillaceae*, *Clostridiaceae* and *Eubacteriaceae*) (Toth et al., 2016). The OXA β-lactamases confer resistance to penicillins, cephalosporins, extended-spectrum cephalosporins (OXA-type ESBLs), and carbapenems (OXA-type carbapenemases). In addition, the β-lactamase inhibitors clavulanic acid and tazobactam exhibit inhibitory activity for some OXA enzymes, but the sulbactam inhibitor does not show inhibitory activity against class D enzymes (Bush and Jacoby 2010; Eiamphungporn et al., 2018). Examples of OXA enzymes include OXA-23 and OXA-24/40 in *A. baumannii* and the constitutively expressed OXA-50 in *P. aeruginosa* (Antunes & Fisher, 2014; Eiamphungporn et al., 2018).

1.2.3. The β-lactamases inhibitors (βLIs)

The β-lactamase inhibitors (βLIs) are compounds with weak antibacterial activity but work primarily by inhibiting β-lactamases produced by antibiotic-resistant bacteria, to render them sensitive to β-lactam antibiotics (Toussaint and Gallagher, 2015). Attempts to identify inhibitors of common β-lactamases began in the mid-1970s to combat β-lactamase-mediated resistance in bacteria. The β-lactamase-inhibitory activity was first detected in a strain of *Streptomyces olivaceus* due to the production of olivanic acids, a family of carbapenem β-lactam compounds (Brown et al., 1976; Butterworth et al., 1979). However, olivanic acids were found to have poor penetration through the cell wall in some pathogens and low stability due to rapid metabolism in the human body; and were therefore not used in clinical treatment (Rolinson, 1991). Shortly after that, a superior clavam β-lactamase inhibitor, clavulanic acid (CA), was discovered (Brown et al., 1976),

followed by penicillanic acid sulfone inhibitors, sulbactam and tazobactam (YTR830) (Figure 1.2) (English et al., 1978; Fisher et al., 1980). All three βLIs are approved for clinical use by the Food and Drug Administration (FDA) in the USA and are commercially available (Table 1.1).

Clavulanic acid is a natural β-lactamase inhibitor and the first to be fully characterized and introduced into clinical medicine. It is produced by the fermentation of *Streptomyces clavuligerus* in broth media (Reading & Cole, 1977). CA binds irreversibly with the serine hydroxyl group present in the active site of β-lactamase, producing a stable acylated intermediate and completely inactivating the enzyme (Figure 1.3); hence it is classified as a "suicide inhibitor" (Rolinson, 1991; Toussaint & Gallagher, 2015). The spectrum of the inhibitor is now recognized to include most class A β-lactamases, including ESBLs and, to a lesser extent, serine carbapenemases and some class D OXA enzymes (Bush and Jacoby, 2010; Viana-Marques et al., 2018).

Sulbactam and tazobactam (Figure 1.2) are other inhibitors that were developed by synthetic routes in the 1980s (Rolinson, 1991). Each follow the same general inactivation-mechanism of CA, and overall, they have a similar spectrum of activity. Sulbactam had less inhibitory activity than CA or tazobactam against class A β-lactamases, but both sulfones are more potent inhibitors of cephalosporinases (Bush et al., 1993). In general, these inhibitors have weak or lack antibacterial activity on their own, but with a notable exception; CA has antibacterial activity against *Haemophilus influenzae* and *Neisseria gonorrhoeae*, sulbactam has modest action against *Acinetobacter* spp., *Bacteroides* spp. and *Burkholderia cepacian*, and tazobactam has

against *Borrelia burgdorferi* (Miller et al., 1978; Higgins et al., 2004; Drawz and Bonomo, 2010; Bush and Bradford, 2016).

Because the βLIs are not potent antibiotics, they are used in combination with β-lactam antibiotics to assist and increase the spectrum of activity of the latter (Tyers & Wright, 2019). For instance, the β-lactam/βLI combinations amoxicillin/clavulanate (Augmentin™) and ticarcillin/clavulanate (Timentin™) significantly expand the antibiotics spectra and improve the clinical efficacy against several β-lactamases-producing bacteria. The two drug combinations are approved by the FDA and are widely available in the market (Drawz & Bonomo, 2010; Geddes et al., 2007). Also, sulbactam is commercially available in combination with ampicillin (Unasyn™) and cefoperazone (Magnex™), providing them with broad-spectrum activity (Williams, 1997; Bush and Bradford, 2016). The ampicillin/sulbactam drug is ideal for polymicrobial infections such as abdominal and gynecological surgical infections (Drawz and Bonomo, 2010). For the third βLI tazobactam, the combination with piperacillin (Zosyn™) has proven clinical efficacy against Gram-positive and Gram-negative pathogens (Drawz and Bonomo, 2010), and the tazobactam/ceftolozane combination is used against antibiotic-resistant *P. aeruginosa* and many ESBL-producing Enterobacteriaceae (Jacqueline et al., 2017).

Other FDA-approved β-lactamase inhibitors are avibactam and vaborbactam (Tyers & Wright, 2019). Unlike the above described inhibitors, avibactam does not belong to the β-lactams class of compounds but is a novel diazabicyclooctane (DBO) (Toussaint and Gallagher, 2015). Instead of the 4-membered β-lactam ring, DBOs have a 5-membered ring with an amide group that targets the active-site serine of β-lactamases (Figure 1.2). Avibactam is the most potent β-lactamase inhibitor to date because only one

to five molecules are required to inhibit one molecule of β-lactamase, compared with >50 for tazobactam and clavulanate (Toussaint & Gallagher, 2015). Avibactam/ceftazidime combination (Avycaz^{TM}) also has a broader spectrum of activity and can inhibit class A penicillinases, ESBLs, serine carbapenemases, class C cephalosporinases, and some class D oxacillinases (Ehmann et al., 2013). Vaborbactam (RPX7009) is a novel boronic acid β-lactamase inhibitor with potent activity against class A β-lactamases, ESBLs, carbapenemases, and some class C enzymes (Eiamphungporn et al., 2018). The FDA recently approved the meropenem/vaborbactam combination (vabomere) in 2017 (Tyers & Wright, 2019).

1.3. Clavams and *Streptomyces clavuligerus*

Streptomyces clavuligerus has the ability to produce a diverse set of β-lactams compounds. It is a natural source for cephamycin C (Ceph-C), the β-lactamase inhibitor clavulanic acid (CA), and the 5*S* clavams. Interestingly, *S. clavuligerus* is the only reported species producing these three kinds of β-lactams (cephamycins, clavulanic acids, and 5*S* clavams) (Brown et al., 1979; Higgens et al., 1974; Reading & Cole, 1977). The other two species, *Streptomyces jumonjinensis* and *Streptomyces katsurahamanus* are known to produce CA along with Ceph-C but not 5*S* clavams (AbuSara et al., 2019; Ward & Hodgson, 1993). In comparison, most species produce either Ceph-C only, such as *Streptomyces griseus, Streptomyces cattleya,* and *Nocardia lactamdurans,* or 5*S* clavams, as in *Streptomyces antibioticus, Streptomyces microflavus (lipmanii), Streptomyces hygroscopicus, Streptomyces platensis, Streptomyces lavendulae* and *Streptomyces brunneogriseus* (Jensen and Paradkar, 1999). The genetics and biochemistry of *S.*

clavuligerus metabolites have been intensively studied for more than 50 years (Ramirez-Malule, 2018). Although cephamycin C can be industrially manufactured by fermentation of *S. clavuligerus* or by synthetic routes, CA is only produced by fermenting the bacteria in suitable broth media (Jensen and Paradkar, 1999; Saudagar et al., 2008).

On the chromosomes of *S. clavuligerus*, the clavulanic acid BGC is situated immediately downstream from the Ceph-C BGC (Figure 1.4A), forming a "super cluster" (Ward & Hodgson, 1993). Both BGCs are principally regulated by the transcription regulator CcaR (OmpR family), which is encoded by the gene *ccaR* situated in the Ceph-C gene cluster (Figure 1.4A) (Perez-Llarena et al., 1997). Although CA and Ceph-C have distinct biosynthetic pathways, their productions generally coincides in fermentation media (Hamed et al., 2013; Romero et al,, 1984). This co-production of metabolites that act synergistically is beneficial for the survival of the organism in its habitat (Challis & Hopwood, 2003).

The 5*S* clavam BGC is ~ 1.4 Mb away from the Ceph-C – CA supercluster on the *S. clavuligerus* chromosome (see Figure 4.1 in chapter 4), and the regulation of 5*S* clavam production is distinct from the coregulated production of Ceph-C and CA (Challis & Hopwood, 2003) (Figures 1.8 – 1.10). The 5*S* clavams have the opposite stereochemistry from CA, which has a 5*R* configuration structure for the bicyclic nucleus (the β-lactam ring and the five-membered oxazolidine ring) (Figure 1.1 B). Therefore, 5*S* clavams do not exhibit inhibitory activity toward β-lactamases, but instead, some display weak antibacterial or antifungal activities (Pruess and Kellett 1983; Jensen 2012). *S. clavuligerus* is unique among clavams producers described so far due to its ability to

produce both stereochemistries (5*R* and 5*S*, Figure 1.1 and 1.5) (Jensen and Paradkar, 1999; Challis and Hopwood, 2003).

Interestingly, a third set of genes, referred to as the "Paralogue" gene cluster, is located on the giant linear plasmid pSCL4, and contains copies of the early genes from the CA gene cluster as well as additional genes for 5*S* clavams (Figure 1.4, Table 1.3) (Tahlan et al., 2004a, 2004b; Tahlan et al., 2007; Zelyas et al., 2008). Therefore, the involvement of three separate gene clusters in the biosynthesis of CA and 5*S* clavams in *S. clavuligerus* has proved challenging for studying the production of this important class of metabolites.

1.3.1. Clavams biosynthesis, the "Early Steps"

The structural similarities between CA and the 5*S* clavams reflect shared elements of a common biosynthetic pathway. The biosynthetic pathways of CA and 5*S* clavams share the first "early steps" of reactions that lead to the formation of clavaminic acid, the last common intermediate (Figure 1.5 and 1.6). These steps are encoded by the genes in CA BGC and the paralogue genes cluster (Figure 1.4). The reactions after clavaminic acid, which are referred as the "late steps", are specific to either CA or 5*S* clavam synthesis, and they are encoded in distinguished gene clusters (Figure 1.4) (Jensen and Paradkar, 1999; Jensen, 2012).

The CA-5*S* clavams pathway begins with the condensation of L-arginine with glyceraldehyde 3-phosphate to form N2-(2-carboxyethyl) arginine; this reaction is catalyzed by carboxyethyl arginine synthase (CEAS) (Khaleeli et al., 1999). In the next step, N2-(2-carboxyethyl) arginine is cyclized by β-lactam synthase (βLS) to form a

monocyclic β-lactam compound, deoxyguanidinoproclavaminate (Bachmann et al., 1998). A hydroxylation reaction is then conducted on deoxyguanidinoproclavaminate by clavaminate synthase (CAS) to form guanidinoproclavaminate, which undergoes a hydrolysis reaction by proclavaminate amidino-hydrolase (PAH) to form proclavaminate (Marsh et al., 1992; Aidoo et al., 1994). The next two successive steps are again catalyzed by CAS. In the first reaction, proclavaminate is cyclized to form the first bicyclic ring structure giving dihydroclavaminic acid, which then undergoes subsequent desaturation to produce clavaminic acid (the last proposed 3*S*, 5*S* common intermediate) (Figure 1.5) (Arulanantham et al., 2006; Salowe et al., 1991). The genes encoding the "early steps", *ceasS2*, *bls2*, *pah2*, and *cas2*, are located in the CA BGC (Figure 1.4), and they are co-expressed and coregulated (Perez-Llarena et al., 1997; Santamarta et al., 2011). Deletion of any of these genes abolishes or significantly decreases the production of clavams (Table 1.2) (Song et al., 2010; Jensen, 2012). Next to *cas2* is the gene *oat2*, which codes for an ornithine acetyltransferase (OAT) (de la Fuente et al., 2004). OAT2 is thought to contribute toward building arginine pools (the first precursor) required for CA-5*S* clavam biosynthesis (Kershaw et al., 2002). The deletion mutant of *oat2* in *S. clavuligerus* decreased CA production. Although its exact role in CA-5*S* clavam biosynthesis is still unknown, it is considered part of the "early genes" cluster (Kershaw et al., 2002; Jensen, 2012). Interestingly, the *S. clavuligerus* genome posseses a second copy of the "early steps" genes; *cas1, ceas1, bls1, pah1,* and *oat1* (Tahlan et al., 2004a, 2004c). While *cas1* resides in the 5*S* clavam BGC, the remaining genes in the paralogue gene cluster are located along with other genes involved in 5*S* clavams biosynthesis on pSCL4 (Figure 1.4). The genetics and biochemistry of both copies of "early genes" are well characterized

and reviewed elsewhere (Jensen and Paradkar, 1999; Song et al., 2010; Jensen, 2012) and are summarized in Tables (1.2 and 1.3). Clavaminic acid is the branch point intermediate that can either be converted to CA or 5*S* clavams (Figure 1.5). The subsequent steps lead to them are known as the "late steps" (Jensen 2012). However, the enzymology and genetics of these late reactions still are not fully characterized.

1.3.2. The "Late Steps" of clavulanic acid biosynthesis

The sequencing of the entire CA gene cluster of *S. clavuligerus* helped examine the types of gene products involved in CA biosynthesis and gives insights into the nature of the steps leading from clavaminic acid to clavulanic acid. The "late steps" reactions are carried out by enzymes encoded by the genes *orf7* through *orf17* located in the CA BGC.

The first devoted "late step" is the conversion of clavaminic acid to N-glycyl-clavaminic acid (NGCA) (Figure 1.6), which is catalyzed by glycyl-clavaminic acid synthase (encoded by *gcaS* or *orf17*) (Jensen et al., 2004a; Arulanantham et al., 2006). Deletion of *gcaS* in *S. clavuligerus* results in block CA production (Jensen et al., 2004a). In the next few steps, N-glycyl-clavaminic acid is converted to 3*R*, 5*R* clavaldehyde, whereby the mechanism still under investigation (Figure 1.6). However, it is predicted that double ring epimerization and oxidative deamination reactions could invert the 3*S*, 5*S* stereochemistry of N-glycyl-clavaminic into the 3*R*, 5*R* clavaldehyde, a stereochemistry essential for the β-lactamase activity (Jensen, 2012). The last reaction comprises the reduction of clavaldehyde to CA, catalyzed by clavaldehyde dehydrogenase (CAD) (Figure 1.6), encoded by *cad* (or *car*) (Figure 1.4) (MacKenzie et al., 2007). Deletion

mutation of *cad* blocks CA production, whereas overexpression of *cad* increases production (Pérez-Redondo et al., 1998; Jensen, 2012).

oppA1 (*orf7*) and *oppA2* (*orf15*) encode similar proteins with sequence similarity to periplasmic oligopeptide-binding protein (MacKenzie et al., 2010). Mutants defective in either *oppA1* or *oppA2* are unable to produce CA, suggesting that these genes are essential for CA production (Lorenzana et al., 2004; Jensen et al., 2004a). The two gene products of *oppA1/2* are proposed to be involved in binding and/or transporting arginine/peptide substrates used in CA biosynthesis. Recently, it was found that the deletion of *oppA2* accumulates the intermediate compound N-acetyl-glycyl-clavaminic acid (AGCA or NAG-clavam) (Jensen et al., 2004a; Álvarez-Álvarez et al., 2018), and with the addition of pure AGCA to *S. clavuligerus* mutants in which genes of the early steps were deleted and the CA formation was blocked, resulted in the recovery of CA production. These results are establishing that AGCA is a late intermediate of CA biosynthesis (between NGCA and clavaldehyde; Figure 1.6), and *oppA2* plays an essential role in the "late" steps (Álvarez-álvarez et al., 2018).

claR (*orf8*) is a pathway-specific regulator gene that encodes a LysR-type transcriptional regulator (Paradkar et al., 1998). It positively regulates the 'late' genes (*orf7 – orf17*) of the CA pathway but with little or no control on the distal genes (*orf18* and *orf19*) (Figure 1.4) (Paradkar et al., 1998; Pérez-Redondo et al., 1998; Martínez-Burgo et al., 2015). Mutants disrupted in *claR* are unable to produce clavulanic acid but can still make the 5*S* clavams and cephamycin C (Paradkar et al., 1998; Pérez-Redondo et al., 1998). Furthermore, *claR*-null mutant accumulates clavaminic acid (the last

intermediate of the "early" pathway), suggesting that *claR* is essential for completing the clavulanic acid pathway by regulating the "late" reactions (Paradkar et al., 1998).

cyp (*orf10*) encodes cytochrome P-450 (CYP) whereas *fd* (*orf11*) encodes ferredoxin, an electron transport protein. Cytochrome P-450 typically carry out oxidative reactions in cooperation with a ferredoxin protein (Khaleeli & Townsend, 2000; Mellado et al., 2002). A mutation in *cyp* leads to loss of CA production, whereas a mutation in *fd* results in a reduction in CA production (Khaleeli & Townsend, 2000; Mellado et al., 2002; Jensen et al., 2004a). CYP and ferredoxin are expected to be responsible for the oxidation-deamination step between N-glycyl clavaminic acid and clavaldehyde in the late steps of clavulanic acid biosynthesis (Figure 1.4 and 1.6). However, their biochemical functions are still unclear (Jensen, 2012).

The next gene, *cpe* (*orf12*), encodes a protein with a C-terminal class A β-lactamase-like domain fused to an N-terminal isomerase/cyclase-like domain (Srivastava et al., 2019; Valegård et al., 2013). However, the protein does not show any β-lactamase activity *in vitro*, but it demonstrates low-level esterase activity towards 3`-O-acetyl cephalosporins and thioester substrate (Valegård et al., 2013). Deletion mutation for *cpe* in *S. clavuligerus* abolishes CA production but not the 5*S* clavams (Jensen et al., 2004a; Srivastava et al., 2019). A recent study showed that both N- and C-terminal domains are required for *in vivo* clavulanic acid production (Srivastava et al., 2019). The role of CPE in clavulanic acid biosynthesis is still unknown, but it may be involved in the epimerization of 3*S*,5*S* clavaminic acid to 3*R*,5*R* clavulanic acid. More details about *cpe* are discussed in Chapter III.

orf13 encodes an amino acid export pump protein, which may be involved in the transport of CA and pathway intermediates from the inside of the cell to the outside (Mellado et al., 2002; Jensen et al., 2004a). Disruption mutation of *orf13* in *S. clavuligerus* severely decreases CA and 5*S* clavams production (Mellado et al., 2002; Jensen et al., 2004a), suggesting a role in the transport of all clavam metabolites out of the cell (Jensen, 2012).

orf14 encodes CBG protein (**C**A **B**iosynthesis **G**NAT) protein, which belongs to the general control non-repressible 5 (GCN5)-N-acetyltransferases (GNAT) family (Iqbal et al., 2010a). Mutation in *orf14* results in almost complete loss of CA production but not 5*S* clavam (Jensen et al., 2004a). It is suggested that the N-acetylated clavaminic acid derivatives in the pathway might be generated from the acetylation activity of CBG during the late reactions of CA biosynthesis (Figure 1.6) (Iqbal et al., 2010; Mellado et al., 2002).

orf16 encodes for a hypothetical protein with unknown function and no significant similarities to any proteins in the database. However, a deletion mutation in *orf16* leads to a complete loss of CA formation with no effect on 5*S* clavams (Jensen et al., 2004a). As in *S. clavuligerus oppA2* defective mutants, *orf16* mutants showed an accumulation of *N*-acetyl-glycyl-clavaminic acid (AGCA) and trace amounts of *N*-glycyl-clavaminic acid metabolites (Jensen et al., 2004a). Altogether, the accumulation of the acetylated clavaminic acid metabolites in *orf15* and *orf16* mutants, and the proposed acetylation role of *orf14*, suggests that clavaminic acid is first converted to *N*-glycyl-clavaminic acid by *gcaS*, then acetylated by *orf14* gene product to give *N*-acetyl-glycyl-clavaminic acid, and subsequently converted to *N*-acetyl-clavaminic acid by *orf15/orf16* gene products (Figure

1.6) (Iqbal et al., 2010; Jensen, 2012; Paradkar, 2013). However, more investigations need to be conducted to prove this hypothesis.

The *pbpA* (*orf18*) and *pbp2* (*orf19*) genes encode proteins similar to penicillin-binding proteins (Ishida et al. 2006). Mutational studies for *pbpA* and *pbp2* showed they are not involved in CA biosynthesis (Jensen et al., 2004a, Jensen, 2012). Further downstream lies *orf20*, which encodes cytochrome P-450 and is also not involved in CA production (Jensen, 2012).

Beyond *orf20* are the open reading frames *orf21* through *orf23*, which have some effects on CA production, but their exact roles in CA synthesis remain uncertain. The deletion of *orf21* (putative sigma factor) showed a relative reduction in Ceph C production but not in CA (Song et al., 2009). The *orf22* (*cagS*; encodes sensor kinase) and *orf23* (*cagR*; encodes response regulator) are a bacterial two-component regulatory system, where defective mutants showed a reduction in both CA and Ceph C (Fu et al., 2019; Jnawali et al., 2008; Song et al., 2009). Besides, *cagS* and *cagR* show other pleiotropic effects in *S. clavuligerus* (Fu et al., 2019). It is worthy to note that the genes *orf18* and *orf20* - *orf23* are not present on the chromosomes of the other CA producers, *S. jumonjinensis* and *S. katsurahamanus* (AbuSara et al., 2019). In addition to the relatively minor or no effects resulting from deleting these genes, these features suggest they are not essential for metabolite production. Therefore, it is proposed that the core CA BGC comprises *ceaS, gcas*, and the intervening genes (Figure 1.4) (AbuSara et al., 2019).

1.3.3. The "Late Steps" of 5*S* clavams biosynthesis

As in CA biosynthesis, the "late steps" in 5*S* clavam biosynthesis are poorly understood. It is generally proposed that a series of deamination and decarboxylation reactions convert clavaminic acid into 2-carboxymethylidene clavam, which by further oxidation and hydrolysis reactions results in the 5*S*-clavams compounds, 2-formyloxymethylclavam (2-FMC), 2-hydroxymethylclavam (2-HMC), clavam-2-carboxylate (C2C), 8-hydroxylalanyl clavam, and alanylclavam (Figure 1.1 B and 1.7) (Baggaley et al., 1997; Egan et al., 1997).

The enzymes involved in the "late steps" of 5*S* clavams biosynthesis are encoded by specific genes found in the clavam and the paralogue BGCs, which are separate from the CA BGC. The clavam BGC contains the *cas1* gene, the paralog for *cas2* in the CA BGC (Marsh et al., 1992). *cas1* and *cas2* are functionally equivalent, but their transcription is regulated by different nutritional conditions. *S. clavuligerus* produces both 5*S* clavams and CA in a complex fermentation medium (soy-based) where both *cas1* and *cas2* are expressed (Mosher et al., 1999). On the other hand, *S. clavuligerus* does not produce 5*S* clavams when fermented in a synthetic medium (starch asparagine) in which only *cas2* is expressed (Mosher et al., 1999; Jensen et al., 2000). Also, a *cas1* mutant showed decreased production in 2HMC and C2C with no effect on alanylclavam (Mosher et al., 1999), suggesting that *cas1* has a important contribution in the synthesis of 5*S* clavams metabolites.

In the clavam BGC, *cas1* is surrounded by 16 genes; *cvm1* - *cvm13, cvmH, cvmP* and *cvmG*, thought to be involved in 5*S* clavams biosynthesis (Figure 1.4) (Mosher et al., 1999; Tahlan et al., 2007). The knock-out mutations in each of these genes show that

three of them, *cvm1, cvm2,* and *cvm5,* are severely affected or completely blocked in the production of 5*S* clavams, while the remaining genes did not demonstrate any effects (Table 1.3) (Mosher et al., 1999; Tahlan et al., 2007). Furthermore, the deletion of *cvm5* results in the accumulation of 2-carboxymethylideneclavam, an intermediate in the 5*S* clavam pathway, suggesting its role in converting 2-carboxymethylideneclavam into 2-formyloxymethylclavam (Figure 1.7) (Tahlan et al., 2007).

The paralogue gene cluster includes more genes involved in the biosynthesis of 5*S* clavams (specially alanylclavam) (Kwong et al., 2012; Tahlan et al., 2007; Zelyas et al., 2008). The central region of the paralogue cluster includes *ceaS1, bls1, pah1*, and *oat1*, the paralogues for the "early steps" genes in CA BGC (Figure 1.4). Disruption of *ceaS1, bls1,* and *pah1* decreased 5*S* clavam and CA production to variable degrees (Table 1.3) (Jensen et al., 2004b; Tahlan et al., 2004a). However, the double mutants defective in both copies of the paralogues *ceaS1/ceaS2, bls1/bls2,* and *pah1/pah2* were totally blocked in 5*S* clavam and CA biosynthesis, except for *oat1/oat2* mutants, which still produced CA and 5*S* clavam metabolites in reduced amounts (Table 1.2) (Jensen et al., 2004b; Tahlan et al., 2004a).

The paralogue gene cluster holds two additional paralogues, *cvm6p* (encoding a putative aminotransferase) and *cvm7p* (encoding a large bi-domain transcriptional regulator) (Figure 1.4), which have similarity with *cvm6* and *cvm7* from the clavam BGC (Figure 1.4). However, in contrast to *cvm6* and *cvm7*, mutation of *cvm6p* and *cvm7p* blocks the production of 5*S* clavam metabolites without affecting CA synthesis (Tahlan et al., 2007). It is thought that *cvm7p* encodes a transcription regulator specific for 5*S* clavam metabolites, while *cvm6p* is proposed to encode an enzyme that deaminates

clavaminic acid to produce the aldehyde intermediate as the first step in the "late" pathway, before proceeding to the synthesis of other 5*S* clavam metabolites (Figure 1.7) (Tahlan et al., 2007). Beyond *cvm7p*, a set of three genes, *snk, res1*, and *res2*, encode atypical two-component regulatory system proteins. Disruption of either *snk* or *res2* abolish the production of 5*S* clavams but not CA production, whereas mutation of *res1* elevates the production of 5*S* clavams (Kwong et al., 2012).

Next to *ceaS1* on the other flank of the paralogue cluster reside four genes, *orfA, orfB, orfC,* and *orfD* (Figure 1.4; Table 1.3). Deletion of each of these abolishes production of the 5*S* clavam metabolite, alanylclavam, without affecting any other clavams 2HMC, 2FMC, C2C or clavulanic acid (Zelyas et al., 2008). Moreover, mutants of *orfC* and *orfD* accumulate a clavam intermediate, 8-hydroxyalanylclavam (Figure 1.1 and 1.7), not detected before in wild-type cells. The 8-hydroxyalanylclavam is proposed to be intermediate in the alanylclavam pathway (Figure 1.7) (Zelyas et al., 2008; Jensen, 2012). Generally, the defective mutant phenotypes indicate that the paralogue BGC may be exclusively involved in 5*S* clavams production and in augmenting CA production in *S. clavuligerus*.

1.3.4. Regulation of clavam production

The CA and 5*S* clavam BGCs in *S. clavuligerus* are regulated hierarchically. At the bottom of the hierarchy, the biosynthetic genes are controlled by pathway-specific transcriptional regulators encoded within the gene clusters. The expression of the transcriptional regulators is tightly controlled by global regulatory mechanisms, which tie

the production of the SMs to the physiological and morphological situation of the bacteria.

As mentioned earlier, *claR* encodes a LysR-type pathway-specific regulator (Paradkar et al., 1998) that positively regulates the "late" genes (*orf7* – *orf17*) and negatively controls *orf18* (*pbpA*) and *orf19* (*pbp2*), with no control of the "early steps" genes in the CA BGC (Figure 1.8) (Paradkar et al, 1998; Martínez-Burgo et al., 2015). *S. clavuligerus* mutants disrupted in *claR* are unable to produce clavulanic acid but accumulate clavaminic acid (the last intermediate of the "early" pathway). This suggests that ClaR regulates the late steps in the clavulanic acid pathway, i.e., those involved in converting clavaminic acid to clavulanic acid (Paradkar et al., 1998; Martínez-Burgo et al., 2015). Also, strains of *S. clavuligerus* carrying multiple copies of *claR* overproduce clavulanic acid (Pérez-Redondo et al., 1998).

ClaR also has some control over the 5*S* clavams biosynthetic genes (Figure 1.9 and 1.10). The amplification of the *claR* gene using multicopy plasmids increases the production of alanylclavam by five- to six-fold compared to wt *S. clavuligerus* (Pérez-Redondo et al., 1998). Recently, Martínez-Burgo et al. (2015) found that the deletion of *claR* in *S. clavuligerus* slightly upregulates the expression of *cvm5* and *cvm7* (~ 2.2-fold) and downregulate *cvm3* (2.4-fold) of the 5*S* clavam BGC, while no significant differences in the expression of other genes was observed. Moreover, the deletion of the *claR* gene upregulated the expression of the *bls1, pah1, cvm6p,* and *orfB* genes (~2.2- to 5.8- fold) in the clavam paralogue gene cluster with no effect on the other genes (Figure 1.10) (Martínez-Burgo et al., 2015). However, further research is required to investigate whether the deletion of the *claR* gene affects the biosynthesis of 5*S* clavams metabolites.

The ClaR regulator negatively controls Ceph-C production. A *S. clavuligerus* strain with multiple copies of the *claR* gene showed a significant reduction in Ceph-C production compared to wild type (Pérez-Redondo et al., 1998). A disruption mutant of *claR* was upregulated in all of the genes in the Ceph-C BGC with a 1.7 average fold change in expression (Martínez-Burgo et al., 2015). However, more work is needed to reveal the regulatory mechanism of ClaR on the Ceph-C BGC.

At the next level of regulation in *S. clavuligerus* is *ccaR*, a gene that encodes a regulatory protein with significant similarity to the OmpR group of regulators belonging to the *Streptomyces*-antibiotic regulatory protein family (SARP) (Perez-Llarena et al., 1997). The *ccaR* gene is situated in the Ceph-C gene cluster (Figure 1.4) and controls the production of both CA and Ceph-C (Perez-Llarena et al., 1997; Alexander & Jensen, 1998; Tahlan et al., 2004c). Deletion mutation of *ccaR* blocks the production of CA and Ceph-C but not the 5*S* clavams (Alexander & Jensen, 1998; Tahlan et al., 2004c). Transcriptional analysis of the CA BGC in a Δ*ccaR* mutant strain shows that the "early" genes *ceaS2-bls2-pah2-cas2*, in addition to *car* and the regulatory gene *claR*, are not expressed (Pérez-Redondo et al., 1999; Tahlan et al., 2004c; Santamarta et al., 2011). This suggests that CcaR regulates CA production directly by regulating the expression of the "early" genes and indirectly by regulating the expression of *claR*, which in turn controls transcription of the “late” genes in the CA pathway (Figure 1.8) (Jensen, 2012).

Clavulanic acid and 5*S* clavam biosynthesis are also subject to multiple higher-level or global regulation systems. One such global regulator is BldG, a putative anti-anti-sigma factor that functions upstream of *ccaR* and *claR* in the CA biosynthesis regulatory cascade (Bignell et al., 2005). A *S. clavuligerus bldG* deletion mutant was blocked in

aerial hyphae formation and in the production of CA, Ceph-C, and 5*S* clavams (Bignell et al., 2005). Transcriptional analysis showed that *ccaR* transcription is eliminated in the Δ*bldG* mutant suggesting that BldG controls the production of CA and Ceph-C by regulating the expression of *ccaR* (Figure 1.8). However, since CcaR does not regulate 5*S* clavam biosynthesis, BldG must control the expression of at least one other gene involved in 5*S* clavam production (Bignell et al., 2005; Jensen 2012).

The *bldA* gene encodes a leucine t-RNA that translates genes containing the TTA codon. This codon is rarely present in *Streptomyces* genomes (Trepanier et al., 2002). The *bldA*-tRNA has a pleiotropic effect and is required to translate genes associated with morphological development and specialized metabolites production in *Streptomyces*, as these genes contain TTA codons (Lawlor et al., 1987). *ccaR*, which regulates both CA and Ceph-C BGCs, contains a TTA codon, and thus it was expected that the expression of the CcaR protein would be dependent on BldA. (Trepanier et al., 2002). Disruption of *bldA* in *S. clavuligerus* results in a lack of aerial hyphae formation, as expected of this phenotype, but the same mutant are not blocked for production of either CA or Ceph-C (Trepanier et al., 2002). This suggests that effective mistranslation of TTA-containing genes can occur in some cases, or an alternative mechanism for translation of the TTA codon is present (Trepanier et al., 2002). Also, this work indicates that the regulation of CA and Ceph-C biosynthesis is independent of the *bldA*-mediated pathway. AdpA, a *bldA*-dependent regulator that contains a TTA codon in the coding region, positively regulates *ccaR* and *claR*. *adpA* deleted mutants produce decreased levels of CA, and *ccaR* and *claR* expression levels (López-García et al., 2010; Takano et al., 2003).

Interestingly, Δ*bldA* mutants of *S. clavuligerus* overproduce some 5*S* clavams compared to the wild type but do not produce alanylclavam (Kwong et al., 2012; Zelyas et al., 2008). Three candidate genes, *res1, res2,* and *orfA* (see section 1.2.4.3 above) from the paralogue gene cluster (Figure 1.4), contain TTA codons and could be regulated by BldA (Figure 1.10) (Kwong et al., 2012). The data suggest that *res2*, like *ccaR*, is expressed in the *bldA* mutant, whereas *res1* and *orfA* are not. However, further investigation could reveal the regulation mechanism of BldA on 5*S* clavams production.

The γ-butyrolactone-type auto-regulators play an essential role in regulating morphological differentiation and SMs production in *S. clavuligerus* (Bibb, 2005). In this type of regulation, the auto-regulators bind to their cognate-binding proteins, called auto-regulator binding proteins, which in turn bind to auto-regulatory sequences (ARE) upstream of the target genes (Kinoshita et al., 1999). *S. clavuligerus* produces a γ-butyrolactone auto-regulator, which binds to the receptor protein ScaR encoded by the *scaR* (or *brp*) gene (Hyun et al., 2004). ScaR acts as a repressor that attaches to the ARE upstream of *ccaR* and suppresses its transcription (Figure 1.8) (Santamarta et al., 2007). *S. clavuligerus scaR* deleted mutants overproduce both CA and Ceph-C, suggesting its negative regulation for SMs biosynthesis (Santamarta et al., 2007). In addition, ScaR negatively controls *adpA* expression, thus in *scaR* mutants, *adpA* transcription is elevated (López-García et al., 2010). The data suggest that ScaR directly controls the expression of *ccaR* and indirectly regulates *ccaR* and *claR* expression via the AdpA pathway (Figure 1.8).

The two-component system (TCS) is another mechanism in *S. clavuligerus* to control SMs production and development (Ferguson, et al., 2016). The system is typically

composed of sensor kinase proteins that detect a specific stimulus, auto phosphorylate, and then transfer the phosphate to another protein called a response regulator. Upon phosphorylation, the response regulators bind to the DNA and activate or repress the transcription of genes. Two *S. clavuligerus* genes located downstream of the CA gene cluster, *cagS* (*orf22*) and *cagR* (*orf23*) (Figure 1.4), encode a putative sensor kinase and response regulator, respectively (Fu et al., 2019; Jnawali et al., 2008; Song et al., 2009). The *cagS/R* TCS participates as a positive regulator of the biosynthesis of CA and Ceph-C (Figure 1.8). The deletion of *cagS* causes some reduction in CA and Ceph-C (Song et al., 2009), while deletion of *cagR* shows reduction in CA and Ceph-C biosynthesis and growth development retardation as well (Fu et al., 2019; Song et al., 2009). The co-overexpression of *cagS/R* causes precocious hyperproduction of spores and increases CA production compared to wild type (Jnawali et al., 2008; Song et al., 2009). Recent transcriptional analyses for *cagR* mutants show a low level of expression for the early genes of CA BGC, *ceaS2, pah2, bls2*, and *oat2*, and the late genes, *claR, car, orf12, orf14, oppA1, oppA2, orf16,* and *gcaS* (Fu et al., 2019; Song et al., 2009). Moreover, the *oat1* gene in the paralogue gene cluster demonstrates reduced transcription levels in *cagR* mutants (Fu et al., 2019). The results suggest that CagR/S regulates the CA BGC (Figure 1.8). Besides CA biosynthesis, the CagR/S TCS has a pleiotropic effect. They regulate genes involved in primary metabolisms, such as glyceraldehyde 3-phosphate (G3P) metabolism and arginine biosynthesis. Notably, both G3P and arginine are precursors of CA (Fu et al., 2019). These data indicate that CagR/S TCS can directly control the biosynthesis of CA and indirectly affect CA production by regulating the metabolism of arginine and G3P.

Another TCS in *S. clavuligerus* is the Snk-Res1/2 system encoded by genes situated next to the *cvm7p* gene (encodes a transcriptional activator of 5*S* clavam biosynthesis) in the paralogue gene cluster. *snk* encodes a sensor kinase protein, and *res1* and *res2* encode response regulators, where Res1 acts as a checkpoint to modulate phosphorylation levels in the TCS. *snk* or *res2* deletion mutants are unable to produce any 5*S* clavams, whereas mutants defective in *res1* overproduced 5*S* clavams with no effects on CA biosynthesis (Kwong et al., 2012). Transcription analyses for Δ*snk* and Δ*res2* mutants show elimination in the expression of *cvm7p* and 5*S* clavam essential genes *cvm1, cvm5, cas1,* and *cvm6p* (Kwong et al., 2012), suggesting that the Snk-Res2 TCS exerts its effect indirectly by regulating the expression of *cvm7p* and that Cvm7P, in turn, controls expression of genes essential for 5*S* clavam biosynthesis (Figure 1.9 and 1.10).

cvm7p encodes a pathway-specific regulator with an N-terminal SARP-like domain and a C-terminal ATPase domain (Tahlan et al., 2007). Cvm7p regulates 5*S* clavam production without affecting CA and Ceph-C production (Tahlan et al., 2007). Disruption of *cvm7p* results in loss of all 5*S* clavams with no effect on CA (Tahlan et al., 2007). Transcriptional analysis for Δ*cvm7p* shows that the expression of *cvm1, cvm5,* and *cvm6p* (genes essential for 5*S* clavam biosynthesis) and *cas1* (paralogue to *cas2* of the early steps of clavams biosynthesis) are abolished (Kwong et al., 2012), suggesting that *cvm7p* controls the 5*S* clavams biosynthesis by regulating the essential genes (Figure 1.9 and 1.10). However, the transcription of *snk, res1*, *res2*, and *orfA* was not affected in the Δ*cvm7p* mutant (Kwong et al., 2012). Further details about the mechanisms of regulations in *S. clavuligerus* are elsewhere (Jensen, 2012; Paradkar, 2013; Ferguson et al., 2016).

1.4. Thesis objectives and goals.

Infectious diseases remain the second leading cause of death worldwide (de Lima Procópio et al., 2012), and the increasing resistance of pathogenic organisms has further complicated the situation. Therefore, the need for novel and more effective antibiotics to combat multidrug-resistant microbial pathogens is a global concern. *Streptomyces* are recognized for their ability to produce a wide range of antimicrobial agents. In this thesis, two species of *Streptomyces*, *S. clavuligerus* and *S. pratensis*, were the subject of my research. *S. clavuligerus* is the primary natural source of CA, the β-lactamase inhibitor, while *S. pratensis*, which holds CA-like and Carb4550-like BGCs, is a promising target to study.

In addition to CA, *S. clavuligerus* produces the 5*S* clavams. However, their biosynthetic pathways are partially understood. In Chapter 3 of this thesis, the *cpe* (*orf12*) and *orf14* genes of the CA BGC in *S. clavuligerus* were investigated for their *in vivo* role in producing CA and 5*S* clavams. Various CPE and ORF14 variants were prepared and used for that purpose. In addition, CPE was examined for its regulatory role in the transcription of essential genes for the production of CA and 5*S* clavams. A complete understanding of CA and 5*S* clavam biosynthesis could provide biological routes for enhancing CA production titers in *S. clavuligerus*.

In Chapter 4, a comparative genetic study was performed between the CA BGC of *S. clavuligerus* and the CA-like BGC in the CA non-producer bacterial species (i.e., *S. pratensis*). The large *nocE* gene within the CA-like BGCs of *S. pratensis* and other CA non-producers is one of the major differences. Therefore, the function of this gene and its role in CA biosynthesis in CA producer *S. clavuligerus* was the subject of this chapter.

The *nocE* gene was deleted and constitutively overexpressed, and its effect on the production of CA, 5*S* clavam, and Ceph-C was examined by LC/MS and/or bioassays. In addition, the impact of *nocE* on the physiology and general metabolism of *S. clavuligerus* was tested on different types of media. Untargeted metabolomics analysis was also performed using extracts from the three strains of *S. clavuligerus* (wt, *nocE*-deleted, and *nocE*-constitutive expressed) to investigate the impact of *nocE* on the general and specialized metabolism of *S. clavuligerus*.

Chapter 5 focused on studying *S. pratensis* since it is predicted to have 27 BGCs, and the CA-like and Carb4550-like BGCs were of particular interest. The OSMAC (one strain many compounds) approach was followed to investigate the ability of *S. pratensis* to produce specialized metabolites. A substance produced by *S. pratensis* was found to show activity similar to β-lactamase inhibitors. Genes expression analysis and genetic manipulation experiments were conducted to try and identify the bioactive substance. Finally, untargeted metabolomics analysis of *S. pratensis* extracts was also carried out to identify the classes of molecules/compounds produced by *S. pratensis* and investigate the bioactive substances and specialized metabolites.

1.5. Figures and tables

1.5.1. Figures

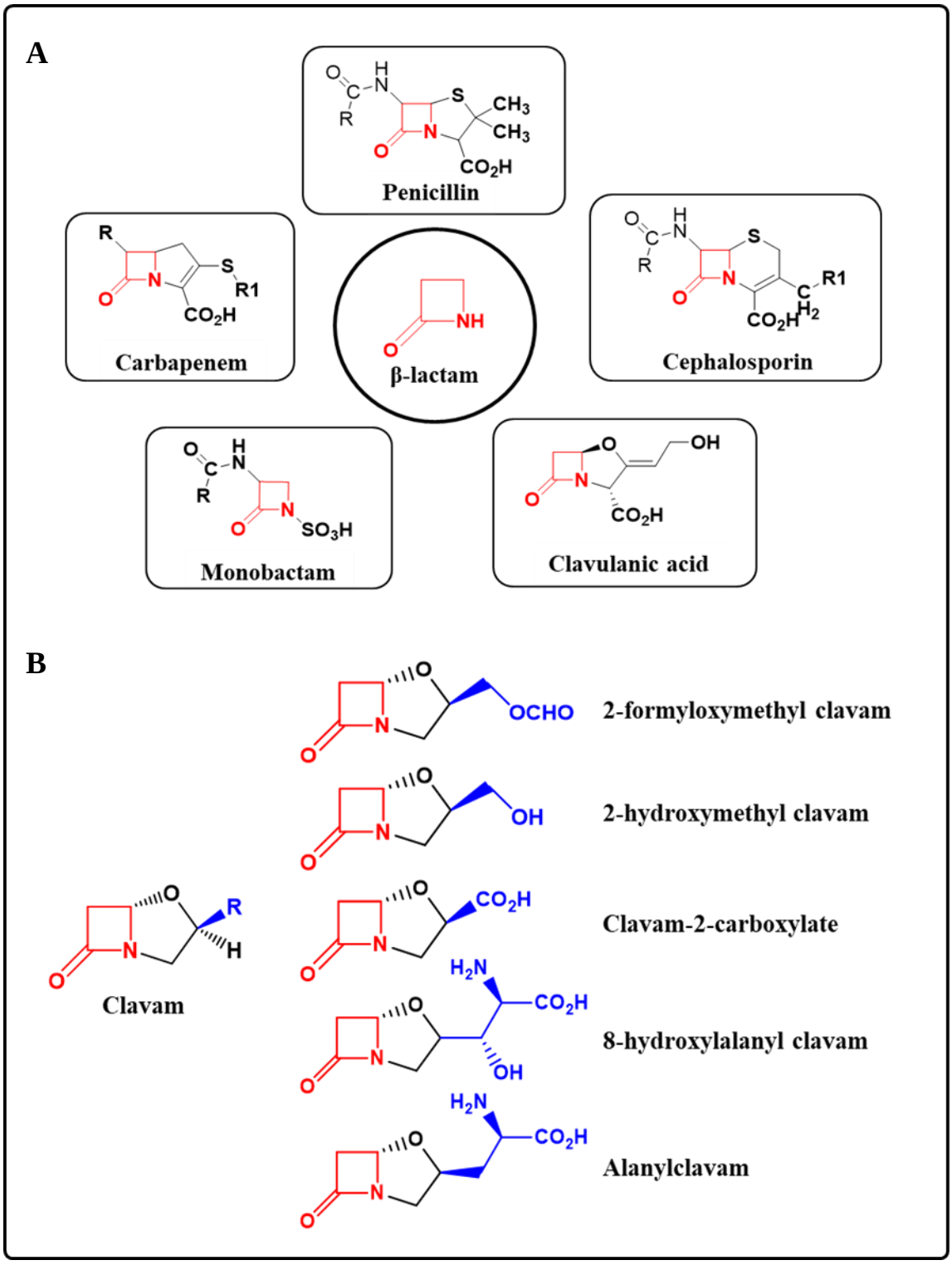


Figure 1.1. Chemical structure of β-lactams. A. The structures of major subfamilies showing the β-lactam ring highlighted in red color. **B.** The chemical structures of 5*S* clavam metabolites. The core structure present in the 5*S* clavams is shown and the side groups (designated by R) present in the different metabolites are shown in blue color.

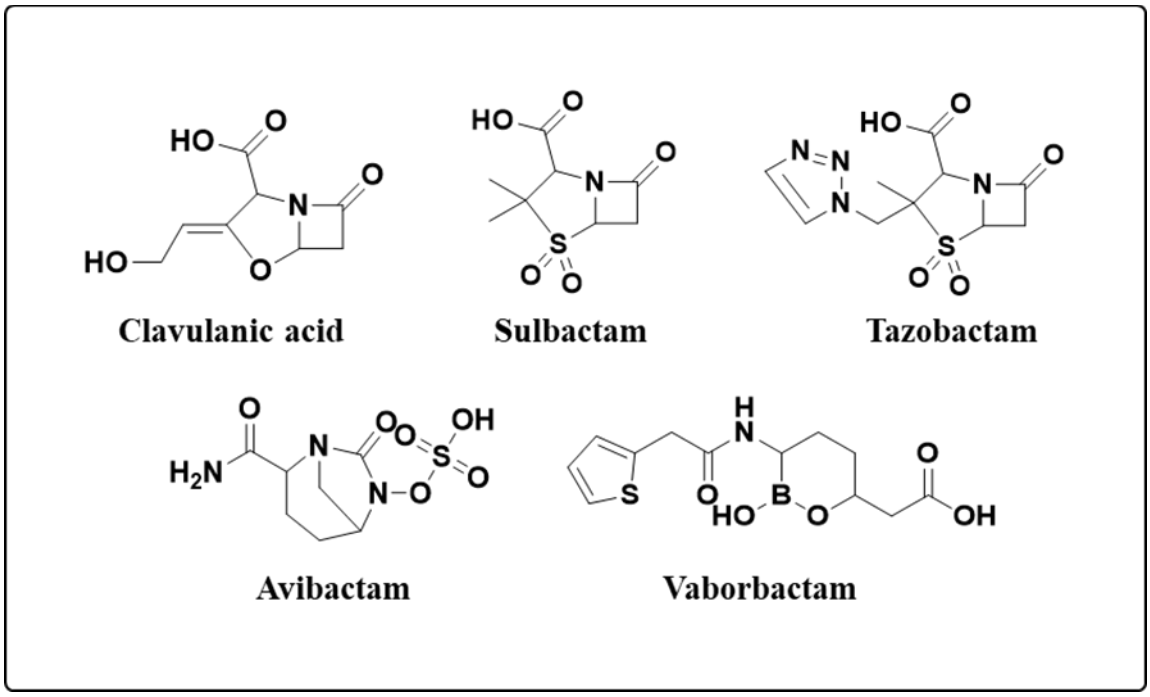


Figure 1.2: Chemical structures of clinically approved β-lactamase inhibitors. Clavulanic acid (natural product), sulbactam and tazobactam (synthetic penicillanic acid sulfones), avibactam (synthetic diazabicyclooctane), and vaborbactam (synthetic boronic).

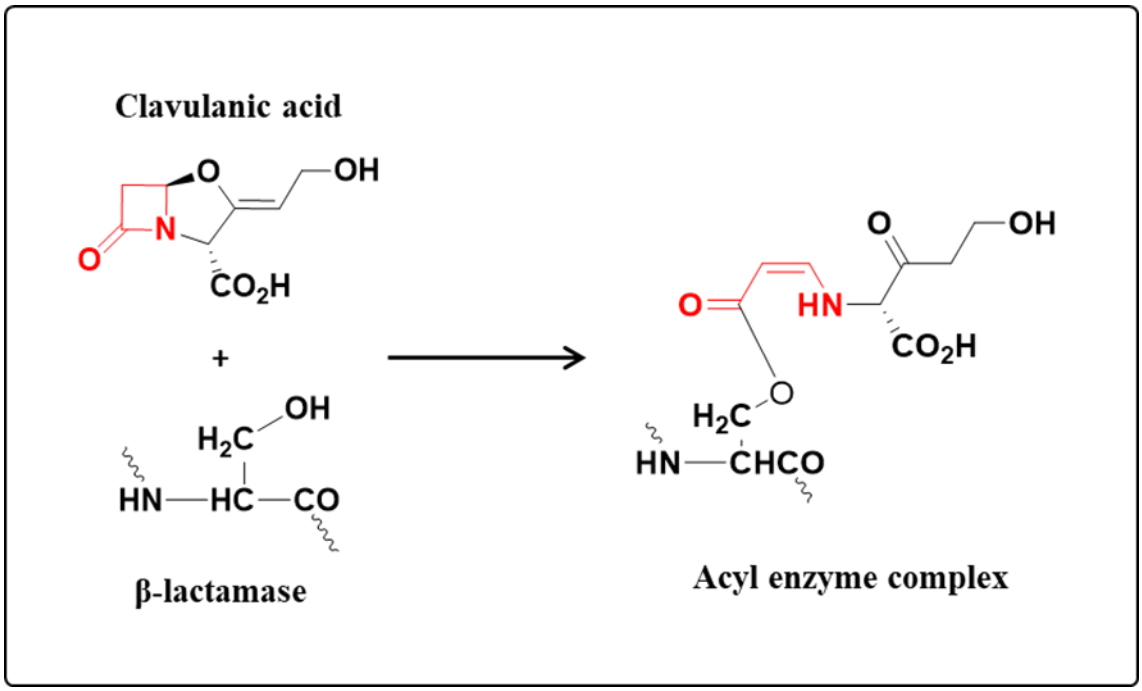


Figure 1.3. Mode of action of clavulanic acid, which reacts with a serine β-lactamase enzyme to form a stable acyl-enzyme complex. Squiggly lines represent other amino acids residues bound to serine in the active site of the enzyme.

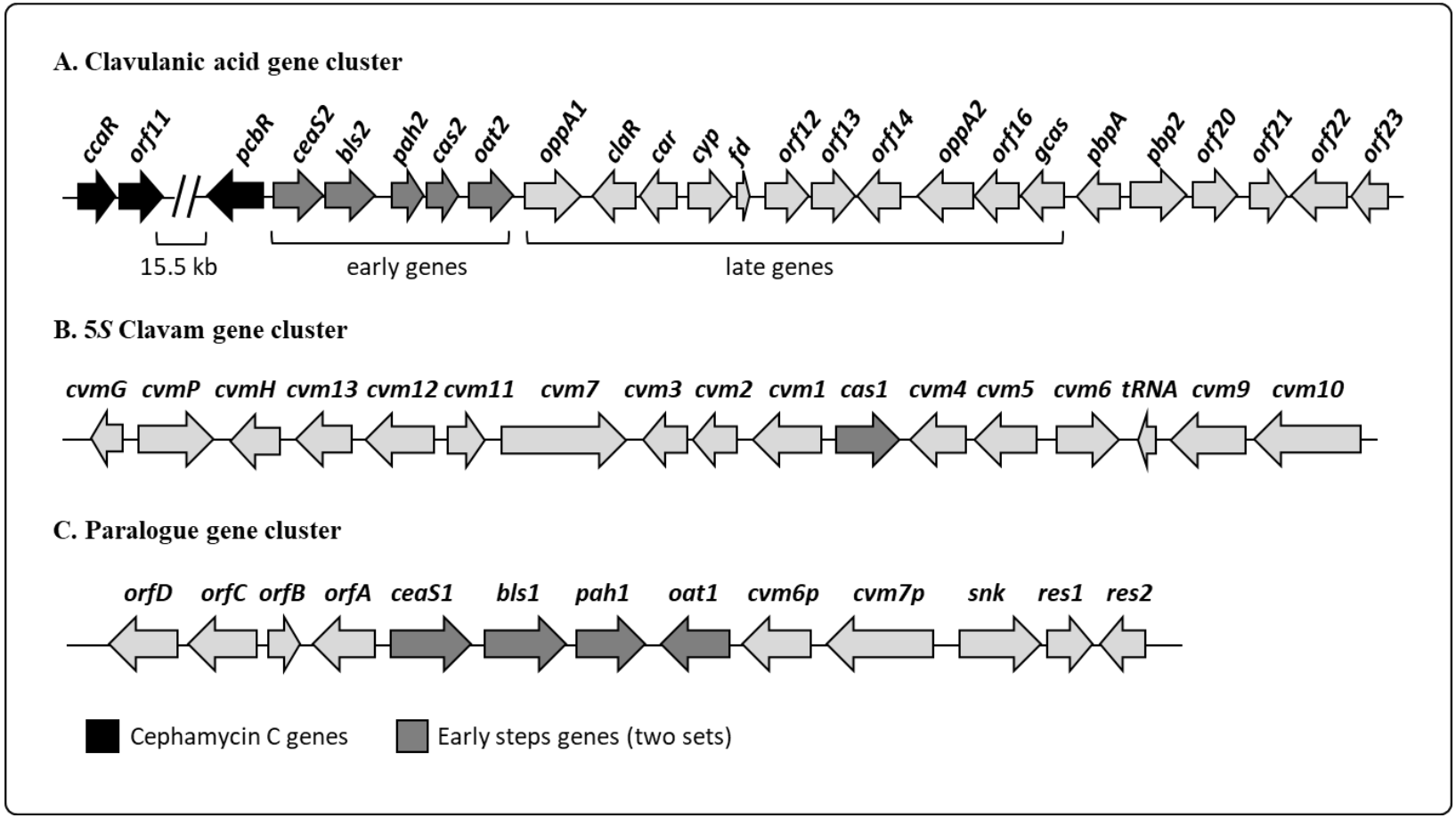


Figure 1.4.: The three biosynthetic gene clusters in *S. clavuligerus* related to the production of clavams (clavulanic acid and 5*S* clavams). Black arrows represent genes within the cephamycin C BGC. The early genes (paralogues) are in dark gray color and distributed in the three gene clusters.

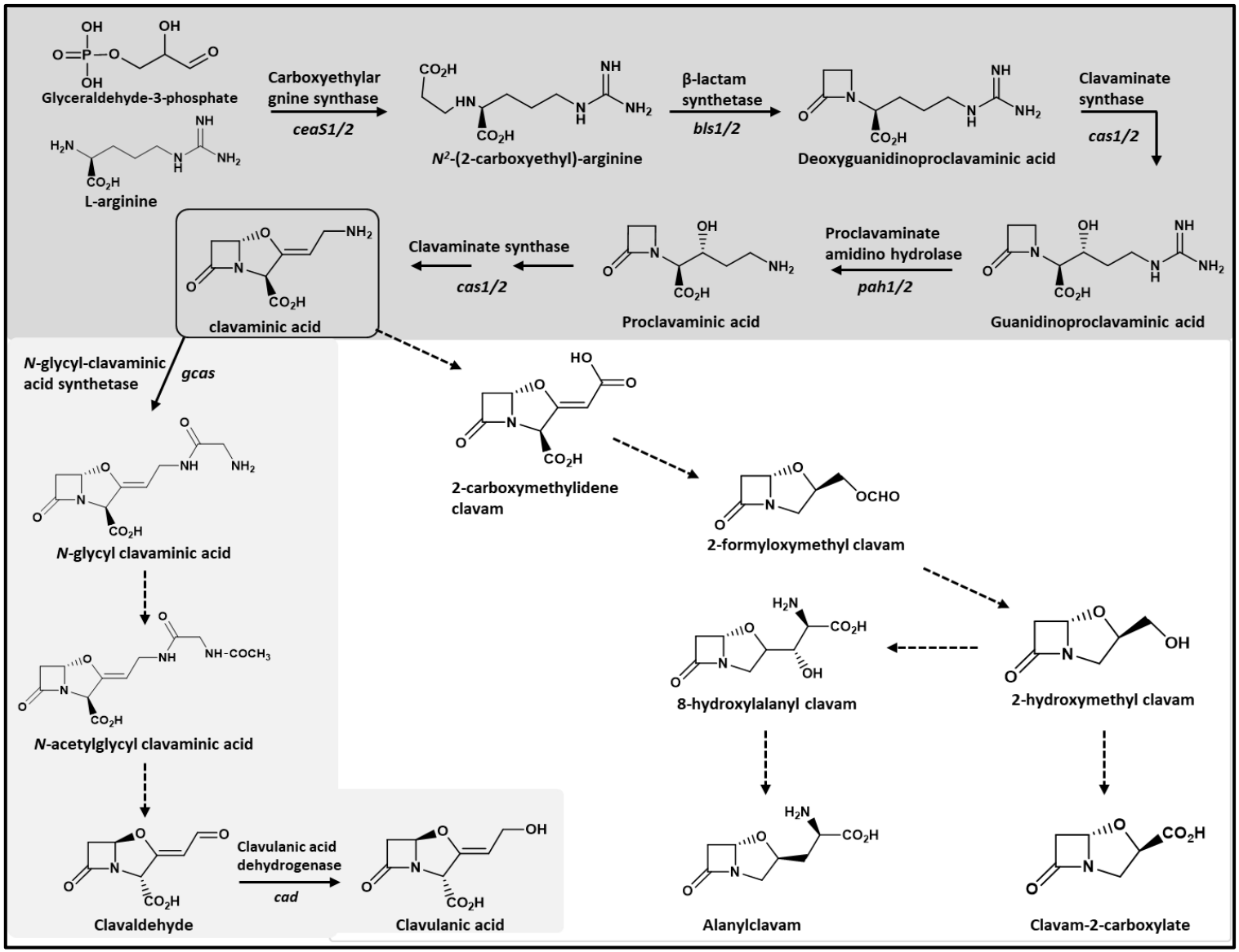

Glyceraldehyde-3-phosphate
L-arginine
Carboxyethylarginine synthase
ceaS1/2
N²-(2-carboxyethyl)-arginine
β-lactam synthetase
bls1/2
Deoxyguanidinoproclavaminic acid
Clavaminate synthase
cas1/2
Guanidinoproclavaminic acid
Proclavamidino hydrolase
pah1/2
Proclavaminic acid
Clavaminate synthase
cas1/2
clavaminic acid
N-glycyl-clavaminic acid synthetase
gcas
2-carboxymethylidene clavam
2-formyloxymethyl clavam
N-glycyl clavaminic acid
N-acetylglycyl clavaminic acid
8-hydroxylalanyl clavam
2-hydroxymethyl clavam
Clavulanic acid dehydrogenase
cad
Clavaldehyde
Clavulanic acid
Alanylclavam
Clavam-2-carboxylate

Figure 1.5. Clavulanic acid and 5*S* clavams biosynthetic pathways in *S. clavuligerus*. The pathway is composed of three parts. The upper grey box represents the early shared steps between the CA and the 5*S* clavams biosynthesis. The lower light grey box (left) and the white one (right) represent the proposed late biosynthetic reactions for CA and 5*S* clavams, respectively. Note that clavaminic acid (boxed) acts as a branch point between clavulanic acid and clavam biosynthesis. The solid arrows represent known reactions and broken arrows indicate uncharacterized steps. The names of the enzymes and their corresponding gene(s) for the known steps are included. Note that the reactions of the early shared steps of the pathway (upper grey box) are catalyzed by enzymes encoded by two sets of genes.

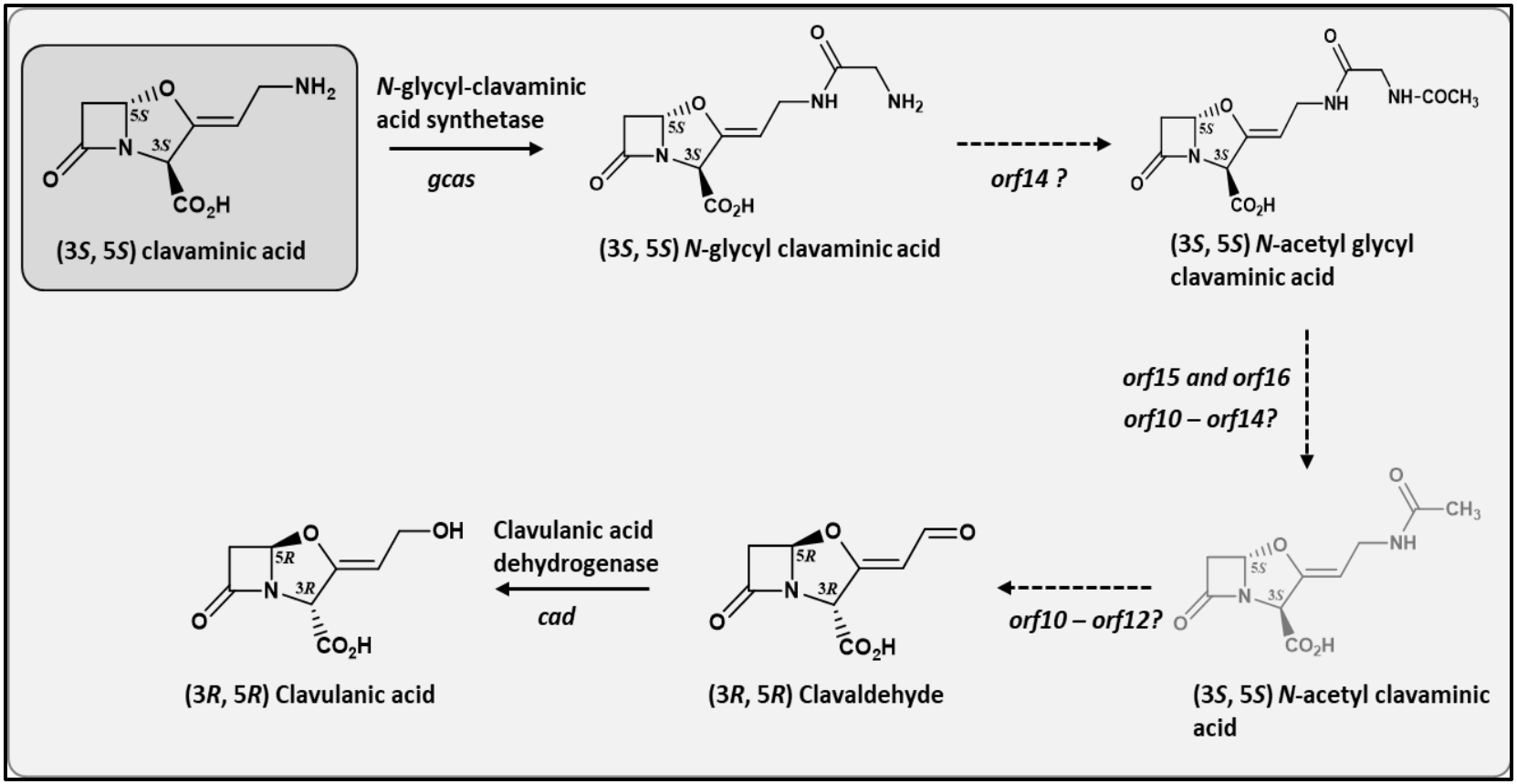


Figure 1.6. The proposed late steps of the clavulanic acid biosynthetic pathway in *S. clavuligerus*. The clavaminic acid (boxed) acts as a branch point between clavulanic acid and clavam biosynthesis. The solid arrows represent known reactions and broken arrows indicate uncharacterized steps. The names of the enzymes and their corresponding gene(s) for the known steps are included.

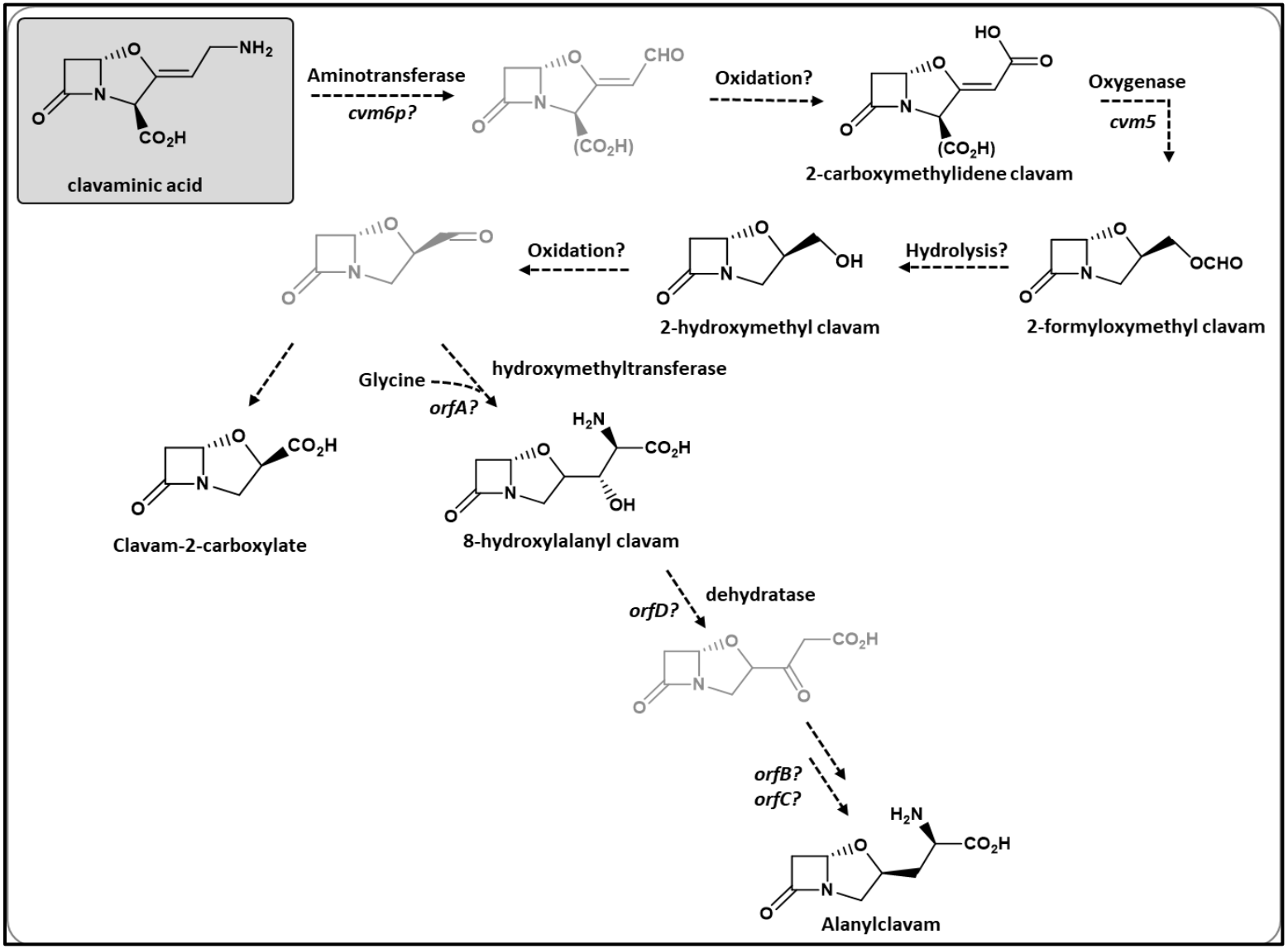

clavaminic acid
Aminotransferase
cvm6p?
Oxidation?
Oxygenase
cvm5
2-carboxymethylidene clavam
Oxidation?
Hydrolysis?
2-hydroxymethyl clavam
2-formyloxymethyl clavam
Glycine
orfA?
hydroxymethyltransferase
Clavam-2-carboxylate
8-hydroxylalanyl clavam
orfD?
dehydratase
orfB?
orfC?
Alanylclavam

Figure 1.7. The proposed late steps of the 5*S* clavam biosynthetic pathway in *S. clavuligerus*. The clavaminic acid (boxed) acts as a branch point between clavulanic acid and 5*S* clavam biosynthesis. Known metabolites are shown in black; proposed metabolites are shown in grey. The solid arrows represent known reactions and broken arrows indicate uncharacterized steps. The names of the proposed enzyme/reaction and their proposed corresponding gene(s) are included.

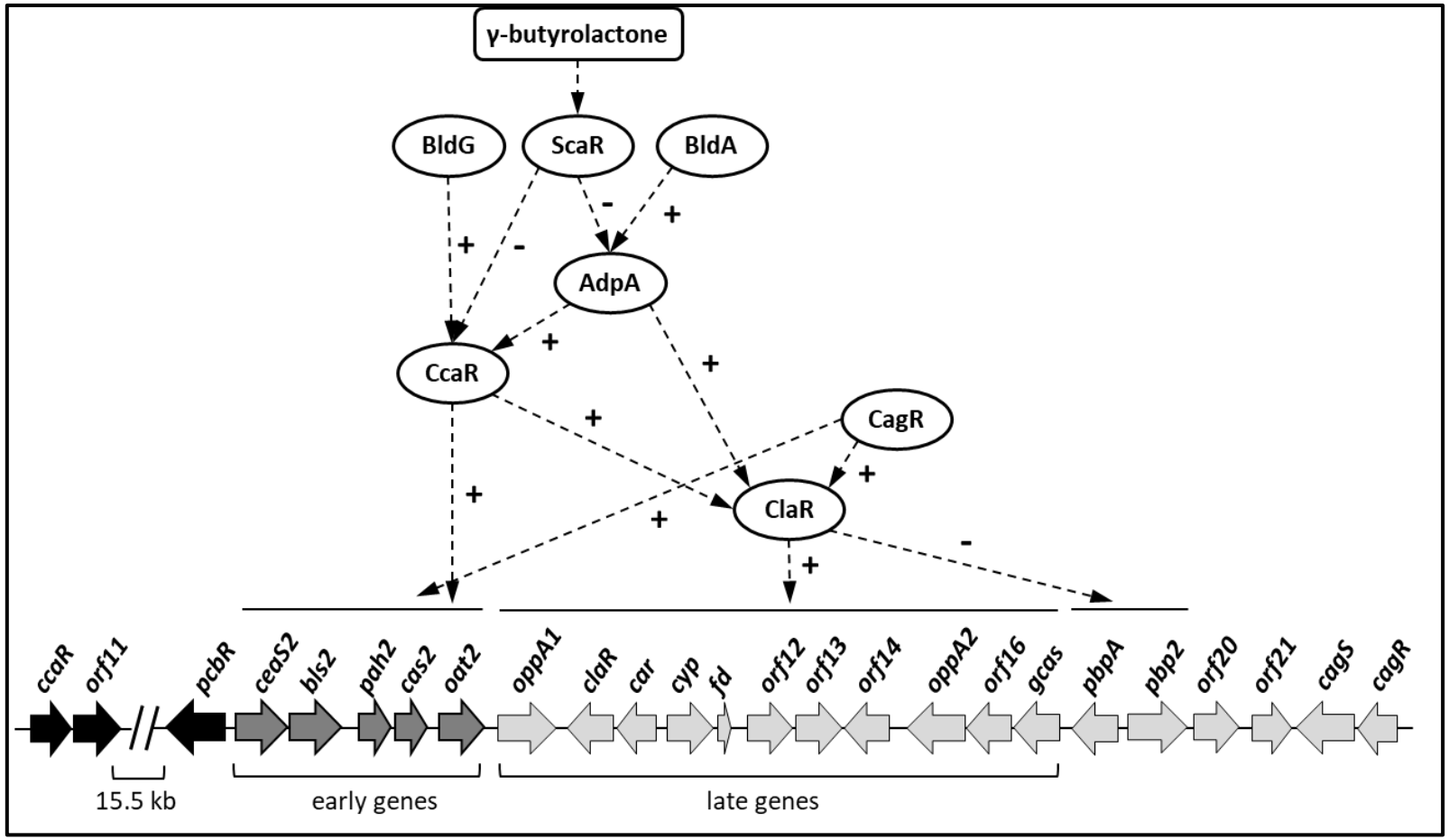


Figure 1.8 A proposed scheme for the regulation of the clavulanic acid BGC based on published studies. Regulator proteins are in oval shapes. Plus (+) sign indicates positive regulation by activating gene expression, and minus (-) sign indicates negative regulation by suppressing transcription.

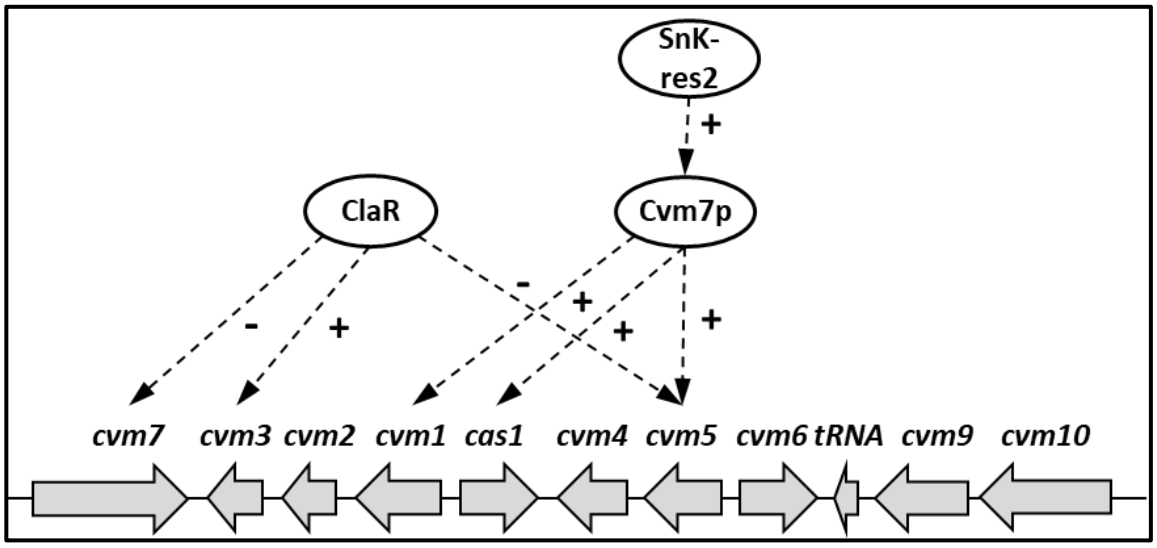


Figure 1.9. A proposed scheme for the regulation of the 5*S* clavam BGC based on published studies. Regulator proteins are in oval shapes. Plus (+) sign indicates positive regulation by activating gene expression, and minus (-) sign indicates negative regulation by suppressing transcription.

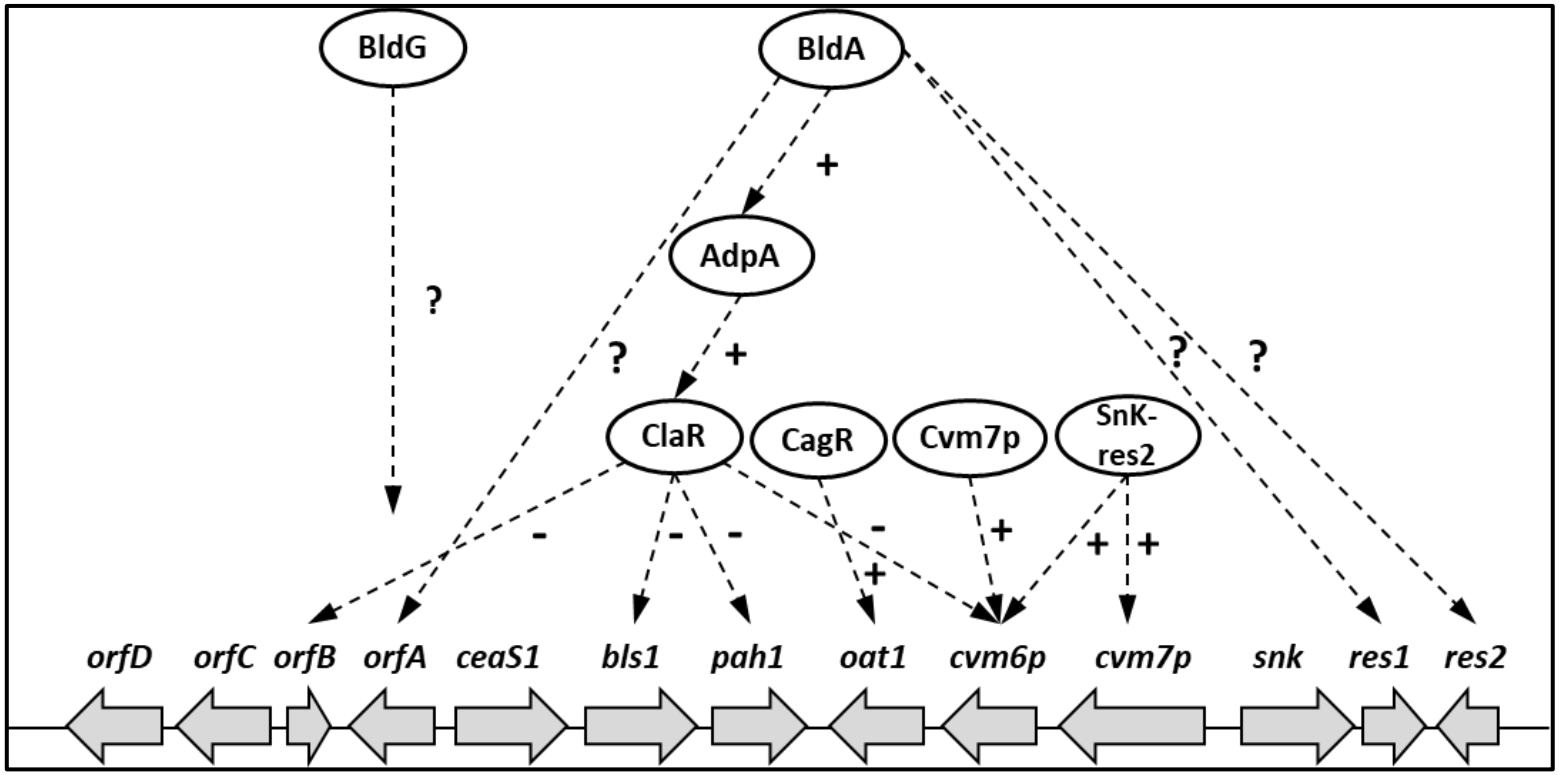


Figure 1.10. A proposed scheme for the regulation of the paralogue gene cluster based on published studies. Regulator proteins are in oval shapes. Plus (+) sign indicates positive regulation by activating gene expression, minus (-) sign indicates negative regulation by suppressing transcription, and question mark (?) sign indicates proposed regulation pathways that need further evidence.

1.5.2. Tables

Table 1.1. β-Lactamase inhibitor characteristics and their combinations. (Adapted from Toussaint and Gallagher (2015); Bush and Bradford (2016); and Viana-Marques et al. (2018).

Inhibitor	Subclass	Spectrum	Combined β-lactam antibiotics	Status	Reference
Clavulanic acid	clavam	Class A narrow spectrum and ESBLs[a]. Some class D enzymes.	Amoxicillin	Approved by FDA[b] and EMA[c]	(Toussaint & Gallagher, 2015; Viana-Marques et al., 2018)
			Ticarcillin		
Sulbactam	Penicillanic acid sulfone	Class A narrow spectrum and ESBLs.	Ampicillin	Approved by FDA and EMA	(Viana-Marques et al., 2018)
			Cefoperazone	Approved by FDA and EMA	(Viana-Marques et al., 2018)
			ETX2514	Phase 1 completed	(Docquier & Mangani, 2018)
Tazobactam	Penicillanic acid sulfone	Class A narrow spectrum and ESBLs. Some class D enzymes.	Piperacillin	Approved by FDA and EMA	(Bush & Bradford 2016)
			Ceftolozane		(Docquier & Mangani, 2018)
			Cefepime	Used in Asia	(Viana-Marques et al., 2018)
Avibactam	Diazabicyclooctane	Class A narrow spectrum, ESBLs, and carbapenemases.	Ceftazidime	Approved by FDA and EMA	(Docquier & Mangani, 2018)

		Some class C and class D enzymes.	Ceftaroline	Phase 2 in progress	(Viana-Marques et al., 2018)
			Aztreonam		
Relebactam	Diazabicyclooctane	Class A narrow spectrum, ESBLs, and carbapenemases. Some class C enzymes.	Imipenem	Phase 3 in progress	(Bush & Bradford 2016)
Vaborbactam	Boronic acid	Class A narrow spectrum, ESBLs, and carbapenemases. Some class C enzymes.	Meropenem	Approved by FDA	(Docquier & Mangani, 2018)
AAI101	Penicillanic acid sulfone	Class A narrow spectrum and ESBLs. Some class C enzymes	Cefepime	Phase 2 in progress	(Papp-Wallace et al., 2011)
Nacubactam (RG6080)	Diazabicyclooctane	Class A narrow spectrum, ESBLs, and carbapenemases. Some class C enzymes	Not selected	Phase 1 complete	(Mushtaq et al., 2019)

[a] ESBLs: Extended-Spectrum β-Lactamases
[b] FDA: U.S. Food and Drug Administration
[c] EMA: European Medicines Agency

Table 1.2. Clavulanic acid (CA) and cephamycin C (Ceph-C) production in *S. clavuligerus* mutants with defects in genes from the clavulanic acid biosynthetic gene cluster. Phenotypes for some *S. clavuligerus* gene mutants that have homologues in clavulanic acid-like gene clusters of non-producers are also included.

Gene	Product (function)	Metabolite production [a]		Reference
		Ceph-C	CA	
ceaS1/2 [b]	Carboxyethylarginine synthase (biosynthesis)	Yes	No	(Pérez-Redondo et al., 1999; Jensen et al., 2000; Tahlan et al., 2004)
bls1/2 [b]	β-Lactam synthetase (biosynthesis)	Yes	No	(Bachmann et al., 1998; Jensen, et al., 2000; Tahlan, et al., 2004)
pah1/2 [b]	Proclavaminic acid amidinohydrolase (biosynthesis)	Yes	No	(Aidoo et al., 1994; Jensen et al., 2004b)
cas1/2 [b]	Clavaminic acid synthase (biosynthesis)	Yes	No	(Jensen et al., 2000; Mosher et al., 1999)
oat1/2 [b]	Ornithine acetyltransferase	Yes	Yes	(de la Fuente et al., 2004; Tahlan, et al., 2004)
oppA1	Oligopeptide transporter	Yes	No	(Jensen et al., 2000; Lorenzana et al., 2004)
claR	Transcriptional activator (regulation)	Yes	No	(Paradkar et al., 1998; Jensen et al., 2000; Martínez-Burgo et al., 2015)
car (cad)	Clavaldehyde reductase or dehydrogenase (biosynthesis)	Yes	No	(Jensen et al., 2000)
cyp (orf10)	Cytochrome P-450 (biosynthesis)	Yes	No	(Li et al., 2000; Jensen et al., 2000; Mellado et al., 2002)
fd (orf11)	Ferredoxin	Yes	70-80% of wt	(Jensen et al., 2004a)
cpe (orf12)	β-Lactamase-like protein (biosynthesis)	Yes	No	(Jensen et al., 2004a; Li et al., 2000; Srivastava et al., 2019)
orf13	Membrane transport protein	Yes	No	(Jensen et al., 2004a)
orf14	Acetyltransferase (biosynthesis)	Yes	No	(Mellado et al., 2002; Jensen et al., 2004a)

oppA2 (orf15)	Oligopeptide transporter (biosynthesis)	Yes	No	(Jensen et al., 2004a; Lorenzana et al., 2004; Alvarez-Alvarez et al., 2018)
orf16	*N*-Acetyltranferase (biosynthesis)	Yes	No	(Jensen et al., 2004a)
gcas (orf17)	*N*-glycyl-clavaminic acid synthetase (biosynthesis)	Yes	No	(Jensen et al., 2004a)
pbpA (orf18)	Penicillin binding protein	NA [c]	NA [c]	(Jensen et al., 2004a)
pbp2 (orf19)	Penicillin binding protein	Yes	Yes	(Jensen et al., 2004a)
orf20	Cytochrome P-450	Yes	Yes	(Jensen, 2012)
orf21	RNA polymerase σ factor (regulation)	Yes	Yes	(Jnawali et al., 2008; Song et al., 2009)
orf22 (*cagS*)	Two-component system histidine kinase (regulation)	Yes	Yes	(Song et al., 2009; Fu et al., 2019)
orf23 (*cagR*)	Two-component system response regulator (regulation)	47% of wt	40% of wt	(Jnawali et al., 2008; Song et al., 2009; Fu et al., 2019)
ccaR	Transcriptional activator (regulation)	No	No	(Alexander & Jensen, 1998; Perez-Llarena et al., 1997)
pcbR	Penicillin binding protein (resistance)	Yes	Yes	(Paradkar et al., 1996)
orf11	Unknown	Yes	Yes	(Alexander and Jensen, 1998)
nocE	Lipases/esterases	Yes	Yes	This study

a >95% level of production when compared to wild type *S. clavuligerus* is reported as "Yes" and <5% production is reported as "No"

b There are two copies each of these genes in the clavulanic acid, clavam and/or paralogue gene clusters of *S. clavuligerus,* and phenotypes of double disruption mutants are reported.

c NA: not applicable. Mutants could not be obtained, and the gene was proposed to be essential for survival in *S. clavuligerus*.

Table 1.3. 5*S* clavams production in *S. clavuligerus* mutants with defects in genes from clavam BGC and paralogue BGC.

Gene	Predicted Function	Effect on 5*S* clavams	Reference
Clavam biosynthetic gene cluster			
cas1	Clavaminic acid synthase isoenzyme	Decrease production in 2HMC and C2C, no effect on alanylclavam	(Mosher et al., 1999)
cvm1	Aldo–keto reductases	No production	
cvm2	Isomerases	severely reduction in alanylclavam and 2HMC, No production of C2C	(Tahlan et al., 2007)
cvm3	Putative flavin reductase	No effect	
cvm4	Acetyltransferase	No effect	
cvm5	Baeyer–Villiger oxidation	No production	
cvm6	Putative aminotransferase	No effect	
cvm7	Pimaricin regulator PimR	No effect	
cvm9	Transcriptional regulator	No effect	
cvm10	Protein Kinase	No effect	
cvm11	Efflux Protein	No effect	
cvm12	Transcriptional regulator	No effect	
cvm13	Asparaginase	No effect	
cvmG	Putative secreted protein	No effect	
cvmP	Arginine deiminase	No effect	
cvmH	Hydrolase	Not done	
Paralogue gene cluster			
ceaS1	Carboxyethylarginine synthase	Decrease production	(Pérez-Redondo et al., 1999; Jensen et al., 2000; Tahlan et al., 2004)
bls1	β-lactam synthetase	Decrease production	(Bachmann et al., 1998; Jensen et al., 2000; Tahlan et al., 2004)
pah1	Proclavaminate amidinohydrolase	Decrease production	(Aidoo et al., 1994)
oat1	Ornithine acetyltransferase	No effect	(de la Fuente et al., 2004; Tahlan, et al.,

			2004)
cvm6p	Putative aminotransferase	No production	(Tahlan et al., 2007)
cvm7p	Large bi-domain transcriptional regulator	No production	
snk	Two-component system	No production	(Kwong et al., 2012)
res1	Two-component system	Increase production	
res2	Two-component system	No production	
orfA	Hydroxymethyltransferase	No alanylclavam production	(Zelyas et al., 2008)
orfB	Amino acids biosynthesis regulator	No alanylclavam production	
orfC	Aminotransferase	No alanylclavam production	
orfD	Threonine dehydratase	No alanylclavam production	

CHAPTER II

Materials and Methods

2.1. Bacterial culture conditions and general procedures

2.1.1. Bacterial strains, cultivation, and maintenance

Escherichia coli strains were routinely cultivated in Luria-Bertani (LB) Lennox medium (Fisher Scientific, Canada) and incubated at 37°C with 200 rpm shaking for 16-18 hours. Low sodium LB medium (1% w/v tryptone, 0.5% w/v yeast extract, 0.25% sodium chloride) was used for *E. coli* strains containing the pHM11a vector. When required, media were supplemented with the appropriate antibiotics as listed in (Table 2.1). *E. coli* strains were maintained on LB agar at 4 °C for short-term storage and -80 °C in 20% v/v glycerol for long-term storage (Sambrook & Russell, 2001). All bacterial strains used in this study are listed in (Table 2.2).

2.1.2. Preparation and transformation of chemically competent cells

Transformation of plasmid and construct DNA was conducted using commercial *E. coli* NEB5α competent cells (New England Biolabs, Canada) according to the manufacturer's instructions. In addition, preparation, and transformation of chemically competent *E. coli* DH5α and ET12567/pUZ8002 strains were carried out as described in the protocol of European Molecular Biology Laboratory (EMBL: www.embl.de), which is based on Inoue et al. (1990).

2.2. General *Streptomyces* procedures

2.2.1. *Streptomyces* strains, cultivation, and maintenance

Streptomyces strains were maintained on International *Streptomyces* Project medium 4 (ISP-4) agar (Difco, USA) or in Trypticase soy broth (TSB: BD Biosciences, Canada) medium. In the case of *S. clavuligerus*, TSB was supplemented with 1% (w/v) soluble starch (TSB-S) (Fisher, USA). For fermentation, metabolite analysis, genomic DNA extraction, and RNA isolation, *Streptomyces* strains were grown in liquid media and incubated at 28 °C with agitation at 250 rpm with stainless-steel springs (Kieser et al., 2000). The *Streptomyces* cultures were supplemented with the appropriate antibiotics when required (Table 2.1). Glycerol stocks were prepared by scraping spores from ISP-4 agar plates into a 1.5 ml microfuge tube containing 20% sterile glycerol before storing at -80 °C (Kieser et al., 2000). All *Streptomyces* strains used in this study are listed in (Table 2.2).

2.2.2. Seeding cultures and fermentation media

Seeding cultures of each *Streptomyces* strain were started from glycerol stocks in 5 ml of TSB or TSB-S media with appropriate antibiotics (if necessary) and incubated at 28 °C for 48 hours in a rotary shaker at 250 rpm with stainless-steel springs. For metabolite production, mycelia from the seed cultures were washed twice with sterile water and used to inoculate 25 ml fermentation medium in Erlenmeyer glass flasks in a 2% (v/v) inoculation (Paradkar and Jensen 1995). Different fermentation media were used in this study: TSB, TSB-S, starch asparagine (SA), soy medium (SM) as described in

Paradkar and Jensen (1995) and Tahlan et al., (2004), MSF (mannitol-soy flour media), and MEY (malt extract-yeast extract medium) as described in Kieser et al. (2000), TBO (tomato Paste-Baby oatmeal media) as described in Higgens et al. (1974), and R5A as described in Rodriguez et al. (2008). The culture flasks were incubated at 28 °C in a rotary shaker at 250 rpm with stainless-steel springs, and culture supernatants were sampled aseptically every 24 hours starting at 48 hours with 1 ml being removed and transferred to clean sterile microfuge tubes to be tested by bioassays or LC/MS. All supernatant samples were stored at -80 °C.

2.2.3. The *Streptomyces* genomic DNA preparation

Chromosomal DNA was isolated from each wild-type strain, *S. clavuligerus* and *S. pratensis*, using the QIAamp DNA Mini Kit (QIAGEN Inc., Canada) with some modification in the protocol. Mycelia from 25 ml cultures (TSB-S for *S. clavuligerus* and TSB for *S. pratensis*) were harvested by centrifugation and washed twice in sterile 10.3% w/v sucrose. Around 0.5 ml of the mycelial pellet was transferred to a 2 ml sterile screw cap tube containing sterile, acid-washed beads (OPS Diagnostics, USA) before adding 200 µl of both ATL and AL buffers from the kit. The samples were homogenized using a SpeedMill PLUS Bead Homogenizer (Analytik Jena AG, Germany) for 2× (3 minutes on and 3 minutes off cycles) protocol. Afterward, samples were centrifuged at 10,000 rpm for 1 minute, and the supernatants were transferred to fresh 1.5 ml microfuge tubes containing 200 µl of 95% ethanol. After a brief vortex, the tubes were centrifuged, and the supernatants were transferred to spin columns and centrifuged at 8,000 rpm for 1

minute. The filtrates were discarded, and 500 µl of AW2 buffer was added to each column before centrifugation at 14,000 rpm for 3 minutes. Filtrates were again discarded, and tubes were centrifuged as above for an additional minute to remove residual buffer and ethanol. The columns were placed into fresh 1.5 ml microfuge tubes, and 100 µl of sterile distilled H_2O was added to each column. The columns were incubated at room temperature for 5 minutes before centrifugation at 8000 rpm for 1 minute to elute the DNA. The samples were stored at -20 °C.

2.2.4. Spore conjugations

Spore conjugations were performed as Kieser et al. (2000) described with some modifications (Tahlan, et al., 2004). A 50 ml culture of the donor *E. coli* strain ET12567/pUZ8002, containing the plasmid to be transferred, was grown in LB supplemented with the appropriate antibiotics (Table 2.1) to an OD_{600} of ~0.4 (between 0.3 and 0.5) (Implen p300 NanoPhotometer, Germany). Cells were centrifuged at 3,500 rpm and washed twice with LB before being resuspended in 0.5 ml of LB. In the meantime, 100 µl *Streptomyces* spores were washed with 0.5 ml of 2× YT medium broth (1.6% w/v tryptone, 1% w/v yeast extract, 0.5% w/v NaCl, in H_2O, pH 7.0), then were heat-shocked for 10 minutes at 50 °C in a water bath. The spores were allowed to cool before mixing with the *E. coli* suspension. The ~ 1 ml mixture of cells was centrifuged at 7,000 rpm for 5 minutes before the supernatant was discarded. The pellet was gently re-suspended in residual broth and the entire suspension was spread onto freshly made AS-1 agar plates (0.01% w/v yeast extract, 0.02% w/v L-alanine, 0.02 w/v % L-arginine,

0.05% w/v L-asparagine, 0.5% w/v soluble starch, 0.25% w/v NaCl, 1.0% w/v Na_2SO_4, 2.0% w/v agar, pH to 7.5 and 1% v/v sterile 1 M $MgCl_2$ added after autoclaving) (Baltz, 1980). The plates were incubated at 28 °C for 16 - 20 hours and were then overlaid with nalidixic acid (40 μg/ml) and the appropriate antibiotics for plasmid transfer selection before being incubated for an additional 4 - 7 days. The exconjugant colonies were picked up and streaked onto Trypticase Soy Agar (TSA) or Trypticase Soy Agar with Starch (TSA-S 1%) with the appropriate antibiotics to get adequate growth. After testing the exconjugants, the new strains were plated onto ISP-4 medium for sporulation and then to prepare the 20% glycerol stocks, which were kept at -80°C for long-term storage.

2.2.5. *Streptomyces* colony PCR method

The *Streptomyces* strains were quadrant-streaked onto nutrient agar (BD Biosciences, Canada) plates with the appropriate antibiotics to grow separated colonies. The plates were incubated at 28 °C for 48 h. A single fresh colony was picked up and crushed in a 1.5 ml microfuge tube before adding 20 μl of sterile dH_2O and mixing. The colony suspension was heated at 100 °C for 10 min, then cooled down in ice for 2 mins. The microfuge tube was centrifuged at 14,000 rpm for 5 min to pellet the cells debris. Around 2.5 μl of the supernatant, which contained the DNA, was added immediately to the prepared PCR mixture.

2.3. Nucleic acids extraction, manipulation, and general procedures

2.3.1. Plasmids and DNA manipulation

Conventional recombinant DNA techniques were carried out following standard protocols (Sambrook & Russell, 2001). All plasmids and constructs used in this study are listed in (Table 2.3). Plasmid DNA was isolated from overnight cultures of *E. coli* using the EZ-10 Spin Column Plasmid DNA kit (Bio Basics Inc, Canada). Alternatively, for routine screening purposes, plasmid DNA was extracted using a modified Birnboim and Doly (1979) method as described in Sambrook and Russell (2001). DNA gel extraction and purification were conducted using the EZ-10 Spin Column DNA Gel Extraction Kit (Bio Basic Canada Inc., Canada). Nucleic acid concentrations were measured using an Implen p300 NanoPhotometer (Implen GmbH, Germany). All restriction enzymes and T4 DNA Ligase enzymes used in this study were obtained from New England BioLabs Ltd. (Canada), and the digestion and ligation reactions were performed according to the manufacturer's instructions.

2.3.2. Primers and polymerase chain reaction (PCR)

All oligonucleotide primers (Supplementary Table S2.1) used in this study for cloning, PCR, reverse transcription PCR (RT-PCR), site-directed mutagenesis, and sequencing were purchased from Integrated DNA Technologies (Coralville, USA). The primer sets for RT-PCR were first optimized for annealing temperatures by performing gradient PCR, using 0.5 ng/µl genomic DNA as a template. For further confirmation, the PCR products for each gene were sequenced. All PCRs were performed using the Taq

DNA polymerase or Phusion High-Fidelity DNA Polymerase kits (ThermoFisher, United States). When required, PCR products were cloned into the pGEM-T Easy vector (Promega, United States) according to the manufacturer's instructions. All the sequencing in this study was conducted at The Centre for Applied Genomics, University of Toronto (Canada).

2.3.3 RNA isolation and RT-PCR

Streptomyces strains were cultured in 50 ml of SA medium, and the mycelia were harvested at 48- and 96-hours time points by centrifugation. Total RNA was isolated from 500 µl of the cell pellet using the innuSPEED Bacteria/Fungi RNA Kit and a SpeedMill PLUS Bead Homogenizer as per the manufacturer's instructions (Analytik Jena AG, Germany). The resulting RNA samples were treated with DNase I (New England Biolabs, Canada) as directed by the manufacturer to remove trace amounts of genomic DNA, after which the DNase-treated RNA samples were quantified and stored at -80 °C. The complementary DNA (cDNA) was synthesized using 500 ng of DNaseI-treated RNA using random hexameric primers and reverse transcriptase enzyme provided with the Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Scientific, United States). RT-PCR was performed using 2 µl of the cDNA from above in a final volume of 20 µl using Taq DNA polymerase kits (ThermoFisher, United States) and gene-specific primers (supplementary table S2.1). To verify the absence of genomic DNA in the RNA samples, control reactions containing DNaseI-treated RNA preparations without reverse transcription were added for each reaction. The resulting PCR products were analyzed by

electrophoresis using a 2% w/v agarose gel and 1× Tris Borate EDTA (TBE) buffer and were visualized by staining with ethidium bromide (Alfa Aesar, USA).

2.4. Cloning, gene deletion, and preparation of *Streptomyces* strains

2.4.1. Preparation of *S. clavuligerus*/pHM11a-*cpe*-His

The gene *cpe* (*orf12*) of *S. clavuligerus* was amplified by PCR using a set of primers that have repetitive histidine codons (6× His) added just before the stop codon (at the C-terminus of the encoded protein) (Supplementary Table S2.1). The Phusion high-fidelity DNA polymerase (Fisher Scientific, Canada) with high GC buffer, 5% v/v DMSO and 1 M betaine were used to perform the amplification reactions according to the manufacturer's instructions. The *cpe*-His tagged PCR product was then A-tailed using Taq DNA polymerase before being cloned into the pGEM®-T Easy vector (Promega, USA) and generate pGEMT/*cpe*-His, which was transferred into *E. coli* NEB5α competent. The DNA construct was isolated from an overnight culture of *E. coli*/pGEMT/*cpe*-His and screened for the presence of positive clones by digesting the plasmids with EcoRI enzyme. The positive plasmids were also confirmed by sequencing.

The positive-confirmed pGEMT/*cpe*-His constructs were digested with NdeI and BamHI to liberate *cpe*-His, then ligated into a similarly digested pHM11a vector to give pHM11a/*cpe*-His. The ligation reaction was conducted using T4 DNA Ligase according to the manufacturer's instructions (New England Biolabs, USA). The new construct DNA was transformed into *E. coli* ET12567/pUZ8002 cells before being introduced into *S. clavuligerus*/Δ*cpe* by intergeneric spores conjugation.

2.4.2. Preparation of *S. clavuligerus*/Δ*orf14*/pHM11a-*orf14* and *S. clavuligerus*/Δ*orf14*/ pSET152-*orf14*

The *S. clavuligerus orf14* gene was amplified by PCR using oligonucleotide primers (Supplementary Table S2.1) and Phusion high-fidelity DNA polymerase (Fisher Scientific, Canada). Polyadenylation was achieved for *orf14* PCR product using Taq DNA polymerase before being cloned into the pGEM®-T Easy vector to generate pGEMT/*orf14*. The construct was transformed into *E. coli* NEB5α competent cells. Restriction enzyme digestion using EcoRI and subsequent gel electrophoresis was performed to screen plasmids for the presence of positive clones, which were then confirmed by sequencing. The resulting confirmed products were digested with NdeI and BamHI to liberate *orf14* which was then ligated into a similarly digested pHM11a vector to give pHM11a-*orf14*. The new construct DNA was transformed into *E. coli* ET12567/pUZ8002 competent cells before being introduced into *S. clavuligerus*/Δ*orf14* for complementation and also to be used in the site-directed mutagenesis study.

To generate pSET152-*orf14*, the plasmid pHMlla-*orf14* was digested with BglII and BamHI to release *orf14* with the promoter *ermE*p* and then was ligated into the BamHI-digested pSET152 plasmid to give pSET152-*orf14*, which was then confirmed by sequencing. The new plasmid was transferred into *E. coli* ET12567/pUZ8002 cells before it was introduced into *S. clavuligerus*/Δ*orf14* by intergeneric conjugation for complementation study.

2.4.3. Preparation of *S. clavuligerus*/pHM11a-*orf14* variants by site-directed mutagenesis

For site-directed mutagenesis, the single amino acid variants of *Sc-orf14* (V142A, V254A, T269A, and V292A) were generated using the QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies, USA) along with mutagenic oligonucleotide primers (Supplementary Table S2.1) and pHM11a-*orf14* construct as a template according to the manufacturer's instructions. Mutagenic primers for the desired mutations were designed online with QuikChange® Primer Design Program (https://www.agilent.com/primerdesignprogram.jsp). All site-directed mutations were verified by DNA sequencing, and plasmids expressing Sc-*orf14* variants (Table 2.3) were introduced into the *S. clavuligerus*/Δ*orf14* mutant for complementation studies. The *S. clavuligerus* Δ*cpe* mutant, the complemented strain *S. clavuligerus*/Δ*cpe*/pSET152-*cpe,* and the five *S. clavuligerus*/pSET152-*cpe* variants strains (S173A, S234A, S27A, L89A, and S206A) were recently made and tested in our lab (Srivastava et al., 2019) and used in this study.

2.4.4. Preparation of the *S. clavuligerus* Δ*nocE* and *ermE*p∗-*nocE* Strains

The *S. clavuligerus nocE* gene mutant was prepared using the meganuclease I-SceI marker-less gene deletion system (Fernández-Martínez & Bibb, 2014). DNA fragments (~1.2 kb each) containing regions immediately upstream and downstream of *nocE* from the *S. clavuligerus* chromosome were amplified using PCR along with engineered primers (Supplementary Table S2.1) and were separately cloned into the

pGEM-T Easy vector. The upstream fragment was released from pGEM-T Easy by digestion with *Hin*dIII and *Eco*RI and was introduced into the same sites of pIJ12738 (Figure 4.3A) to give pIJ12738/*nocE*-UP. The downstream fragment was then introduced into the EcoRI and XbaI sites of pIJ12738/*nocE*-UP to give pIJ12738/*nocE*-UPDN, which functioned as the *nocE* disruption construct (Table 2.3, Figure 4.3B). pIJ12738/*nocE*-UPDN was conjugated into *S. clavuligerus* to obtain the apramycin-resistant single crossover strain, which was confirmed using genomic DNA PCR (Supplementary table S2.1). The plasmid pIJ12742 expressing the I-SceI meganuclease (Table 2.3, Figure 4.3A) was then conjugated into *S. clavuligerus* pIJ12738/*nocE*-UPDN to obtain apramycin and thiostrepton resistant exconjugants, which were made to undergo sporulation at 28 °C without any selection to facilitate double homologous recombination and loss of pIJ12738 from the chromosome. Spore stocks were prepared and re-streaked onto ISP-4 plates without selection and incubated for 5 days at 37 °C to promote the loss of temperature-sensitive pIJ12742. This led to the isolation of the apramycin and thiostrepton-sensitive *S. clavuligerus* Δ*nocE* mutant, which was verified using genomic DNA PCR (Figure 4.3C; Supplementary Table S2.1). To prepare an *S. clavuligerus* strain constitutively expressing *nocE*, the *ermE*p∗ promoter (Bibb et al., 1985) was inserted upstream of the gene in the *S. clavuligerus* chromosome. A 1.1-kb DNA fragment from the 5' end of the gene was amplified by PCR (Supplementary Table S2.1) and was cloned into pGEM-T Easy. The insert was re-isolated as a NdeI and EcoRI fragment and was ligated with similarly digested pIJ8668-*ermE*p∗ to give pIJ8668-*ermE*p∗-*nocE* (Table 2.3, Supplementary Figure S4.1), which was introduced into wt *S. clavuligerus* by

conjugation. This resulted in the *S. clavuligerus ermE*p∗-*nocE* strain, which was confirmed using genomic DNA PCR (Supplementary Table S2.1) and was used to examine the effect of constitutively expressing *nocE* in *S. clavuligerus*.

2.4.5. Preparation of *S. pratensis*/Sc-*cpe*, *S. pratensis*/Sc-*orf14*, and *S. pratensis*/Sc-*cpe-orf14*

The construct pSE152-*cpe* (Srivastava et al., 2019), which is carrying the *cpe* (*orf12*) gene from *S. clavuligerus*, was transferred into *S. pratensis* (wt) through intergeneric conjugation to give *S. pratensis/Sc-cpe*. In addition, the construct pHM11a-*orf14*, which has the gene *orf14* from *S. clavuligerus* (see section 2.4.2.), was moved into both *S. pratensis* (wt) and *S. pratensis/Sc-cpe* through intergeneric conjugation to give *S. pratensis/Sc-orf14* and *S. pratensis/Sc-cpe-orf14*, respectively. The three strains were tested for the production of bioactive substances in broth and solid media.

2.4.6. Preparation of *S. pratensis* Δ*cas2* by insertional inactivation

Two sets of primers (Supplementary Table S2.1) were designed to amplify two regions in the gene *cas2* of the CA-like gene cluster in *S. pratensis*. The PCR products for the two regions, Sp-*cas2*-KO-1 (448 bp) and Sp-*cas2*-KO-2 (418 bp), were gel purified and then were digested with HindIII before they were cloned into similarly digested plasmid pIJ773 to construct pIJ773/Sp-*cas2*-KO-1 and pIJ773/Sp-*cas2*-KO-2, respectively. The two constructs were confirmed by sequencing and moved to *S. pratensis*

(wt) by conjugation to achieve single crossover insertional inactivation in Sp-*cas2* gene (Figure 5.10A) (Kieser et al., 2000).

The same as above was conducted for *carE* gene from the Carb4550-like gene cluster of *S. pratensis*. Two sets of primers (Supplementary Table S2.1) were used to amplify Sp-*carE*-KO-1 (299 bp) and Sp-*carE*-KO-2 (479 bp) regions, the PCR products were gel purified and cloned into pIJ773 to give pIJ773/Sp-*carE*-KO-1 and pIJ773/Sp-*carE*-KO-2, respectively, and they were confirmed by sequencing.

2.4.6.1. Confirmation of the insertional inactivation in *cas2* gene of *S. pratensis*

After conjugation, the ex-conjugant isolated colonies were picked up and streaked on TSA (+ Apr^{25} +Nal^{40}) plates. PCR was conducted to screen for mutants with successful insertional inactivation for Sp-*cas2* gene. Two sets of primers were used for that purpose; the first set (cas2-conf-F and cas2-conf-R) (Supplementary Table S2.1) was used to confirm *cas2* disruption in *S. pratensis*, and the second set of primers (cas2-conf-F and T3) (Supplementary Table S2.1) was used to verify that the constructs pIJ773/Sp-*cas2*-KO-1/2 were successfully integrated with the bacterial chromosome at Sp-*cas2* gene. The PCR products were sent for sequencing for further confirmation.

The successful *S. pratensis* Δ*cas2* mutants were streaked onto soy media plates to test the production of the bioactive substances. The agar plug bioassays were conducted on TSA plates on day seven of the growth as described in Section 2.6.2. Since the single crossover insertional inactivation is unstable and to be confident that the mutation has not reverted to wild type, the second run of colony PCR was carried out for the cells on the

agar plugs that had been tested in the bioassay. The confirmed positive strains were streaked on ISP-4 (+Apr^{25}) plates for sporulation and preparing glycerol stocks.

2.5. *Streptomyces* growth measurements

2.5.1. Growth curve measurements of *S. clavuligerus*/Δ*nocE* and *S. clavuligerus*/*ermE*p*-*nocE* in liquid media

Growth curve measurements were performed for *S. clavuligerus* (wt), *S. clavuligerus*/Δ*nocE,* and *S. clavuligerus*/*ermE*p*-*nocE*. Three types of media, soy, starch asparagine, and TSB-S, were inoculated in triplicate from 40 h seed cultures. One milliliter of sample was collected from each flask every 24 hours until 144 hours. A simplified diphenylamine colorimetric method based on DNA extraction and quantification was used for growth curve measurements as described previously in Burton (1968) and Zhao et al. (2013). The cell pellet from 0.5 ml cultures were washed twice with 0.5 ml sterile dH_2O and resuspended in 1 ml of diphenylamine reagent [1.5% (w/v) diphenylamine, 1.5% (v/v) concentrated H_2SO_4 prepared in 100 ml glacial acetic acid, and 500 µl of 1.6% aqueous acetaldehyde] (Burton, 1968). The mixture of cells and diphenylamine reagent were incubated at 60 °C for 1 hour. The tubes were centrifuged, and 150 µl of each sample's supernatants were transferred into a 96 – well microtiter plate. The DNA concentrations were measured based on absorbance at 595 nm using a multifunctional microtiter plate reader (Synergy Hybrid Reader, BIOTEK, USA). The statistical analysis (ANOVA repeated measure) was performed using R 3.4.3 (Snee, 1972).

2.5.2. Growth assessment of *S. clavuligerus*/Δ*nocE* and *S. clavuligerus*/*ermE*p*-*nocE* strains on solid media.

To assess the growth characteristics of *S. clavuligerus* (wt), *S. clavuligerus*/Δ*nocE,* and *S. clavuligerus*/*ermE*p*-*nocE* on solid media, spore suspensions with a concentration of 4×10^4 spores/µl were diluted in 10-folds, and 5 µl from each strain and each dilution were inoculated in spots (Figure 4.7) on 4 different types of media plates, SA, ISP-4, TSA-S, and Minimal medium-Starch (MM-S; 0.05% (w/v) L-asparagine, 0.05% (w/v) K_2HPO_4, 0.02% (w/v) $MgSO_4.7H_2O$, 0.001% (w/v) $FeSO_4.7H_2O$, 1% (w/v) starch, and 1.8% (w/v) agar). The plates were incubated for up to 7 days, and colony pictures were taken every 24 h to observe the growth (Figure 4.7).

2.6. Metabolite detection by bioassays

2.6.1 Disc-diffusion bioassays

Liquid bioassays using the disc-diffusion method were performed to detect the production of CA, Ceph-C, or 5*S* clavams from the supernatant of *Streptomyces* cultures, and different indicator microorganisms were used for that purpose. The production of CA was detected by the zone of growth inhibition of *Klebsiella pneumoniae* on media agar plates. TSA medium was prepared in Petri dishes by adding 100 µl of an overnight culture of *K. pneumoniae* (grown in TSB) and penicillin G (6 µg/ml). Sterile 10 mm filter paper discs (Whatman, UK) were aseptically placed on the agar, and 10 µl of supernatants were spotted onto the filter paper discs. As a control, additional TSA plates were set up the same way but without adding penicillin G to the medium. The plates were incubated

overnight at 37 °C right-sides-up, and the zones of inhibition were measured and recorded. For a large number of samples, a 22 × 22 cm plastic bioassay tray was used instead of Petri dishes. In some experiments for testing different parameters, *Enterobacter cloacae* KM31 was used instead of *K. pneumoniae*, and ampicillin (100 μg/ml) instead of penicillin G.

To detect the production of cephamycin C, the *E. coli* ESS, a supersensitive strain to β-lactam compounds, was used as an indicator microorganism. TSA plates were prepared by adding 100 μl of an overnight culture of *E. coli* ESS. Sterile 10 mm filter paper discs were placed on top of the agar, and 10 μl of supernatants were added to them. The plates were then incubated overnight at 37 °C, and the zones of inhibition were measured and recorded.

To detect the production of clavams (2-hydroxymethylclavam and alanylclavam) *Bacillus* sp. ATCC 27860 was used as the indicator organism (Pruess & Kellett, 1983; Zelyas et al., 2008). In this bioassay, Davis-Mingioli agar medium (DMM; 0.8% (w/v) D-glucose, 0.05% (w/v) sodium citrate.$3H_2O$, 0.7% (w/v) K_2HPO_4, 0.3% (w/v) KH_2PO_4, 0.01% (w/v) $MgSO_4.7H_2O$ and 0.1% (w/v) $(NH_4)SO_4$, pH7) was prepared and inoculated with 100 μl overnight culture of *Bacillus* sp. ATCC 27860 grown in TSB. As a control, another DMM plate was supplemented with 200 μg/ml with methionine to antagonize the activity of alanylclavam. Sterile filter paper discs were placed on top of the agar, and 30 μl of supernatant samples were added onto the filter paper discs. The plates were incubated overnight at 37 °C right-sides-up, and the zones of inhibition were measured and recorded.

2.6.2. Agar plug diffusion bioassay

Streptomyces pratensis was streaked onto appropriate solid agar plates and incubated for seven days at 28°C. Agar-plugs or cylinders were cut aseptically with a sterile cork borer and deposited onto the agar surface of TSA plate previously inoculated with the indicator microorganism and the appropriate antibiotics (if needed) as mentioned in Section 2.6.1. The TSA plates with plugs were incubated overnight at 37°C right-side-up. The antimicrobial activity of the substances secreted by *S. pratensis* was detected by forming a growth inhibition zone around the agar plug; the zones of inhibitions were measured and recorded.

2.6.3. Agar-plot diffusion bioassay with using cellophane membrane

Streptomyces pratensis was streaked onto appropriate solid agar plates and incubated at 28 °C. On day seven, agar-plots (~28 mm diameter) were cut aseptically and placed onto a sterile cellophane membrane (75 mm diameter; membranes provided by Dr. Bignell, Memorial University of Newfoundland) on top of the agar surface of the TSA plate. The cellophane membrane allows the metabolites in the agar-plots to diffuse into the TSA medium while preventing the bacterial mycelia from penetrating the agar. The TSA plates with cellophane and *S. pratensis* agar-plots were incubated at 28 °C right-side-up for 48 h, after which the cellophane membrane and agar-plots were removed. Ten ml of melted TSA (0.8% agar) plus appropriate indicator microorganism and antibiotics were then poured over the TSA and solidified before being incubated overnight at 37 °C. The antimicrobial activity of the substances secreted by *S. pratensis* was detected by the

formation of the inhibition zone on the place of the agar-plot. As a control, the supernatant of *S. clavuligerus* (Sc) and *Streptomyces cattleya* (Scat; producer of β-lactam thienamycin), and clavulanic acid solution (CA; clavulanate sodium 10 μg) were applied in disc-diffusion bioassays.

2.7. Protein crosslinking, extraction, and detection

Streptomyces clavuligerus/cpe-His_6 strain was cultured in 100 ml (×6 flasks) of SA fermentation medium. One flask of *S. clavuligerus* wt was included in the experiment as a control. All flasks were incubated at 28°C for 48 h with continuous shaking. Mycelia pellets were collected by centrifugation at 3500 rpm for 7 minutes. One ml of the supernatants from each flask were stored at – 80 °C to be tested for CA production.

2.7.1. Protein crosslinking

Protein crosslinking using formaldehyde was performed as described in Chowdhury et al. (2009) and Lougheed et al. (2014) with some modifications. As an optimization step, four different concentrations of formaldehyde 1%, 2%, 3%, and 5% (v/v final concentration) were added to 3 ml of cell pellets for a total volume of 10 ml suspended pellets were then incubated for 15 min at room temperature with gentle rocking. The crosslinking reaction was quenched by adding 2.5 M of ice-cold glycine to a final concentration of 0.125 M. Afterwards, cells were centrifuged, washed with sterile cold dH_2O, and subjected to protein extraction.

2.7.2. Protein extraction and purification using nickel affinity resins

After crosslinking, cellular protein extraction was conducted as basically described in Ferguson et al. (2016). The mycelial cells were resuspended in PBS (0.8% w/v NaCl, 0.02% w/v KCl, 0.14% w/v Na_2HPO_4, 0.02% KH_2PO_4, pH 7.4) + 0.01% SDS. Three ml of each mycelial suspension was transferred to 3 ml screw cap cryovials (Fisher Scientific, Canada) and sonicated on ice using a QSonica 56 sonicator (Q125-110, VWR, Canada) with a 5/64-inch probe (Fisher Scientific, Canada). The sonication program consisted of 6 cycles of 15 seconds on and 15 seconds off for a total of 3 minutes.

After sonication, the cells debris was removed by centrifugation at 4000 rpm for 10 minutes, and the supernatant containing soluble proteins was stored on ice. Proteins of interest were purified using HisPur^{TM} Nickle-nitrilotriacetic acid (Ni-NTA) resin system (Thermo Fisher Scientific, USA) according to the manufacturer's recommendations with some changes. Ni-NTA resin is specifically designed to purify recombinant proteins fused to the 6× histidine (6×His) tag expressed in bacteria.

In the first step, 3 ml of 2× Equilibration Buffer (1×: 20 mM sodium phosphate, 300 mM sodium chloride and 10 mM imidazole, pH 7.4) was added to the same amount (3 ml) of sample supernatants. At the same time, 100 µl of Ni-NTA agarose resin were washed with 500 µl 1× Equilibration Buffer and resuspend again in fresh 100 µl 1× Equilibration Buffer before adding them to the samples' supernatants. The mixtures were incubated at room temperature for 30 min with well rocking. Next, the tubes were centrifuged at 700 ×g for 2 min, and the supernatants were saved at -80 °C. The resin was washed twice with two resin-bed volumes of Washing Buffer (20 mM sodium phosphate,

300 mM sodium chloride and 25 mM imidazole, pH 7.4). In the last step, the *cpe*-His tagged proteins were eluted by adding 200 µl of Elution Buffer (20 mM sodium phosphate, 300 mM sodium chloride and 250 mM imidazole, pH 7.4), then the tubes were centrifuged at 700 ×g for 2 min, and carefully transferred to fresh tubes to be analyzed by Western blot, or the samples were saved at -80 °C.

2.7.3. Protein analysis by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot.

To visualize proteins, the eluted protein samples were subjected to SDS-PAGE on a 12% (w/v) polyacrylamide gel. A total of 20 µl of each sample was loaded onto an SDS-PAGE (resolving gel: 12% w/v acrylamide/bis-acrylamide, 50% w/v 1.5 M Tris-HCl-pH 8.8, 0.01% w/v each of ammonium persulfate (APS) and SDS, and 0.1% v/v Tetramethylethylenediamine (TEMED); stacking gel: 3.2% w/v acrylamide/bis-acrylamide, 25% w/v 0.5 M Tris-HCl-pH 6.8, 0.01% w/v each of APS and SDS and 0.1% v/v TEMED). The thermo scientific PageRuler Plus Prestained Protein ladder (5 µl each) were loaded and used as protein size markers.

The electrophoresis was run using a Bio-Rad Mini-PROTEAN® Tetra System (Bio-Rad, Canada) at 150 V for ~1.5 h with 1× Tris-glycine electrophoresis buffer (50 mM Tris-HCl pH 8.3, 380 mM glycine, and 0.1% w/v SDS). Western blot analysis was then performed to transfer the proteins from SDS-PAGE gels to Amersham™ Hybond™-ECL nitrocellulose membranes (GE Healthcare, Canada), using the Bio-Rad Trans-Blot® Cell according to manufacturer's instructions. The membranes were washed with Tris-

buffered saline-Tween (TBS-T; 50 mM Tris-HCl pH 7.6, 150 mM NaCl, and 0.5% v/v Tween-20) and were blocked overnight at 4˚C in blocking buffer (TBS-T with 10% w/v non-fat milk). The membranes were probed using the primary antibody "anti-6×His" Monoclonal Antibody (Thermo Fisher Scientific Canada), used at a 1:1000 dilution. The secondary antibody was an anti-Mouse IgG2b Secondary Antibody HRP (Thermo Fisher Scientific Canada) used at a 1:2000 dilution. Signals were visualized using an ImageQuant™ LAS 4000 Digital Imaging System (GE Healthcare Canada).

2.8. Specific experiments and analysis for *S. pratensis* project

2.8.1. Time-course production of the bioactive substances from *S. pratensis*

Streptomyces pratensis spores (~50 µl) were streaked onto soy medium (SM) and beef extract-starch (BES) medium [0.6 % (w/v) beef extract, 2% (w/v) soluble starch, 1.8% (w/v) agar] plates, and were incubated at 28 °C for up to 11 days. Time-course agar plugs bioassays were performed every 24 h starting from day 2 of the incubation. The agar bioassays were conducted on TSA plates, and *K. pnuemoniae* were used as indicator organisms as described in Section 2.6.2. The zones of growth inhibition were measured and recorded. Agar plugs from blank SM and BES plates were used as a negative control.

2.8.2. Fermentation of *S. pratensis* strains in different types of broth media

Wild-type *S. pratensis* and the strains *S. pratensis/Sc-cpe*, *S. pratensis/Sc-orf14*, *S. pratensis/Sc-cpe-orf14* were cultured in 25 ml of 6 different types of media: MEY, MSF, R5A, SA, SM, TBO, and TSB without antibiotics. The cultures were incubated at 28 °C

(see the whole protocol in Section 2.2.2). Supernatant samples were collected aseptically every 24 hours starting at 48 hours, and the disc-diffusion bioassays were conducted to investigate the production of bioactive substances. All supernatant samples were stored at -80 °C.

2.8.3. Growing of *K. pneumoniae* on SM agar with *S. pratensis*

A plate of SM agar was streaked with *S. pratensis* and incubated at 28 °C. On day seven of incubation, the agar was cut into two halves and placed upside down in a sterile plate (Figure 5.3), and 15 µl of Penicillin G (6 µg/µl final concentration) were spread on one half while the other half remained without antibiotics. As a control, the same was conducted for a blank SM plate without *S. pratensis*. Five microliter of *K. pneumoniae* overnight culture were inoculated at three spots onto each half of the agar media. The plates were incubated overnight at 37 °C. Coomassie blue stain (50%) was added to the agars and contrasted with the *K. pneumoniae* colonies to be scored visually.

2.8.4 The bacteriostatic effect assay for the bioactive substances of *S. pratensis*

An agar-plot bioassay against *K. pneumoniae* was conducted for *S. pratensis* growing on SM agar as described in Section 2.6.3. As positive controls, disc-diffusion bioassays were performed using clavulanic acid solution (CA; clavulanate sodium 10 µg) and supernatant from *S. clavuligerus* culture (Sc). After the overnight incubation, agar-plugs were taken aseptically from the zone of growth inhibition of *K. pneumoniae* and placed onto the agar surface of TSA plate. As a control, agar-plugs from the *K.*

pneumoniae growth area were taken and placed in the same TSA plate. The plate was incubated at 37 °C for 5 days to monitor the re-growth of *K. pneumoniae* and determine the bacteri(cidal/ostatic) effect of the bioactive substances produced by *S. pratensis*. To confirm that the re-growth of *K. pneumoniae* on the plugs was due to the bacteriostatic effect of the bioactive substances produced by *S. pratensis*, and not because of resistance mechanisms developed by *K. pneumoniae* itself, a subsequent agar bioassay was conducted using the re-grown *K. pneumoniae* as an indicator microorganism.

2.8.5. Testing the bacteriostatic and bactericidal effects of some antibiotics

Disc diffusion bioassays were carried out against *K. pneumoniae* according to the standard protocol for the Kirby-Bauer disk diffusion susceptibility test (Hudzicki, 2009). A TSA plate was inoculated with an overnight culture of *K. pneumoniae*, and four different antibiotics discs were placed on the surface of the agar, the two bactericidal antibiotics gentamycin (GM, 10 μg) and streptomycin (S, 10 μg), and the bacteriostatic antibiotics tetracycline (T, 30 μg) and chloramphenicol (C, 30 μg). After overnight incubation at 37°C, agar plugs were taken aseptically from the zones of no growth and placed on the surface of a fresh TSA plate, which was then incubated for five days to monitor the re-growth of *K. pneumoniae*.

2.8.6. RNA isolation from *Streptomyces pratensis* cultured on solid agar media

Wild type *Streptomyces pratensis* spores (50 μl) were spread on the surface of an SM plate and incubated at 28 °C. On day 7 of incubation, the agar plate was cut into two

halves, and agar bioassays were performed to confirm the production of the bioactive substances before the isolation of RNA. One half of the SM agar was placed onto a large TSA plate (22×22 cm) with *K. pneumoniae* only and no antibiotics. The other half was placed onto TSA plate with *K. pneumoniae* and PenG (60 µg/ml), and the two TSA plates were incubated overnight at 37 °C. After confirming the zone of growth inhibition (ZOI) around the half plot of SM agar, the mycelial layer of *S. pratensis* was gently and carefully scraped from each half and resuspended in 2 ml of diethyl pyrocarbonate (DEPC) treated dH_2O (0.1% v/v). The mycelial suspensions were vortexed and centrifuged at 4000 rpm, and the pellet was resuspended again with 500 µl of DEPC treated dH_2O and used for RNA extraction following the standard protocol as described in Section 2.3.3.

2.9. Liquid chromatography-mass spectrometry (LC-MS and LC-MS/MS) analysis

Streptomyces strains were grown for fermentation studies, and 96-h culture supernatants were analyzed for clavulanic acid and 5*S* clavams production using high-performance liquid chromatography-mass spectrometry [HPLC/MS (TOF)] (1260 Infinity LC-6230 TOF LC/MS, Agilent Technologies, USA) as described previously (Srivastava et al., 2019). Supernatants in microcentrifuge tubes were centrifuged at 12,000 rpm for 5 minutes before being filtered using 0.2 µm PTFE membrane filters (VWR, USA). One hundred microliters of the filtered supernatant were derivatized in the dark for 15 minutes with 25 µl of 20.6% imidazole, adjusted to pH 6.8 using 5M HCl (Bird, Bellis, & Gasson, 1982). The derivatized samples were analyzed on an XTerra MS C18 column (2.1 × 150

mm, 3.5 µm, 125 Å; Waters Scientific, USA) at a 0.25 ml/min flow rate. The mobile phase consisted of solvent A (10 mM ammonium bicarbonate, pH 10) and solvent B (acetonitrile) used in a binary gradient system as follows: 100% solvent A for 5 min, linear gradient to 85% solvent A over 20 min, 85% solvent A for 5 min, linear gradient to 100% solvent A over 1 min, and 100% solvent A for 9 min. Eluant was monitored at 311 nm to detect the imidazole derivatives. A control blank sample (HPLC water + imidazole) was set up using the same amount of supernatant with 25 µl filter sterilized HPLC water. The diode array detector (DAD) was set to detect UV absorption at 311 nm, 210 nm, 280 nm, 317 nm, and 350 nm wavelengths, and the electrospray ionization (Dual ESI) MS was in positive ion mode. Chromatographic data were analyzed using MassHunter Workstation version B.05.01 software (Agilent Technologies, USA).

2.10. Metabolomics and molecular networking

Untargeted metabolomics was conducted using bacteria grown on solid media. One hundred microliters of a standardized spore stock (4×10^4 spores/µl) of each species were used to inoculate agar plates in duplicate, and each plate was extracted using 15 ml of methanol or ethyl acetate. Two milliliters of each extract were sent to Dr. Pieter Dorrestein's lab at University of California, San Diego, USA, where the LC-MS/MS analyses was gratefully performed by Dr. Kapil Tahlan. The two milliliter samples were dried, resuspended in 130 µl of 70% (v/v) methanol containing 0.2 µM of amitriptyline (internal standard), and transferred to a 96-well plate, which was centrifuged at 2000 rpm for 15 min at 4 °C. One hundred microliters of each sample were then transferred to a

new 96-well plate for LC-MS/MS analysis. Samples were analyzed using a Vanquish UHPLC System coupled Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Scientific, United States). Chromatographic separation was performed in mixed mode (allowing weak anion/cation exchange) on a Scherzo SM-C18 column (2 × 250 mm, 3 µm, 130 Å; Imtakt, United States) maintained at 40 °C. Ten microliters of each sample were injected for analysis, and the mobile phase consisted of (A) 0.1% formic acid in water and (B) 0.1% formic acid in acetonitrile. Chromatography was performed at a flow rate of 0.5 ml/min using the following program: 0–5 min, 98% A; 5–8 min, gradient of 98–50% A (or 50% B); 8–13 min, gradient 50–100% B; 13–14.00 min, 100% B; 14–14.10 min, 100–2% B; 14.10–18 min, 2% B.

Mass spectrometry was performed using a heated electrospray ionization source (heater temperature, 370 °C and capillary temperature, 350 °C) in either positive or negative ionization mode (± 3000.0 V; S-lens RF, 55; sheath gas flow rate, 55; and auxiliary gas flow rate, 20). MS^1 and MS^2 scans (at 200 *m/z*) were acquired from 0.48 to 16.0 min at a resolution of 35,000 and 17,500, respectively, for the 100–1500 *m/z* range. The automatic gain control (AGC) target value and maximum injection time were set at 5×10^5 and 150 ms. Four MS^2 scans in data-dependent mode were acquired for most abundant ions per duty cycle, with a starting value of 70 *m/z* and an exclusion parameter of 10 s. Higher-energy collision-induced dissociation was performed with a normalized collision energy of 20, 35, and 50 eV. The apex trigger mode was used (2–7 s), and the isotopes were excluded. Inclusion lists of ions for molecules observed in *Streptomyces* extracts were generated from the Dictionary of Natural Products1 and the StreptomeDB

(Lucas et al., 2013) and were used for prioritizing the acquisition of their MS^2 when observed. The raw LC-MS/MS data files were converted to .mzXML format using ProteoWizard (Adusumilli & Mallick, 2017).

2.11. MS data annotation and analysis

Molecular networks were generated using positive and negative ionization mode data in the Global Natural Products Social Molecular Networking (GNPS; Wang et al. 2016). The resulting networks were visualized in Cytoscape (Shannon et al., 2003), allowing nodes associated with uninoculated media controls to be removed. Annotations were first obtained by matching spectra in public libraries (Wang et al. 2016), including NIST173, Metlin, and GNPS public libraries. Library annotations were manually validated using mirror plots (maximum ion mass accuracy = 5 ppm) corresponding to level 2 annotation based on the Minimum Standard Initiative (Spicer et al., 2017). This work was conducted with the immense help from my lab mate Arshad Sheikh.

To generate a heat map using the *S. clavuligerus* wt, Δ*nocE*, and *ermE*p∗-*nocE* strains, feature-based detection, and alignment of positive mode ionization data were performed (parameters: MS^1 noise level of 25000, MS^2 noise level of 1000) using the MZmine 2 toolbox (v2.39) (Pluskal et al., 2010). Chromatograms were built using the ADAP module (parameters: min group size in # of scans = 4, group intensity threshold = 700,000, min highest intensity = 100,000, max *m/z* tolerance = 10 ppm), which were then deconvoluted (parameters: S/N threshold = 10.0, min feature height = 7000000, coefficient/area threshold = 60.0, peak duration range = 0.01–0.5 min, RT wavelet range

= 0.01– 0.1 s). Fragmentation spectra were paired with deconvoluted peaks using 0.02 Da and 0.2 min windows, and LC-MS features were annotated using the Peak-Grouping module (parameters: deisotope = true, remove features without isotope pattern = false, minimal intensity for interval selection = 0.1, minimal intensity overlap = 0.7, minimal correlation = 0.7). Features were aligned in the JoinAligner module (parameters: ppm tolerance = 7, weight for *m/z* = 75.0, retention time tolerance = 0.5 min, weight for RT = 25.0; require same charge state = false, require same ID = false, compare isotope pattern = false). The aligned peaklist was filtered with the row filter module to keep features with at least two isotopic ions, two occurrences, and at least one MS^2 spectrum before gap filling (parameters: intensity = 5%, ppm window = 5, retention time tolerance = 0.15). The aligned peaklist containing 3149 features was exported as a .CSV file, and the spectral data as .MGF files using the GNPS Export module for further processing. The signal intensities of the features (.CSV) were normalized to that of an internal standard (*m/z* 278.189; retention time, 9.2 min), and only 1684 features with an intensity 3-fold higher than in experimental controls (uncultivated media) were retained. MetaboAnalyst4.0 (Chong et al., 2018) was used to perform the hierarchal clustering, which was visualized as a heat map.

2.12. Bioinformatics analysis

Protein sequences analysis and classification were carried out using InterPro (https://www.ebi.ac.uk/interpro) and Phyre2 (Kelley et al., 2015). Homologues of the proteins in this study were identified using NCBI BLAST (online version, blastp, default

settings) (Altschul et al., 1990) with the *S. clavuligerus* amino acid sequences. Amino acid sequence alignments were generated using ClustalW within Geneious 8.1.9 (Biomatters Ltd., New Zealand). The secretory signals for *nocE* homologues were predicted using the SignalP-5.0 Server (Almagro Armenteros et al., 2019).

Phylogenetic trees were constructed from the alignments using the maximum likelihood method in the MEGA 7 program (Kumar et al., 2016). Bootstrap analyses were performed with 1000 replicates in each algorithm. The biosynthetic gene clusters encoding specialized metabolites were identified in the investigated *Streptomyces* using AntiSMASH 4.0 with the default cluster search algorithm (Blin et al., 2017).

2.13. Tables

Table 2.1. Stock and experimental concentrations of antibiotics used throughout this study.

Antibiotic	Final concentration (µg/ml)		
	E. coli	***S. clavuligerus***	***S. pratensis***
Ampicillin	100	100	N/A*
Apramycin	50	50	25
Chloramphenicol	25	N/A*	N/A*
Hygromycin	50 (in liquid media) 100 (in solid media)	100 (in liquid media) 200 (in solid media)	100 (in liquid media) 200 (in solid media)
Kanamycin	50	50	N/A*
Nalidixic Acid	40	40	40
Thiostrepton	N/A*	5	N/A*

*N/A: Not applicable -the antibiotic was not used with this organism.

Table 2.2. Bacterial strains used in this study.

Bacteria/Strain	Genotype/Description	Reference or source
***E. coli* and indicator bacteria strains**		
E. coli ET12567/pUZ8002	Non-methylating strain *(dam⁻ dcm⁻ hsdM)*, Cml^R, carrying helper plasmid pUZ8002 Kan^R.	(MacNeil et al., 1992; Paget et al., 1999)
E. coli NEB5α	DH5α derived cloning host	New England Biolabs
E. coli ESS	Indicator strain for cephamycin C bioassay	A. L. Demain, Drew University, Madison, USA
K. pneumoniae ATCC 15380	Indicator strain for clavulanic acid bioassay	(Reading & Cole, 1977)
Enterobacter cloacae KM31	Indicator strain produces β-lactamases enzymes	(Podder et al., 2014)
Bacillus sp. ATCC 27860	Indicator strain for 5*S* clavams bioassay	A. L. Demain, Drew University, Madison, USA
***Streptomyces* strains**		
S. clavuligerus ATCC27064	Wild type capable of normal development and production of CA, Ceph-C, and 5*S* clavams	American type culture collection (ATCC)
Δ*cpe*	*S. clavuligerus cpe* (*orf12*) deletion mutant; gene replaced by 81 bp markerless in-frame scar sequence	(Srivastava et al., 2019)
Δ*cpe*/pSET-*cpe*	*S. clavuligerus* Δ*cpe* mutant expressing *cpe* on pSET152 plasmid	(Srivastava et al., 2019)
Δ*cpe*/pHM-*cpe*	*S. clavuligerus* Δ*cpe* mutant expressing *cpe* on pHM11a plasmid	(Srivastava et al., 2019)
*S. clavuligerus/*pSET152-*cpe*-S173A	Δ*cpe* mutant strain expressing S173A variant of *cpe*	(Srivastava et al., 2019)
*S. clavuligerus/*pSET152-*cpe*-S234A	*Δcpe* mutant strain expressing S234A variant of *cpe*	(Srivastava et al., 2019)
*S. clavuligerus/*pSET152-	*Δcpe* mutant strain expressing	(Srivastava et al.,

cpe-S27A	S27A variant of *cpe*	2019)
*S. clavuligerus/*pSET152-*cpe*-L89A	*⊿cpe* mutant strain expressing L89A variant of *cpe*	(Srivastava et al., 2019)
S. clavuligerus Δ*cpe/*pSET152-*cpe*-S206A	*⊿cpe* mutant strain expressing S206A variant of *cpe*	(Srivastava et al., 2019)
S. clavuligerus Δ*orf14*	Δ*orf14* deletion mutant (an Apr^R cassette in the same orientation as *orf14*)	(Jensen et al., 2004)
S. clavuligerus Δ*orf14/*pHM11a-*orf14*	Δ*orf14* complemented with pHM11a-*orf14*	This study
S. clavuligerus Δ*orf14/*pSET152-*orf14*	*S. clavuligerus* Δ*orf14* complemented with pSET152-*orf14*	This study
S. clavuligerus Δ*orf14/*pHM11a-*orf14*-V142A	*S. clavuligerus* Δ*orf14* expressing V142A variant for *orf14*	This study
S. clavuligerus Δ*orf14/*pHM11a-*orf14*-V254A	*S. clavuligerus* Δ*orf14* expressing V254A variant for *orf14*	This study
S. clavuligerus Δ*orf14/*pHM11a-*orf14*-T269A	*S. clavuligerus* Δ*orf14* expressing T269A variant for *orf14*	This study
S. clavuligerus Δ*orf14/*pHM11a-*orf14*-V292A	*S. clavuligerus* Δ*orf14* expressing V292A variant for *orf14*	This study
S. clavuligerus Δ*nocE*	*nocE* null mutant	This study
S. clavuligerus/pIJ8668-*ermE*p*/*nocE*	Strain constitutively expressing *nocE*	This study
Streptomyces pratensis ATCC 33331	Wild type; clavulanic acid and carbapenem non-producer	ATCC
S. pratensis/Sc-cpe	*S. pratensis* with heterologous expressing pSET152-*Sc-cpe*	This study
S. pratensis/Sc-orf14	*S. pratensis* with heterologous expressing pHM11a-*Sc-orf14*	This study
S. pratensis/Sc-cpe-orf14	*S. pratensis* with heterologous expressing pSET152-*Sc-cpe* and pHM11a-*Sc-orf14*	This study

S. pratensis Δ*cas2*-1	Δ*cas2* mutant strain with insertional inactivation using pIJ773/*Sp-cas2*-KO-1	This study
S. pratensis Δ*cas2*-2	Δ*cas2* mutant strain with insertional inactivation using pIJ773/*Sp-cas2*-KO-2	This study
Streptomyces cattleya	Wild type, thienamycin, and Ceph-C producer	ATCC

Table 2.3. Plasmids and DNA constructs used in this study.

Plasmid	Antibiotic resistance marker(s)	Genotype/Description	Reference/ source
pGEMT® T - Easy	Amp^R	Cloning vector for PCR product	Promega
pHM11a	Hyg^R	Integrative *Streptomyces* expression vector with strong constitutive promotor, P_E	(Motamedi et al., 1995)
pHM11a/*cpe*-His	Hyg^R	Expression plasmid pHM11a containing *cpe* from *S. clavuligerus* with a C-terminal 6×His tag	This study
pHM11a/*orf14*	Hyg^R	Expression plasmid pHM11a containing *orf14* from *S. clavuligerus*	This study
pHM11a-*orf14*-V142A	Hyg^R	pHM11a containing *orf14* from *S. clavuligerus* with SDM at V142A	This study
pHM11a-*orf14*-V254A	Hyg^R	pHM11a containing *orf14* from *S. clavuligerus* with SDM at V254A	This study
pHM11a-*orf14*-T269A	Hyg^R	pHM11a containing *orf14* from *S. clavuligerus* with SDM at T269A	This study
pHM11a-*orf14*-V292A	Hyg^R	pHM11a containing *orf14* from *S. clavuligerus* with SDM at V292A	This study
pSET152	Apr^R	Integrative *Streptomyces* cloning vector	(Bierman et al., 1992)
pSET152-*orf14*	Apr^R	pSET152 containing *orf14* from *S. clavuligerus*	This study
pIJ8668-*ermE*p*	Apr^R	Plasmid contains constitutive promoter *ermE*p* and *aac(3)IV*	(Tahlan et al., 2017)
pIJ8668-*ermE*p∗- *nocE*	Apr^R	pIJ8668-*ermE*p∗ containing a portion of the 5' end of *nocE* from *S. clavuligerus*	This study
pIJ12738	Apr^R	Conjugative vector containing I-SceI site and *aac(3)IV*.	(Fernández-Martínez & Bibb, 2014)
pIJ12738/*nocE*-	Apr^R	pIJ12738 containing regions	This study

UP-DN		upstream and downstream of *nocE* from *S. clavuligerus*	
pIJ12742	Apr^R, Thi^R	Vector containing *ermE**p, *I-SceI* gene, *oriT*, *to* terminator, *tipAp*, RBS, *tsr* and the temperature-sensitive replication origin of pSG5	(Fernández-Martínez & Bibb, 2014)
pIJ773	Apr^R	pBluescript II SK (+)-based plasmid containing the apramycin resistance cassette *aac(3)IV* and the *oriT* of RP4 (=RK2).	(Gust et al., 2003)
pIJ773/*Sp-cas2*-KO-1	Apr^R	pIJ773 plasmid with the region 34 – 484 of *cas2* gene of *S. pratensis* inserted at HindIII site.	This study
pIJ773/*Sp-cas2*-KO-2	Apr^R	pIJ773 plasmid with the region 413 – 831 of *cas2* gene of *S. pratensis* inserted at HindIII site	This study
pIJ773/*Sp-carE*-KO-1	Apr^R	pIJ773 plasmid with the region 9 – 307 of *carE* gene of *S. pratensis* inserted at HindIII site	This study
pIJ773/*Sp-carE*-KO-2	Apr^R	pIJ773 plasmid with the region 9 – 487 of *carE* gene of *S. pratensis* inserted at HindIII site	This study

2.14. Supplementary Materials

Supplementary Table S2.1. Sequences of oligonucleotide primers used in the current study and their details.

Name	Sequence (5' – 3')	Product size (bp)	Description
nocE-KO-UP-F2 nocE-KO-UP-R2	AAGCTTCCCTGGCTGAAACCCTATGG GAATTCGCGCTTGGATCTGCTCAAAG	1200	Primers for amplification of the upstream region of *nocE* from *S. clavuligerus* to prepare pIJ12738-*nocE*-UP-DN
nocE-KO-DN-F nocE-KO-DN-R2	GAATTCCTGCCGTCGATGAAGTCCTT TCTAGACACCAAGGCGATCCTCTACC	1220	Primers for amplification of the downstream region of *nocE* from *S. clavuligerus* to prepare pIJ12738-*nocE*-UP-DN
nocE-KO-UP-F2 nocE-KO-DN-R2	AAGCTTCCCTGGCTGAAACCCTATGG TCTAGACACCAAGGCGATCCTCTACC	2421	Primers for confirming upstream and downstream regions of *nocE* in *S. clavuligerus* pIJ12738-*nocE*-UP-DN
Sc-nocE-F2 Sc-nocE-R1	GTCGAGAAGCTCCCGTACCA CGGTAGCCGTGGACCATCTT	1787	Primers for detection of *nocE* in *S. clavuligerus* pIJ12738-*nocE*-UP-DN
nocE-UPDN-ID-F nocE-UPDN-ID-R	GTCTGAACCACTTTCGCAGC GTGAAGTGGCATGGCGAATC	439	Primers for confirming the presence of upstream and downstream regions of *nocE* in *S. clavuligerus* Δ*nocE*
Sc-nocE-F1 nocE-ID-R	GCCGACGAGAAGGACGGTTA CAGCTTGTTGGTGAAGGTGC	156	Primers for confirming deletion of *nocE* in *S. clavuligerus* Δ*nocE*
nocE-KN-F nocE-KN-R	CATATGGAATTTCCCCGGACTCC GAATTCACCTCACCACCGGTCAGAT A	1050	Primers for amplification of the 5' end of *nocE* from *S. clavuligerus* to prepare pIJ8668-*ermE*p*-*nocE*
ermEp-F nocE-K-R	GATATCGGTACCAGCCCGAC GCGCTTGGATCTGCTCAAAG	578	Primers for confirming the insertion of *ermE*p* in *S. clavuligerus ermE*p*-*nocE*
Sc-nocE-F1 nocE-ID-R	GCCGACGAGAAGGACGGTTA CAGCTTGTTGGTGAAGGTGC	156	Primers for RT-PCR of *nocE* from *S. clavuligerus*

KTA1 (blm – F) KTA2 (sc-c-Ter-His)	ATACATATGATGAAGAAAGCTG ATAGGATCCTCAGTGGTGGTGGTGG TGGTGTCGCCGGGCGGCTTC	1377	Primers for amplifying *cpe* gene and adding 6x histidine tag at the C-terminus.

RT-PCR primers

ceaS2-F ceaS2-R	ATCGACTTCGTTCTGACCCG GGTGTCGTTCGGGAAGATGT	213	Primers for RT-PCR of *ceaS2* from *S. clavuligerus*
cas2-O73 cas2-O74	GCAAGCGGCTGGTGATGG GGTCTCCGAGGACAGGTAGTGC	143	Primers for RT-PCR of *cas2* from *S. clavuligerus*
oat2-O83 oat2-O84	CACCGTCCTCGCCTCCAC CGTTCTCCTCGCCCTCCAG	176	Primers for RT-PCR of *oat2* from *S. clavuligerus*
oppA1-O85 oppA1-O86	CGGGGTACGGGGAGTGG CGGAGGAAGTTCCAGGTGTA	126	Primers for RT-PCR of *oppA1* from *S. clavuligerus*
claR-O77 claR-O78	CGGGCGGCGGTTCTT TCGTCGAGCAGGGGTTCC	123	Primers for RT-PCR of *claR* from *S. clavuligerus*
car-F car-R	GTCTACCAGGCCACGAAGTT GATCCGCTGCTCGTACATCT	168	Primers for RT-PCR of *car* from *S. clavuligerus*
cyp-O79 cyp-O80	ACGAACTCGACGGCTATCTG ACATCGGGACCATCTCCTC	132	Primers for RT-PCR of *cyp* from *S. clavuligerus*
orf12-O89 orf12-O90	GGCGATGGGGCTGCTGAC GTGCGCGACGGGGTGGTA	160	Primers for RT-PCR of *cpe* (*orf12*) from *S. clavuligerus*
orf13-O91 orf13-O92	CTGCGCTGGCTGCTGGTGTA CTGCCGCCGGGAGATGC	174	Primers for RT-PCR of *orf13* from *S. clavuligerus*
orf14-O93 orf14-O94	CGAACGACGACGAAACG CAGCGAGCCGACCATGT	107	Primers for RT-PCR of *orf14* from *S. clavuligerus*
Sc-oppA2-F Sc-oppA2-R	CCCACCTATCTCATCCCGC CATCAGATGGTCGAAGTCGGA	153	Primers for RT-PCR of *orf15* (*oppA2*) from *S. clavuligerus*

orf16-F orf16-R	TTCCTGGCCGACATGACCAA CCGTACTTGCGCAGCAGATT	155	Primers for RT-PCR of *orf16* from *S. clavuligerus*
gcas-O81 gcas-O82	GGTCAACTGGAGCCTGTGTA CCGCGAACTTGGCATAGTC	101	Primers for RT-PCR of *gcas2* from *S. clavuligerus*
pbpA-F pbpA-R	CAAGTACCAGCGCACCTACA CGCTCAATACGCTGTCGAAC	113	Primers for RT-PCR of *pbpA* from *S. clavuligerus*
hrdB-4F (NF) hrdB-4R (NF)	CGCGGCATGCTCTTCCT AGGTGGCGTACGTGGAGAAC	109	Primers for RT-PCR of *hrdB* from *S. clavuligerus*
cas1-F cas1-R	AGCCGAACTACGTCATGCTG CCGTAGAGCGGTTTGACCTG	211	Primers for RT-PCR of *cas1* (SCLAV_ 2925) from *S. clavuligerus*
cvm1-F cvm1-R	GTACTACCAGCACTGGACGG TCGGAGAGACCGAGCCTG	105	Primers for RT-PCR of *cvm1* (SCLAV_2926) from *S. clavuligerus*
cvm2-F cvm2-R	GACTACTTCGCCGAGGACG AATCCAGTTGACGGACCACA	141	Primers for RT-PCR of *cvm2* (SCLAV_2927) from *S. clavuligerus*
cvm5-F cvm5-R	ACTTCCACACCGAGGGTTTC TCATGTGGTCGAGCATCGC	162	Primers for RT-PCR of *cvm5* (SCLAV_2923) from *S. clavuligerus*
cvm6p-F cvm6p-R	GACACTCGGTCACTTCCACA GTGAAGTAGACGCGCTGGA	111	Primers for RT-PCR of *cvm6p* (SCLAV_ p1078) from *S. clavuligerus*
cvm7p-F cvm7p-R	CCGTATCTGGGGCAACTCAC CCTGCTCAAAACGGTTCGC	188	Primers for RT-PCR of *cvm7p* (SCLAV_ p1079) from *S. clavuligerus*
snk-F snk-R	CGTGATTTTCCCGCCGGTAT GCGGAATCCCCACTCCTTG	150	Primers for RT-PCR of *snk* (SCLAV_ p1080) from *S. clavuligerus*
res1-F res1-R	GATCCGTCCCGACGATTCTG TCTTGGGCAGGAAACCGATG	93	Primers for RT-PCR of *res1* (SCLAV_ p1081) from *S. clavuligerus*
res2-F res2-R	ATGGCAGGAGTGAGGGTAGT ATCCGGATGTCCACGATCAC	173	Primers for RT-PCR of *res2* (SCLAV_ p1082) from *S. clavuligerus*
orfA-F orfA-R	GTTCTTACCCTTCGACCGCC CGTCCAGATAGATCACGTCGG	95	Primers for RT-PCR of *orfA* (SCLAV_ p1072) from *S. clavuligerus*

orfB-F orfB-R	AAGGTGGTCGAGGGAGGTAT CAGATAGATCCGCACGGTGA	117	Primers for RT-PCR of *orfB* (SCLAV_ p1071) from *S. clavuligerus*
orfC-F orfC-R	CACTGATCGTCAACACCCCC GAAGACGAAGTCGGCGTACA	134	Primers for RT-PCR of *orfC* (SCLAV_ p1070) from *S. clavuligerus*
orfD-F orfD-R	TGACGGTCACGGTGTGTATG CGATCAGGTCGTCCGTGAAG	99	Primers for RT-PCR of *orfD* (SCLAV_ p1069) from *S. clavuligerus*

S. pratensis **project primers**

cas2-F-KO-1 cas2-R-KO-1	ATTATTAAGCTTGAACTCCTGGAACT CGCCTC ATTATTAAGCTTAGGCCAGCATCACG TAGTTC	448	Primers for cloning Sp-*cas2* (region 34 – 484) at HindIII site in pIJ773
cas2-F-KO-2 cas2-R-KO-2	ATTATTAAGCTTGACACTGCTGGAGT TCCACA ATTATTAAGCTTCGTACGGAAGTTGT CGACGA	418	Primers for cloning Sp-*cas2* (region 413 – 831) at HindIII site in pIJ773
cas2-conf-F cas2-conf-R	CATCGACTGCTCCTCACTCC CGTACGGAAGTTGTCGACGA	824	Primers for confirmation of *cas2*-like gene deletion in *S. pratensis*
cas2-conf-F T3	CATCGACTGCTCCTCACTCC ATTAACCCTCACTAAAGGGA	----	Primers for confirmation of *cas2*-like gene deletion in *S. pratensis*
carE-F-KO-1 carE-R-KO-1	ATTATTAAGCTTTCCGATCCAGCCTT CTCAGA ATTATTAAGCTTCCGTGTAGAGGTCG GTGATG	299	Primers for cloning Sp-*carE* (region 9 – 307) at HindIII site in pIJ773
carE-F-KO-1 carE-R-KO-2	ATTATTAAGCTTTCCGATCCAGCCTT CTCAGA ATTATTAAGCTTCGGTGTAACGCAGC ATGAAG	479	Designed for cloning Sp-*carE* (region 9 – 487) at HindIII site in pIJ773
carE-conf-F carE-conf-R	CGAACGAGCGAGAAAGAGGT CACTCCTCGCAGGTGAAGAG	556	Primers for confirmation of *carE* gene deletion in *S. pratensis*

RT-PCR primers for *S. pratensis* gene clusters

Sf-cas2-F Sf-cas2-R	CGTCTACCACGACGTGTACC AGGCCAGCATCACGTAGTTC	128	Primers for RT-PCR of *cas2* from *S. pratensis*
Sf-ceaS2-F Sf-ceaS2-R	ACGACATCTTCCCCAACGAC GAGATGAAGCTGGGACCGAC	166	Primers for RT-PCR of *ceaS2* from *S. pratensis*
Sf-orf12-F Sf-orf12-R	ATGATCGCGATGAGCGACAA CAGCCGATCTCGAAGACCTC	152	Primers for RT-PCR of *orf12* from *S. pratensis*
Sf-hrdB-F Sf-hrdB-R	CGAGTTCGGAGACCTGATCG CCGTAGACCTTGCCGATCTC	192	Primers for RT-PCR of *hrdB* from *S. pratensis*
Sf-carE-F Sf-carE-R	ACATCACCGACCTCTACACG GTGTAACGCAGCATGAAGCC	199	Primers for RT-PCR of *carE* from *S. pratensis*
Sf-carM-F Sf-carM-R	CTGCTCACCCTGCAGATCG TGGAACTCGTTGCTCCGAC	194	Primers for RT-PCR of carM from *S. pratensis*
Sf-carP-F Sf-carP-R	CATCTGGTCCACGAGTACGG CTTGTCGAGACGCATCACCT	129	Primers for RT-PCR of carP from *S. pratensis*
Sf-carI-F Sf-carI-R	TCCTGATCCGGACCATTGC TACGGTGAAACTGACCGACG	197	Primers for RT-PCR of hrdB from *S. pratensis*

Site-directed mutagenesis primers

Orf14-F142A-F Orf14-F142A-R	CCCGGCCGCCGCCGCCGCCGCG CGCGGCGGCGGCGGCGGCCGGG	--------	Primers to generate site-directed mutation at Phe 142 to Ala.
Orf14-T269A-F Orf14-T269A-R	GGTGCACCACGGCCATGCCCTGGAG CTCCAGGGCATGGCCGTGGTGCACC	--------	Primers to generate site-directed mutation at Thr 269 to Ala.
Orf14-V254A-F Orf14-V254A-R	GGTGGTCTTGGAGGCGCTGGTGTAC CC GGGTACACCAGCGCCTCCAAGACCA CC	--------	Primers to generate site-directed mutation at Val 254 to Ala.

Orf14-V292A-F Orf14-V292A-R	GCTCGTGCCGCAGCGCGTACTCCAG ATTG CAATCTGGAGTACGCGCTGCGGCAC GAGC	--------	Primers to generate site-directed mutation at Val 292 to Ala.

CHAPTER III

Molecular investigations into the role of *orf12* (*cpe*) and *orf14* in clavam metabolite biosynthesis in *Streptomyces clavuligerus*

3.1 Abstract

This study investigated the molecular features of two genes, *cpe* (*orf12*) and *orf14* from the CA BGC of *S. clavuligerus*, and their roles in CA and 5*S* clavam biosynthesis. Previous studies reported that the inactivation of the *cpe* gene blocked the production of CA but not that of the 5*S* clavams in *S. clavuligerus*. My study showed that the deletion of *cpe* did not affect the transcription of CA biosynthetic genes, suggesting it does not have any regulatory role on CA production. However, the overexpression of *cpe* in two complemented strains Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe* showed induction of 2HMC production (a 5*S* clavam metabolite) in SA media, conditions under which the metabolite is not produced by the wt strain. Transcription analyses showed that the expression of some genes essential for 5*S* clavam biosynthesis (*cas1, cvm1, cvm2, cvm5, cvm6p, cvm7p, orfA,* and *orfB*) were upregulated in the complemented strains in comparison to wt *S. clavuligerus*, suggesting that the product of *cpe* has indirect positive effect on the regulator Cvm7P, and subsequently on the production of 2HMC. LC-MS analyses for different CPE variants showed that the amino acid residues Ser173, Ser234, and Leu89 are essential for CPE to perform its role in CA or 2HMC biosynthesis. Interestingly, the substitution of Ser206 with Ala almost abolished the production of 2HMC but not CA, suggesting that the enzymatic activity of CPE exerted on CA biosynthesis is different

from 2HMC biosynthesis. We proposed that CPE performs a hydrolysis step to convert 2FMC into 2HMC when *S. clavuligerus* overexpresses *cpe* in SA medium.

The *orf14* gene, which encodes an N-acetyltransferase, is proposed to be responsible for the formation of the N-acetylated clavaminic acid derivatives during the later stages of the CA biosynthetic pathway. In this study, site-directed mutagenesis was performed for certain amino acid residues in the N-terminal and C-terminal domains. The residues Phe142, Val254, Thr269, and Val292 were replaced with Ala and tested for CA production. The four plasmids containing *orf14* variants successfully complemented the *S. clavuligerus* Δ*orf14* mutant, and the CA production was restored in all of them. The results suggest that these amino acid residues are not essential for the production of CA. In this study we investigated some of the molecular features of two essential genes (*orf12* and *orf14*) of CA BGC on the biosynthesis of CA and 5*S* clavams.

3.2 Introduction

cpe (*orf12*) and *orf14* are parts of the "late genes" from the clavulanic acid (CA) BGC of *S. clavuligerus*. They are of interest due to their unknown function but are thought to be involved in the conversion of clavaminic acid into clavaldehyde during CA biosynthesis (Figure 1.6) (Iqbal et al., 2010a; Srivastava et al., 2019). The two genes are essential for the production of CA; deletion of either one of them abolishes CA biosynthesis without affecting Ceph-C or 5*S* clavams production (Jensen et al. 2004a; Srivastava et al., 2019).

The amino acid sequence encoded by *cpe* (SCLAV_4187) shows similarity to class A β-lactamases, therefore it has been proposed to be somehow involved in the opening of the β-lactam ring during the stereochemical inversion of (3*S*, 5*S*) clavaminic acid to (3*R*, 5*R*) clavulanic acid. However, the specific role of *cpe* in CA production is still unclear (Figure 1.6) (Jensen, 2012; Valegård et al., 2013). CPE is composed of 458 amino acids and has a molecular mass of 49.8 KDa. Recently, the crystal structure of CPE revealed that the protein contains a previously unrecognized N-terminal domain. The N-terminal domain (residues 1-127) resembles those present in steroid isomerases and polyketide cyclases and is fused to a C-terminal domain (residues 128–458) resembling a class A β-lactamase type fold (Valegård et al., 2013). In addition, both domains are required for CA biosynthesis but their exact roles in the process are not known (Srivastava et al., 2019).

Other CA producer species such as *S. jumonjinensis* and *S. katsurahamanus* encode CPE proteins that share ~68% amino acid identity with CPE of *S. clavuligerus*. The CA non-producing species such as *Streptomyces pratensis* and *Saccharomonospora*

viridis, which have CA-like BGC, encode CPE orthologs to *S. clavuligerus* CPE with 58.8% and 48.8% identity, respectively (Srivastava et al., 2019). Heterologously expressed and purified CPE (*S. clavuligerus* gene) does not show any β-lactamase activity against β-lactams compounds such as nitrocefin, clavulanic acid, benzylpenicillin, and ampicillin. In addition, CPE does not display any penicillin-binding activity (Valegård et al., 2013), suggesting that CPE is not a part of any intrinsic β-lactam resistance mechanism in *S. clavuligerus*. Instead, CPE has a low level of deacetylase/esterase activity towards cephalosporin C leading to the formation of deacetylcephalosporin C (Valegård et al., 2013), thus ORF12 was designated as CPE for "Cephalosporin Esterase" (Srivastava et al., 2019). In addition, CPE shows acetyltransferase activity when incubated with deacetylcephalosporin C and cephaloglycin, which leads to the formation of cephalosporin C and deacetylcephaglycin, respectively. However, the cephalosporin esterase activity of CPE is likely to be a nonessential function but may reflect promiscuity based on substrate similarity (Valegård et al., 2013). The crystal structure of the *S. clavuligerus* CPE found that two molecules of CA can bind non-covalently to the protein, and several amino acid residues from both the N- and the C-terminal domains participated in this binding (Valegård et al., 2013). Site-directed mutagenesis studies on these residues and others located on the active site of the CPE protein revealed that some of them (Lys89, Ser173, Lys176, Ser234, Tyr359, Lys375, Ser378, and Arg418) are essential for functional CPE and consequently for *in vivo* CA production (Srivastava et al., 2019).

The protein encoded by *orf14* (SCLAV_4185) is a member of tandem GCN5-related N-acetyltransferases (GNAT) family and contains NAT domains (Iqbal et al., 2010). *orf14* is essential for CA biosynthesis (Jensen et al., 2004a), and it is proposed that

the N-acetylated clavaminic acid derivatives in CA pathway might be generated from the acetylation activity of ORF14 (Figure 1.6) (Iqbal et al., 2010a; Mellado et al., 2002). N-glycyl-clavaminic acid (NGCA) resulting from GcaS activity (Figure 1.6) is thought to be acetylated by ORF14 to give N-acetyl-glycyl-clavaminic acid (AGCA), which is subsequently converted to N-acetyl-clavaminic acid (NACA) by the activity of OPPA2/ORF16 proteins (Figure 1.6) (Jensen et al., 2004a; Jensen, 2012; Paradkar, 2013). However, Álvarez-Álvarez et al. (2018) suggest that ORF14 may act downstream of AGCA formation since the addition of purified AGCA to Δ*orf14* mutant culture does not lead to CA formation.

ORF14 (also called CBG: CA biosynthesis GNAT protein), comprises 339 amino acids and has molecular mass of 37 kDa (Iqbal et al., 2010). The GNAT group of enzymes catalyze the transfer of an acetyl/acyl group from acetyl/acyl-CoA to an acceptor amine (Baumgartner et al., 2021). Crystallographic studies on ORF14 reveal that one molecule of acetyl-CoA (AcCoA) binds to the N-terminal GNAT domain, whereas the C-terminal domain is unoccupied. However, mass spectrometric analyses for ORF14 demonstrated that a second acyl-CoA molecule can bind to ORF14, which most likely occurs in the C-terminal GNAT domain (Hamed et al., 2013; Iqbal et al., 2010), since both the N- and C-terminal domains of ORF14 possess the characteristic GNAT superfamily mixed α,β-fold (Hamed et al., 2013). In addition, it was shown that CoA derivatives (succinyl-CoA and myristoyl-CoA) can bind to the ORF14 monomer without displacing the already-bound AcCoA, suggesting that the C-terminal domain of ORF14 is directly involved in the acetyl transfer, whereas the AcCoA tightly bound to the N-terminal domain might have a structural role (Iqbal et al., 2010).

3.3. Objectives

The ambiguous function of CPE (ORF12) and its role in CA biosynthesis requires further investigation. One of the main goals of this chapter was to investigate if CPE has any regulatory role in the transcription of CA and 5*S* clavam biosynthetic genes. I also attempted to examine the *in vivo* influence of expressing different CPE variants on the production of CA and 5*S* clavams to understand how the protein exerts its effect on clavam biosynthesis in *S. clavuligerus*. In addition, I aimed to investigate the essentiality of *orf14* during CA biosynthesis by performing complementation studies using an *S. clavuligerus orf14* deletion mutant. My goal was to determine the importance of the substrate binding sites in ORF14 by preparing and expressing different variants of ORF14 in complementation studies.

3.4. Results and discussion

3.4.1. Fermentation of the *S. clavuligerus* Δ*cpe* mutant and complemented strains

The recently prepared *S. clavuligerus* Δ*cpe* mutant and the two complemented *S. clavuligerus*/Δ*cpe*/pHM11a-*cpe* and *S. clavuligerus*/Δ*cpe*/pSET152-*cpe* strains from our lab were used in this study (Table 2.2; Srivastava et al., 2019). Originally, *cpe* was first cloned into the pHM11a vector for complementation studies, and a second pSET152-*cpe* construct was prepared for further site directed mutagenesis work due to the smaller size of the latter vector (Srivastava et al., 2019). To assess the ability of these bacterial strains to produce CA and 5*S* clavams, they were subjected to fermentation, bioassays, and liquid chromatography-mass spectrometry (LC-MS) analysis. Wild type *S. clavuligerus*, Δ*cpe*, Δ*cpe*/pHM-*cpe*, and Δ*cpe*/pSET-*cpe* were cultivated in SA and SM media, and

supernatants were collected at 48 and 96 h time points. Bioassays using *Klebsiella pneumoniae* as an indicator (Figure 3.1) showed zones of growth inhibition around discs infused with supernatants from the wt and two complemented Δ*cpe* strains, but not from Δ*cpe* mutant by itself. Zones of inhibition were due to CA production in the successfully complemented strains Δ*cpe*/pHM11a-*cpe* and Δ*cpe*/pSET152-*cpe* strains (Figure 3.1). As reported previously, the deletion of *cpe* in *S. clavuligerus* blocks CA production (Jensen et al., 2004; Srivastava et al., 2019), reflecting its essential role in CA biosynthesis. All bioassays were with duplicate cultures of each strain at two-time points (Supplementary Table S3.1), and CA solution (10 μg) was used as positive control.

3.4.2. LC-MS analysis for assessing the production of clavulanic acid and 5*S* clavams in the *S. clavuligerus* Δ*cpe* mutant and complemented strains.

To further investigate the effect of the *cpe* deletion and gene complementation on CA and 5*S* clavams production, LC-MS analysis was performed on the supernatant samples taken after 96 h of growth in SA and SM media. Derivatization for the supernatant samples was carried out by adding imidazole prior to injection, and CA production was detected by absorbance at 311 nm. Supernatants from the wt and the Δ*cpe*/pSET-*cpe*, Δ*cpe*/pHM-*cpe* strains showed peaks for CA at ~19 min retention time while no peak was detected in samples from the Δ*cpe* strain (Figure 3.2A). Mass spectral peaks corresponding to imidazole-derivatized CA $[M+H]^+$ (*m/z* = 224) and the fragmented product $[M\text{-imidazole}]^+$ (*m/z* = 156) could be clearly identified in the wt and the complemented strains (Figure 3.2B and C, and Supplementary Table S3.2). The LC-

MS analysis confirmed the bioassay results and was consistent with previous reports (Jensen et al., 2004, Srivastava et al., 2019).

It was previously reported that *S. clavuligerus* produces 5*S* clavams when fermented in SM but not in SA media (Mosher et al., 1999; Jensen et al., 2000). Interestingly, the LC analyses for supernatants from the Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe* showed the presence of an additional peak with an earlier retention time (~17 min) (Figure 3.2A), which was predicted to be 2HMC (Jensen et al., 2000; Tahlan et al., 2007). The MS analyses revealed the presence of a peak corresponding to imidazole-derivatized 2HMC fragmented product $[M\text{-imidazole}]^+$ (*m/z* = 144) (Figure 3.2B and C, and supplementary Table S3.2). However, the imidazole-derivatized product of 2HMC $[M+H]^+$ (*m/z* = 212) was not detected, which could be because of its low concentration in the sample or due to differences in the set-up parameters of the LC-MS machine from those previously reported (Tahlan et al., 2007). Moreover, no other prominent mass ions belonging to the 5*S* clavams (2FMC, C2C, 8-hydroxy alanylclavam, or alanylclavam) were detected in the analysis, suggesting that the overexpression of *cpe* in the Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe* strains induced only the production of 2HMC in SA medium. The LC-MS analyses for imidazole derivatized wt *S. clavuligerus* SA supernatants did not demonstrate any peak or mass ion for 5*S* clavams including 2HMC (Figure 3.2), which agrees with previous reports (Paradkar and Jensen, 1995; Mosher et al., 1999; Jensen et al., 2000). Just to confirm, LC-MS analyses were conducted on wt *S. clavuligerus* SM supernatants, conditions known to foster CA and 5*S* clavam production. The peaks corresponding to CA (~19 min) and 2HMC (~17 min) were detected (Figure 3.3A). Peaks corresponding to imidazole-derivatized CA $[M+H]^+$ (*m/z* = 224) and its

fragmented product $[M\text{-imidazole}]^+$ (*m/z* = 156), and imidazole-derivatized 2HMC $[M+H]^+$ (*m/z* = 212) and the fragmented product $[M\text{-imidazole}]^+$ (*m/z* = 144) were clearly detectible (Figure 3.3B). The retention times for 2HMC from wt SM supernatants are the same as those observed in the Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe* complemented strains grown in SA media, further suggesting that the metabolite was present in the two latter samples.

3.4.3. Bioassays for detecting 2HMC production in the *S. clavuligerus* Δ*cpe* complementation strains

Supernatants from SA cultures were further characterized by bioassays for the production of 2HMC metabolite using *Bacillus* sp. ATCC 27860 on two DMM plates (Figure 3.4). The first DMM plate (without methionine) was to demonstrate the bioactivity of 5*S* clavams (mainly 2HMC and alanylclavam), while the second plate was supplemented with methionine to antagonize alanylclavam activity (if any), and hence show only 2HMC activity (Figure 3.4). Zones of growth inhibition were detected in both DMM plates for *cpe* complemented samples (Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe*), while no growth inhibition was noticed around the wt and Δ*cpe* discs in both plates (Figure 3.4). As mentioned earlier, Paradkar and Jensen (1995) showed that wt *S. clavuligerus* does not produce any 5*S* clavams in SA medium, which is consistent with our bioassay result. The bioactivity noticed on the DMM (+methionine) plate indicates that Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe* produced 2HMC that inhibited the growth of the indicator microorganism and supports the LC-MS analysis described above. Therefore, the overexpression of *cpe*

in the *S. clavuligerus* Δ*cpe* mutant somehow induces the production of the 2HMC in SA medium, an observation not reported previously.

3.4.4. Examining the influence of CPE on the expression of clavulanic acid biosynthesis genes.

CPE has been shown to bind non-covalently with CA under *in vitro* conditions (Valegård et al., 2013), but the relevance of this interaction is still not clear as the protein did not catalyze any associated reaction. CPE is located in the cytoplasm of *S. clavuligerus* (Srivastava et al., 2019), and Δ*cpe* mutants are completely blocked in CA production (Figure 3.1 and 3.2). This raises the possibility that the protein could have a role as a cytoplasmic sensor/receptor for CA or related metabolites to indirectly regulate production under *in vivo* conditions. To test this hypothesis, RT-PCR was conducted with RNA samples isolated from wt *S. clavuligerus* and Δ*cpe* strains grown for 96 h in SA medium, which only supports the production of CA and not the 5*S* clavams. Random hexameric primers along with reverse transcriptase were used to generate cDNA, and specific primer pairs (Supplementary Table S2.1) were used to determine if the genes of interest were transcribed in the tested strains. We tested the expression level of the first gene from each transcriptional unit as shown in Figure 3.5A (*ceaS2, cas2, oat2, oppA1, claR, car, cyp, cpe, orf13, orf14, oppA2, orf16, gcaS,* and *pbpA*) from the clavulanic acid gene cluster of *S. clavuligerus* in the Δ*cpe* and compared it with that from the wt strain. RT-PCR analysis showed that only expression of the *cpe* gene was altered in the comparison, which was expected, while that of all other CA BGC genes tested was not affected (Figure 3.4B). The gene *cpe* (*orf12*) is 1,377 bp in length (including the stop

codon), starting with an ATG codon with a potential RBS sequence GGCCG located 10 bp upstream (Li et al., 2000; Jensen et al., 2004), and it is situated 232 nucleotides downstream of *orf11* (*fd*). On the other side, the stop codon of *cpe* (*orf12*) overlaps with the start codon of *orf13* (a gene encoding a predicted efflux pump protein) (Srivastava et al., 2019). Both *orf12* and *orf13* have been found to be transcribed together as a bicistronic mRNA (Li et al., 2000; Santamarta et al., 2011), and so the deletion of *cpe* located in the 5′ region of the operon can potentially affect the expression of the downstream *orf13* gene and lead to polar effects (Srivastava et al., 2019). However, our RT-PCR analysis demonstrated that the Δ*cpe* mutation is not associated with any transcriptional polarity as the expression *orf13* was unaffected in the strain (Figure 3.4B). Moreover, *orf13* and *orf14* have convergent transcription directions (Figure 3.5A), where there is a 48 bp overlap between the 3' ends of the two genes (Srivastava et al., 2019). Thus, there is a potential for some interference there also (transcription and/or translation), but our analysis showed that the deletion of *cpe* does not in any way influence the transcription of other genes from the CA BGC in *S. clavuligerus* (Figure 3.5B).

Further analysis was performed on the transcriptional levels of *cpe* and the genes in its immediate vicinity (*orf13* and *orf14*) in the *S. clavuligerus* Δ*cpe* complementation strains (Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe*) that produce 2HMC in SA medium. Total RNA samples were isolated from *S. clavuligerus*, wt, Δ*cpe*, Δ*cpe*/pSET-*cpe*, and Δ*cpe*/pHM-*cpe* stains grown for 96 h in SA medium and the RT-PCR was conducted as described above. The transcription level of *cpe* in the complemented strains (Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe*) was elevated as compared to wt (Figure 3.6), which was

expected as the two integrative plasmids carry strong constitutive promoter *ermE*p* driving the expression of *cpe* (Bierman et al., 1992; Motamedi et al., 1995). The transcription of *orf13* and *orf14* was the same in all strains tested (Figure 3.6), suggesting that production of 2HMC by Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe* strains in SA is mainly related to the independent overexpression of *cpe* gene. The transcription of *hrdB*, encoding the constitutively expressed sigma factor in *S. clavuligerus* was monitored as a control and was detected at similar levels in all RNA samples tested.

3.4.5. The influence of CPE on the expression of select genes from the 5*S* clavam and paralogue gene clusters

In order to explain the production of 2HMC in the Δ*cpe* complementation strain in SA medium, the effect of CPE on the expression of essential genes from the 5*S* clavam and the paralogue gene cluster was also examined (Figure 1.4 and Table 1.3). Total RNA isolated after 96 h of growth in SA medium was isolated from the *S. clavuligerus* wt, Δ*cpe*, Δ*cpe*/pSET-*cpe*, and Δ*cpe*/pHM-*cpe* strains and subjected to RT-PCR analysis as described above. We first assessed the expression level of *cas1, cvm1, cvm2,* and *cvm5* from the 5*S* clavam BGC of *S. clavuligerus* (Figure 3.7A), which have been shown to be essential for producing 5*S* clavams, mainly 2HMC and C2C (Mosher et al., 1999; Tahlan et al., 2007). The tested genes (*cas1, cvm1, cvm2,* and *cvm5*) were not transcribed in the wt and Δ*cpe* in SA, while their transcription was detected in the Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe* strains (Figure 3.7B). The results suggest that the overexpression of *cpe* has a positive effect on transcription of some 5*S* clavam genes, which might explain the production of 2HMC by the strains in SA medium. *cas1* encodes clavaminic acid

synthase isoenzyme (Table 1.2 and 1.3), and its expression was reported to be nutritionally regulated, where it is not transcribed when *S. clavuligerus* grown in SA medium (Mosher et al., 1999). Deletion of *cas1* decreased the production level of 2HMC and C2C without affecting the alanylclavam due to the presence of a paralogous copy of the gene in the CA BGC (Mosher et al., 1999). Therefore, the expression of *cas1* in Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe* strains in SA medium suggests an important role for it in the production of 2HMC specifically. *cvm1, cvm2,* and *cvm5* are also essential for producing 5*S* clavams (Table 1.3) (Tahlan et al., 2007); the deletion of any one of them completely abolishes the production of 5*S* clavams, including 2HMC. Therefore, the production of 2HMC in Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe* strains is strongly related to the expression of these genes. Overall, the expression of 5*S* clavam genes (*cas1, cvm1, cvm2,* and *cvm5*) is necessary to produce 2HMC, and the CPE overexpression somehow directly or indirectly affects their expression.

The genes essential for producing 2HMC, and in general 5*S* clavams, are also present in paralogue gene cluster located in the pSCL4 plasmid (Table 1.3). The transcription of regulators and the essential genes from the paralogue gene cluster were also tested using RT-PCR and strains grown on SA media (Figure 3.8A). While *cvm7p* and *cvm6p* transcripts were not detected in RNA from the wt and Δ*cpe* strain, they were present in Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe* strains (Figure 3.8B), suggesting the positive effect of *cpe* overexpression on their transcription. *cvm7p* encodes a pathway-specific regulator protein for 5*S* clavam biosynthesis (Tahlan et al., 2007) and positively regulates the expression of 5*S* clavams essential genes *cvm1, cvm5, cas1,* and *cvm6p* (Figure 1.9 and 1.10). *cvm6p,* which encodes aminotransferase protein, is essential for 5*S*

clavam production, where Cvm6P was proposed to be involved in the first step of biosynthesis by deamination of clavaminic acid (Figure 1.7) (Tahlan et al., 2007). Based on the results presented, we propose that the overexpression of *cpe* indirectly activates *cvm7p*, which in turn induces the expression of *cvm6p, cas1, cvm1, cvm5,* and maybe other genes not tested in this study to produce 2HMC (Figure 3.9).

The RT-PCR analysis demonstrated that the genes *snk, res1,* and *res2* from the paralogue gene cluster (encoding phosphorylation system; Table 1.3) are not expressed in all the SA RNA samples tested (Figure 3.8), indicating they are not involved in regulating 2HMC production. The Snk-Res1/2 system was reported to be expressed and responsible for the production of 5*S* clavams when *S. clavuligerus* is cultivated in SM (Kwong et al., 2012), where it positively regulates the expression of Cvm7P, which in turn controls the expression of genes essential for 5*S* clavam biosynthesis (Figure 1.9 and 1.10) (Kwong et al., 2012). Our results show that the Snk-Res1/2 system is not transcribed when the bacteria were grown in SA, indicating that the activation of *cvm7p* and consequently the production of 2HMC was independent of the Snk-Res1/2-mediated pathway (Figure 3.9). Generally, the sensor kinase protein of the TCS receives an environmental or nutritional cue and then transfers the signal to a soluble response regulator protein (Romagnoli & Tabita, 2007), which is thought to occur in the case of the Snk-Res1/2 system in SM but not in SA medium, where the signal is not present under the latter conditions (Figure 3.9).

Transcription analyses were also performed on the alanylclavam biosynthetic genes *orfA, orfB, orfC,* and *orfD,* which are located as one group within the paralogue gene cluster (Figure 3.8A). RT-PCR results showed that *orfA* and *orfB* are comparatively upregulated in Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe* as compared to the wt and Δ*cpe*

strains in SA medium, while the expression of *orfC* and *orfD* was not detected in the four tested strains under similar conditions (Figure 3.8B). The production of alanylclavam requires the four genes and deletion of any one of them (*orfA, orfB, orfC,* or *orfD*) eliminates the production of the metabolite (Zelyas et al., 2008). However, the production of alanylclavam or 8-OH-alanylclavam (the proposed intermediate in the alanylclavam biosynthesis) was not detected in our LC-MS analyses, which is consistent with our RT-PCR results. The upregulation in *orfA* and *orfB* expression is not related to any regulatory activity of Cvm7P, since the deletion of *cvm7p* does not alter the expression of *orfA* (Kwong et al., 2012). However, it has been proposed that *orfA* might be under the control of other regulators such as BldA (Kwong et al., 2012; Zelyas et al., 2008).

To summarize, the production of 5*S* clavams in *S. clavuligerus* is subject to a complex regulatory hierarchy involving at least three systems that include an atypical two-component system, the transcriptional regulator Cvm7P, and global regulators such as BldA. In SM medium, a stimulus induces Snk-Res1/2 TCS that upregulates Cvm7P, which in turn regulates the expression of essential genes for 5*S* clavam production (Figure 3.9). In SA medium, which likely lacks the stimulus for the Snk-Res1/2 TCS, the CPE protein indirectly activates *cvm7p,* and consequently other genes involved in 5*S* clavam biosynthesis, thereby leading to 2HMC production (Figure 3.9). What is intriguing is that the production of other 5*S* clavams was not detected under the same conditions (Figure 3.9), something that needs to be addressed along with the details of the regulatory mechanism occurring in SA medium.

3.4.6. Protein crosslinking and detection of CPE by western blot

In order to investigate if CPE interacts with other proteins to influence the regulation of 5*S* clavam biosynthetic genes, CPE was 6×His-tagged at the C-terminus (CPE-His_6) and expressed in *S. clavuligerus* using the constitutive *ermE*p* promoter in pHM11a (Table 2.3). *S. clavuligerus*/*cpe*-His_6 and *S. clavuligerus*/pHM11a (empty vector control) were cultured in SA medium for 48 h, and the mycelial pellets were lysed by sonication to extract the cytoplasmic contents. Heterologously expressed CPE-His_6 previously purified from *E. coli* (Srivastava et al., 2019) was used as a control (Figure 3.10). The western blot analysis using anti-6×His monoclonal antibodies detected the presence of one major protein band in the cytoplasm extract of *S. clavuligerus*/*cpe*-His_6 and purified CPE-His_6 from *E. coli*, with a size ~50 kDa, which is close to the expected size range 50 – 55 kDa of CPE-His_6 (Figure 3.10A). As expected, no band was detected for *S. clavuligerus*/pHM11a control sample (Figure 3.10A). The results agree with those obtained by Srivastava et al. (2019), which showed that the CPE is not a secreted protein, and as most of the β-lactamases, it is localized in the cytoplasm. The presence of CPE in the cytoplasm supports its role as a biosynthetic protein more than as part of the self-resistance mechanism in *S. clavuligerus*.

To investigate if CPE exerts its function by binding to other proteins in the cell, *in vivo* protein-protein crosslinking was conducted using formaldehyde (a non-specific crosslinker, Section 2.7.1), followed by western-blot analyses. Following crosslinking using 1-5% of formaldehyde (final concentration) (Supplementary Figure S3.1), cellular proteins were extracted and subjected to SDS-PAGE and western blot analysis. A single band corresponding to CPE-His_6 was observed in all formaldehyde concentrations, with

no differences between them. The analysis was performed again using 1 % formaldehyde as previously reported by Chowdhury et al. (2009), to induce protein-protein cross-linking and again a single band for CPE-His_6 (~ 50 KDa) was detected in the treated sample (1 % formaldehyde) and in the untreated sample (Figure 3.10B). The results suggest that CPE does not bind to any other cellular proteins or regulators, and that the protein might function as an independent cytoplasmic enzyme.

3.4.7. The effect of different variants of CPE on the production of 2HMC

Based on the reported crystal structure of CPE, site-directed mutagenesis studies were recently conducted to change several amino acid codons in the *cpe* gene to make different variants for assessing *in vitro* cephalosporin esterase activity or *in vivo* CA production (Valegård et al., 2013; Srivastava et al. 2019). In this study, we chose five of those CPE variants (Δ*cpe*/pSET-*cpe*-S173A, Δ*cpe*/pSET-*cpe*-S234A, Δ*cpe*/pSET-*cpe*-S27A, Δ*cpe*/pSET-*cpe*-L89A, and Δ*cpe*/pSET-*cpe*-S206A; Supplementary Table S3.3) to test the production of 2HMC by the Δ*cpe* in SA medium. Ser173 and Ser234 were selected because they are present in the conserved SXXK and SDN motifs of the CPE active site, respectively, which also includes a third KTG motif (Galleni et al., 1995). Residues from these motifs form the Ser173/Lys176/Ser234/Lys375 catalytic tetrad (Pratt & McLeish, 2010; Valegård et al., 2013) and substitution of Ser173 or Ser234 with Ala in CPE blocked the *in vivo* production of CA (Srivastava et al., 2019) and significantly reduced *in vitro* esterase activity (~100-fold in Ser173Ala) (Valegård et al., 2013). In comparison, Lys89 is located within an N-terminal active site pocket of CPE and can bind with a CA molecule by hydrogen bonding (Valegård et al., 2013). The substitution of

Lys89 with Ala led to complete loss of CA production in *S. clavuligerus* (Srivastava et al., 2019). In the current study, the Δ*cpe*/pSET-*cpe*-S27A and Δ*cpe*/pSET-*cpe*-S206A variants were selected as controls since the production of CA is not affected in them (Srivastava et al., 2019). Also, the Ser27 (from the N-terminal domain) and Ser206 (from the C-terminal domain) residues are not part of any conserved motif and do not interact with CA directly based on the reported crystal structure of the protein (Valegård et al., 2013).

Wild-type *S. clavuligerus*, the Δ*cpe* mutant, Δ*cpe*/pSET152-*cpe*, and the five CPE variants (Δ*cpe*/pSET-*cpe*-S173A, Δ*cpe*/pSET-*cpe*-S234A, Δ*cpe*/pSET-*cpe*-S27A, Δ*cpe*/pSET-*cpe*-L89A, and Δ*cpe*/pSET-*cpe*-S206A) were cultured in SA and SM, and supernatant samples were collected after 48 h and 96 h of growth. Bioassays against *K. pneumoniae* were performed to test the production of CA (Figure 3.11) and zones of growth inhibition were detected at both time points for wt, Δ*cpe*/pSET152-*cpe*, Δ*cpe*/pSET-*cpe*-S27A, and Δ*cpe*/pSET-*cpe*-S206A SA and SM samples (Figure 3.11). No zones of inhibition were observed in Δ*cpe*, Δ*cpe*/pSET-*cpe*-S173A, Δ*cpe*/pSET-*cpe*-S234A, and Δ*cpe*/pSET-*cpe*-L89A supernatants (Figure 3.11), indicating the complete loss of CA production in strains containing the respective CPE variants (Supplementary Table S3.3) as reported previously by Srivastava et al. (2019). Supernatants from SA cultures were also subjected to LC-MS analysis to examine the production of CA and 2HMC as described earlier. CA was detected in wt, Δ*cpe*/pSET-*cpe*, Δ*cpe*/pSET-*cpe*-S27A, and Δ*cpe*/pSET-*cpe*-S206A supernatants, whereas it was not present in Δ*cpe*, Δ*cpe*/pSET-*cpe*-S173A, Δ*cpe*/pSET-*cpe*-S234A, and Δ*cpe*/pSET-*cpe*-L89A samples (Figure 3.12A and B, Table 3.1). The results are consistent with the bioassays and

revealed the importance of the respective amino acid residues (Ser173, Ser234, and Leu89) in CPE to perform its role in CA production. The 2HMC was not detected in wt, Δ*cpe*, and Δ*cpe*/pSET-*cpe*-S234A samples (Figure 3.12A and C), and it was detected with a very weak signal in Δ*cpe*/pSET-*cpe*-S173A, Δ*cpe*/pSET-*cpe*-L89A and unexpectedly in Δ*cpe*/pSET-*cpe*-S206A (the variant which has no effect on CA production). In comparison, 2HMC was clearly detected in Δ*cpe*/pSET-*cpe* and the CPE variant control Δ*cpe*/pSET-*cpe*-S27A (Figure 3.12A and C; Table 3.1).

In addition to their essential role in CA production and esterase activity, the Ser173 and Ser234 residues seem to have important roles in the production of 2HMC in the *cpe* overexpression strains. While the substitution of Ser173 with Ala almost completely eliminated the production of 2HMC (with some traces of the mass ion *m/z*=144 remaining), the replacement of Ser234 with Ala completely abolished the production of the metabolite (Figure 3.12C). These results suggest that substitution of Ser234 more drastically affects 2HMC production as compared to Ser173 (Table 3.1). However, the effect was the opposite in the case of esterase activity of CPE, where the substitution of Ser173 more severely impacted the activity of the protein as compared to substitution of Ser234 (Valegård et al., 2013). Both Ser173 and Ser234 residues are located in the conserved SXXK and SDN motifs of the CPE active site, respectively. Moreover, the equivalent residues to Ser173 and Ser234 in class A β-lactamases (e.g., TEM-1) and some PBPs, which have high similarity to CPE, are reported to be essential for their binding and catalysis functions (Maveyraud et al., 1998; Vandavasi et al., 2016). Also, the substitution of equivalent residues to Ser173 and Ser234 in EstB (a protein with

esterase function from *Burkholderia gladioli*) results in a complete loss of its activity (Wagner et al., 2009).

Interestingly, the substitution of Ser206 with Ala significantly reduced the production of 2HMC in Δ*cpe*/pSET-*cpe*-S206A without affecting the production of CA (Figure 3.12 and Table 3.1), suggesting that the enzymatic activity exerted by CPE toward CA biosynthesis could be different from that for 2HMC production. This hypothesis needs further investigation with other essential amino acids residue. Ser206 is located in the C-terminal of the CPE protein and does not interact with CA directly as per the crystal structure (Valegård et al., 2013). Although Ser206 is not part of any known motif (Srivastava et al., 2019), it is conserved in the CPE orthologs found in other CA BGC-containing bacteria (*S. jumonjinensis*, *S. katsurahamanus*, *S. pratensis*, and *S. viridis*) (Srivastava et al., 2019) suggesting a possible important role in the function of this protein.

The substitution of Lys89 with Ala in Δ*cpe*/pSET-*cpe*-L89A almost completely abolished the production of 2HMC. Lys89 is found in the active site pocket of CPE and can bind with CA by hydrogen bond (Valegård et al., 2013). The N-terminal domain of CPE (residues 1–127), which is important for CA production (Srivastava et al., 2019), is structurally similar to steroid isomerase enzymes from other bacteria species that also contains the equivalent residues to Lys89 (Valegård et al., 2013). Therefore, it is proposed that Ser173 and Ser234 promote a nucleophilic attack on a still unknown substrate to form the primary CPE-substrate complex, while other essential residues from the N-terminal domain (including Lys89) might be involved in stabilizing the intermediate followed product formation (Srivastava et al., 2019).

The results revealed that CPE could have different enzymatic activities. One of them attains the proposed hydrolysis step to convert 2-formyloxymethylclavam (2FMC) into 2HMC in the 5*S* clavams biosynthesis pathway (Figure 1.7). The deletion of *cpe* did not affect the production of 2HMC when bacteria were grown in SM medium, suggesting that other proteins catalyze the hydrolysis step in the cell. However, BLAST results did not reveal any similar protein for CPE in the *S. clavuligerus* genome or the 5*S* clavams producers (*S. antibioticus, S. hygroscopicus, S. platensis, S. lavendulae,* and *S. brunneogriseus*) (Jensen and Paradkar, 1999). Therefore, we hypothesize that when *S. clavuligerus* grows in SM, where the 5*S* clavam genes are expressed, one of the gene products carries out the hydrolysis of 2FMC into 2HMC, while in SA medium, where 5*S* clavams genes, or some of them, are not active, the CPE (when it is overexpressed) performs the hydrolysis step. It seems that this activity, as the esterase function described by Valegård et al. (2013), is not the main biological function of CPE.

Further studies on the potential functional residues described in Valegård et al. (2013) and Srivastava et al. (2019) are required to fully describe CPE's enzymatic mechanism in biosynthesis of CA or 5*S* clavams in both synthetic SA medium and complex SM. The complexity of understanding the function of CPE and the different activities observed in it reflects "moonlighting" activities for this protein, an idea that warrants further investigation.

3.4.8. Investigations into the involvement of *orf14* from the CA BGC in clavam metabolite production in *S. clavuligerus*.

In order to study the *orf14* gene and its participation in the CA and 5*S* clavam biosynthesis, the gene was first amplified from wt *S. clavuligerus* genome and successfully cloned into two integrative *Streptomyces* cloning vectors, pHM11a and pSET152 (Bierman et al., 1992; Motamedi et al., 1995) to give pHM11a-*orf14* and pSET152-*orf14*, respectively (Table 2.3). The two constructs were separately introduced into the *S. clavuligerus orf14* deleted mutant (Δ*orf14*) by conjugation for complementation studies. The empty vectors (pHM11a and pSET152) were also introduced into the Δ*orf14* mutant separately for use as controls. The *S. clavuligerus* wt, Δ*orf14*, Δ*orf14*/pHM-*orf14*, Δ*orf14*/pSET-*orf14*, and control strains (Δ*orf14*/pHM11a and Δ*orf14*/pSET152) were cultured in SA and SM media and supernatants were collected after 48, 96, and 120 h of growth. Bioassays using *K. pneumoniae* showed zones of growth inhibitions for Δ*orf14*/pHM-*orf14* samples grown in SM (Table 3.2; Figure 3.13), but not SA medium, indicating that *orf14* complementation was successful and CA was restored in SM only using the pHM11a vector system (Table 3.2). Supernatant samples for Δ*orf14*/pSET-*orf14* grown in SA and SM both showed inhibitory activity against *K. pneumoniae* (Table 3.2; Figure 3.13), indicating that the complementation was successful, and CA was produced in both SA and SM tested. Zones of growth inhibitions were not noted for the Δ*orf14* mutant samples (Table 3.2), which agrees with results obtained by Jensen et al. (2004), in which the deletion of *orf14* almost completely abolished the production of CA. However, Mellado et al. (2002) showed that the disruption of *orf14* caused partial loss of CA production in compared to wild type. No

zones of inhibitions were reported for Δ*orf14* mutant carrying the empty vectors pHM11a and pSET152 (Table 3.2), indicating that the restoration of CA production in the complemented strains was due to the reintroducing of *orf14*, which could now be used for further detailed studies described in Section 3.4.9.

To further test for clavam metabolite production, LC-MS analyses were performed for SA supernatants samples collected from 96-h culture of the different strains. CA was detected in wt and Δ*orf14*/pSET-*orf14* samples but not in Δ*orf14*/pHM-*orf14* and Δ*orf14* (Figure 3.13). The LC-MS results are consistent with those obtained from bioassays results and confirm that the complementation was achieved in Δ*orf14*/pSET-*orf14* but not in Δ*orf14*/pHM-*orf14* when grown in SA medium.

To investigate whether overexpression of *orf14* would induce the production of 2HMC or any other 5*S* clavams in SA medium as seen for *cpe* (*orf12*), the LC-MS spectra of Δ*orf14*/pSET-*orf14* were inspected for *m/z* values corresponding to 2HMC (Figure 3.13B) and the other 5*S* clavams (C2C, 2FMC, and alanylclavam). Analysis did not show any peak related to 5*S* clavams, indicating that the overexpression of *orf14* in the complemented strain (Δ*orf14*/pSET-*orf14*) does not affect the 5*S* clavams biosynthesis pathway.

3.4.9. Examining the effect of different variants of ORF14 on the production of clavulanic acid

Although we did not see 2HMC production in the *orf14* complementation studies in SA medium, the previously reported crystal structure showed that ORF14 is a member of the tandem GNAT subfamily. It was also shown that a single molecule of acetyl

coenzyme A (AcCoA) binds to the N-terminal domain via a predominantly hydrophobic pocket (formed in part Trp106 and Phe142) and other acyl-CoA derivatives could bind to its C-terminal domain (Iqbal et al., 2010). In addition, the N-terminal binding site does not contain residues nearby predicted to be involved in general acid/base catalysis. The C-terminal binding site is larger to accommodate derivatives with a potential oxyanion hole formed by the backbone amides of Met268 and Thr269 along with other hydrophobic residues (Val254, Leu289, Val292, and Leu293), suggesting that it may accommodate CoA analogs with hydrophobic character (Iqbal et al. 2010). Therefore, Trp106 and Phe142 from the N-terminal domain and Val254, Met268, Thr269, Leu289, Val292, and Leu293 residues in the C-terminal domain (Figure 3.14) of ORF14 were selected for mutagenesis studies to examine their *in vivo* contributions during CA production in *S. clavuligerus*.

The pHM11a vector containing *orf14* (pHM11a-*orf14*) was subjected to site-directed mutagenesis using the QuikChange II Site-Directed Mutagenesis Kit as described in the Materials and Methods chapter, in which the codons of the respective amino acids were replaced by the codon for alanine (Figure 3.14). The site-directed mutagenesis was accomplished for four targeted residues (Phe142Ala, Val254Ala, Thr269Ala, and Val292Ala) and confirmed by sequencing (Figure 3.14B). Unfortunately, the mutagenesis for the remaining residues was not confirmed. The *orf14* variants were then assessed for their ability to complement the *S. clavuligerus* Δ*orf14* mutant. The *S. clavuligerus* wt, Δ*orf14*, Δ*orf14*/pHM-*orf14*, and four *orf14* variants (Δ*orf14*/pHM-*orf14*-F142A, Δ*orf14*/pHM-*orf14*-V254A, Δ*orf14*/pHM-*orf14*-T269A, and Δ*orf14*/pHM-*orf14*-V292A) were cultivated in SM medium, and samples were collected at 48- and 96-h time points.

CA production was assessed by bioassays, and the zones of growth inhibitions were measured and reported (Figure 3.15; Supplementary Table S3.4). All ORF14 variants successfully complemented Δ*orf14* and produced CA, indicating that the substitutions of Phe142, Val254, Thr269, or Val292 with Ala does not affect CA biosynthesis.

Since substitutions in ORF14 did not lead to loss in CA production, additional studies were not pursued. The N-terminal domain of ORF14 strongly binds to AcCoA, which is more deeply situated as compared to other reported GNAT proteins (Vetting et al., 2003). Therefore, replacing Phe142 with Ala in that hydrophobic pocket of ORF14 might not exert enough change that affects the protein's function, suggesting that the substitution of more than one amino acid in the hydrophobic pocket may be required to make a difference. In the C-terminal domain, which was proposed to be the catalytic region, the substitutions of Val254, Thr269, or Val292 with Ala did not affect the production of CA. This suggests that additional amino acids in the predicted binding pocket such as Met268, Leu289, Leu293 (or others) might play a more critical role in the function of ORF14 than the ones we tested. Therefore, it is possible that the substitution of more than one amino acid is required to see alteration in the function of ORF14 or that the protein could have other role in CA biosynthesis than the proposed acetylation reaction. It was reported that the addition of AGCA (the already acetylated glycyl-clavaminic acid) to Δ*orf14* mutant culture does not lead to CA production (Álvarez-Álvarez et al., 2018), suggesting that ORF14 contributes to the pathway after the acetylation step. This hypothesis requires further investigation. In conclusion, it is apparent that the phenotypes observed during complementation studies using *orf14* were completely different from those observed for *cpe*, where there was a clear correlation with

catalytic residues and CA production, with the concomitant production of 2HMC in SA medium in some cases.

3.5. Conclusion

Although CPE is located in the cytoplasm of *S. clavuligerus*, and the Δ*cpe* mutants are entirely blocked in CA production, deletion of *cpe* didn't show any regulatory effect on the transcription of CA biosynthesis genes (Figure 3.5). However, we demonstrated that CPE, when it is overexpressed, has a positive impact on the regulation of essential genes (*cas1, cvm1, cvm2,* and *cvm5)* in clavam BGC, and *cvm7p* and *cvm6p* in the paralogue BGC, but did not affect the transcription of *snk, res1,* and *res2* genes. This effect of *cpe* overexpression results in the production of 2HMC when the corresponding strains *S. clavuligerus*/Δ*cpe*/pSET-*cpe* and *S. clavuligerus*/Δ*cpe*/pHM-*cpe* cultured in SA. This result has not been reported before.

In SM, the production of the 5*S* clavams was mediated by the induction of the two-component system Snk-Res1/2, as reported by (Kwong et al., 2012). Our results indicate that the production of the 2HMC was independent of the Snk-Res1/2 system in SA media (Figure 3.9).

In addition, our site-directed mutagenesis results revealed that CPE could exert enzymatic activity toward 2HMC different from that for CA. Therefore, we proposed that CPE, when it is overexpressed, carries out the hydrolysis step to convert 2-formyloxymethylclavam (2FMC) into 2HMC in the 5*S* clavams biosynthesis pathway (Figure 1.7), and it seems that this activity is not the primary biological function of CPE.

3.6. Figures and tables

3.6.1. Figures

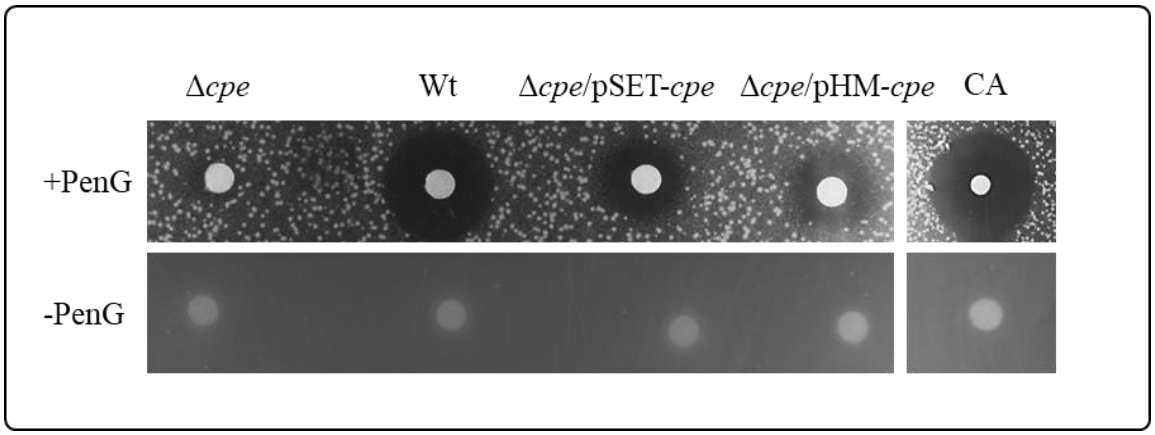
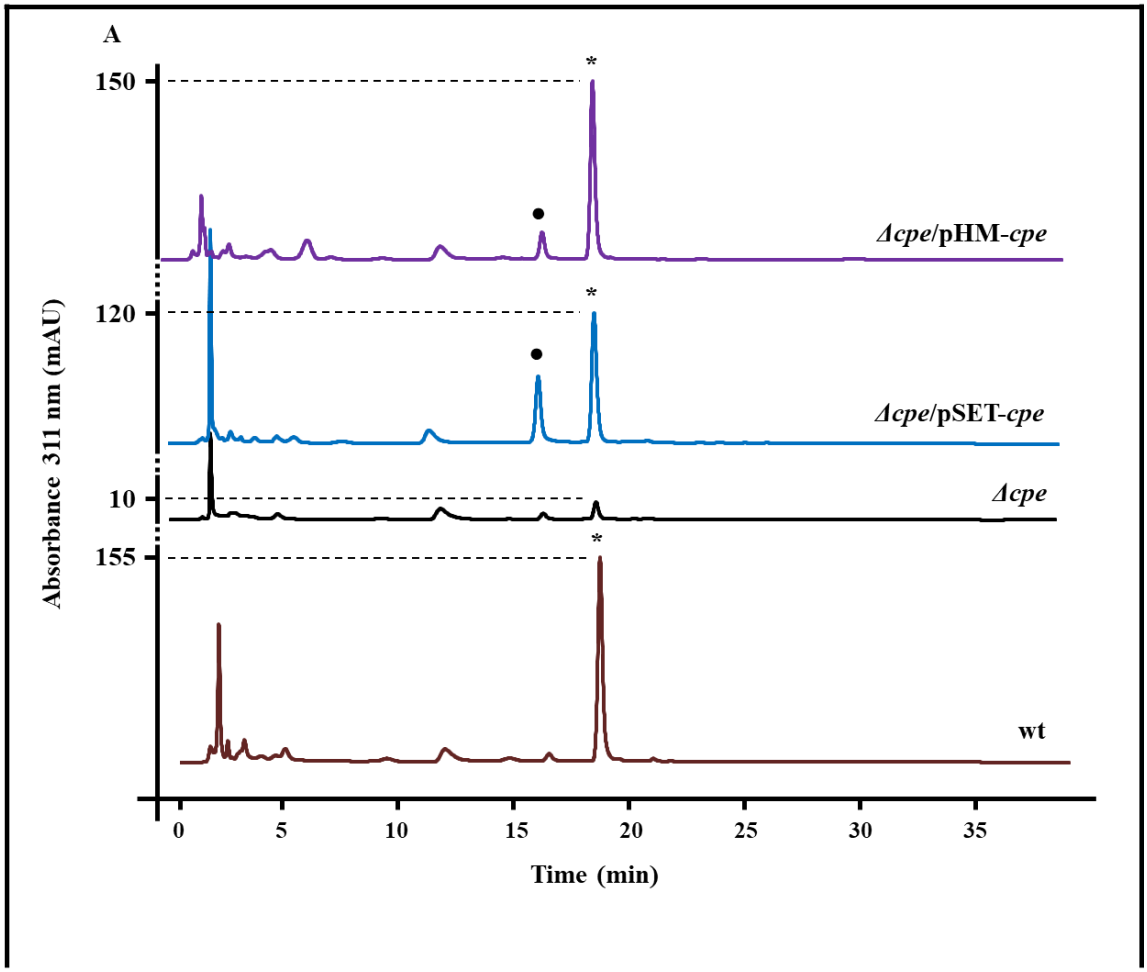
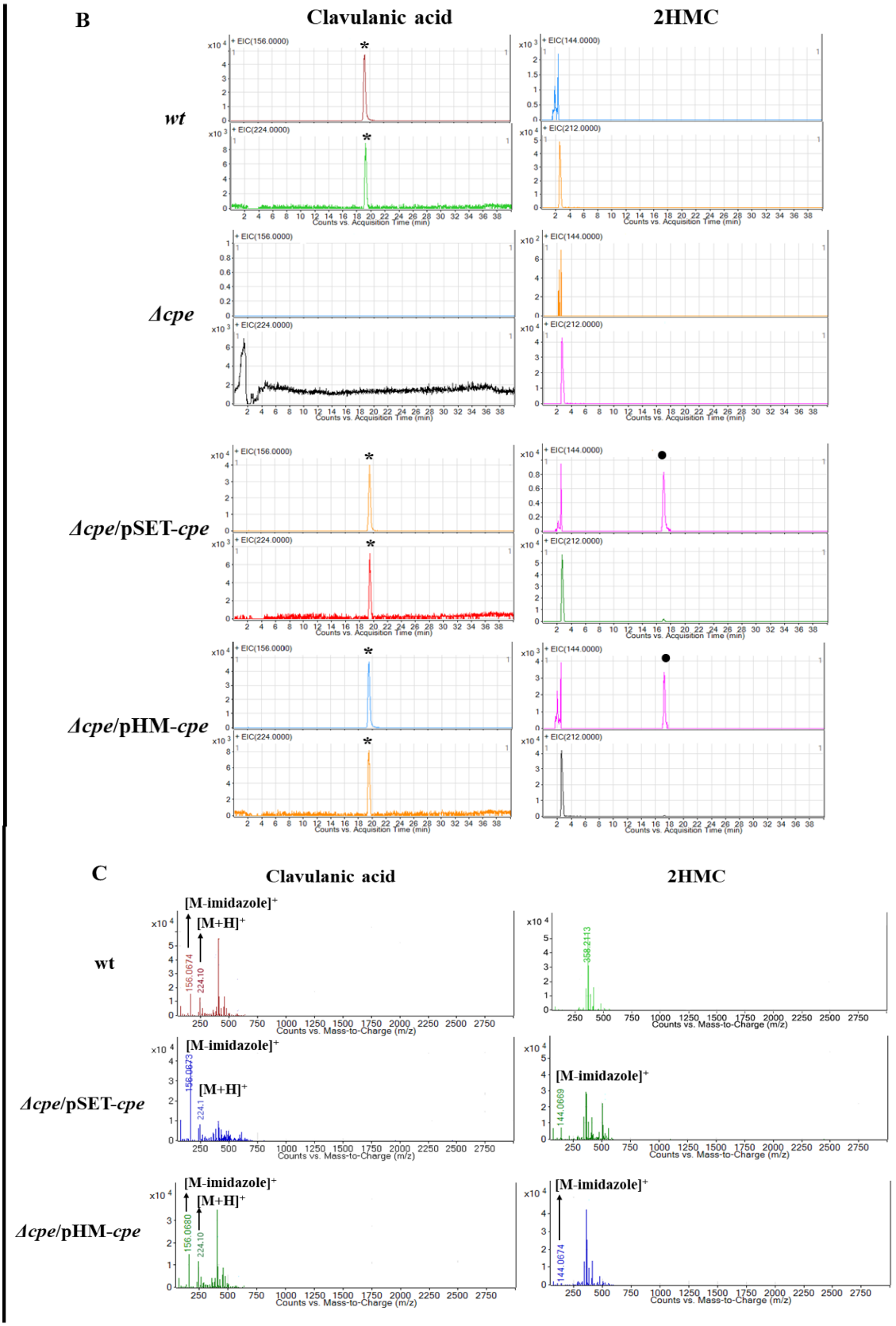


Figure 3.1. Clavulanic acid bioassay for *cpe* (*orf12*) complemented *S. clavuligerus* strains. Supernatants from liquid cultures for wt *S. clavuligerus*, *S. clavuligerus*/Δ*cpe*, Δ*cpe*/pSET152-*cpe*, Δ*cpe*/pHM11a-*cpe*, were tested for CA production. Each culture was grown for 96 h in SA medium and bioassays were performed against *K. pneumoniae* on TSA plates with or without PenG. The picture shows results for one of the two replicates completed for each culture. CA: clavulanic acid solution (10 μg).


A
150
120
10
155
*
Δcpe/pHM-cpe
Δcpe/pSET-cpe
Δcpe
wt
Absorbance 311 nm (mAU)
0
5
10
15
20
25
30
35
Time (min)


B
Clavulanic acid
2HMC
wt
Δcpe
Δcpe/pSET-cpe
Δcpe/pHM-cpe
+ EIC(156.0000)
+ EIC(224.0000)
+ EIC(144.0000)
+ EIC(212.0000)
Counts vs. Acquisition Time (min)
C
Clavulanic acid
2HMC
wt
Δcpe/pSET-cpe
Δcpe/pHM-cpe
[M-imidazole]+
[M+H]+
156.0674
224.10
358.2413
156.0673
224.1
144.0669
156.0680
224.10
144.0674
Counts vs. Mass-to-Charge (m/z)

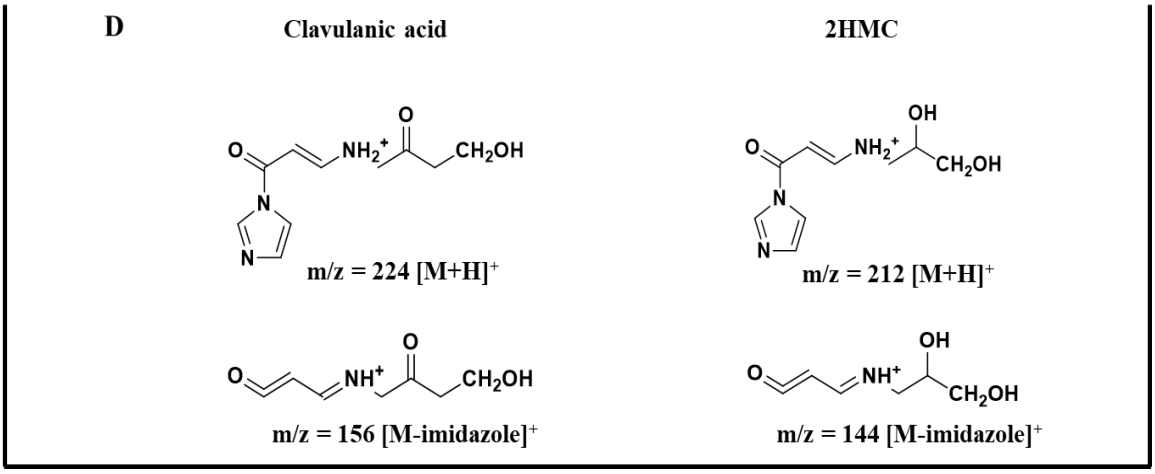


Figure 3.2. Detection of 2-hydroxymethyl clavam (2HMC) in *cpe* overexpression strains. LC-MS analysis of 96-h SA culture supernatants after imidazole derivatization using the ammonium bicarbonate buffer system. Cultures from wt *S. clavuligerus*, Δ*cpe*, Δ*cpe*/pSET-*cpe*, and Δ*cpe*/pHM-*cpe* strains were used to assess CA and 2HMC metabolite production. (A) Liquid chromatography profiles showing the elution of the peaks corresponding to imidazole-derivatized clavulanic acid (indicated by the star symbol *) and the 2HMC (indicated by the black dot ●). (B) The extracted ion chromatograms (EIC) for the mass spectra showing the major peaks corresponding imidazole-derivatized clavulanic acid $[M+H]^+$ (*m/z* = 224) and the fragmented product $[M\text{-imidazole}]^+$ (*m/z* = 156) (indicated by the star symbol), and the imidazole- derivatized 2HMC fragmented product $[M\text{-imidazole}]^+$ (*m/z* = 144) (indicated by the black dot) which were detected in supernatants from the strains shown in (A). (C) The mass spectra chromatograms for the major peaks shown on (B) corresponding to imidazole-derivatized clavulanic acid and 2HMC. (D) Proposed chemical structures for the derivatized clavulanic acid and 2HMC.

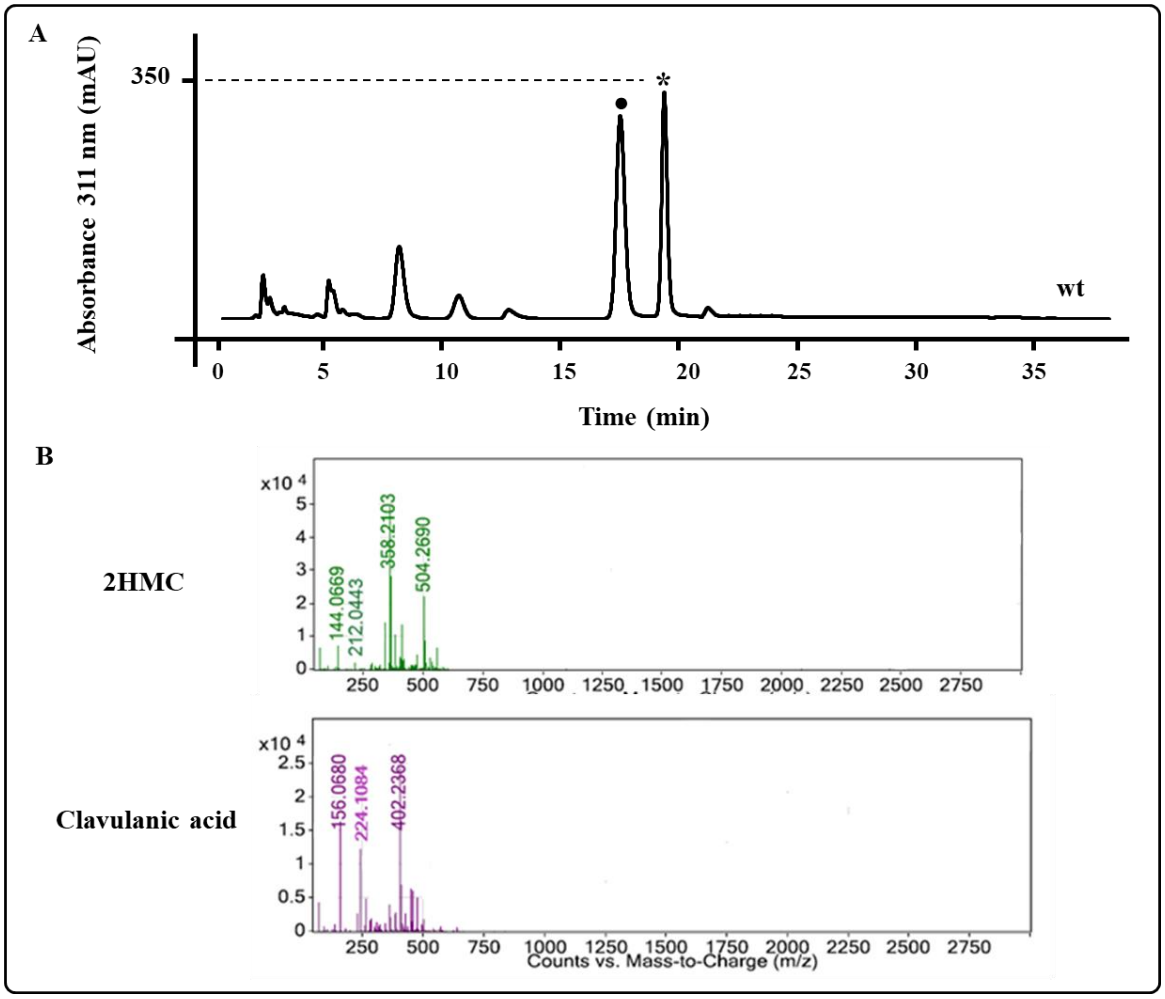
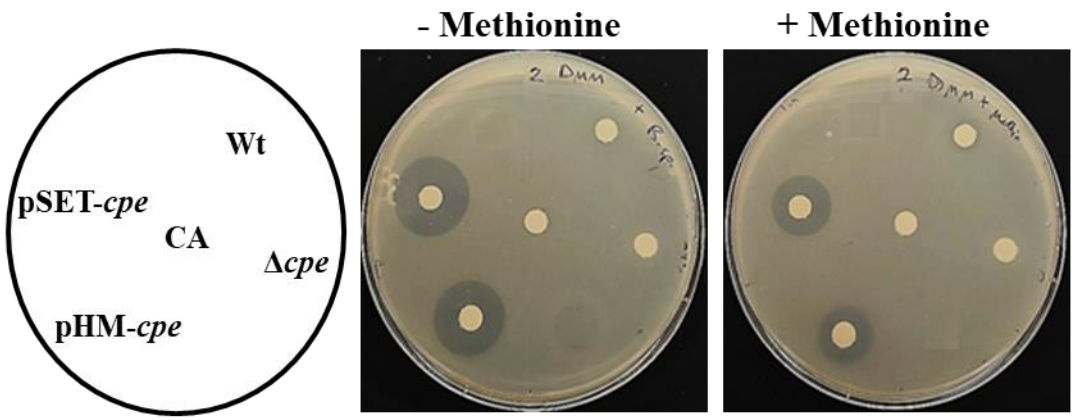


Figure 3.3. Detection of clavulanic acid and 2-hydroxymethyl clavam (2HMC) in soy medium. LC-MS analysis for 96-h soy medium (SM) culture of wt *S. clavuligerus* supernatant after imidazole derivatization using the ammonium bicarbonate buffer system. (A) Liquid chromatography profiles showing the elution of the peaks corresponding to imidazole-derivatized clavulanic acid (indicated by the star symbol *) and the 2HMC (indicated by the black dot ●). (B) The mass spectra of the major peaks corresponding imidazole-derivatized 2HMC $[M\text{-imidazole}]^+$ (*m/z* = 212) and the fragmented product $[M\text{-imidazole}]^+$ (*m/z* = 144), and the imidazole- derivatized clavulanic acid $[M+H]^+$ (*m/z* = 224) and the fragmented product $[M\text{-imidazole}]^+$ (*m/z* = 156).



	ZOI of *Bacillus sp.* (mm)	
	- Methionine	+ Methionine
***S. clavuligerus* (wt)**	0	0
Δ*cpe*	0	0
Δ*cpe*/pSET-*cpe*	22	15
Δ*cpe*/pHM-*cpe*	20	13
Clavulanic acid	0	0

Figure 3.4. 5*S* clavam bioactivity assay. The production of 2HMC by Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe* in SA medium was tested against *Bacillus* sp. ATCC27860. The bioassay was performed on Davis-Mingioli medium. To exclude the bioactivity of alanylclavam, one plate (right) was supplemented with 200 μg/ml methionine, which has an antagonistic effect against alanylclavam. The zones of inhibition (ZOI) were detected in the two plates. Samples from wt *S. clavuligerus* and the Δ*cpe* mutant were also tested. Clavulanic acid (CA) solution (10 μg) was included as a control. The table shows the measurements of ZOIs diameters from the two plates.

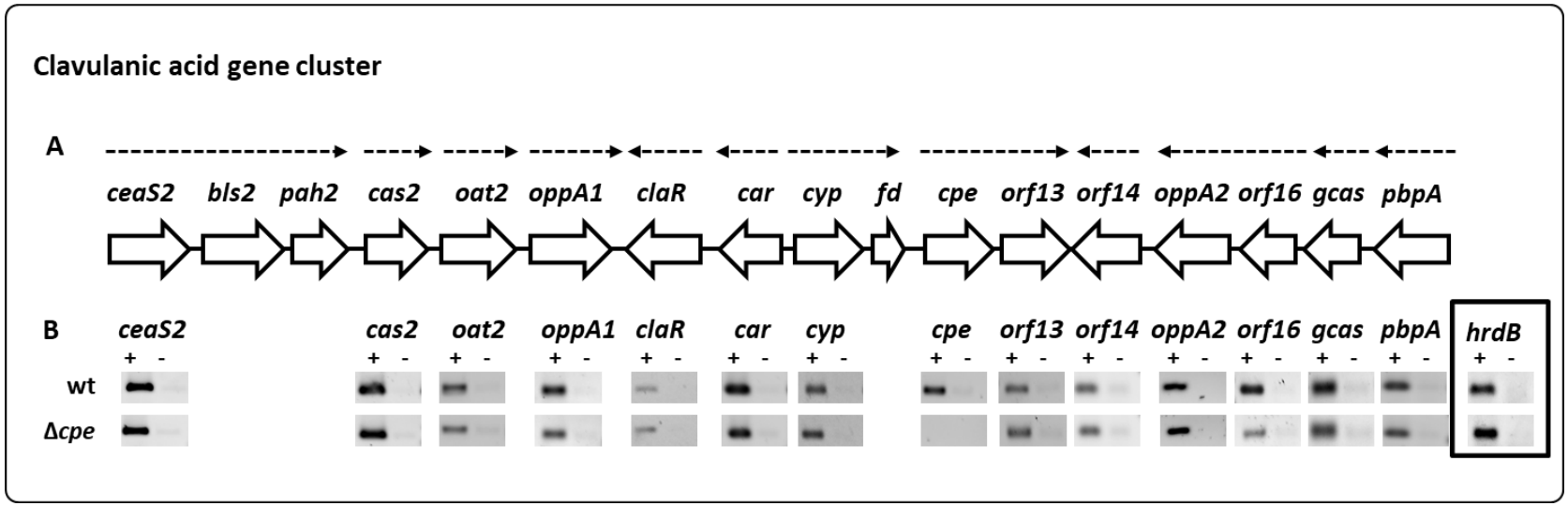


Figure 3.5. Transcriptional analysis of genes from the clavulanic acid biosynthetic gene cluster (BGC) in wt *S. clavuligerus* and the Δ*cpe* mutant. (A) The overall architecture of the CA BGC is shown with each hollow arrow representing a gene and the arrowhead its orientation. The known transcriptional units are also indicated, and the broken lines represent transcripts (B). The first gene from each transcriptional unit in (A) was selected for analysis to determine its comparative expression level in the two respective strains. RNA isolated from wt *S. clavuligerus* and the Δ*cpe* mutant after 48 h of growth in SA medium was used for RT-PCR (+) analysis. As controls, treated RNA samples were used directly in PCR without RT or cDNA synthesis (-). The expression of the constitutively expressed *hrdB* gene (extreme right boxed panel) was used as internal control to normalize expression levels between different samples/strains.

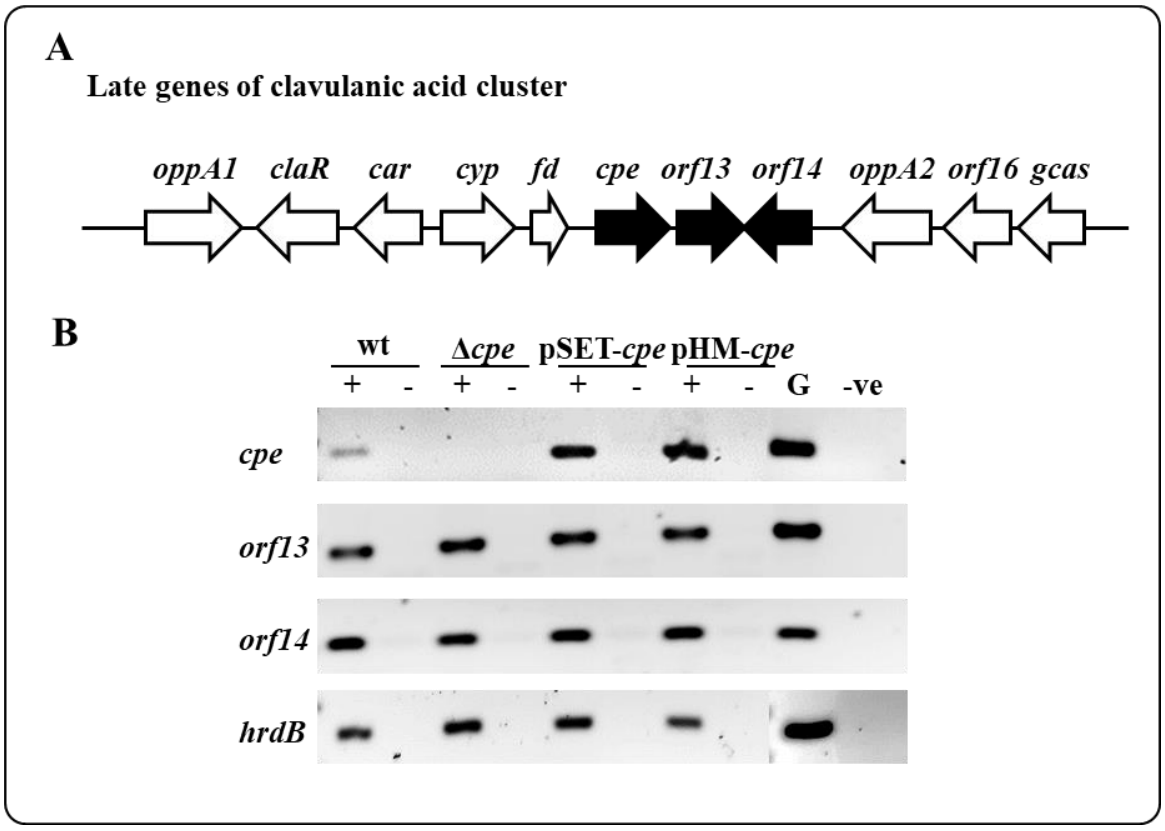


Figure 3.6. Transcriptional analysis of genes from clavulanic acid gene cluster in wt *S. clavuligerus*, Δ*cpe* mutant, and the two complemented strains (Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe*). (A) The relative arrangement of genes responsible for the late steps of clavulanic acid biosynthesis in *S. clavuligerus*. (B) The expression of *cpe* and the two downstream genes *orf13* and *orf14* genes (black arrows in A) were tested in the four strains. RNA isolated after 96 h of growth in SA medium was used for RT-PCR (+) analysis. As controls, treated RNA samples were used directly in PCR without RT or cDNA synthesis (-). PCR with genomic DNA (G, positive control) or without the addition of any template (-ve), negative control) were also conducted. The constitutively expressed *hrdB* gene expression was used as internal control to normalize expression levels between different samples/strains.

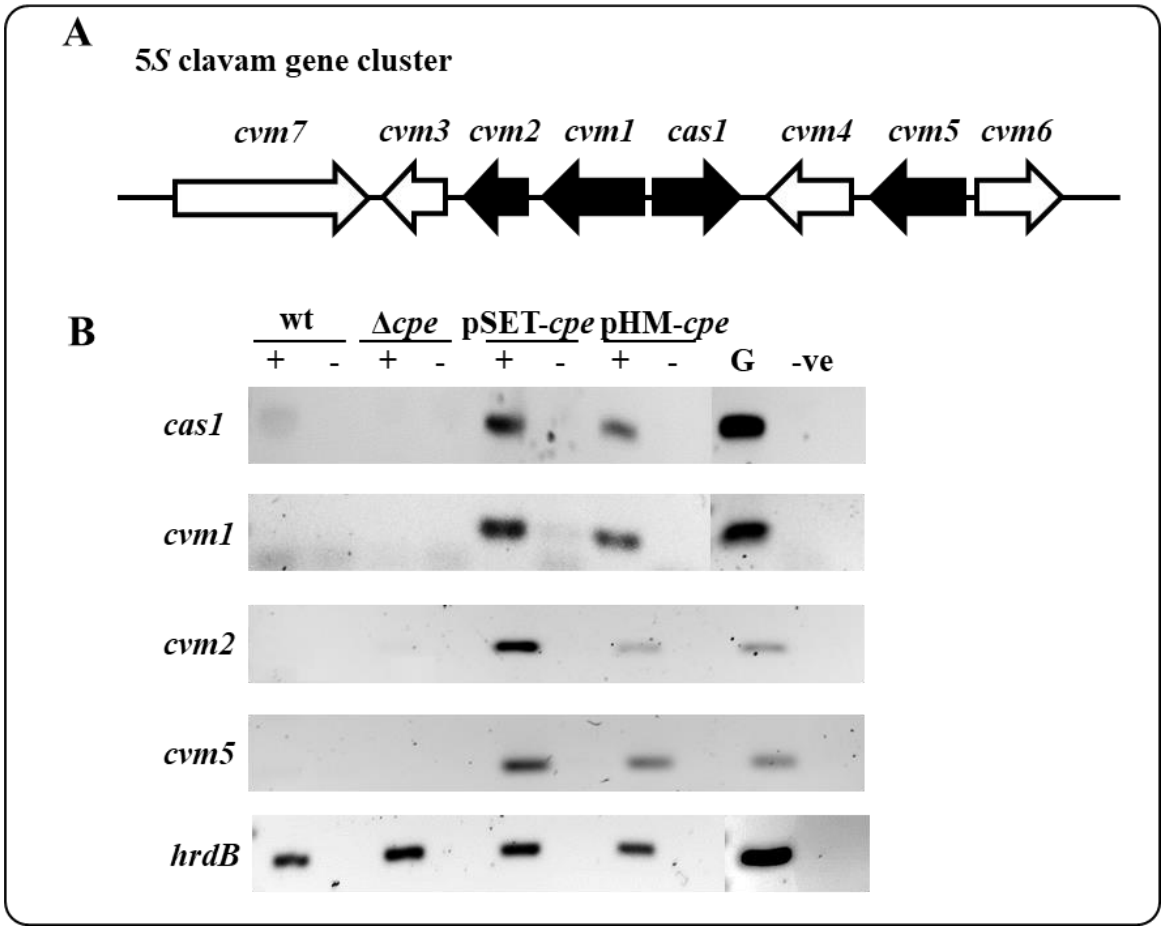


Figure 3.7. Transcriptional analysis of essential genes from 5*S* clavam gene cluster in wt *S. clavuligerus*, Δ*cpe* mutant, and the two complemented strains Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe*. (A) The architecture of the 5*S* clavam gene cluster; black arrows represent genes with important roles in 5*S* clavam biosynthesis in *S. clavuligerus*, and their expression was therefore tested. (B) The expression of *cas1*, *cvm1*, *cvm2*, and *cvm5* genes (black arrows in A) was tested in the four strains. RNA isolated after 96 h of growth in SA medium was used for RT-PCR (+) analysis. As controls, treated RNA samples were used directly in PCR without RT or cDNA synthesis (-). PCR with genomic DNA (G) or without (-ve) were also conducted. The constitutively expressed *hrdB* gene was used as internal control to normalize expression levels between different samples/strains.

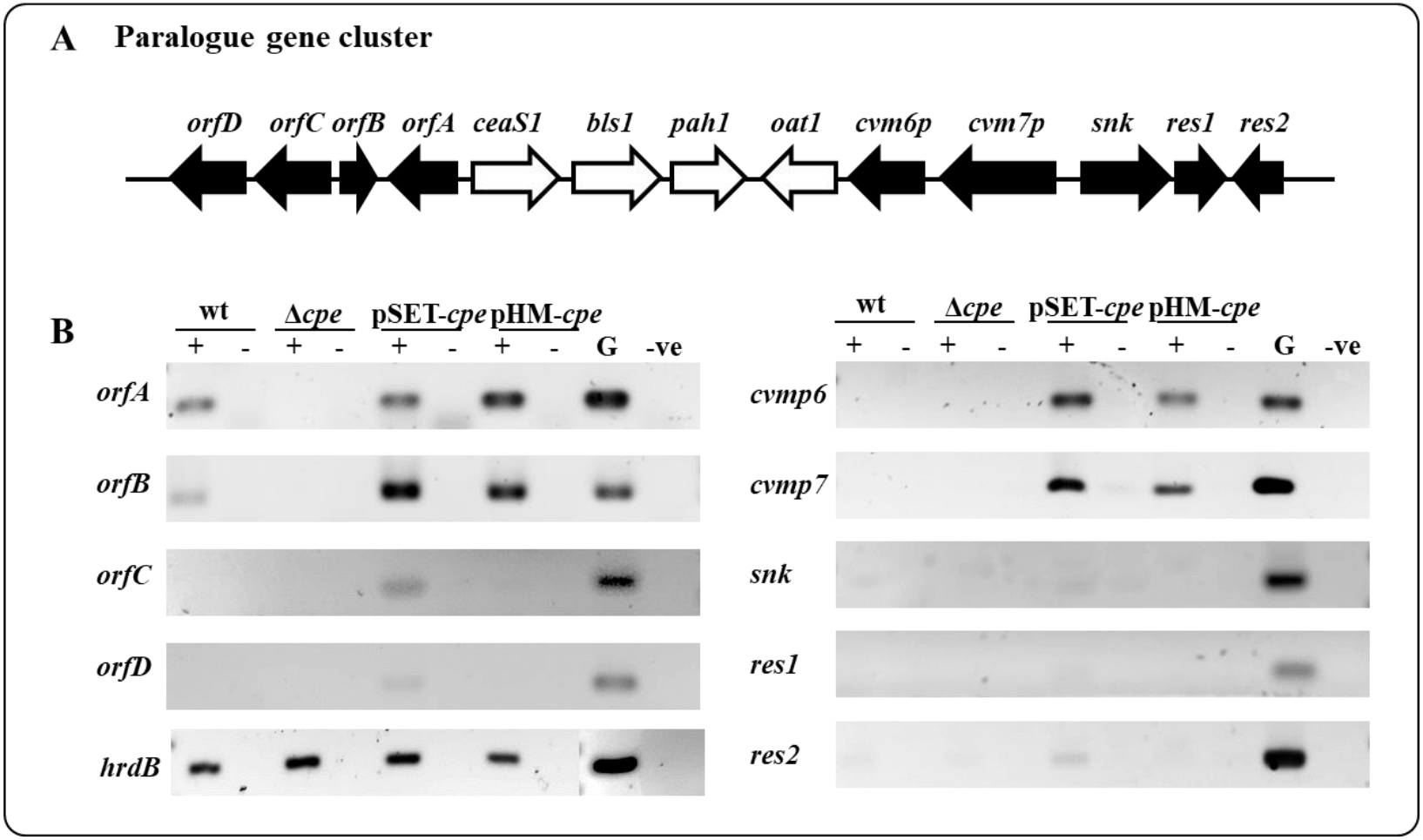

A
Paralogue gene cluster
orfD
orfC
orfB
orfA
ceaS1
bls1
pah1
oat1
cvm6p
cvm7p
snk
res1
res2
B
wt
Δcpe
pSET-cpe
pHM-cpe
+ - + - + - + - G -ve
orfA
orfB
orfC
orfD
hrdB
wt
Δcpe
pSET-cpe
pHM-cpe
+ - + - + - + - G -ve
cvmp6
cvmp7
snk
res1
res2

Figure 3.8. Transcriptional analysis of essential genes from paralogue gene cluster in wt *S. clavuligerus*, Δ*cpe* mutant, and the two complemented strains Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe*. (A) The architecture of the paralogue gene cluster; black arrows represent genes with important role in clavam biosynthesis in *S. clavuligerus* and were tested for their expression. (B) The level of RNA expression was tested in the four strains for *orfA, orfB, orfC,* and *orfD* genes (left panel), which are essential for alanylclavam production, *cvmp6* and *cvmp7* (right panel) for clavam production, and the two-component system genes snk, *res1,* and *res2* (right panel). RNA isolated after 96 h of growth in SA medium was used for RT-PCR (+) analysis. As controls, treated RNA samples were used directly in PCR without RT or cDNA synthesis (-). PCR with genomic DNA (G) or without (-ve) were also conducted. The constitutively expressed *hrdB* gene was used as internal control to normalize expression levels between different samples/strains.

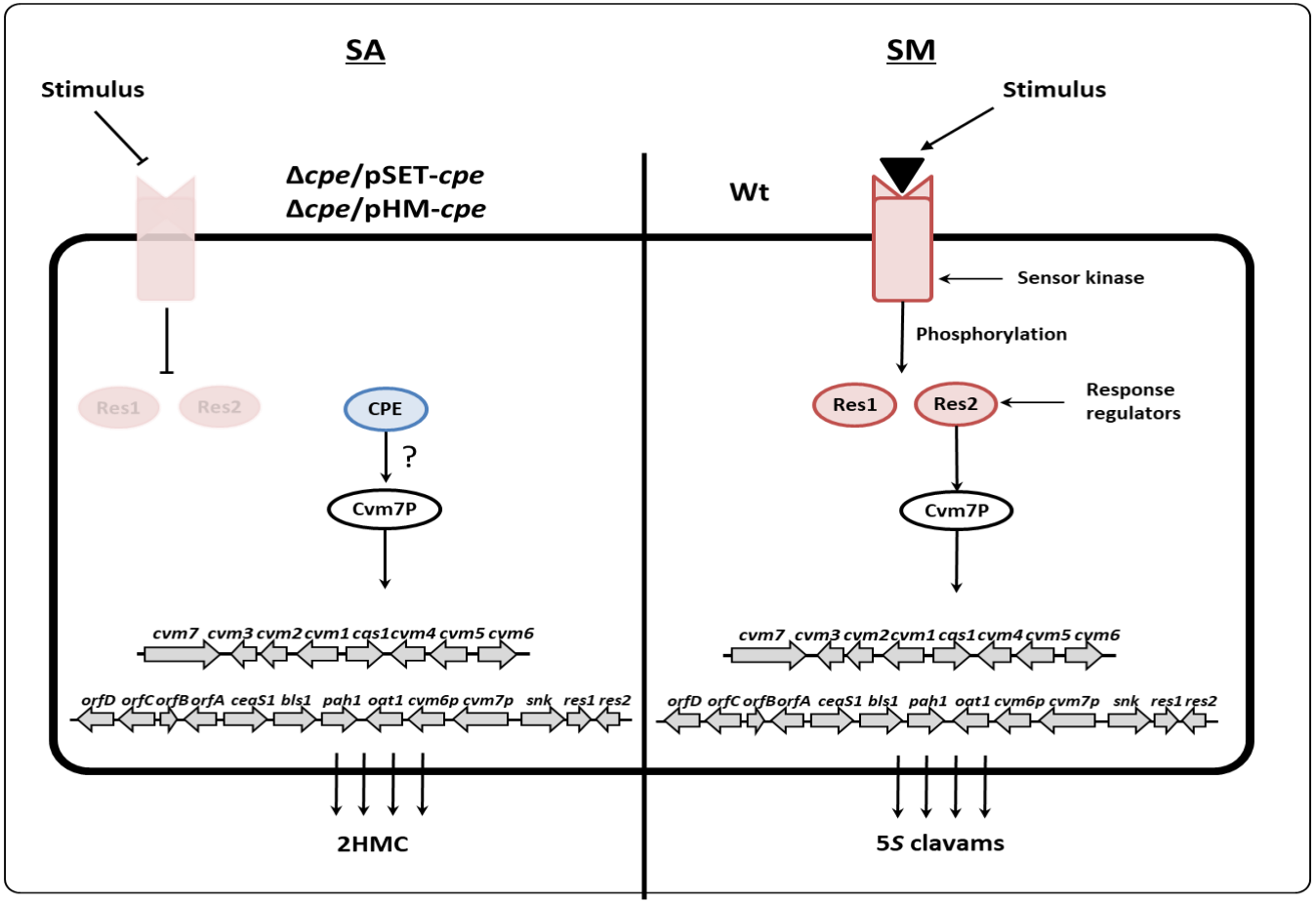


Figure 3.9: Schematic diagram for the proposed mechanism of 2HMC production in starch asparagine medium. The right side shows the proposed mechanism for the production of 5*S* clavams by wt *S. clavuligerus* grown in soy medium (SM). The left side of the figure shows the proposed mechanism for the production of 2HMC by *cpe* overexpression strains Δ*cpe*/pSET-*cpe* and Δ*cpe*/pHM-*cpe* grown in SA medium.

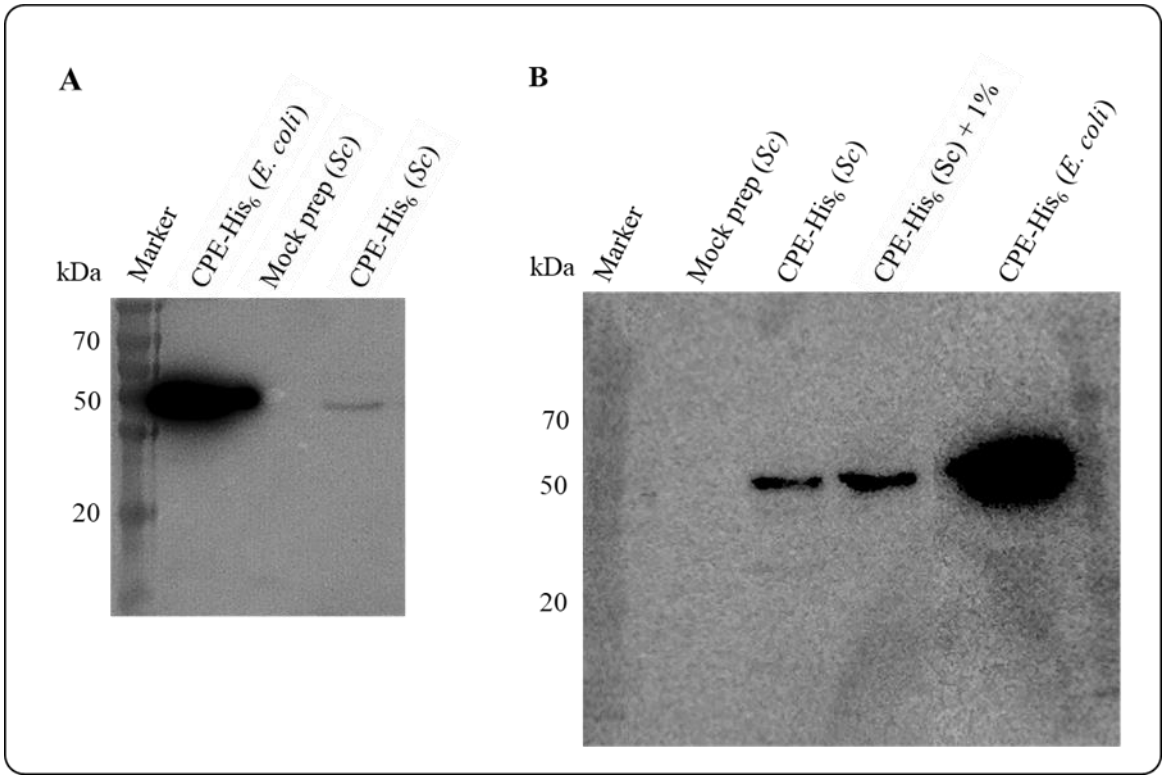


Figure 3.10: CPE detection and crosslinking. (A) CPE-His_6 tagged was expressed in *S. clavuligerus* for western blot analysis. Mycelial lysates from *S. clavuligerus*/*cpe*-His_6 and *S. clavuligerus*/pHM11a empty vector (Mock prep) cultures were prepared by sonication. The supernatant samples which contain the cellular extracts were separated by 12% SDS-PAGE. A previously purified CPE-His_6 tagged protein from *E. coli* cells was used as positive control. The Mock prep (from *S. clavuligerus*/pHM11a) was used as a control to account for any non-specific antibody binding. Western blot analysis was conducted along with the anti-6×His Tag Monoclonal Antibodies (B) Bacterial cultures of *S. clavuligerus/cpe*-His_6 were treated with 0 or 1% formaldehyde to induce protein-protein binding. CPE-His_6 proteins in this gel were purified using a Ni-NTA gravity elution column. The size of the band corresponding to CPE-His_6 in all lanes was approximately 50-55 kDa, and the prestained protein ladder (Marker) was used as a reference for estimating molecular weight. (Sc); *S. clavuligerus*.

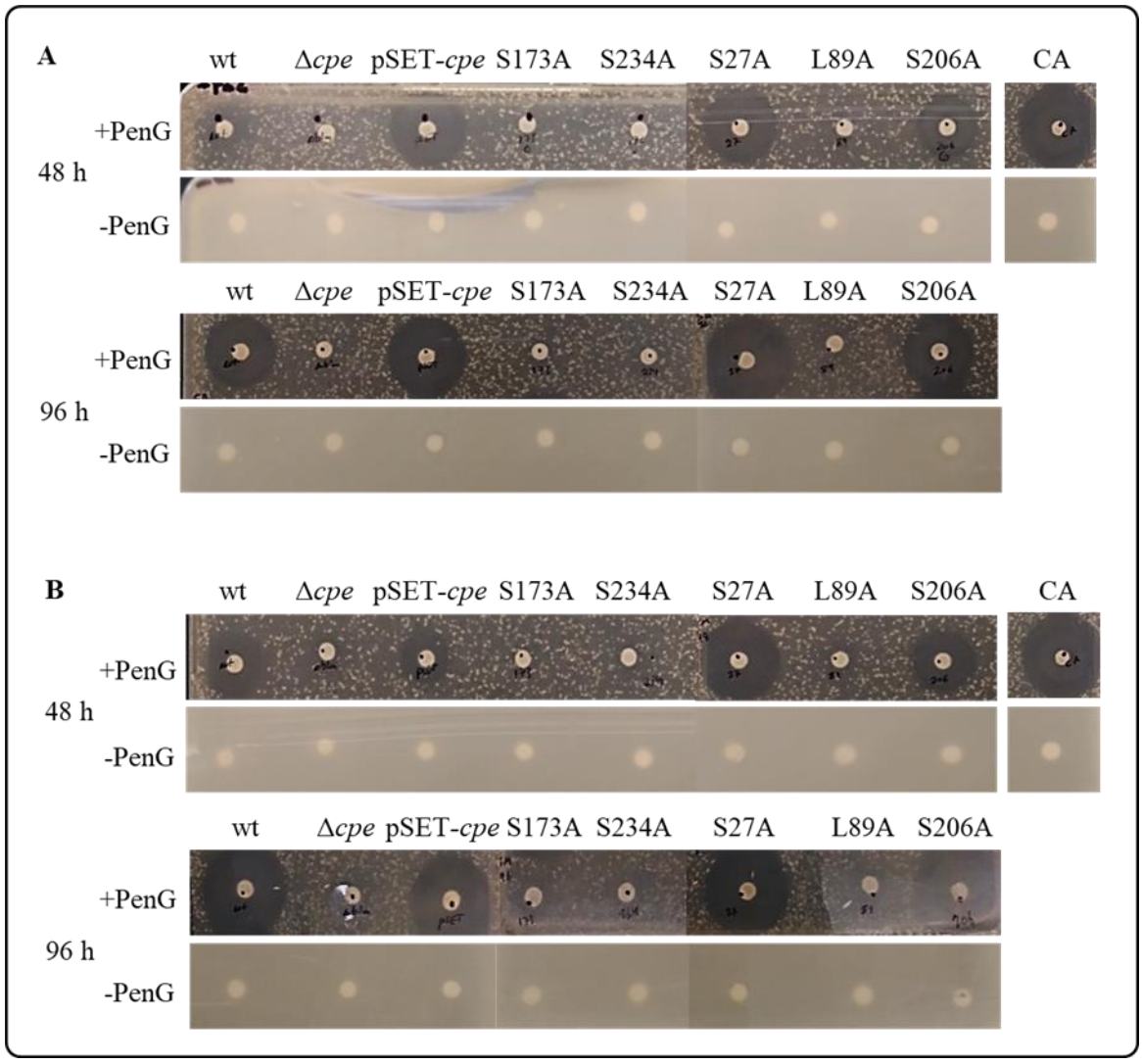
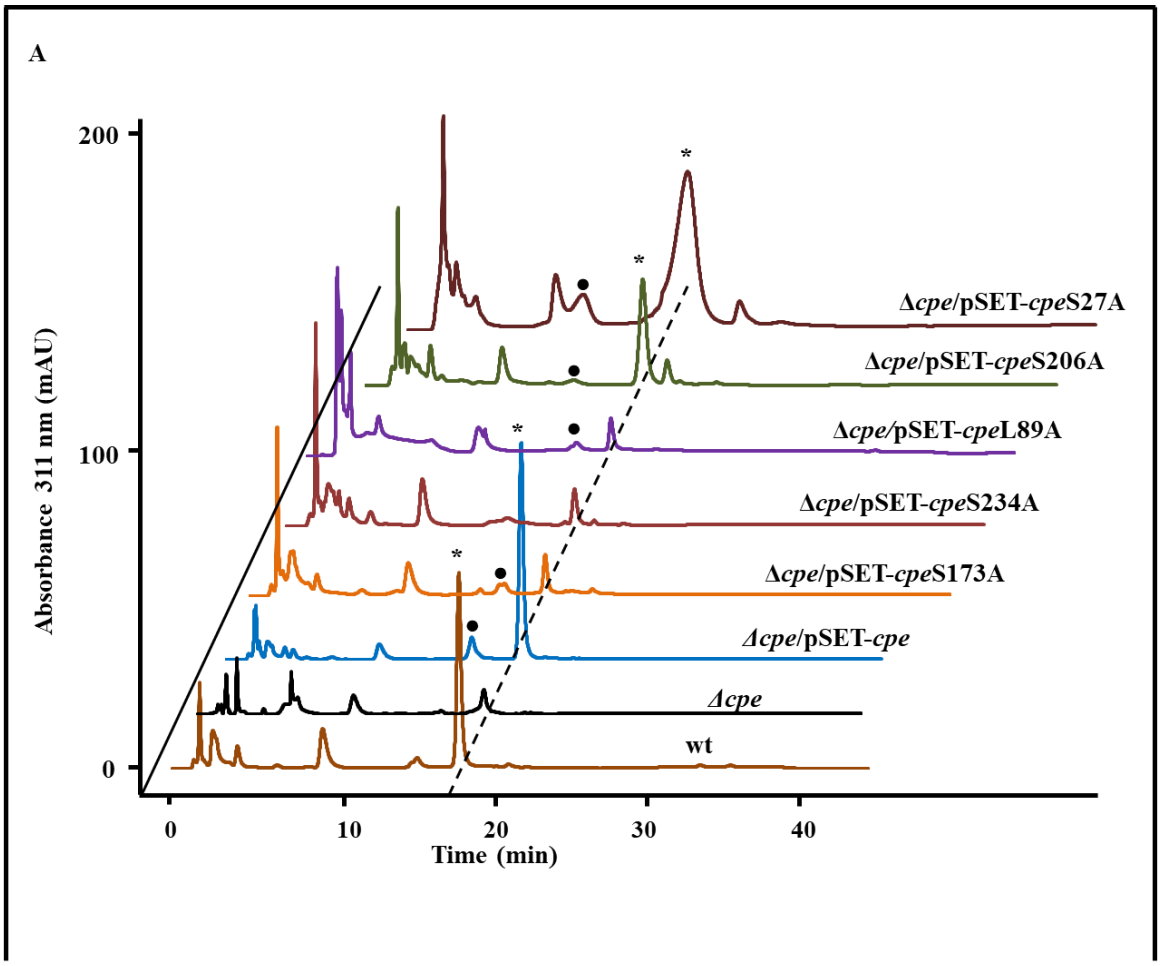
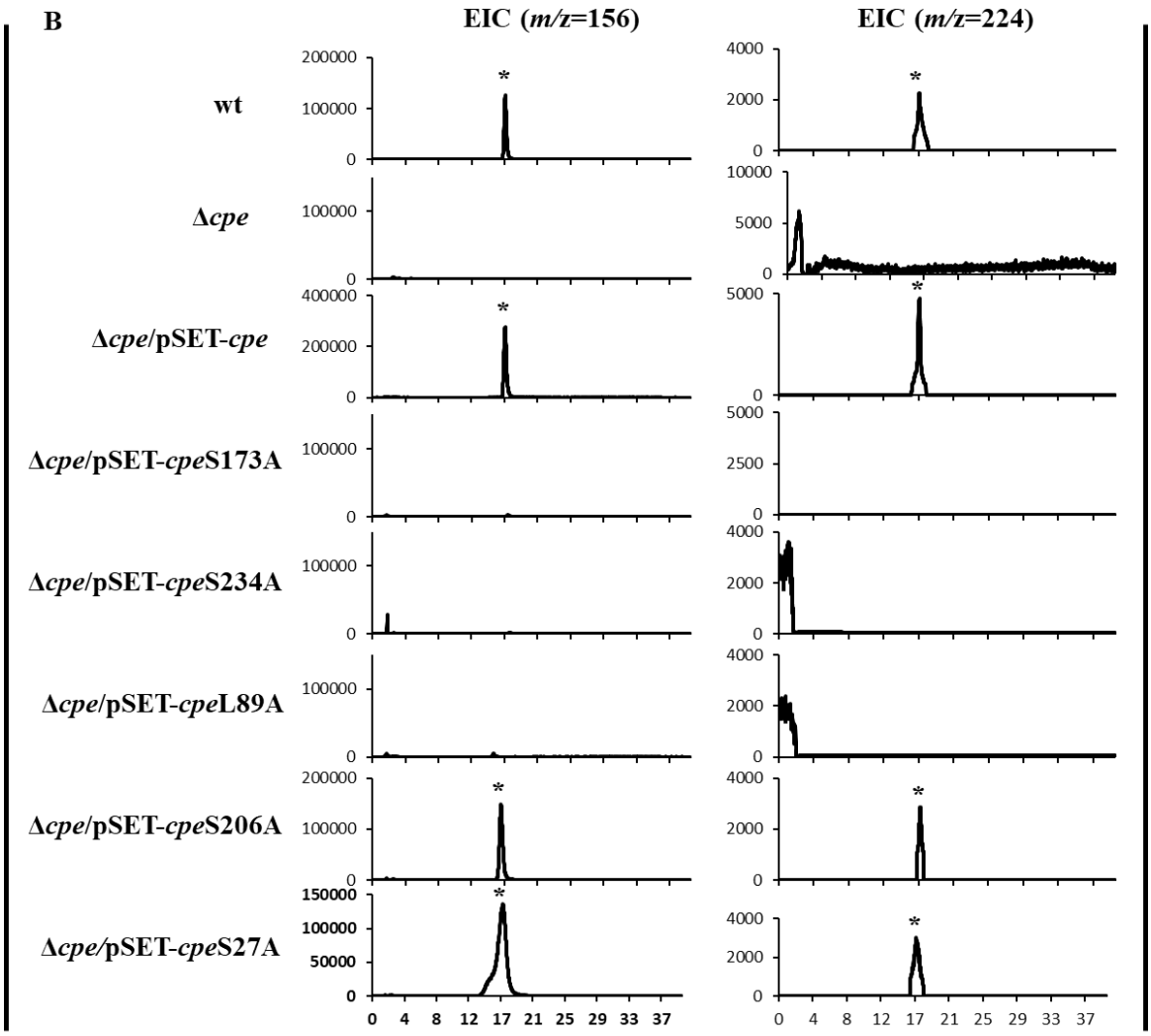


Figure 3.11: Clavulanic acid bioassay results for wt, Δ*cpe*, Δ*cpe*/pSET152-*cpe*, and CPE variant strains. Supernatants from liquid cultures for wt *S. clavuligerus*, Δ*cpe*, Δ*cp*e/pSET152-*cpe*, Δ*cpe*/pSET152-*cpe*-S173A, Δ*cpe*/pSET152-*cpe*-S234A, Δ*cpe*/pSET152-*cpe*-S27A, Δ*cpe*/pSET152-*cpe*-L89A, Δ*cpe*/pSET152-*cpe*-S206A were tested for CA production. Each culture was grown for 48 and 96 h in SA medium (A) or SM (B). The bioassays were performed against *K. pneumoniae* on TSA medium with penicillin G (+PenG) or without (-PenG). The picture shows results for one of the two replicates completed for each culture.


A
200
100
0
Absorbance 311 nm (mAU)
0
10
20
30
40
Time (min)
Δcpe/pSET-cpeS27A
Δcpe/pSET-cpeS206A
Δcpe/pSET-cpeL89A
Δcpe/pSET-cpeS234A
Δcpe/pSET-cpeS173A
Δcpe/pSET-cpe
Δcpe
wt


B
EIC (m/z=156)
EIC (m/z=224)
wt
Δcpe
Δcpe/pSET-cpe
Δcpe/pSET-cpeS173A
Δcpe/pSET-cpeS234A
Δcpe/pSET-cpeL89A
Δcpe/pSET-cpeS206A
Δcpe/pSET-cpeS27A
0 4 8 12 17 21 25 29 33 37

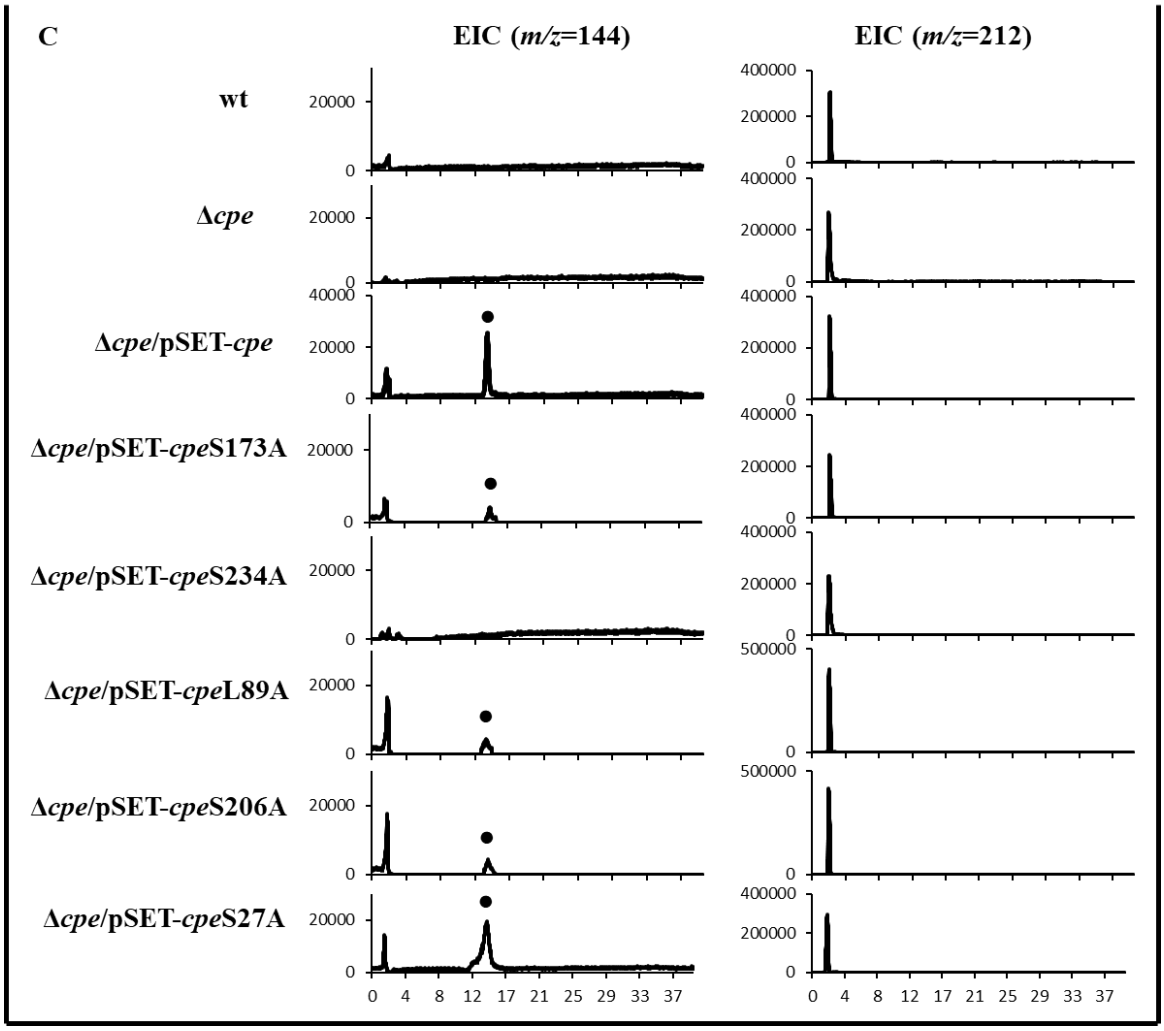


Figure 3.12: Detection of 2-hydroxymethyl clavam (2HMC) in CPE variant strains. LC-MS analysis of 96-h SA culture supernatants after imidazole derivatization using the ammonium bicarbonate buffer system. Cultures from *S. clavuligerus* wt, Δ*cpe*, Δ*cpe*/pSET152-*cpe*, and the CPE different variants strains Δ*cpe*/pSET-*cpe*-S173A, Δ*cpe*/pSET-*cpe*-S243A, Δ*cpe*/pSET-*cpe*-L89A, Δ*cpe*/pSET-*cpe*-S206A, and Δ*cpe*/pSET-*cpe*-S27A were used to assess CA and 2HMC metabolites production. (A) Liquid chromatography profiles showing the elution of the peaks corresponding to imidazole-derivatized clavulanic acid (indicated by the star symbol *) and 2HMC (indicated by the black dot ●). (B) The extracted ion chromatograms (EIC) for the mass spectra showing the major peaks corresponding to imidazole-derivatized clavulanic acid $[M+H]^+$ (*m/z* = 224) (right panel) and the fragmented product $[M\text{-imidazole}]^+$ (*m/z* = 156) (left panel)

which were detected in supernatants from the strains shown in (A). (C) The extracted ion chromatograms (EIC) for the mass spectra corresponding imidazole- derivatized 2HMC $[M+H]^+$ (*m/z* = 212) and the fragmented product $[M\text{-imidazole}]^+$ (*m/z* = 144) which were detected in supernatants from the strains shown in (A).

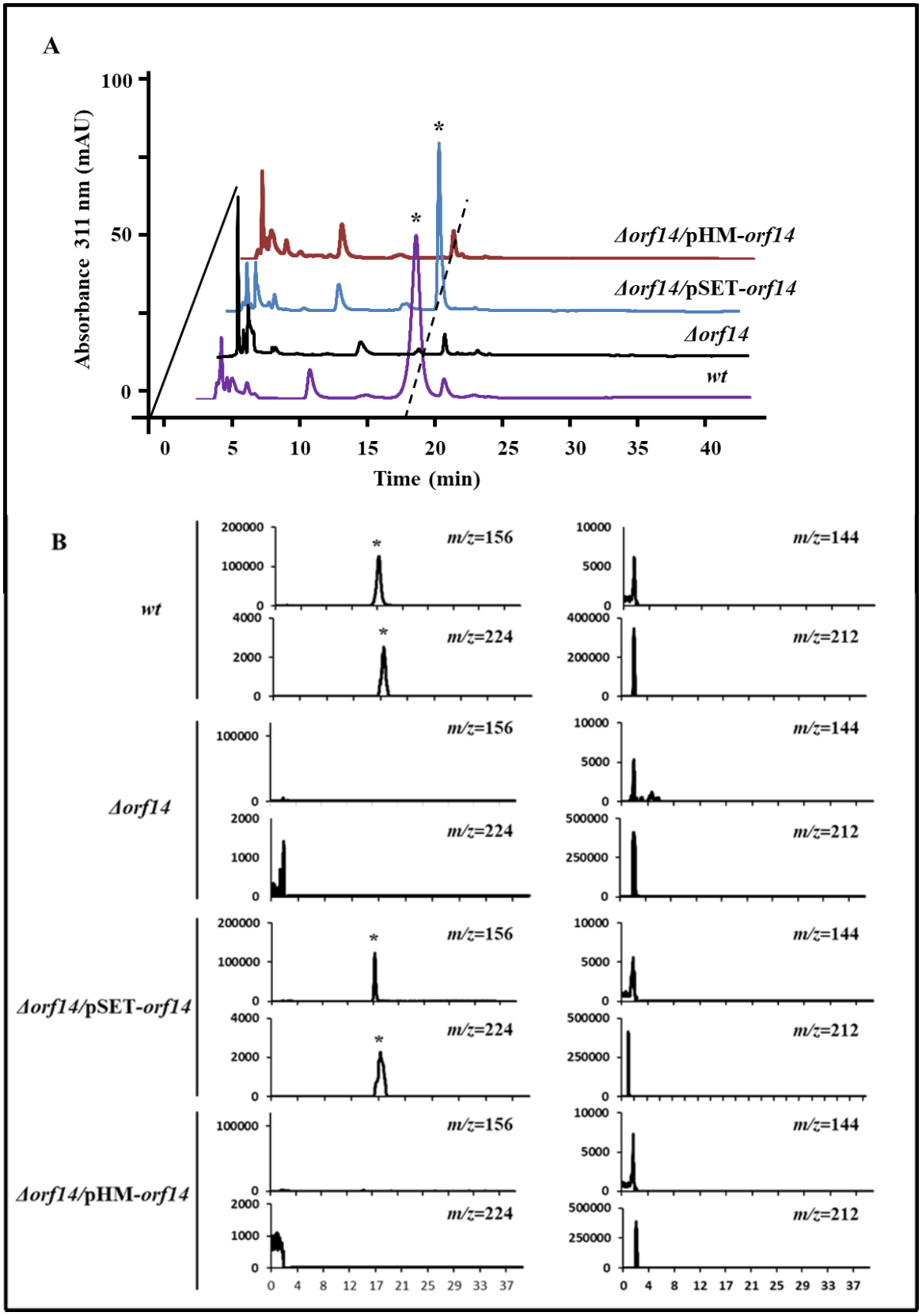

A
Absorbance 311 nm (mAU)
100
50
0
0
5
10
15
20
25
30
35
40
Time (min)
Δorf14/pHM-orf14
Δorf14/pSET-orf14
Δorf14
wt
B
wt
Δorf14
Δorf14/pSET-orf14
Δorf14/pHM-orf14
m/z=156
m/z=224
m/z=144
m/z=212
0 4 8 12 17 21 25 29 33 37

Figure 3.13: Detection of clavulanic acid in *orf14* deletion mutant and complementation strains. LC-MS analysis of 96-h SA culture supernatants after imidazole derivatization using the ammonium bicarbonate buffer system. Cultures from *S. clavuligerus* wt, Δ*orf14*, Δ*orf14*/pSET152-*orf14*, and Δ*orf14*/pHM11a-*orf14* were used to assess CA and 2HMC production. (A) Liquid chromatography profiles showing the elution of the peaks corresponding to imidazole-derivatized clavulanic acid (indicated by the star symbol *). (B) The extracted ion chromatograms (EIC) for the mass spectra showing the major peaks corresponding imidazole- derivatized clavulanic acid $[M+H]^+$ (*m/z* = 224) and $[M\text{-imidazole}]^+$ (*m/z* = 156) (left panel), and the extracted ion chromatograms (EIC) for imidazole-derivatized 2HMS $[M+H]^+$ (*m/z* = 212) and $[M\text{-imidazole}]^+$ (*m/z* = 144) (right panel) which were detected in supernatants from the strains shown in (A). The peaks for the respective *m/z* values are indicated by the star symbol (*).

A

S. clavuligerus strains	ORF14 variants (codon substitution)	Accomplishment
Δ*orf14*/pHM11a-*orf14*	Wt (none)	NA
Δ*orf14*/pHM-*orf14*-F142A	Phe142Ala (TTC→GCC)	Yes
Δ*orf14*/pHM-*orf14*-V254A	Val254Ala (GTC→GCC)	Yes
Δ*orf14*/pHM-*orf14*-T269A	Thr269Ala (ACC→GCC)	Yes
Δ*orf14*/pHM-*orf14*-V292A	Val292Ala (GTG→GCG)	Yes
Δ*orf14*/pHM-*orf14*-W106A	Trp106Ala (TGG→GCG)	No
Δ*orf14*/pHM-*orf14*-M268A	Met268Ala (ATG→GCG)	No
Δ*orf14*/pHM-*orf14*-L289A	Leu289Ala (CTG→GCG)	No
Δ*orf14*/pHM-*orf14*-L293A	Leu293Ala (CTG→GCG)	No

B

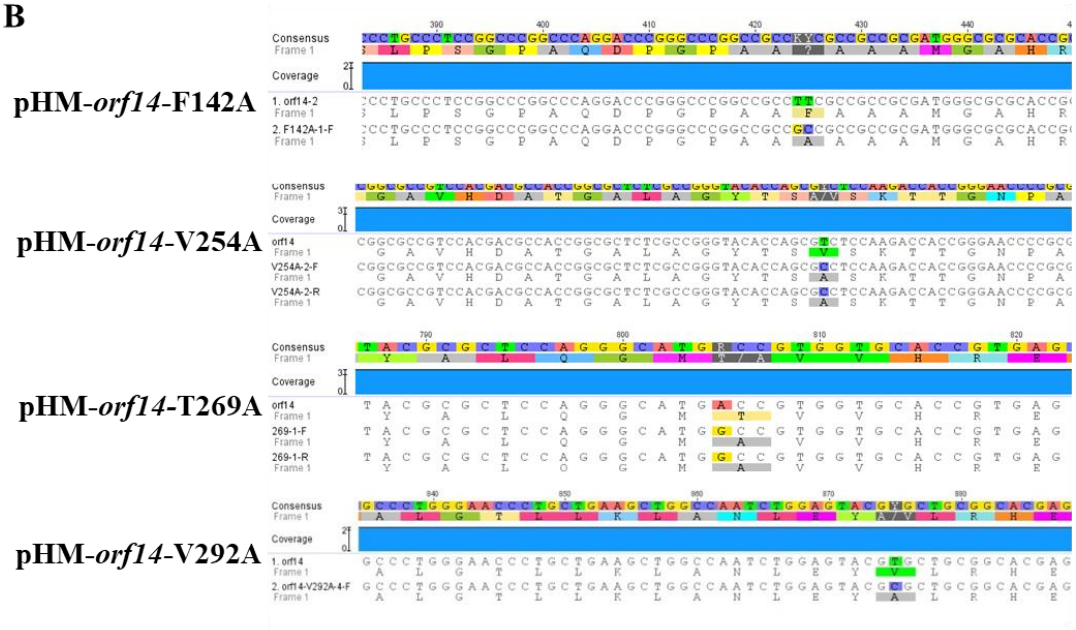


Figure 3.14. Site-directed mutagenesis on *orf14* of *S. clavuligerus*. (A) The table shows the ORF14 variants prepared by substitution of the respective amino acid residues with alanine. **(B)** The sequencing confirmation for the successful, accomplished site-directed mutagenesis on pHM11a vector carrying *orf14* variants. NA: not applicable.

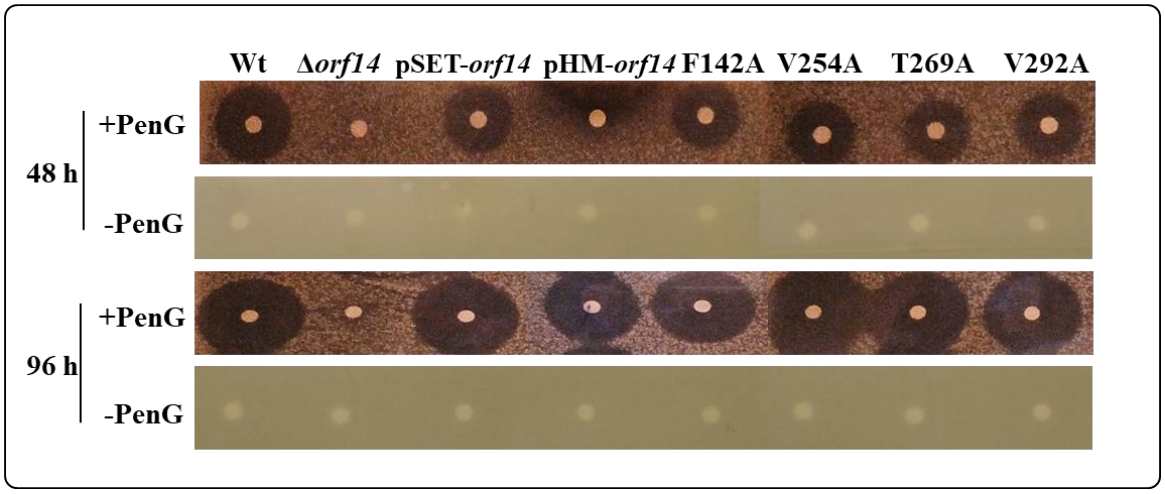


Figure 3.15: Clavulanic acid bioassay results for *S. clavuligerus* strains carrying ORF14 variants. Supernatants from liquid cultures for wt *S. clavuligerus*, Δ*orf14*, pSET152-*orf14*, pHM11a-*orf14*, pHM11a-*orf14*-F142A, pHM11a-*orf14*-V254A, pHM11a-*orf14*-T269A, pHM11a-*orf14*-V292A, were tested for CA production. Each culture was grown for 48 and 96 h in SM medium. The bioassays were performed against *K. pneumoniae* on TSA media with or without PenG. The picture shows results for one of the two replicates completed for each culture.

3.6.2. Tables

Table 3.1: The LC/MS assessment of clavulanic acid and 2-hydroxymethylclavam production in *S. clavuligerus* and Δ*cpe* mutants expressing different variants of CPE grown in SA.

***S. clavuligerus* strains**	**CPE variants (Codon substitution)**	**LC/MS detection for SA samples**	
		CA	**2HMC**
wt	NA	+++	-
Δ*cpe*	NA	-	-
Δ*cpe*/pSET152-*cpe*	Wt (none)	++++	++++
Δ*cpe*/pSET152-*cpe*-S173A	Ser173Ala (TCG → GCG)	-	+
Δ*cpe*/pSET152-*cpe*-S234A	Ser234Ala (AGC → GCC)	-	-
Δ*cpe*/pSET152-*cpe*-L89A	Lys89Ala (AAG → GCG)	-	+
Δ*cpe*/pSET152-*cpe*-S206A	Ser206Ala (AGC → GCC)	+++	+
Δ*cpe*/pSET152-*cpe*-S27A	Ser27Ala (TCC → GCC)	+++	++++

Table 3.2: Clavulanic acid bioassay for *orf14* deleted mutant and complemented strains grown in SA medium and SM. The bioassays were conducted against *K. pneumoniae* on TSA plates. The measurements of the zones of growth inhibitions are reported in millimetres.

Bacterial strain	**Zone of growth inhibition (mm)**					
	48 h		**96 h**		**120 h**	
	SA	**SM**	**SA**	**SM**	**SA**	**SM**
***S. clavuligerus* (wt)**	21	25	23	26	24	27
Sc*/Δ*orf14	0	0	0	0	0	0
***Sc*/Δ*orf14*/pHM11a**	0	0	0	0	0	0
***Sc*/Δ*orf14*/pHM-*orf14* (1)**	0	13	0	19	0	19
***Sc*/Δ*orf14*/pHM-*orf14* (2)**	0	12	0	17	0	16
***Sc*/Δ*orf14*/pHM-*orf14* (3)**	0	11	0	17	0	17
***Sc*/Δ*orf14*/pSET152**	0	0	0	0	0	0
***Sc*/Δ*orf14*/pSET-*orf14* (1)**	18	18	23	25	29	27
***Sc*/Δ*orf14*/pSET-*orf14* (2)**	16	17	22	25	27	28
***Sc*/Δ*orf14*/pSET-*orf14* (3)**	17	18	23	26	26	28
CA solution	~35					

3.7. Supplementary materials

3.7.1. Supplementary Figures

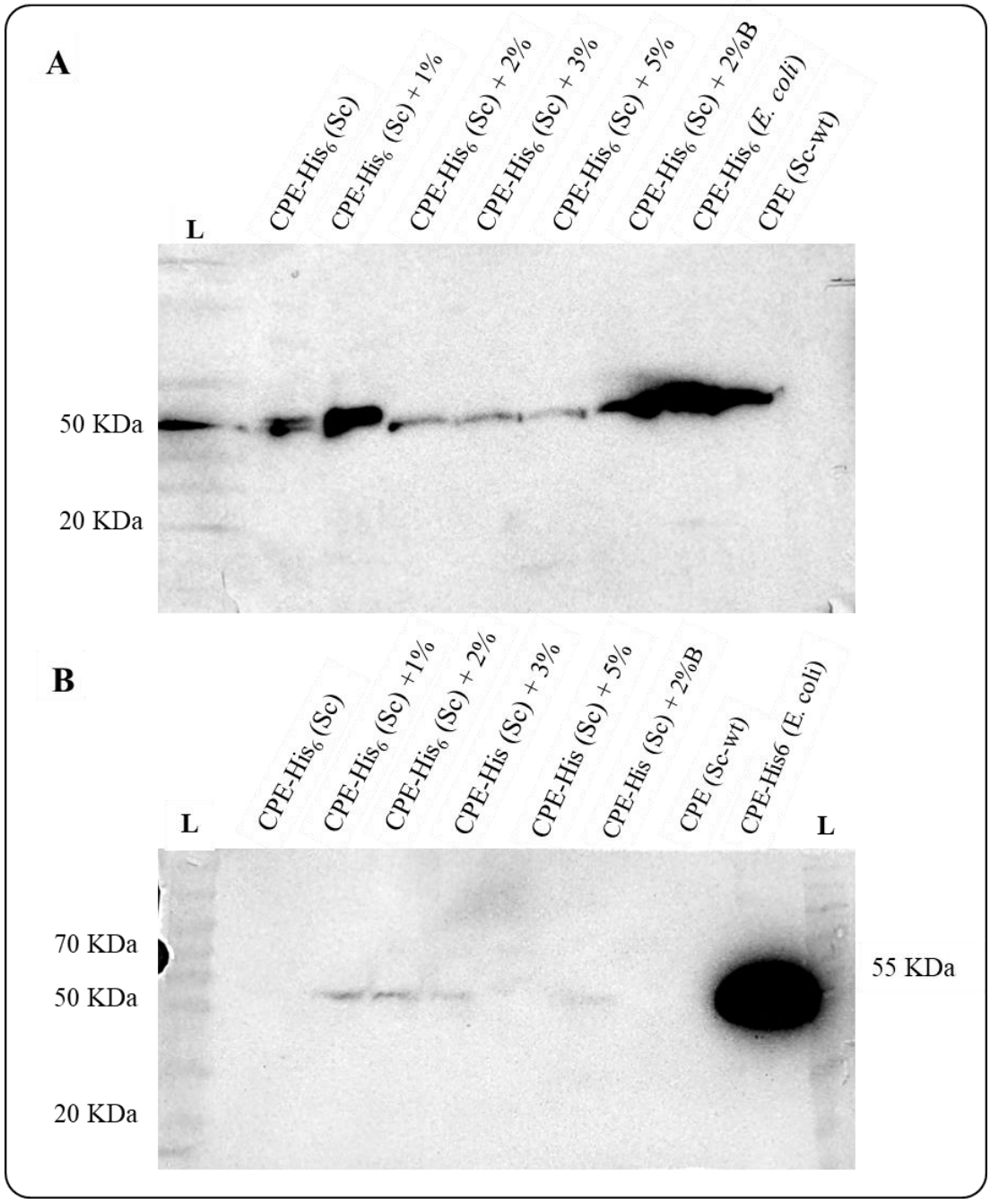


Figure S3.1. CPE crosslinking optimization. Bacterial cultures of *S. clavuligerus*/*cpe*-His_6 were treated with different concentrations of formaldehyde (0, 1, 2, 3, and 5%) to induce protein-protein crosslinking. Mycelial lysates were prepared by sonication then centrifugation. The Ni-NTA resin system was used to purify CPE-His_6 from supernatants. Samples from the supernatants **(A)** and lysates pellets **(B)** were separated by 12% SDS-PAGE. Western blot analysis was conducted, and the anti-6×His tag monoclonal antibodies were used at a 1:1000 dilution. Positive control of CPE-His_6 from *E. coli* cells was used. L: PiNK Plus prestained protein ladder, Sc (wt); a culture of wt *S. clavuligerus* was used as control. CPE-His_6 + 2%-B; a sample of 2% formaldehyde was added 30 minutes before collecting the mycelia (see Section 2.7 in Materials and methods).

3.7.2. Supplementary Tables

Table S3.1: Clavulanic acid bioassay for two *cpe* (*orf12)* complemented strains grown in SA medium at two time points. The bioassays were conducted against *K. pneumoniae* on TSA plates with penicillin G.

Bacterial strains	Zone of growth inhibition	
	48 h	96 h
***S. clavuligerus* (wt) (1)**	13 mm	29 mm
***S. clavuligerus* (wt) (2)**	12 mm	27 mm
***S. clavuligerus*/Δ*cpe* (1)**	0 mm	0 mm
***S. clavuligerus*/Δ*cpe* (2)**	0 mm	0 mm
***S. clavuligerus*/Δ*cpe*/pSET152-*cpe* (1)**	12 mm	25 mm
***S. clavuligerus*/Δ*cpe*/pSET152-*cpe* (2)**	11 mm	24 mm
***S. clavuligerus*/Δ*cpe*/pHM11a-*cpe* (1)**	11 mm	25 mm
***S. clavuligerus*/Δ*cpe*/pHM11a-*cpe* (2)**	11 mm	24 mm
CA solution	34 mm	34 mm

Table S3.2: The ion abundance values for the corresponding extracted ion chromatogram for imidazole-derivatized CA $[M+H]^+$ (*m/z* = 224), the fragmented product [M-imidazole]$^+$ (*m/z* = 156) and the 2HMC fragmented product [M-imidazole]$^+$ (*m/z* = 144) for SA supernatant samples.

Strain	m/z	abundant
Wt	144.0	----
	156.0	~48000.0
	224.0	~8500.0
Δ*cpe*	-----	-----
Δ*cpe*/pSET-*cpe*	144.0	~8000.0
	156.0	~40000.0
	224.0	~6750.0
Δ*cpe*/pHM-*cpe*	144.0	~3200.0
	156.0	~48000.0
	224.0	~8100.0

Table S3.3: The assessment of clavulanic acid production in *S. clavuligerus* and Δ*cpe* mutants expressing different variants of CPE grown in SA and SM. The bioassays were conducted against *K. pneumoniae* on TSA plates with penicillin G.

***S. clavuligerus* strains**	**Production of CA**	
	SA	**SM**
wt	Yes	Yes
Δ*cpe*	No	No
Δ*cpe*/pSET152-*cpe*	Yes	Yes
Δ*cpe*/pSET152-*cpe*-S173A	No	No
Δ*cpe*/pSET152-*cpe*-S234A	No	No
Δ*cpe*/pSET152-*cpe*-S27A	Yes	Yes
Δ*cpe*/pSET152-*cpe*-L89A	No	No
Δ*cpe*/pSET152-*cpe*-S206A	Yes	Yes

Table S3.4: Clavulanic acid bioassays measurements for *orf14* variants grown in SM media at two time points. The bioassays were conducted in duplicates against *K. pneumoniae* on TSA plates with penicillin G.

***S. clavuligerus* strains**	**Zone of growth inhibition (mm)**	
	48 h	**96 h**
S. clavuligerus (wt)	27	38
Δorf14	0	0
Δorf14/pSET-orf14	25	39
Δorf14/pHM-orf14	17	32
Δorf14/pHM-F142A (1)	24	40
Δorf14/pHM-F142A (2)	22	40
Δorf14/pHM-V254A (1)	20	38
Δorf14/pHM-V254A (2)	22	37
Δorf14/pHM-T269A (1)	25	35
Δorf14/pHM-T269A (2)	27	35
Δorf14/pHM-V292A (1)	27	37
Δorf14/pHM-V292A (2)	24	37
CA solution	~ 40	

CHAPTER IV

The investigation of *nocE* and its impact on the physiology and metabolism of *Streptomyces clavuligerus*

4.1. Abstract

A comparative genomic study for CA/CA-like BGCs between the CA producer (*S. clavuligerus*) and non-producers (*Streptomyces pratensis*, *Saccharomonospora. viridis*, and *Streptomyces* sp. M41) showed the presence of a large gene, *nocE*, within the CA-like BGCs of the non-producers. In contrast, the *nocE* homologue in *S. clavuligerus* is located distantly from the CA BGC. A bioinformatics analysis revealed that homologues of *nocE* are present in more Actinomycete species that possess a BGC for at least one kind of β-lactam antibiotic. NocE proteins belong to the SGNH/GDSL-hydrolase superfamily, member of which have domains with esterase or lipase activities. The deletion of *nocE* or the constitutive expression did not affect the production of clavulanic acid, cephamycin-C, or 5*S* clavams, indicating that *nocE* does not have any role in the production of any β-lactam metabolites in *S. clavuligerus*. However, the deletion of *nocE* significantly affected the growth of *S. clavuligerus* in starch asparagine medium but not in tryptic soy medium. Furthermore, untargeted metabolomics analysis using the GNPS demonstrated that *nocE* has some role in the general metabolism of *S. clavuligerus* but not in the specialized metabolism.

4.2. Introduction

CA and Ceph-C are industrially produced by fermenting *S. clavuligerus* (Jensen and Paradkar, 1999; Saudagar et al., 2008). Another two species, *Streptomyces jumonjinensis* and *Streptomyces katsurahamanus* are also identified for their ability to produce both CA and Ceph-C. Unlike *S. clavuligerus*, *S. jumonjinensis* and *S. katsurahamanus* do not produce 5*S* clavams and do not have the associated BGCs (Ward and Hodgson, 1993; Jensen and Paradkar, 1999; Saudagar et al., 2008). Genome sequencing studies have revealed the presence of CA-like BGCs in several other actinomycetes. *Streptomyces pratensis* (formerly called *flavogriseus*) ATCC 33331 and *Saccharomonospora viridis* DSM 43017 have CA-like BGCs, but neither has been shown to produce CA under laboratory conditions, suggesting that the BGCs are fully or partially silent. The CA-like BGCs in *S. pratensis* and *S. viridis* are composed of blocks of conserved genes in the same order as in the *S. clavuligerus* CA BGC but assembeled in different organization (Figure 4.1) (Álvarez-Álvarez et al., 2013).

One of the intriguing features of the CA-like BGCs in *S. pratensis* and *S. viridis* is the presence of a large *nocE* gene (Figure 4.1) (Jensen, 2012; R. Álvarez-Álvarez et al., 2013). This gene, Sfla_0550 (4,248 bp) in *S. pratensis*, and Svir_33350 (4,068 bp) in *S. viridis*, is located in the middle of CA-like clusters (Figure 4.1B), but its role in the gene cluster is still unknown (Jensen, 2012; R. Álvarez-Álvarez et al., 2013). *nocE* was initially named for its similarity to the *nocE* gene in the biosynthetic gene cluster of the β-lactam nocardicin A in *Nocardia uniformis* subsp. *tsuyamanensis* (Gunsior et al., 2004). Nocardicin A is a monocyclic β-lactam that shows moderate activity against a broad spectrum of Gram-negative bacteria (Demain and Elander, 1999). A similar gene to *nocE*

was also found in *S. clavuligerus*, but it is located outside the CA BGCs (Figure 4.1A) and the function of this gene has not been investigated.

The amino acid sequence of the *S. clavuligerus* NocE protein showed that it belongs to the SGNH hydrolase superfamily of proteins. Members of this family contain domains that act as esterases and lipases but have little sequence homology to true lipases (Akoh et al., 2004). SGNH-hydrolase is a subgroup of the GDSL family and was further classified due to four strictly conserved residues, Ser-Gly-Asn-His, in four conserved blocks in the protein (Mølgaard et al., 2000; Akoh et al., 2004). This group of enzymes was also identified in Actinomycetales, Ascomycota, and Nematoda, indicating that these enzymes are most conserved among soil-inhabiting organisms (Bielen et al., 2009).

The disruption of *nocE* in *N. uniformis* does not affect the production of nocardicin A (Davidsen & Townsend, 2009), but the role of *nocE* in CA producers and non-producers has not been examined yet. As I will address in this study, the presence of *nocE* genes in many actinomycetes that also possess BGCs for β-lactam metabolites makes this gene an interesting target to study and investigate its potential role in specialized metabolite biosynthesis or any other physiological role in *S. clavuligerus.*

4.3. Objectives

The presence of *nocE* within CA-like BGCs in *S. pratensis* and *S. viridis* (CA non-producers) is noticeable (Figure 4.1), as the gene is similar to one from the nocardicin A monobactam BGC of *N. uniformis*. The existence of NocE homologues in other β-lactam-producing *Streptomyces* has been reported, but their functions have not

been examined to date. Therefore, the main goal of this study was to investigate the role of *nocE* in the industrially important bacterium *S. clavuligerus*.

I first aimed to predict the function of the NocE protein by using bioinformatics analyses to identify the conserved domain(s) that are present. Then, I tried to understand the relationship between *nocE* and some of the β-lactams BGCs in bacteria by performing a comparative bioinformatics analysis between the CA producers and non-producers (but that carry CA-like BGCs). In addition, I aimed to characterize the role of *nocE* in *S. clavuligerus;* therefore, a *nocE* deletion mutant and *nocE* constitutive expression strains were successfully prepared and tested by bioassays and LC-MS for the production of β-lactam metabolites. Moreover, the effect of the *nocE* deletion and constitutive expression on the growth of *S. clavuligerus* was examined in different types of media. Finally, untargeted metabolomics analysis was conducted for extracts from three strains of *S. clavuligerus* (wt, *nocE* deletion, and *nocE* constitutive expression) using GNPS and MolNetEnhancer to investigate the impact of *nocE* on the general and specialized metabolism in *S. clavuligerus*.

4.4. Results and Discussion:

4.4.1. Comparative study of clavulanic acid BGCs between producers and non-producers.

S. clavuligerus, *S. jumonjinensis*, and *S. katsurahamanus* are the only species reported to produce CA and Ceph-C, and hereafter in this study they are called "CA producers". Recent genome sequencing projects revealed the existence of CA gene clusters infrequently in the genomes of many other actinomycetes, and these CA or CA-like clusters have varying similarity to the CA BGC of *S. clavuligerus*; for examples, see Table 4.1. However, some of these clusters are inactive or silent under laboratory conditions, and the molecular basis for this lack of activity is not known. In this study, a comparative analysis was first performed for the CA/CA-like BGCs between the CA producers and some of the CA non-producers using the bioinformatics tools Geneious8 and AntiSMASH 4.0 (Blin et al., 2017).

S. pratensis, *S. viridis*, and *Streptomyces* sp. M41, hereafter called "CA non-producers," are species that have CA-like BGCs with 42%, 58%, and 54% similarities, respectively, to the CA BGC in *S. clavuligerus*. However, *S. pratensis* and *S. viridis* do not show any production of CA when fermented in various types of media (Jensen, 2012; Álvarez-Álvarez et al., 2013), and *Streptomyces.* sp. M41 has not yet been tested for CA production.

While the CA BGC is located side by side with the Ceph-C cluster in the CA producers species (Figure 4.1A), our BGCs analysis showed that *S. pratensis*, *S. viridis*, and *Streptomyces*. sp M41 do not possess Ceph-C, 5*S* clavams, paralogue genes, or alanylclavam gene clusters as in *S. clavuligerus* (Figure 4.1A). The results support that *S.*

clavuligerus is unique among the β-lactam producers so far described in its ability to produce CA, Ceph-C, and 5*S*-clavams (Jensen, 2012). Interestingly, *S. pratensis* has a β-lactam BGC for carbapenem MM4550 with 65% similarity to the one in *Streptomyces argenteolus* ATCC11009 (a carbapenem MM4550 producer) (Figure 4.1A) (Li et al., 2014). More details about this BGC are covered in the next chapter of this thesis. The antiSMASH analysis for *S. viridis* and *Streptomyces*. sp M41 did not identify any other β-lactam BGC in their genomes.

The CA-like BGCs in the CA non-producers contain all of the genes for CA biosynthesis (Figure 4.1B). While these genes in *Streptomyces*. sp M41 are organized exactly like those in *S. clavuligerus*, the CA genes in *S. pratensis* and *S. viridis* are in blocks of conserved genes that are in the same order as those of the *S. clavuligerus* cluster, but the blocks are assembled in a different organization (Figure 4.1B). In addition, the non-producers are missing a group of genes: *orf18* through *orf23*, which are part of the *S. clavuligerus* cluster (Figure 4.1B; for their functions, see Table 1.2). The *orf18* (*pbpA), orf19 (pbp2),* and *orf20* genes do not have a role in CA production; deletion of these genes did not affect the biosynthesis of CA or Ceph-C (Jensen et al., 2004; Jensen, 2012). The *orf21* to *orf23* genes showed some effects on CA production, but their exact function in CA biosynthesis remains uncertain (Fu et al., 2019; Jnawali et al., 2008; Song et al., 2009). However, *S. jumonjinensis* and *S. katsurahamanus* (CA producers) do not possess these extra genes (*orf18* and *orf20* - *orf23*) in their clusters, but they still can produce CA, suggesting that these genes are not required for CA biosynthesis, and their absence in the CA non-producers BGCs is not the reason behind the non-production of CA.

One more major difference is that the CA-like BGCs of the non-producers contain three additional genes, *pcbR*, *orf11*, and *nocE*, which are not present in the CA BGCs of *S. clavuligerus* (Figure 4.1B). While these extra genes are situated in the middle of CA-like BGCs of *S. pratensis* and *S. viridis*, only *orf11* and *nocE* genes were found in the *Streptomyces*. sp M41 genome located immediately next to the CA-like cluster (Figure 4.1B). The two genes *pcbR* and *orf11* were named because of their similarities to those included in the Ceph-C BGC of *S. clavuligerus*. *S. clavuligerus pcbR* encodes a PBP involved in β-lactam resistance (Paradkar et al., 1996), whereas *orf11* encodes a predicted protein of unknown function. Previous reports have shown that disruption of neither *pcbR* nor *orf11* in *S. clavuligerus* affected Ceph-C or CA production (Paradkar et al., 1996; Alexander and Jensen, 1998), suggesting that they are not required for the biosynthesis of the respective metabolites. The third gene, *nocE* (Sfla_0550 with 4,248 bp) in *S. pratensis* and (Svir_33350 with 4,068 bp) in *S. viridis*, is the largest gene in the CA-like BGCs. The homologue *nocE* (SCLAV_5162 with 4,029 bp) in *S. clavuligerus* is located distantly (~ 1.19 Mb) from the CA-Ceph-C supercluster (Figure 4.1A). In the following sections of this chapter, we tried to decipher the function of *nocE* in *S. clavuligerus*, the most important industrial CA producer.

4.4.2. NocE characterization and comparative study

The *nocE* gene encodes a predicted protein with 1,343 amino acids and possesses two conserved domains. The first domain is near the C-terminus (amino acids 1,040 – 1,339) (Figure 4.2A) and belongs to the SGNH-hydrolase esterase or GDSL-like lipase superfamily. The folding in this domain enables it to act as an esterase and lipase but it

has little sequence homology to true lipases (Upton and Buckley, 1995; Akoh et al., 2004). In *Streptomyces*, proteins containing GDSL-like motifs were found in extracellular lipases from *Streptomyces rimosus* R6-554W (Vujaklija et al., 2002), and the ones with an esterase domain type were found in *Streptomyces scabies* and hydrolyze a specific ester bond in suberin, a plant lipid (Wei et al., 1995). The SGNH-hydrolase enzymes were also reported in *Streptomyces coelicolor* A3(2), *Streptomyces exfoliatus* M11, *Streptomyces griseus*, and others (Bielen et al., 2009; Servín-González et al., 1997).

Next to the SGNH/GDSL domain in the NocE protein is a non-catalytic carbohydrate-binding (CBM6-CBM35-CBM36-like) domain (amino acids 925 – 1030) (Van Bueren et al., 2005) (Figure 4.2A). CBMs are non-catalytic carbohydrate-binding modules present in some hydrolase enzymes produced by bacteria. These modules are connected via linker peptides to the catalytic modules in the enzymes. The advantage of CBMs is to enhance the catalytic activity of the enzyme by mediating prolonged and intimate association with its target substrate (McCartney et al., 2006). The most common associated modules are enzymes such as xylanases, cellulase, chitinase, lichenases, β-agarases and deacetylases (Boraston et al., 2004). The SGNH/GDSL lipase/esterase enzymes display various functional properties. In growing *Streptomyces*, the activities of these enzymes were reported to increase constantly until reaching the late stationary phase (Bielen et al., 2009; Vujaklija et al., 2002). Therefore, and despite there being no direct evidence, it has been proposed that lipases and other hydrolytic enzymes might be necessary for the production of specialized metabolites in *Streptomyces* (Bielen et al., 2009; Horinouchi, 2002).

Our analysis of NocE sequences present in the public database (NCBI BLASTP) showed that homologues of *S. clavuligerus* NocE could be found with high percent identity (>50%) in several other species of Actinomycetes (Table 4.1). From the top 50 hits in identity, we found 11 NocE matches (Table 4.1) whose corresponding species' genomes are fully sequenced and deposited in NCBI. The genomes of these species were analyzed by the AntiSMASH online program searching for specialized metabolite BGCs. Interestingly, our analysis revealed that all species possess at least one or two BGCs for β-lactam antibiotics, and they were either CA, Ceph-C, CA and Ceph-C together, or CA and carbapenem MM4550 (Carb4550) BGCs together (Table 4.1). None of these species was found to have Carb4550 BGC by itself. These observations suggested that NocE could have some relation with β-lactam metabolites. Similar to *Streptomyces.* sp. M41, three more bacterial species, *Streptomyces* sp. NRRL S-325, *Streptomyces* sp. NRRL B-24051, and *Streptomyces* sp. AD193-02, were found to have the *nocE* gene situated next to the CA-like BGCs, while in the remaining seven species, *nocE* was positioned far away from the β-lactam BGCs. The details about the 11 species' NocEs, the CA producers' (*S. jumonjinensis* and *S. katsurahamanus*) NocEs, and the non-producers' (*S. pratensis* and *S. viridis*) NocEs are in Table 4.1. To date, there are no reports indicating whether these 11 species produce any β-lactam metabolites, and the function of *nocE* in these species has not been investigated. However, whether these species have the ability to produce the respective metabolites is an interesting question that warrants further investigation.

For further comparative investigation, a phylogenetic analysis using the complete amino acid sequences of NocE proteins was performed (Figure 4.2B). The phylogenetic analysis showed that the NocE proteins from the known CA producers *S. clavuligerus, S.*

jumonjinensis, and *S. katsurahamanus* form a distinct clade, and those for *S. jumonjinensis* and *S. katsurahamanus* are closer to each other than NocE of *S. clavuligerus*. Supporting that, we found that the protein products of the genes surrounding *nocE* in *S. jumonjinensis* and *S. katsurahamanus* are similar to each other but not with those in *S. clavuligerus* (Supplementary Table S4.1), suggesting that the "*nocE*-cluster" in *S. jumonjinensis* and *S. katsurahamanus* share the same ancestor. Interestingly, the bacterial species that possess BGCs for both CA and Carb4550 (*S. pratensis, Streptomyces* sp. PAMC26508, *Streptomyces* sp. NRRLB-24051, and *Streptomyces* sp. NRRLS-325) form a separate clade from the others (Figure 4.2B). Also, the species that have Ceph-C BGCs only (*Streptomyces* sp. HNM0039, *Streptomyces* sp. CNT302, *Streptomyces* sp. CNS615, and *Streptomyces*. sp. CNR698) form another separate clade from the others (Figure 4.2B and Table 4.1). These observations indicate that the members of each clade descended from the same ancestor and supported the proposed relation between NocE and the β-lactams antibiotics. The remaining NocE homologues belong to bacterial species that hold CA-like BGCs only: *Kitasatospora albolonga* (formerly called *Streptomyces albolongus,* Labeda et al. 2017), *Streptomyces* sp. AD193-02, *Streptomyces* sp. M41, *Kitasatospora papulosa,* and *S. viridis*. NocE of the nocardicin A BGC from *N. uniformis* was included in the phylogenetic analysis and was clustered with NocE of *S. viridis* (Figure 4.2B), suggesting they share the same evolutionary history.

Many of the GDSL lipases/esterase are extracellular enzymes that can be secreted outside the cells (Vujaklija et al., 2002). Therefore, the amino acid sequences of NocE homologues were analyzed for secretory signal peptides using SignalP-5.0 (Almagro

Armenteros et al., 2019). Out of 17 NocE homologues tested, 13 NocE proteins were predicted to have conserved N-terminal Sec-signal type I sequence ($p>0.6$) (Table 4.2). Remarkably, NocE of *S. clavuligerus* was found to have a highly conserved N-terminal Sec-signal sequence with $p > 0.9$ (Table 4.2).

4.4.3. *nocE* transcription, deletion, and constitutive expression in *S. clavuligerus*

To determine if *nocE* has any role during CA production in the model β-lactam producer *S. clavuligerus*, and if it is active during *S. pratensis* growth, the expression of the *nocE* gene was examined in both species (Figure 4.2C). Also, the expression of two genes essential for CA biosynthesis, *ceaS2* and *cas2*, were tested along with *nocE* as controls. The two genes were previously reported to be expressed during growth in both species (Paradkar and Jensen, 1995; Jensen et al., 2000; Álvarez-Álvarez et al., 2013). RT-PCR analysis of RNA samples isolated from wt *S. clavuligerus* grown in SA medium demonstrated that Sc-*nocE* is transcribed along with Sc-*ceaS2* and Sc-*cas2* during CA production (Figure 4.2C). Also, the RT-PCR analysis for *S. pratensis* RNA samples showed that the Sp-*nocE*, Sp-*ceaS2*, and Sp-*cas2* genes are also expressed in the CA-like BGC, and they are not silent. These results support the hypothesis that *nocE* might have some role during CA biosynthesis in *S. clavuligerus*. To investigate further, gene manipulation experiments were conducted in *S. clavuligerus*, where the *nocE* gene was either deleted or overexpressed.

A marker-less deletion mutant of the *nocE* gene was generated using the meganuclease I-SceI system (Figure 4.3; Fernández-Martínez and Bibb, 2014). The two constructs, pIJ12738, which has an I-SceI meganuclease recognition site (IRS), and

pIJ12742, which contains the *I-SceI* gene for the expression of I-SceI meganuclease, were used in this approach as described in the "Materials and Methods" chapter (Figure 4.3A). The upstream (UP) and downstream (DN) regions for *nocE* were cloned together in pIJ12738 to give pIJ12738/*nocE*-UPDN, which was then introduced into *S. clavuligerus* by conjugation. As the plasmid is unable to replicate in *Streptomyces*, the resulting apramycin resistant exconjugants would harbour the plasmid integrated to the chromosome via homologous recombination (Figure 4.3B). The exconjugants (*S. clavuligerus*/pIJ12738/*nocE*-UPDN) were then used to introduce pIJ12742, which has the expression cassette for the I-SceI meganuclease that cuts the chromosome at the IRS site (Figure 4.3B). As a repairing mechanism, a second homologous recombination will generate both the *nocE* deletion and the wild-type genotypes. The deletion of *nocE* was confirmed by genomic DNA PCR (Figure 4.3C), and three *S. clavuligerus*/Δ*nocE* mutants were obtained and used to test the production of CA and Ceph-C. A strain constitutively expressing *nocE* was prepared by cloning a fragment of *nocE* gene downstream of the promoter *ermE*p* in the pIJ8668 integrative vector (see Materials and Methods). The resulting construct pIJ8668-*ermE*p*-*nocE* (Supplementary Figure S4.1) was introduced into *S. clavuligerus* by conjugation to give *S. clavuligerus*/*ermE*p*-*nocE*, in which the *ermE*p∗ promoter is inserted upstream of *nocE* in the chromosome. The *ermE*p∗-*nocE* strains were confirmed by genomic DNA PCR.

4.4.4. The impact of *nocE* on β-lactam metabolite (CA, Ceph-C, and 5*S* clavams) productions in *S. clavuligerus*

The wt *S. clavuligerus*, Δ*nocE,* and *ermE*p*-*nocE* strains were cultured in SA, SM, and TSB-S liquid media in triplicate, and supernatant samples were collected at 24, 48, and 96-h time points. CA disc diffusion bioassays were performed against *K. pnuemoniae* with PenG for supernatant samples from SA and SM. As shown in Figure 4.4A and Supplementary Table S4.2, zones of growth inhibition were observed around discs infused with supernatants from all strains (wt, Δ*nocE,* and *ermE*p*-*nocE*), which indicates that neither the deletion of *nocE* nor constitutively expressing it affected the production of CA, and the *nocE* gene is not required for CA biosynthesis. As a control, the bioassays were conducted without adding PenG to the TSA testing plate, and no growth inhibition zones were detected (Figure 4.4A).

To investigate the role of *nocE* on Ceph-C production, Ceph-C bioassays were performed against *E. coli* ESS for supernatant samples from SM and TSB-S cultures. Zones of growth inhibition were noticed in all samples for wt, Δ*nocE,* and *ermE*p*-*nocE* strains as demonstrated in Figure 4.4B and Supplementary Table S4.3. The results revealed that the *nocE* gene is not involved in Ceph-C biosynthesis also.

For further confirmation and to test the role of *nocE* on 5*S* clavams biosynthesis, LC-MS analysis was carried out for supernatant samples of 96-h cultures in SM for wt *S. clavuligerus*, Δ*nocE,* and *ermE*p*-*nocE* strains. Derivatization of the supernatant samples was carried out by adding imidazole prior to injection, and the CA and 2HMC production were detected using absorbance at 311 nm. As shown in Figure 4.5A, the peaks for CA and 2HMC appeared at their respective retention times in the three tested samples (wt,

Δ*nocE,* and *ermE*p*-*nocE*). The mass spectra analysis of the CA and 2HMC peaks demonstrated the major peaks corresponding imidazole derivatized CA $[M+H]^+$ (*m/z* = 224) and the fragmented product $[M\text{-imidazole}]^+$ (*m/z* = 156), and to imidazole derivatized 2HMC $[M+H]^+$ (*m/z* = 212) and the fragmented product $[M\text{-imidazole}]^+$ (m/z = 144) (Figure 4.5B). The LC-MS results clearly showed that the deletion or the constitutive expression of the *nocE* gene did not affect CA or 5*S* clavams biosynthesis in *S. clavuligerus,* raising the possibility that the gene might be associated with primary metabolism.

4.4.5. The role of *nocE* in *S. clavuligerus* growth

The predicted lipase/esterase-like domain present in NocE is also found in hydrolytic secreted enzymes from other *Streptomyces* species (Vujaklija et al., 2002; Wei et al., 1995). Furthermore, closer examination of the predicted NocE amino acid sequence from *S. clavuligerus* suggested that it is also a secreted protein, as it contains a highly conserved N-terminal Sec-signal sequence ($p > 0.9$) (Table 4.2). These findings further ruled out the direct involvement of NocE in CA production, which occurs in the cytoplasm, and suggested that NocE might have some other exocellular hydrolytic function instead. Therefore, we tested the effect of *nocE* deletion or constitutive expression on the growth of *S. clavuligerus.* The wt *S. clavuligerus*, Δ*nocE,* and *ermE*p*-*nocE* were grown under different nutritional conditions using TSB-S (rich), SM (complex fermentation), or SA (defined fermentation) media. The samples were collected every 24-h for a six-day incubation period. The growth levels (Figure 4.6) were determined based on DNA content measurements according to the protocol of Zhao et al. 2013 (see

Materials and Methods). Interestingly, the growth of the Δ*nocE* mutant was significantly reduced in each medium tested, whereas that of the *ermE*p*-*nocE* strain was enhanced in only SA medium when compared to the wt strain (Figure 4.6).

Streptomyces bacteria demonstrate a complex life cycle when growing on solid media; therefore, we assessed the effect of *nocE* on the growth of *S. clavuligerus* when cultured on solid agar media. Spores suspensions of wt *S. clavuligerus*, Δ*nocE,* and *ermE*p*-*nocE* strains at different dilutions from 2×10^5 to 2×10^2 spores (Figure 4.7) were inoculated in spots over four different types of media plates: SA, minimal essential medium (MM), ISP-4, and TSA-S. The growth was monitored up to 120-h (Figure 4.7). Differences in growth between the bacterial strains were noticed on SA and MM media plates (Figure 4.7 A and B), where Δ*nocE* showed less growth than wt and *ermE*p*-*nocE*. The differences in growth were noticed after 48-h incubation in $2x10^5$ spores-inoculated spots, 72-h in 2×10^2 and 2×10^3 spores-inoculated spots, and 96-h and 120-h in 2×10^2 spores-inoculated spots (Figure 4.7A). On the MM plate, the differences in growth were only noticed after 48-h incubation in 2×10^5 and 2×10^3 spores-inoculated spots (Figure 4.7B). In general, as the number of inoculated spores and the incubation period increased, the differences in growth between the strains became less distinguished. No difference in growth was noticed when ISP-4 or TSB-S agar plates were used (Figure 4.7 C and D). The strains Δ*nocE* and *ermE*p*-*nocE* typically developed aerial hyphae on all types of media and sporulated very well on ISP-4 (sporulation media) when compared to wt.

4.4.6. The impact of *nocE* on the metabolome of *S. clavuligerus*

Although NocE is still an uncharacterized protein, it could function in the hydrolysis and breakdown of esters of organic compounds that affect the metabolism and the growth of *S. clavuligerus*. To examine the influence of *nocE* on primary metabolism in *S. clavuligerus*, the wt, Δ*nocE,* and *ermE*p∗-*nocE* strains were grown on TSA-S medium (Figure 4.7D), which demonstrated no differences in the growth between the three strains, and SA agar, which showed differences in the growth between the strains (Figure 4.7A). Each plate was extracted using methanol or ethyl acetate and subjected to LC-MS/MS for untargeted metabolomics analysis as described in the Materials and Methods. A heat map was constructed using feature-based detection and alignment of positive mode ionization data for wt, Δ*nocE,* and *ermE*p∗-*nocE* strains (Figure 4.8). The analysis showed distinct metabolite differences for *S. clavuligerus* when cultured on TSA-S and SA media. Also, it demonstrated marked differences in overall metabolite levels between the respective wt, Δ*nocE,* and *ermE*p∗-*nocE* strains (Figure 4.8), indicating that *nocE* has some role in the general metabolism of *S. clavuligerus.* For further analysis and to identify the classes of molecules found in the extracts of the three strains, the MS/MS spectral data were analyzed and annotated by the Network Annotation Propagation (NAP) tool in GNPS, which is a data-driven platform for the storage, analysis, and knowledge dissemination of MS/MS spectra. Then the spectral data were analyzed using the MolNetEnhancer to identify the classes of the molecules/compounds found in the samples. The analysis outputs were visualized and interpreted as colored networks in Cytoscape 3.8 (Figure 4.9). More than 14000 spectral features of molecules

were detected in the analysis, and ~46% of them were found to belong to 16 superclasses of compounds listed in Figure 4.9 and supplementary Table S4.4. The total number of spectral features for the group "Lipids and lipid-like molecules" was the highest among the other superclasses with 2287 spectral features of molecules, followed by the "Organoheterocyclic compounds" with 2188 spectral features detected (Table S4.4). In terms of bacterial strains, the *nocE*-deleted *S. clavuligerus* contained a higher number of "Benzenoids" and "Lipids and lipids-like molecules" in comparison to wt and *ermE*p∗-*nocE* strains (Figure 4.9B). In contrast, the number of "Organoheterocyclic compounds" and "Phenylpropanoids and polyketides" detected in Δ*nocE* were less than those in wt and *ermE*p∗-*nocE* (Figure 4.9B). Interestingly, in the *nocE* constitutively overexpressed strain *ermE*p∗-*nocE,* the number of spectral features for "Organic acids and derivatives", "Organic oxygen compounds", and "Organoheterocyclic compounds" were significantly higher than those in wt and Δ*nocE* (Figure 4.9B), indicating that the overexpression of *nocE* affects these kinds of compounds. Therefore, the deletion and overexpression of *nocE* had an impact on the metabolism of *S. clavuligerus*.

To identify the specialized metabolites produced by the three bacterial strains, wt, Δ*nocE* and *ermE*p∗-*nocE,* the MS/MS data obtained from both positive and negative ionization mode were used to build a molecular network. Specialized metabolites were annotated by matching spectra against public libraries in GNPS. The resulting networks were visualized and interpreted in Cytoscape 3.8 (Figure 4.10). Each node in the figure represents one fragmentation spectrum from a detected metabolite. The edges connecting nodes indicate the relative similarity of MS/MS data between nodes. In the current

analysis, >22,000 molecular nodes were obtained using MS-based metabolomics and GNPS analysis (Figure 4.10), but only 8% could be annotated by matching spectra with available libraries. In the analysis, SMs previously reported to be produced by *S. clavuligerus* were assessed first. The ions corresponding to CA ($[M-H]^-$, *m/z* 198.039), Ceph-C ($[M-H]^-$, *m/z* 445.104), tunicamycin A ($[M+H]^+$, *m/z* 831.424), and naringenin ($[M-H]^-$, *m/z* 271.062), were successfully detected in the network in the three strains wt, Δ*nocE*, and *ermE*p∗-*nocE* tested (Figure 4.10; Table 4.3), which indicated that neither deletion nor overexpression of *nocE* affects the production of the known secreted SMs in *S. clavuligerus*. Two common SMs, desferrioxamine E and ectoine, were also detected in the extracts of wt, Δ*nocE*, and *ermE*p∗-*nocE*. The GNPS showed the ions corresponding to desferrioxamine E (Nocardamine, $[M + H]^+$, *m/z* 601.356) and ectoine ($[M + H]^+$, *m/z* 143.082) (Figure 4.10; Table 4.3). Desferrioxamines are nonpeptide hydroxamate siderophores in bacteria that exhibit antitumor activity (Barona-Gómez et al., 2004; Kalinovskaya et al., 2011). Ectoine is a commonly produced metabolite that helps bacteria survive extreme osmotic stress (Sadeghi et al., 2014). Since the desferrioxamines and ectoine are produced by many Actinomycetes and are involved in general cellular growth/survival processes (Challis, 2005; Czech et al., 2018), finding them was not surprising, but it validates the sensitivity of our analysis. Also, it indicates that *nocE* has no role with these widely produced SMs.

In the GNPS analysis, we identified 41 SMs with high confidence (>0.6 cosine score; Table 4.3), and most of them were previously reported to have some type of bioactivities such as antibacterial, anticancer, or antiparasitic, etc. (Table 4.3). Thirty-one

of these SMs were detected in all three strains wt, Δ*nocE*, and *ermE*p∗-*nocE*, three were found in two strains (wt + Δ*nocE* or Δ*nocE* + *ermE*p*-*nocE*), and 7 SMs were detected in either wt, Δ*nocE*, or *ermE*p*-*nocE* alone, as shown with details in Table 4.3. Interestingly, many of the SMs are compounds that were thought to be only produced by plants (Table 4.3). The BGCs corresponding to most of the SMs reported in this study are still not identified, and thus there is the potential for new avenues of research on *S. clavuligerus* specialized metabolism stemming from this study.

4.5. Conclusions

This study demonstrated genomic differences between the CA BGC of *S. clavuligerus* and CA-like BGCs in the non-producer species *S. pratensis*, *S. viridis* and *Streptomyces* sp M41. One of the differences is the existence of the *nocE* gene within CA-like BGCs, while *nocE* in *S. clavuligerus* is located distant from the CA BGC and has an unknown function. In addition to previously reported bacterial species (Jensen and Paradkar, 1999, Jensen et al., 2012; Alvarez-Alvarez et al., 2013), we have shown that several Actinomycetes species have the potential to produce CA as their genomes contain CA/CA-like BGCs, though further studies are needed to determine whether such species can produce this β-lactamase inhibitor. In addition, we demonstrated that these CA/CA-like BGC-containing bacteria also possess *nocE* homologues in their genomes, which are evolutionarily descended from the same ancestor.

The NocE protein belongs to the SGNH/GDSL-hydrolases family of proteins with esterase or lipase activity. However, it has been hypothesized that lipases and other

hydrolytic enzymes could play a role in the production of the specialized metabolites in *Streptomyces* (Horinouchi, 2002; Bielen et al., 2009). This study demonstrated neither the deletion nor the overexpression of *nocE* affected the production of CA, Ceph-C, or 5*S* clavams. Moreover, the untargeted metabolomics analysis showed little difference in the detection of the specialized metabolites between the wt, Δ*nocE*, and the overexpressed *ermE*p*-*nocE* strains.

The deletion of *nocE* demonstrated a significant effect on the growth of *S. clavuligerus* in SA medium in comparison to wt, which supported the hypothesis that *nocE* might have a role in the primary metabolism in *S. clavuligerus*. Furthermore, the metabolomics profile for the three strains wt, Δ*nocE*, and *ermE*p*-*nocE* showed distinct differences between them. Moreover, the GNPS analysis for the types of compounds detected in the extracts of the three strains wt *S. clavuligerus*, Δ*nocE*, and *ermE*p*-*nocE* demonstrated variety in the class and numbers of these compounds, which may explain the distinctions in growth and general metabolomics profiles among the three bacterial strains. However, the high number of "Lipids and lipids-like molecules" detected in Δ*nocE* extract support the proposed function of NocE as a lipase enzyme that breaks down and catabolizes lipid molecules.

4.6. Figures and tables

4.6.1. Figures

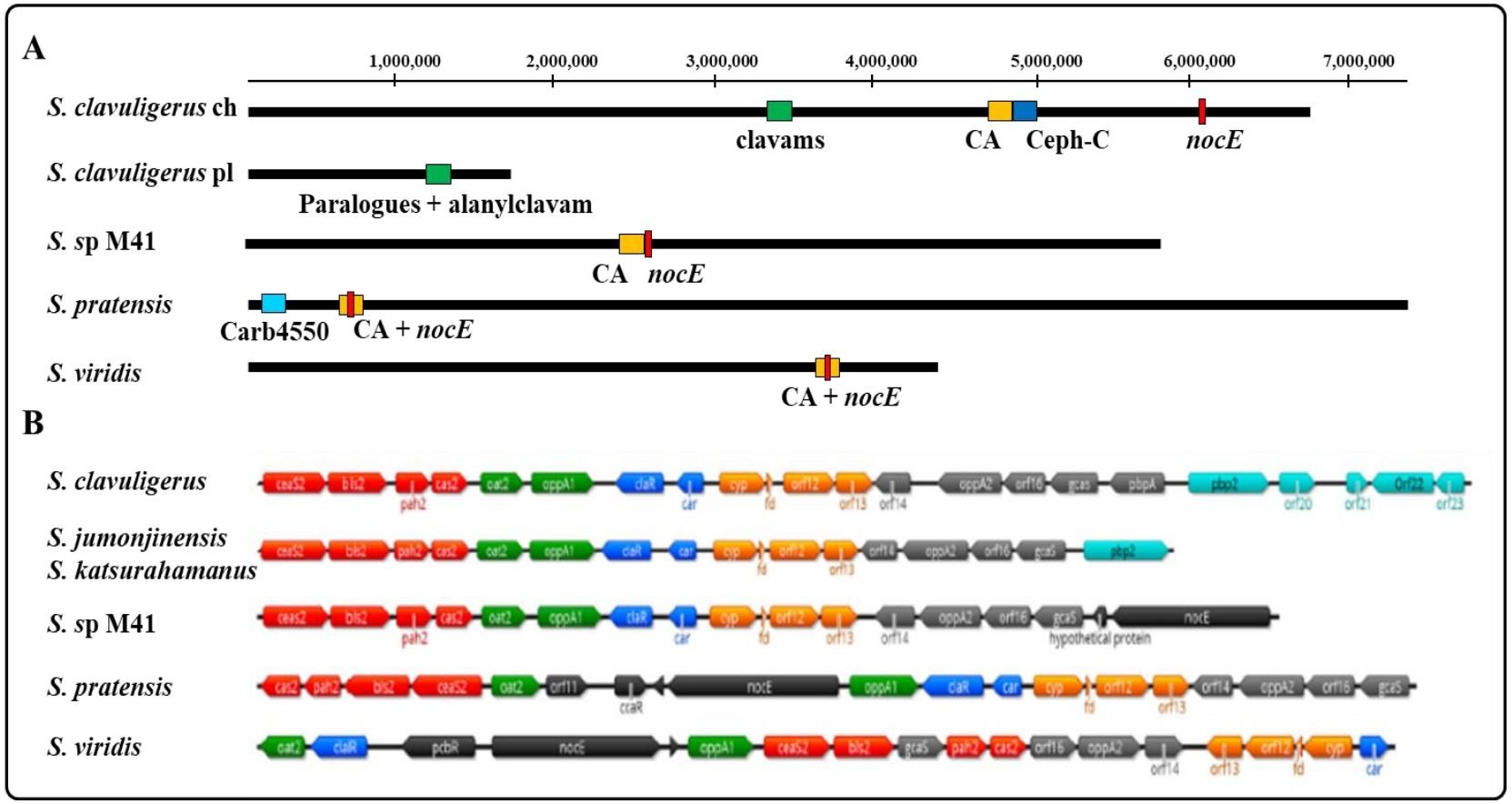


Figure 4.1. Genetic comparison between the CA producing/non-producing species. (A) Biosynthetic gene cluster mapping shows the location of β-lactam BGCs on the chromosomes/plasmid of *S. clavuligerus* (CA producer) and the *Streptomyces* sp M41, *S. pratensis*, and *S. viridis* (CA non-producers). Also, the figure shows the location of the *nocE* gene (red) corresponding to CA/CA-like BGCs. **(B)** The architecture of CA/CA-like BGCs from the CA producer species; *S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus*, in comparison to the non-producers; *Streptomyces* sp M41, *S. pratensis, and S. viridis* showing their gene content and relative organization. Genes are color-coded based on known or predicted transcriptional units. It shows the *nocE* gene in the CA-like BGCs of the non-producers. CA: clavulanic acid, Carb4550: carbapenem MM4550, Ceph-C: cephamycin C, ch: chromosome, pl: plasmid.

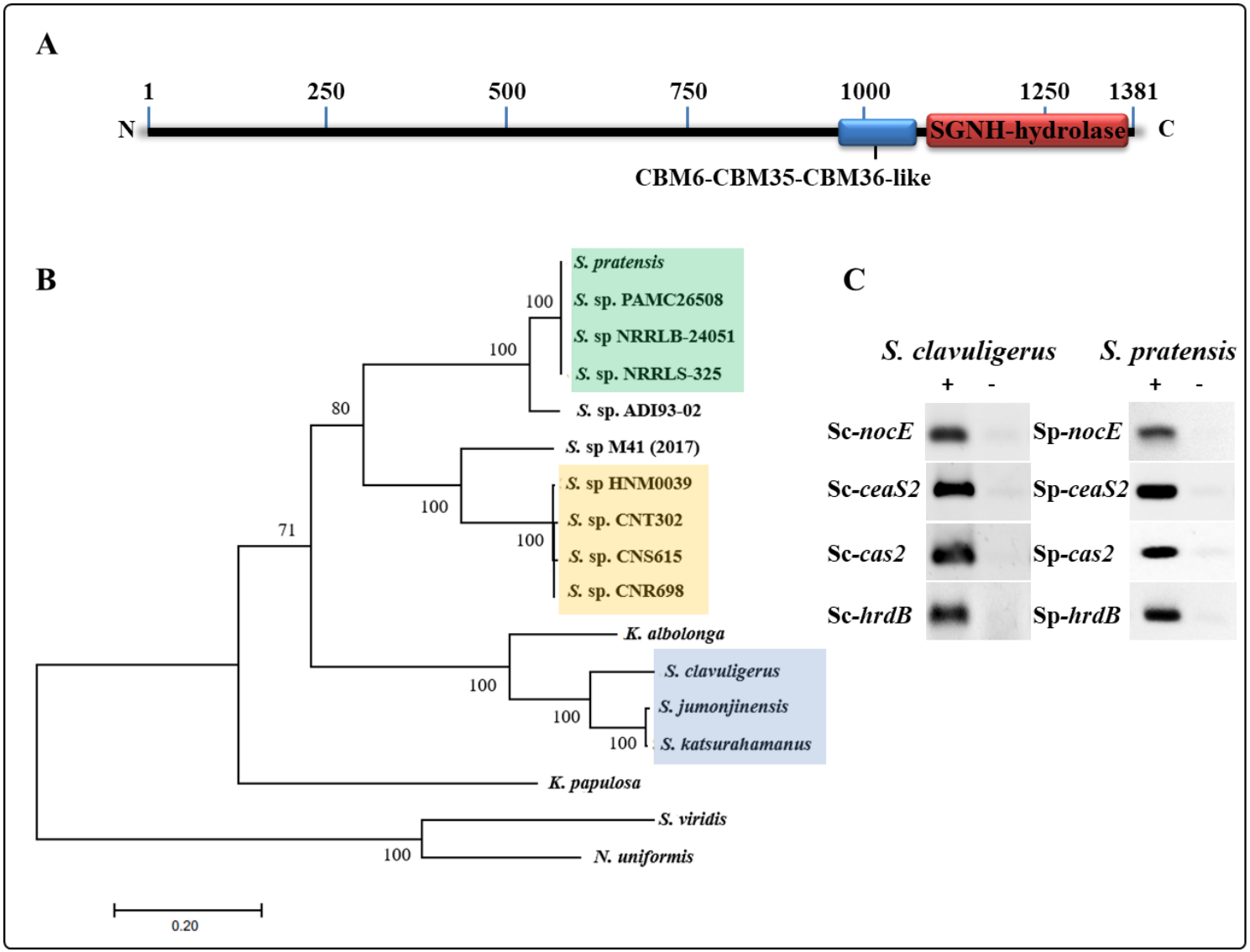
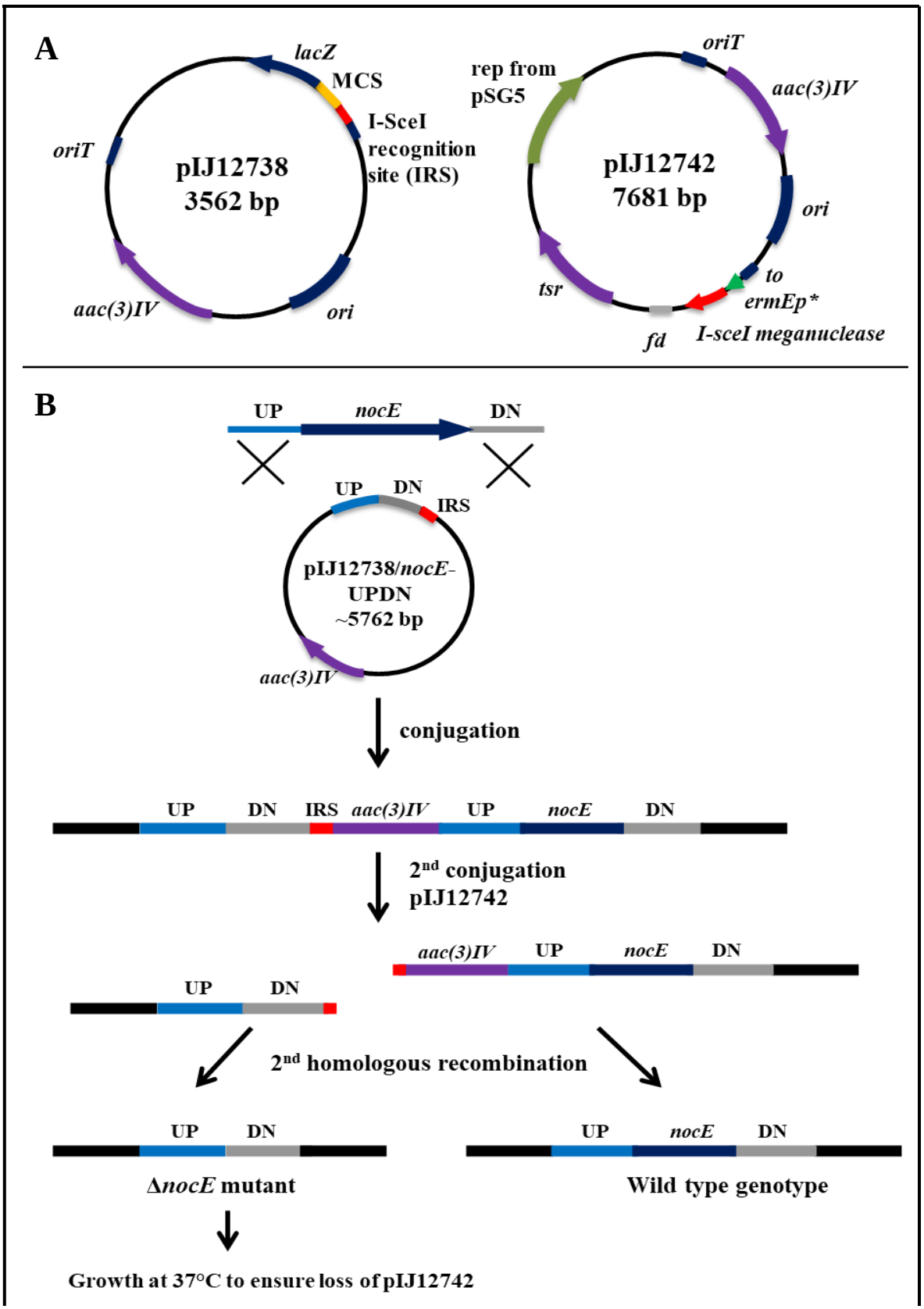

A
1
250
500
750
1000
1250
1381
N
C
SGNH-hydrolase
CBM6-CBM35-CBM36-like
B
S. pratensis
S. sp. PAMC26508
S. sp NRRLB-24051
S. sp. NRRLS-325
S. sp. ADI93-02
S. sp M41 (2017)
S. sp HNM0039
S. sp. CNT302
S. sp. CNS615
S. sp. CNR698
K. albolonga
S. clavuligerus
S. jumonjinensis
S. katsurahamanus
K. papulosa
S. viridis
N. uniformis
100
100
80
100
71
100
100
100
100
0.20
C
S. clavuligerus
S. pratensis
+
-
+
-
Sc-nocE
Sp-nocE
Sc-ceaS2
Sp-ceaS2
Sc-cas2
Sp-cas2
Sc-hrdB
Sp-hrdB

Figure 4.2. NocE features. (**A)** The schematic diagram for the NocE protein from *S. clavuligerus* showing the main two domains, SGNH hydrolase, and the carbohydrate-binding CBM6-CBM35-CBM36-like domain**. (B)** A phylogenetic tree based on NocE proteins was built using MEGA 7.0 for 17 species of bacteria. The NocE proteins' phylogeny for CA and Ceph-C producer bacteria form one clade (blue box), the bacteria with Ceph-C BGC their NocE proteins form a distinct clade from others (orange box), and the clade in the green box includes the bacteria that carry CA-like and carb4550 BGCs. The remaining species (*S*. sp. AD193-02, *S*. sp. M41, *K. albolonga, K. papulosa*, and *S. viridis*) possess CA-like BGC only. *N. uniformis,* which has *nocE* in the nocardicin A BGC, was included in the tree. The protein accession numbers are in Table 4.1. The tree was constructed using the maximum likelihood algorithm, and bootstrap values ≥50% for 1000 repetitions are shown. The scale bar indicates the number of amino acid substitutions per site. **(C)** RT-PCR analysis (+) of RNA isolated from *S. clavuligerus* and *S. pratensis,* showing the expression of Sc-*nocE* during CA production in *S. clavuligerus* and the expression of Sp-*nocE* during *S. pratensis* growth in SM. Transcription of *ceaS2* and *cas2* of the respective species was used as a reporter for CA-BGC expression. The constitutively expressed Sc-*hrdB* and Sp-*hrdB* were used as controls. Negative controls (−) consisted of RNA samples subjected to PCR without undergoing reverse transcription.


A
lacZ
MCS
I-SceI
recognition
site (IRS)
oriT
pIJ12738
3562 bp
aac(3)IV
ori
rep from
pSG5
oriT
aac(3)IV
pIJ12742
7681 bp
ori
to
ermEp*
tsr
fd
I-sceI meganuclease
B
UP
nocE
DN
UP
DN
IRS
pIJ12738/nocE-
UPDN
~5762 bp
aac(3)IV
conjugation
UP
DN
IRS
aac(3)IV
UP
nocE
DN
2nd conjugation
pIJ12742
aac(3)IV
UP
nocE
DN
UP
DN
2nd homologous recombination
UP
DN
ΔnocE mutant
UP
nocE
DN
Wild type genotype
Growth at 37°C to ensure loss of pIJ12742

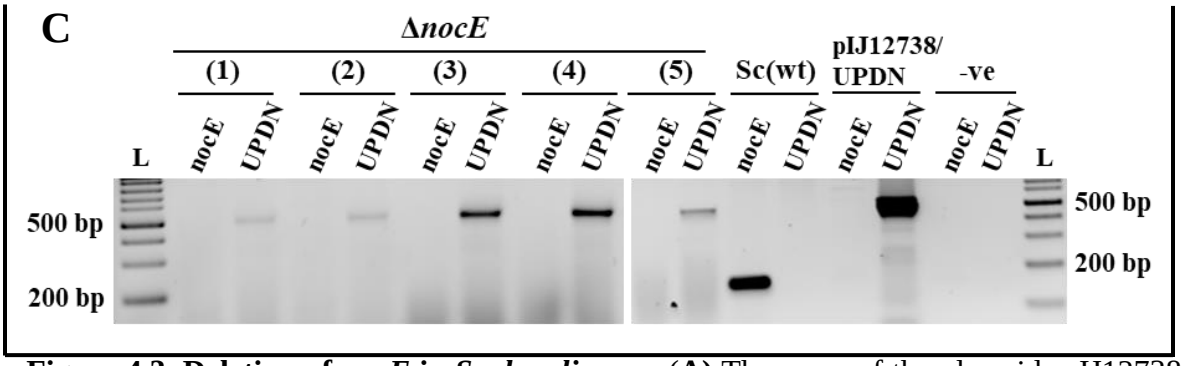


Figure 4.3. Deletion of *nocE* in *S. clavuligerus*. (**A)** The maps of the plasmids pIJ12738 and pIJ12742 used in this study. (**B)** The schematic diagram for the protocol that was followed to obtain markerless *nocE* deleted *S. clavuligerus*. The pIJ12738 containing the I-SceI recognition sequence (IRS, shown in red) and the flanking regions of the *nocE* gene (upstream region (UP) in the blue and downstream part (DN) in grey) was conjugated into wt *S. clavuligerus*. Apramycin-resistant exconjugants were selected as recipients of the second vector pIJ12742, which expresses the I-SceI meganuclease gene. After the second conjugation, the plate was overlaid with thiostrepton to select for exconjugants. I-SceI creates double-strand breaks at its introduced recognition sequence, and the only genomes to survive are those that undergo homologous recombination to reconstitute an intact chromosome. The recombination could result in the wild-type genotype or the Δ*nocE* mutant genotype. Exconjugants were analyzed by PCR to confirm the required mutant genotype. **(C)** The gel electrophoresis of PCRs was conducted for five exconjugants that are Thio and Apr sensitive**.** Two sets of primers were used (see Supplementary Table S2.1), the first set to amplify the *nocE* gene and thus confirm the gene's absence or deletion in our samples. The second set amplifies the internal region from the upstream to the downstream (UPDN) region and confirms the *nocE* gene deletion. Samples from wt *S. clavuligerus* and pIJ12739/*nocE*-UPDN were used as controls. *aac(3)IV*, apramycin resistance gene; *ori*, *E. coli* origin of replication; *oriT*, the origin of transfer; MCS, multiple cloning site; *lacZ*, a fragment of *lacZ*; rep, origin of replication from pSG5; *tsr*, thiostrepton resistance gene; *fd*, transcriptional terminator from phage *fd*; *to*, transcriptional terminator; *ermE*p*, mutated constitutive promoter from *ermE*.

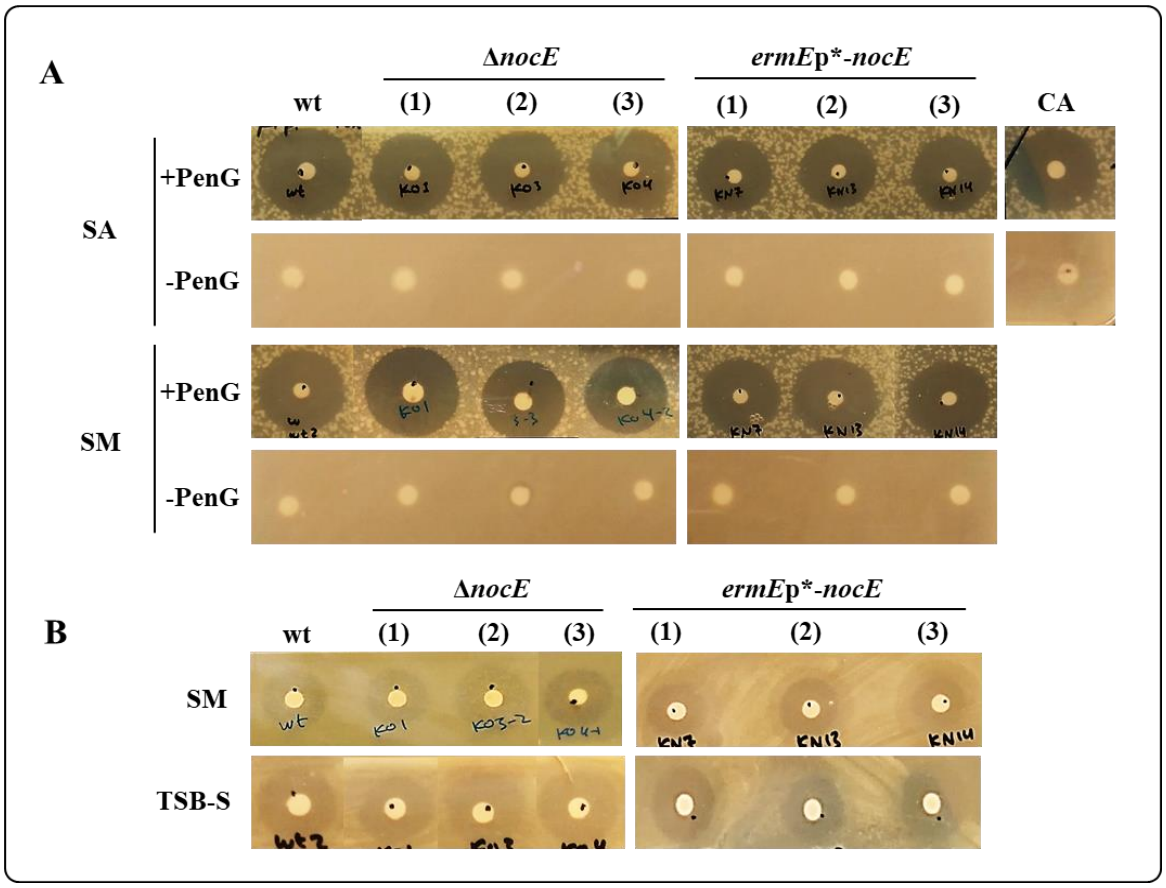


Figure 4.4. Clavulanic acid and cephamycin C bioassays in *nocE* deleted and overexpressed strains. Supernatants of liquid cultures for wt *S clavuligerus*, Δ*nocE*, and *ermE*p*-*nocE* in three types of media SA, SM, and TSB-S were used to test the production of CA and cephamycin C. (**A)** For CA detection the bioassays were performed against *K. pneumoniae* on TSA plates with or without PenG for triplicate samples of Δ*nocE* and *ermE*p*-*nocE* strains grown in SA and SM. The zones of growth inhibition were noticed in all samples tested in the plate with PenG. Supernatant from the wt *S. clavuligerus* and the CA solution (10 μg) were used as controls. (**B)** For Ceph-C detection, the bioassays were performed against *E. coli* ESS (the β-lactam antibiotic sensitive strain) for triplicate samples of Δ*nocE* and *ermE*p*-*nocE* strains grown in SM and TSB-S. The zones of growth inhibition were observed in all samples. Supernatant from the wt *S. clavuligerus* was used as control. CA, clavulanic acid; SA, Starch asparagine media; TSB-S, tryptic soy broth – starch 1% media; +PenG, penicillin G was added to TSA plate; -PenG, without penicillin G.

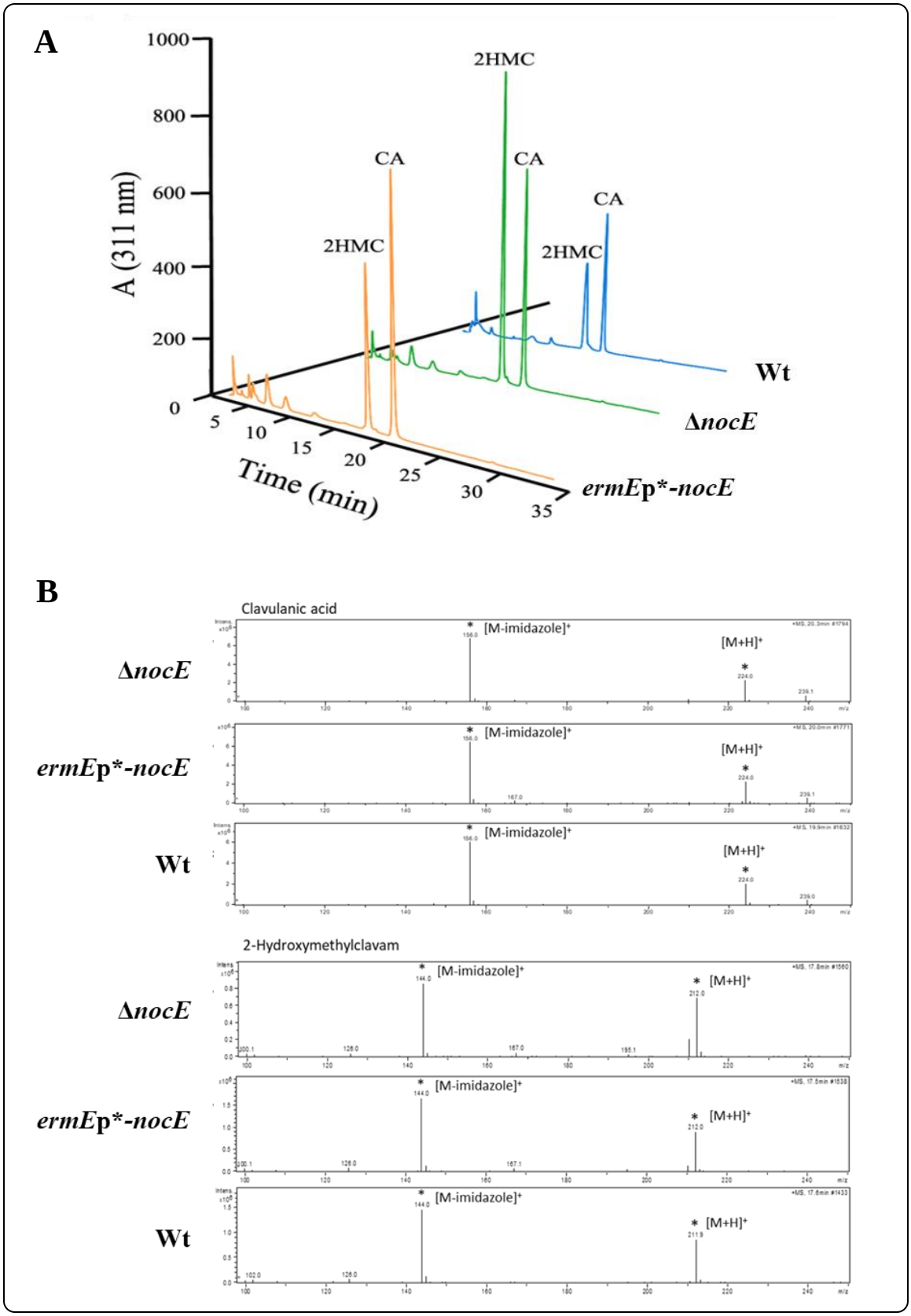

A
1000
800
600
400
200
0
A (311 nm)
2HMC
CA
2HMC
CA
2HMC
CA
Wt
ΔnocE
ermEp*-nocE
5
10
15
20
25
30
35
Time (min)
B
Clavulanic acid
ΔnocE
ermEp*-nocE
Wt
[M-imidazole]+
[M+H]+
2-Hydroxymethylclavam
ΔnocE
ermEp*-nocE
Wt
[M-imidazole]+
[M+H]+

Figure 4.5. Metabolite detection for clavulanic acid and 2-hydroxymethylclavam (2HMC). (**A)** LC/MS analysis of imidazole derivatized 96-h soy culture supernatants from the wt *S. clavuligerus* (blue), Δ*nocE* (green), and *ermE*p∗-*nocE* (orange) strains to assess CA and 5*S* clavam metabolite production. (**B)** The mass spectra analysis of the CA and 2HMC peaks. The major peaks corresponding to imidazole derivatized CA $[M+H]^+$ (*m/z* = 224) and the fragmented product $[M\text{-imidazole}]^+$ (*m/z* = 156), and imidazole derivatized 2HMC $[M+H]^+$ (*m/z* = 212) and the fragmented product $[M\text{-imidazole}]^+$ (*m/z* = 144) the peaks are pointed by (*).

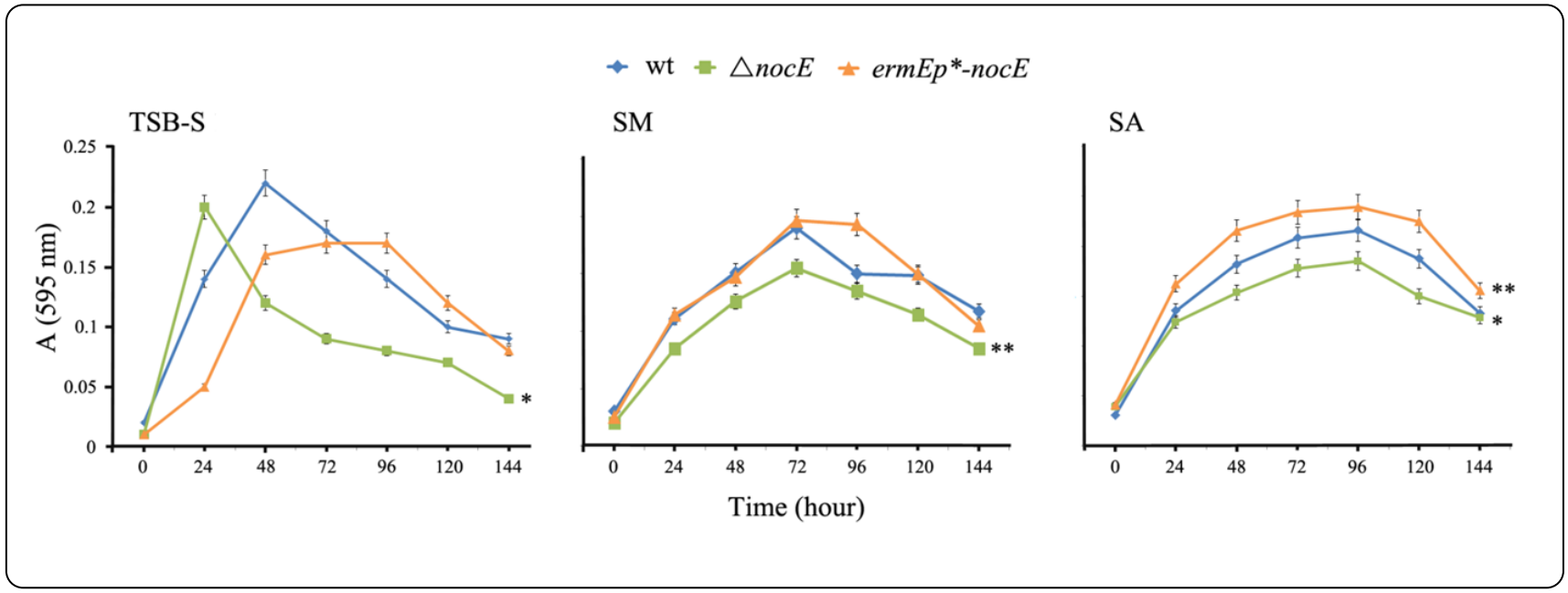
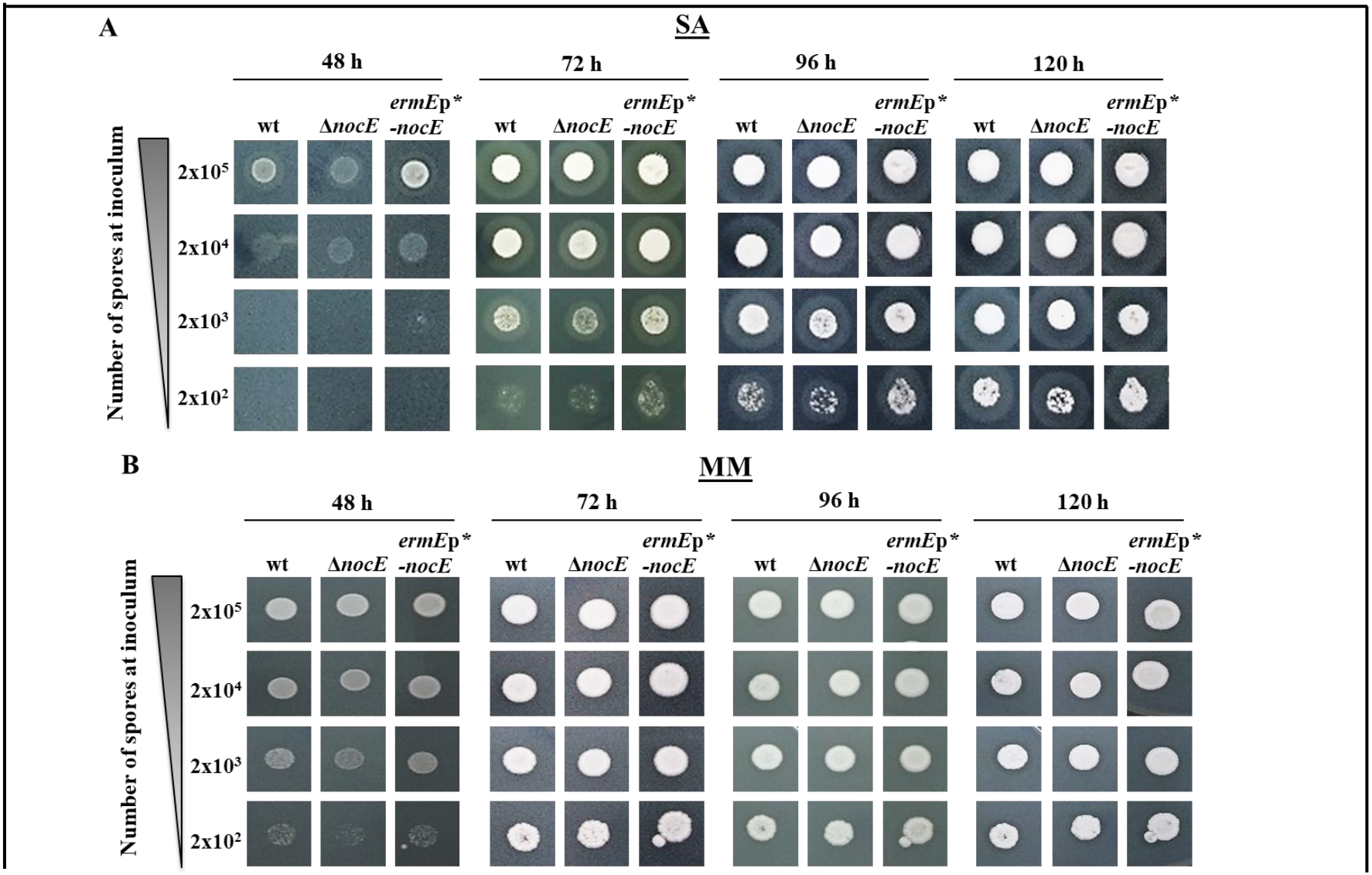


Figure 4.6. Cellular growth curves of Δ*nocE* (green squares) and *ermE*p-nocE* (orange triangles) mutants in comparison to wt *S. clavuligerus* (blue diamonds) using three different types of media: TSB-S, soy, and starch asparagine. Growth curves were calculated based on DNA extraction and quantification using a simplified diphenylamine colorimetric method (see Materials and Methods). * $p<0.05$ and ** $p<0.001$.


A
SA
48 h
72 h
96 h
120 h
wt
ΔnocE
ermEp*-nocE
Number of spores at inoculum
2x10^5
2x10^4
2x10^3
2x10^2
B
MM

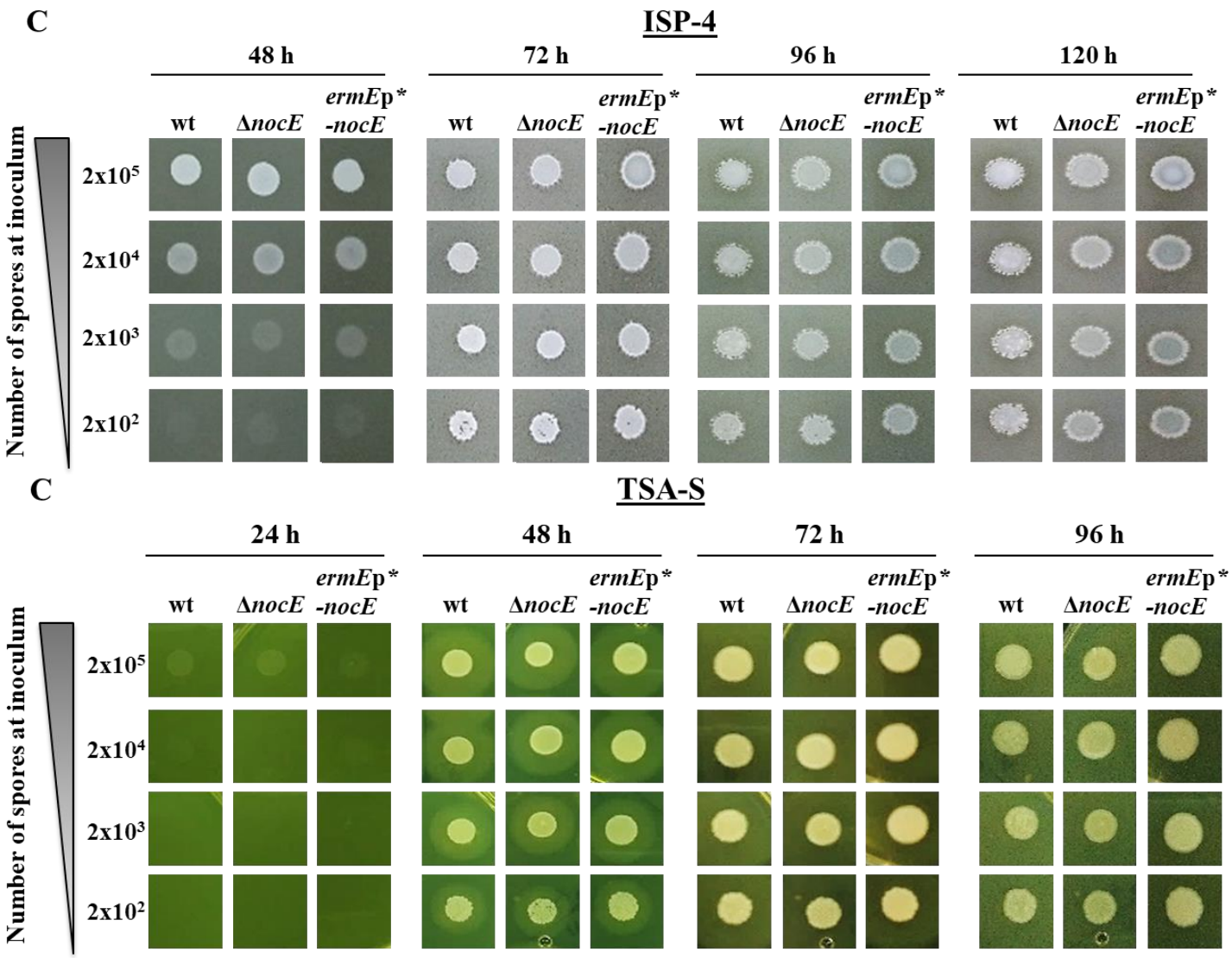

C
ISP-4
48 h
72 h
96 h
120 h
wt
ΔnocE
ermEp*-nocE
Number of spores at inoculum
2x10^5
2x10^4
2x10^3
2x10^2
C
TSA-S
24 h
48 h
72 h
96 h
wt
ΔnocE
ermEp*-nocE
Number of spores at inoculum
2x10^5
2x10^4
2x10^3
2x10^2

Figure 4.7. Growth characteristics of the wt *S. clavuligerus*, Δ*nocE*, and *ermE*p∗-*nocE* strains on different solid media. **(A)** On starch asparagine agar plate (SA). **(B)** On minimal essential medium (MM). **(C)** On ISP-4 medium. **(D)** On TSB-S media. The spores for each strain were inoculated in spots over the agar media in descending fold numbers of 2×10^5, 2×10^4, 2×10^3, 2×10^2 spores. The growth of the bacterial spots was observed for five days.

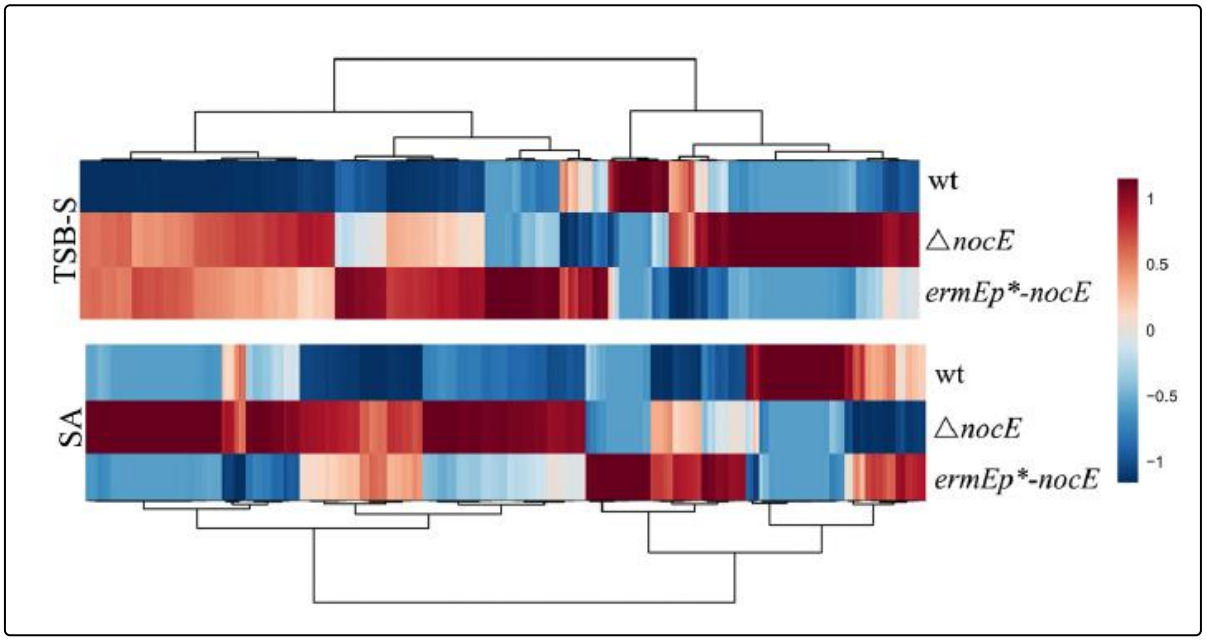
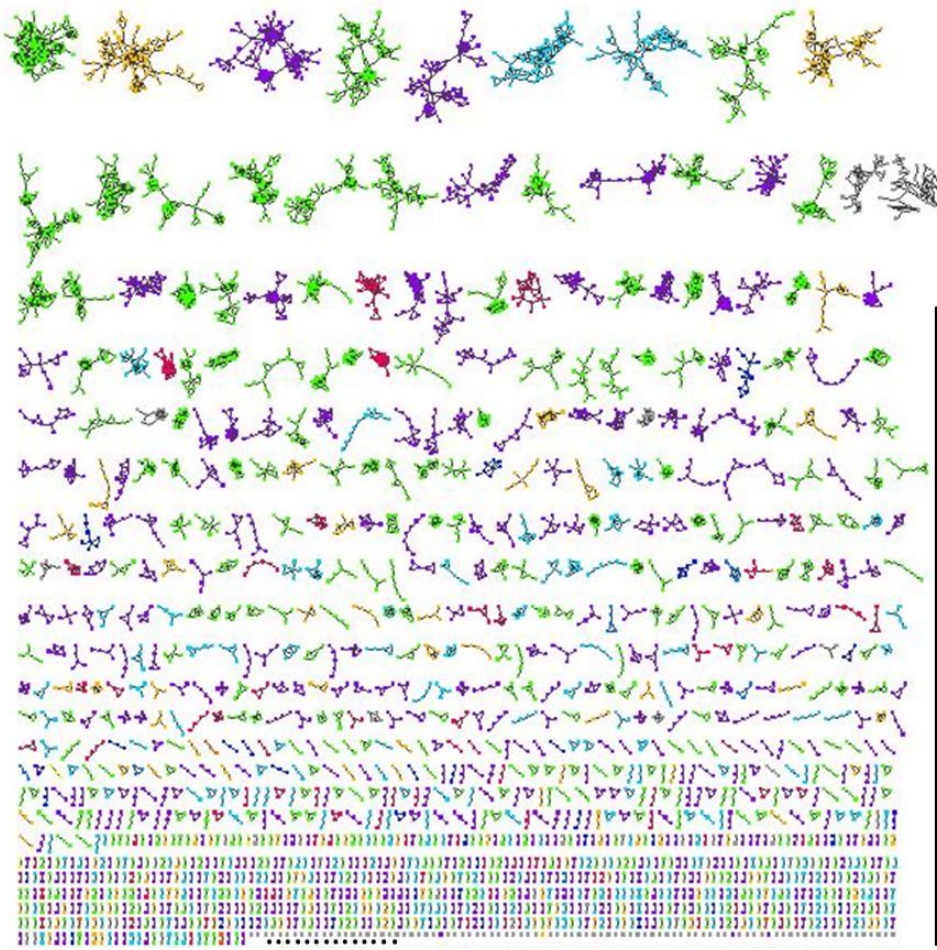


Figure 4.8. Comparative metabolomics of the wt *S. clavuligerus*, Δ*nocE,* and *ermE*p∗-*nocE* strains grown on two different media SA and TSA-S. The heat map was constructed by hierarchical clustering of ~1000 statically significant features to show overall differences between the three strains. The dendrograms on the top and bottom indicate groups of similarly expressed features resulting from the hierarchical clustering analysis.

A



- Alkaloids and derivatives
- Benzenoids
- Homogeneous non-metal compounds
- Hydrocarbons
- Lipids and lipid-like molecules
- Mixed metal/non-metal compounds
- Nucleosides, nucleotides, and analogues
- Organic 1,3-dipolar compounds
- Organic acids and derivatives
- Organic nitrogen compounds
- Organic oxygen compounds
- Organohalogen compounds
- Organoheterocyclic compounds
- Organometallic compounds
- Organosulfur compounds
- Phenylpropanoids and polyketides
- no matches

B

Superclass of molecules	Number of spectral features in each strain		
	Wt	Δ*nocE*	*ermE*p*-*nocE*
Alkaloids and derivatives	8	6	8
Benzenoids	297	315	276
Homogeneous non-metal compounds	5	5	4
Hydrocarbons	3	4	4
Lipids and lipid-like molecules	1447	1483	1456
Mixed metal/non-metal compounds	1	1	1
Nucleosides, nucleotides, and analogues	12	9	12
Organic 1,3-dipolar compounds	1	1	1
Organic acids and derivatives	510	522	562
Organic nitrogen compounds	29	26	23
Organic oxygen compounds	66	79	96
Organohalogen compounds	5	4	6
Organoheterocyclic compounds	1464	1446	1563
Organometallic compounds	1	1	1
Organosulfur compounds	0	1	0
Phenylpropanoids and polyketides	237	220	238
Total	4086	4123	4251

Figure 4.9. Metabolomics analysis for the main groups of molecules in wt *S. clavuligerus*, Δ*nocE*, and ermEp*-*nocE*. (A) The molecular networking for spectral features detected by untargeted LC-MS/MS and GNPS analysis in the extractions of TSA-S and SA cultures for the three strains wt *S. clavuligerus*, Δ*nocE*, and ermEp*-*nocE*, and visualized by Cytoscape 3.8 program. The networks show >14000 nodes representing molecules/compounds of 16 superclasses/groups; their names are in the figure legend. **(B)** The table shows the number of spectral features detected for each superclass/group of molecules in each bacterial strain, after *in silico* annotation using NAP in GNPS.

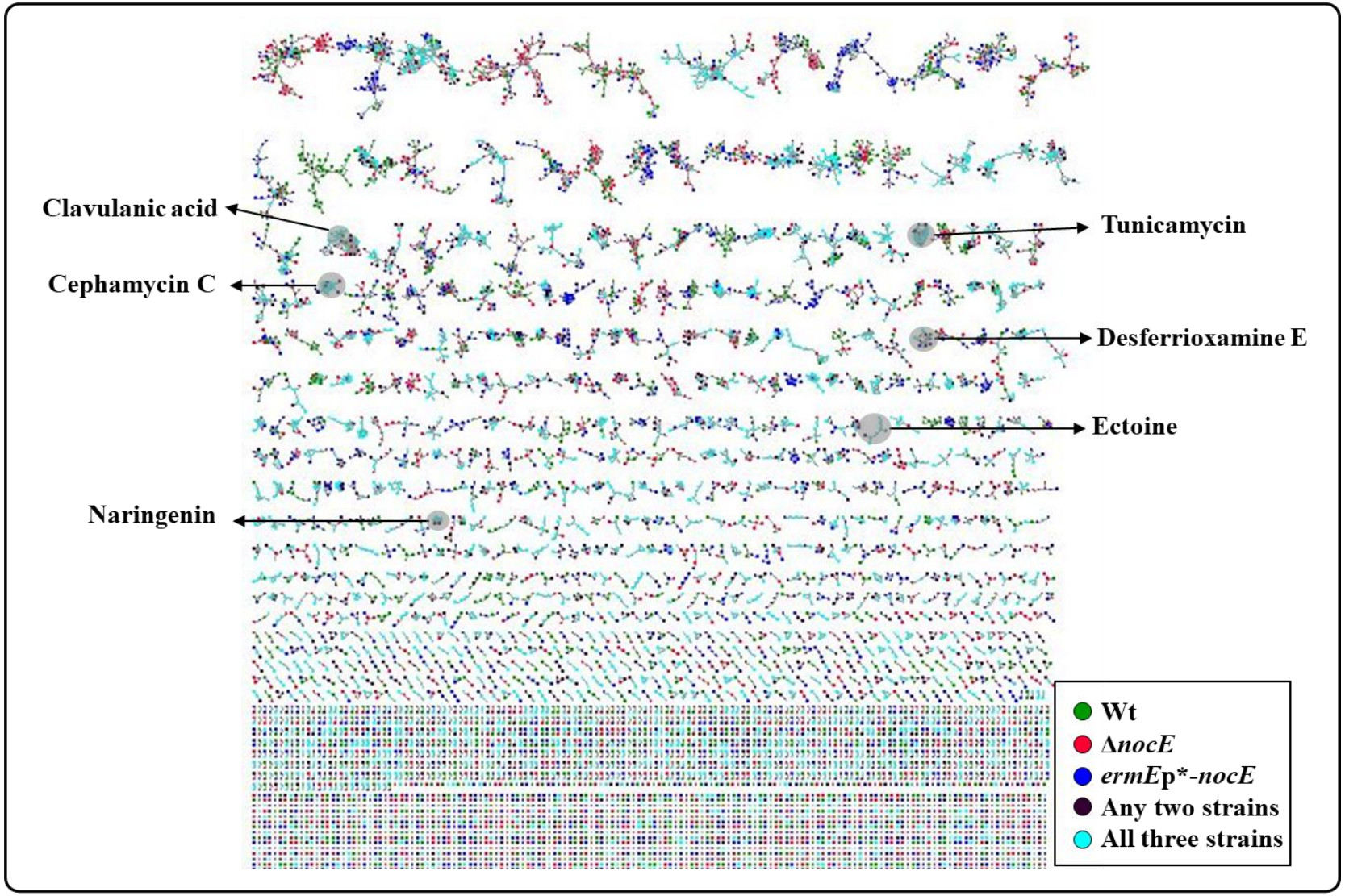

Clavulanic acid
Cephamycin C
Naringenin
Tunicamycin
Desferrioxamine E
Ectoine
Wt
ΔnocE
ermEp*-nocE
Any two strains
All three strains

Figure 4.10. Metabolic network constructed using *S. clavuligerus* wt, Δ*nocE*, and *ermE*p∗-*nocE* strains culture extracts (culture conditions and details are described in the "Materials and Methods"). The network is color-coded according to the source organism (bottom right legend), where each node depicts a mass spectrum and edges represent the relationship between different nodes. The clusters of nodes related to metabolites (CA, Ceph-C, Tunicamycin, Naringenin, Desferriocamine E, and Ectoine) which have been reported to be produced by *S. clavuligerus* are indicated by black arrows and their names.

4.6.2. Tables

Table 4.1. Comparative analysis of NocE in *S. clavuligerus* with other orthologues in Actinomycete species that are predicted to have β-lactam antibiotic biosynthesis gene clusters.

Bacteria	NocE/(SGNH/GDSL hydrolase)				Refseq (NCBI)	β-lactam antibiotics BGCs (similarity%)[a]	MIBiG BGC-ID[b]
	Query cover	E value	Identity (%)	Accession			
S. jumonjinensis NRRL 5741	99%	0.0	84%	WP_153520536.1	NZ_VCLA00000000.1	CA (54%)	BGC0845
						Ceph-C (84%)	BGC0319
S. katsurahamanus T-272	99%	0.0	83%	WP_153484997.1	NZ_VDEQ00000000	CA (54%)	BGC0845
						Ceph-C (84%)	BGC0319
Kitasatospora albolonga YIM 101047	98%	0.0	70%	WP_084748002.1	NZ_CP020563.1	CA (29%)	BGC0845
Streptomyces sp. CNT302	100%	0.0	53%	WP_026281417.1	KB898270.1	Ceph-C (52%)	BGC0319
Streptomyces sp. HNM0039	100%	0.0	53%	WP_108907594.1	NZ_CP029188.1	Ceph-C (52%)	BGC0319
Streptomyces sp. CNS615	100%	0.0	53%	WP_026165521.1	NZ_AQPE00000000.1	Ceph-C (42%)	BGC0319
Streptomyces sp. NRRL S-325	100%	0.0	53%	WP_051800379.1	NZ_JOIW00000000.1	CA (20%)	BGC0845

						Carb4550 (65%)	BGC0842
Streptomyces sp. NRRL B-24051	100%	0.0	53%	WP_030879880.1	NZ_JOAE00000000.1	CA (20%)	BGC0845
						Carb4550 (65%)	BGC0842
Streptomyces sp. AD193-02	97%	0.0	54%	WP_124279445.1	NZ_RPGU00000000.1	CA (20%)	BGC0845
Streptomyces sp. PAMC26508	99%	0.0	54%	AGJ58755.1	NC_021055.1	CA (20%)	BGC0845
						Carb4550 (65%)	BGC0842
S. pratensis ATCC33331	100%	0.0	53%	ADW02015.1	NC_016114	CA (42%)	BGC0845
						Carb4550 (65%)	BGC0842
Streptomyces sp. CNR698	100%	0.0	53%	WP_027733979.1	NZ_AZXC00000000.1	Ceph-C (42%)	BGC0319
Streptomyces sp. M41(2017)	98%	0.0	52%	WP_107482601.1	MWFK01000001.1	CA (54%)	BGC0845
Kitasatospora papulosa	99%	0.0	50%	WP_030122804.1	NZ_JNYQ00000000.1	CA (20%)	BGC0845
Sac. viridis DSM 43017	100%	0.0	40%	WP_015787608.1	CP001683.1	CA (58%)	BGC0845

[a] The β-lactam BGCs found in this bacterial genome with its similarity percent to the reference BGC. CA: Clavulanic acid, Ceph-C: Cephamycin C, Carb4550: Carbapenem MM4550.

[b] The number of the reference BGC in MIBiG. MIBiG BGC-ID: Minimal information about biosynthetic gene cluster – identification number.

Table 4.2. The secretory signal peptide prediction for NocE homologues from Actinomycetes species. The value 1 has the highest probability to have a signal peptide.

Bacteria	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other [a]
Streptomyces clavuligerus	0.9264	0.0677	0.0033	0.0026
Streptomyces jumonjinensis	0.03	0.003	0.0054	0.9616
Streptomyces katsurahamanus	0.0529	0.0051	0.0082	0.9338
Streptomyces pratensis	0.6467	0.225	0.1155	0.0128
Saccharomonospora viridis	0.1142	0.0075	0.0134	0.8649
Kitasatospora albolonga	0.9474	0.0252	0.0239	0.0034
Streptomyces sp. CNT302	0.5737	0.0358	0.3893	0.0012
Streptomyces sp. HNM0039	0.6058	0.025	0.3686	0.0007
Streptomyces sp. CNS615	0.6063	0.027	0.3659	0.0008
Streptomyces sp. NRRL S-325	0.6467	0.225	0.1155	0.0128
Streptomyces sp. NRRL B-24051	0.6467	0.225	0.1155	0.0128
Streptomyces sp. AD193-02	0.7801	0.0586	0.1464	0.0149
Streptomyces sp. PAMC26508	0.0672	0.0163	0.0073	0.9092
Streptomyces sp. CNR698	0.9462	0.0384	0.0143	0.0012
Streptomyces sp. M41(2017)	0.6797	0.0447	0.2743	0.0013
Kitasatospora papulosa	0.9393	0.0581	0.0019	0.0006
Nocardia uniformis	0.9881	0.0051	0.0059	0.001

[a]: The sequence does not have any kind of secretory signal peptide.

Table 4.3. Specialized metabolites (SMs) detected with high confidence in wt *S. clavuligerus*, Δ*nocE*, and *ermE*p∗-*nocE* strains using MS-based metabolomics and GNPS analysis.

Name	Detected in	Observed m/z [Adduct]	Molecular formula (Weight, g/mol)	Cosine score[a]	Bioactivity	Reference
(-)-Carveol	Wt, Δ*nocE*, *ermE*p**nocE*	135.117 $[M -H_2O+H]^+$	$C_{10}H_{16}O$ (152.237)	0.96	plant-associated metabolite	(Bouwmeester et al., 1998)
(-)-Caryophyllene oxide	Wt, Δ*nocE*, *ermE*p**nocE*	221.19 $[M+H]^+$	$C_{15}H_{24}O$ (220.356)	0.89	plant-associated metabolite	(Ghelardini et al., 2001)
Bisucaberin	Wt, Δ*nocE*, *ermE*p**nocE*	401.239 $[M+H]^+$	$C_{18}H_{32}N_4O_6$ (400.476)	0.65	macrocyclic siderophore	(Hou et al., 1998)
Cephamycin C	Wt, Δ*nocE*, *ermE*p**nocE*	445.104 $[M-H]^-$	$C_{16}H_{22}N_4O_9S$ (446.431)	0.83	antibiotic	(Nagarajan et al., 1971)
Clavulanic acid	Wt, Δ*nocE*, *ermE*p**nocE*	198.039 $[M-H]^-$	$C_8H_9NO_5$ (199.162)	0.96	β-lactamase inhibitor	(Reading & Cole, 1977)
Costunolide	Wt, Δ*nocE*, *ermE*p**nocE*	233.154 $[M+H]^+$	$C_{15}H_{20}O_2$ (232.323)	0.86	plant-associated metabolite	(De Kraker et al., 2002)
Cuminyl alcohol	Wt, Δ*nocE*, *ermE*p**nocE*	133.101 $[M -H_2O+H]^+$	$C_{10}H_{14}O$ (150.221)	0.94	plant-associated metabolite	(Bartoňková & Dvořák, 2018)
Desferrioxamine E	Wt, Δ*nocE*, *ermE*p**nocE*	599.342 $[M-H]^-$	$C_{27}H_{48}N_6O_9$ (600.714)	0.89	nonpeptide hydroxamate siderophores	(Álvarez-álvarez et al., 2017)
Desmethylenyl nocardamine	Wt, Δ*nocE*, *ermE*p**nocE*	587.34 $[M+H]^+$	$C_{26}H_46N_6O_9$ (586.7)	0.89	siderophores	(Lee et al., 2005)
Ectoine	Wt, Δ*nocE*, *ermE*p**nocE*	143.081 $[M+H]^+$	$C_6H_{10}N_2O_2$ (142.158)	0.86	osmolyte	(Sadeghi et al., 2014)
Ethirimol	Wt, Δ*nocE*, *ermE*p**nocE*	182.118 $[M-C_2H_4+H]^+$	$C_{11}H_{19}N_3O$ (209.29)	0.70	antifungal	(Lewis et al., 2016)
Hydroxyvalerenic acid	Wt, Δ*nocE*, *ermE*p**nocE*	499.307 $[2M-H]^-$	$C_{15}H_{22}O_3$ (250.338)	0.92	plant-associated metabolite	(Wong et al., 2018)

Imazapic	Wt, Δ*nocE*, *ermE*p**nocE*	258.124 $[M -H_2O+H]^+$	$C_{14}H_{17}N_3O_3$ (275.308)	0.69	herbicide	(Melland & McLaren, 1998)
Isoalantolactone	Wt, Δ*nocE*, *ermE*p**nocE*	215.143 $[M -H_2O+H]^+$	$C_{15}H_{20}O_2$ (232.323)	0.88	apoptosis inducer, antifungal	(Khan et al., 2012)
Kahweol	Wt, Δ*nocE*, *ermE*p**nocE*	315.196 $[M+H]^+$	$C_{20}H_{26}O_3$ (314.42)	0.65	plant-associated metabolite	(Fumimoto et al., 2012)
L-Saccharopine	Wt, Δ*nocE*, *ermE*p**nocE*	277.154 $[M+H]^+$	$C_{11}H_{20}N_2O_6$ (276.2863)	0.75	fungi and plant-associated metabolite	(Arruda & Barreto, 2020)
Naringenin	Wt, Δ*nocE*, *ermE*p**nocE*	271.062 $[M-H]^-$	$C_{15}H_{12}O_5$ (272.256)	0.90	antibacterial, antifungal, anticancer	(Álvarez-Álvarez et al., 2015)
Narirutin/ Isonaringenin	Wt, Δ*nocE*, *ermE*p**nocE*	271.02 [271.02]	$C_{27}H_{32}O_{14}$ (580.5)	0.63	plant-associated metabolite	(Rouseff et al., 1987)
Neoandrographolide	Wt, Δ*nocE*, *ermE*p**nocE*	479.264 $[M-H]^-$	$C_{26}H_{40}O_8$ (480.598)	0.64	anti-inflammatory	(Sharma et al., 2019)
Parthenolide	Wt, Δ*nocE*, *ermE*p**nocE*	249.149 $[M+H]^+$	$C_{15}H_{20}O_3$ (248.317)	0.77	plant-associated metabolite	(Long et al., 2013)
Pentostatin	Wt, Δ*nocE*, *ermE*p**nocE*	135.066 $[M+2H]^{2+}$	$C_{11}H_{16}N_4O_4$ (268.273)	0.87	anticancer	(Dillman, 2004)
Quadrone	Wt, Δ*nocE*, *ermE*p**nocE*	247.134 $[M-H]^-$	$C_{15}H_{20}O_3$ (248.32)	0.72	antitumor	(Wijeratne et al., 2003)
Tunicamycin A	Wt, Δ*nocE*, *ermE*p**nocE*	831.424 $[M+H]^+$	$C_{38}H_{62}N_4O_{16}$ (830.926)	0.88	antibiotic	(Kenig & Reading, 1979)
Tunicamycin B	Wt, Δ*nocE*, *ermE*p**nocE*	845.439 $[M+H]^+$	$C_{39}H_{64}N_4O_{16}$ (844.953)	0.88	antibiotic	(Kenig & Reading, 1979)
Tunicamycin C putative	Wt, Δ*nocE*, *ermE*p**nocE*	817.408 $[M+H]^+$	$C_{37}H_{60}N_4O_{16}$ (816.899)	0.92	antibiotic	(Kenig & Reading, 1979)
Tunicamycin I	Wt, Δ*nocE*, *ermE*p**nocE*	803.392 $[M+H]^+$	$C_{36}H_{58}N_4O_{16}$ (802.872)	0.91	antibiotic	(Martínez-Burgo et al., 2019)
Tunicamycin I-CH_2	Wt, Δ*nocE*,	789.376	$C_{35}H_{56}N_4O_{16}$	0.94	antibiotic	(Martínez-Burgo et

putative	*ermE*p**nocE*	$[M+H]^+$	(788.836)			al., 2019)
Tunicamycin IX	Wt, Δ*nocE*, *ermE*p**nocE*	859.455 $[M+H]^+$	$C_{40}H_{66}N_4O_{16}$ (858.447)	0.68	antibiotic	(Martínez-Burgo et al., 2019)
Uvaol	Wt, Δ*nocE*, *ermE*p**nocE*	425.36 $[M -H_2O+H]^+$	$C_{30}H_{50}O_2$ (442.7)	0.80	antitumor, antioxidant	(Marquez-Martin et al., 2006)
Valerenic acid	Wt, Δ*nocE*, *ermE*p**nocE*	217.159 $[M-H_2O+H]^+$	$C_{15}H_{22}O_2$ (234.339)	0.87	plant-associated metabolite	(Circosta et al., 2007)
Zerumbone	Wt, Δ*nocE*, *ermE*p**nocE*	161.133 $[M-C_3H_6O + H]^+$	$C_{15}H_{22}O$ (218.33)	0.96	antioxidant, anti-inflammatory	(Kalantari et al., 2017)
Holomycin	Wt, Δ*nocE*	214.994 $[M+H]^+$	$C_7H_6N_2O_2S_2$ (214.257)	0.77	antibiotic, antitumor	(Kenig & Reading, 1979)
Atractylenolide III	Δ*nocE*, *ermE*p**nocE*	231.138 $[M -H_2O +H]^+$	$C_{15}H_{20}O_3$ (248.32)	0.61	anticancer	(Liu et al., 2020)
Sophocarpine	Δ*nocE*, *ermE*p**nocE*	150.136 $[M -C_5H_7ON +H]^+$	$C_{15}H_{22}N_2O$ (246.354)	0.66	antiviral, anti-inflammatory	(Gao et al., 2012)
(-)-Indolactam V	Wt	274.191 $[M -CO+H]^+$	$C_{17}H_{23}N_3O_2$ (301.39)	0.89	protein kinase C activator	(Abe, 2018)
Dihydroartemisinin	*ermE*p**nocE*	267.159 $[M -H_2O+H]^+$	$C_{15}H_{24}O_5$ (284.35)	0.67	antiparasitic	(Arinaitwe et al., 2009)
Marrubiin	*ermE*p**nocE*	315.195 $[M -H_2O+H]^+$	$C_{20}H_{28}O_4$ (332.4)	0.72	plant-associated metabolite	(Popoola et al., 2013)
Theobromine	*ermE*p**nocE*	181.072 $[M+H]^+$	$C_7H_8N_4O_2$ (180.164)	0.84	plant-associated metabolite	(Ashihara et al., 2013)
Pinocembrin	Δ*nocE*	179.033 $[M-C_6H_6+H]^+$	$C_{15}H_{12}O_4$ (256.25)	0.68	plant-associated metabolite	(Lan et al., 2016)
Protocatechuic acid	Δ*nocE*	153.019 $[M-H]^-$	$C_7H_6O_4$ (154.12)	0.71	antioxidant, anticancer	(Liu et al., 2002)
Strobilactone A	Δ*nocE*	265.148	$C_{15}H_{22}O_4$	0.68	anticancer,	(Shiono et al.,

		$[M-H]^-$	(266.337)		antibiotic	2007)

[a] Cosine score: a scoring scheme that determines the similarity of two MS/MS spectra. The similar compounds have similar fragmentation patterns, which is computed as a cosine score from 1 (identical fragmentation spectra) to 0 (completely different spectra).

4.7. Supplementary Materials

4.7.1. Supplementary Figures

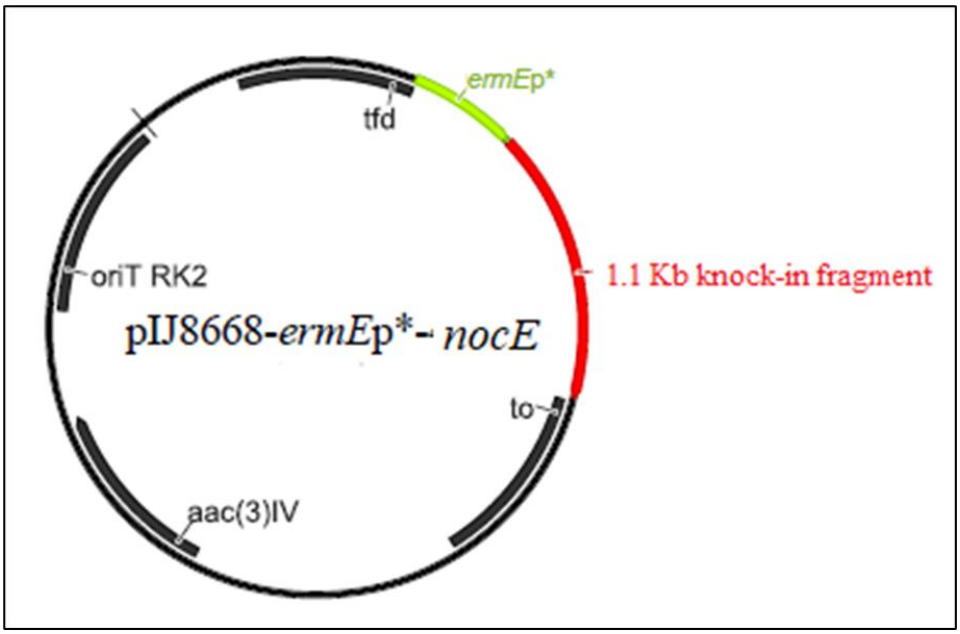


Figure S4.1: Plasmid map of pIJ8668-*ermE*p*-*nocE*. The 1.1 kb fragment of *nocE* (red) inserted downstream of *ermE*p*. *ermE*p*, mutated constitutive promoter from *ermE* (green); *tfd*, transcription terminator of phage fd; *to*, transcription terminator of phage λ; *ori*T RK2, origin of transfer from RK2; *aac*(3)*IV*, apramycin resistance cassette.

4.7.2. Supplementary tables

Table S4.1. The predicted protein products of the genes surrounding *nocE* (upstream and downstream) in CA producers species *S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus*, and CA-like BGC holding strain *Streptomyces* sp. M41. The similar proteins are in the same colors.

S. clavuligerus	***S. jumonjinensis***	***S. katsurahamanus***	***Streptomyces* sp M41**
Transcriptional regulator	Predicted L-lactate dehydrogenase	Predicted L-lactate dehydrogenase	Clip protease
Hypothetical protein	Hypothetical protein	Predicted L-lactate dehydrogenase	Bifunctional exonuclease DNA polymerase
Peptidase	Transcriptional regulator TrmB family/LuxR family	Hypothetical protein	AfsR/SARP regulator
Isochorismatase	Xaa-Pro amino peptidase	Transcriptional regulator TrmB family/LuxR family	Glycotransferase
Hypothetical protein	Hypothetical protein	Xaa-Pro amino peptidase	Hypothetical protein
NocE	**NocE**	**NocE**	**NocE**
Hypothetical protein	Hypothetical protein	Hypothetical protein	Hypothetical protein
Hypothetical protein	Hypothetical protein	ClpB protein	ATP-grasp domain (83% Sc-GcaS)
Hypothetical protein	ClpB protein	Hypothetical protein	GNAT-family N-acetyltransferase (72% Sc- Orf16)
Hypothetical protein	Hypothetical protein	Heat-shock protein GrpE	ABC transporter (80% Sc-Orf15)
Glucose dehydrogenase	DnaJ-class chaperon cbpA	Chaperon protein Dnak	GNAT-family (75% Sc-Orf14)
Hypothetical protein	Heat-shock protein GrpE		
	Chaperon protein Dnak		

Table S4.2. The diameters (in mm) of zones of growth inhibition of *K. pneumoniae* by CA-containing supernatants from triplicate cultures of Δ*nocE*, *ermE*p*-*nocE* and wt *S. clavuligerus* strains in two types of media SA and SM.

Strains	ZOI (mm) against *K. pneumoniae*			
	SA samples		SM samples	
	+PenG	-PenG	+PenG	-PenG
S. clavuligerus (wt)	31	0	29	0
Δ*nocE* (1)	30	0	30	0
Δ*nocE* (2)	30	0	28	0
Δ*nocE* (3)	30	0	28	0
*ermE*p*-*nocE* (1)	30	0	29	0
*ermE*p*-*nocE* (2)	30	0	29	0
*ermE*p*-*nocE* (3)	30	0	29	0
CA (+ ve control)	30	0	30	0

Table S4.3. The diameter (in mm) of zones of growth inhibition of *E. coli* ESS from cephamycin C-containing supernatants from triplicate cultures of Δ*nocE*, *ermE*p*-*nocE* and Wt *S. clavuligerus* strains in two types of media SM and TSB-S.

Strains	ZOI (mm) against *E. coli* ESS	
	SM samples	TSB-S samples
S. clavuligerus (wt)	20	19
Δ*nocE* (1)	21	21
Δ*nocE* (2)	22	23
Δ*nocE* (3)	22	22
*ermE*p*-*nocE* (1)	18	21
*ermE*p*-*nocE* (2)	18	18
*ermE*p*-*nocE* (3)	18	20

Table S4.4. The superclasses of molecules detected by in silico annotation in GNPS, for the extracts of wt *S. clavuligerus*, Δ*nocE*, and *ermE*p**nocE* strains.

Superclass	Number of Spectral Features							
	Wt	Δ*nocE*	*ermE*p* -*nocE*	Wt + Δ*nocE*	Wt + *ermE*p* -*nocE*	Δ*nocE* + *ermE*p*-*nocE*	Wt + Δ*nocE* + *ermE*p*-*nocE*	Total
Alkaloids and derivatives	3	1	2	0	1	1	4	12
Benzenoids	124	142	100	20	23	23	130	562
Homogeneous non-metal compounds	2	1	0	1	1	2	1	8
Hydrocarbons	0	0	0	0	0	1	3	4
Lipids and lipid-like molecules	323	381	304	127	177	155	820	2287
Mixed metal/non-metal compounds	0	0	0	0	0	0	1	1
Nucleosides, nucleotides, and analogues	5	1	3	0	1	2	6	18
Organic 1,3-dipolar compounds	0	0	0	0	0	0	1	1
Organic acids and derivatives	129	132	151	35	56	65	290	858
Organic nitrogen compounds	10	6	4	3	2	3	14	42
Organic oxygen compounds	14	20	31	4	10	17	38	134
Organohalogen compounds	0	0	1	0	1	0	4	6
Organoheterocyclic compounds	252	253	306	120	184	165	908	2188
Organometallic compounds	0	0	0	0	0	0	1	1
Organosulfur compounds	0	1	0	0	0	0	0	1
Phenylpropanoids and polyketides	65	46	62	24	26	28	122	373
no matches	1942	1798	1966	221	343	302	1093	7665
Total	2869	2782	2930	555	825	764	3436	14161

CHAPTER V

The production of a bioactive substance by *Streptomyces pratensis* ATCC 33331

5.1. Abstract

Streptomyces bacteria are prolific producers of specialized metabolites, which are used in various medicinal and agricultural applications. Genome sequencing studies have shown that these bacteria harbor a large reservoir of gene clusters, responsible for producing specialized metabolites. However, many of these clusters are silent or weakly expressed in *Streptomyces*; therefore, their potential to produce beneficial specialized metabolites remains unknown. Examination of *S. pratensis* ATCC33331 revealed that it contains 27 predicted biosynthesis gene clusters (BGCs), two of which have similarities to the clavulanic acid and carbapenem MM4550 BGCs in *Streptomyces clavuligerus* and *Streptomyces argenteolus,* respectively. Both metabolites are potent inhibitors of some β-lactamase enzymes; however, they have not been reported to be produced by *S. pratensis*. In this study, we reported that *S. pratensis* produces a bioactive substance (BS) when cultivated on soy (SM) or beef extract-starch (BES) agar media plates. The BS showed a bacteriostatic growth effect against *Klebsiella pneumoniae* when combined with penicillin G, suggesting that it has a β-lactamase inhibitory activity. A transcriptional study showed that Sp-*cpe* (*orf12*) in the CA-like BGC and the essential genes *carE, M, I,* and *P* in Carb4550-like BGC are not expressed. In addition, the heterologous expression of Sc-*cpe* and Sc-*orf14* in *S. pratensis* had no impact on the production of the BS. Also, the deletion of Sp-*cas2* from the CA-like BGC in *S. pratensis* did not affect the production of the BS. Global natural products social molecular networking (GNPS) analysis using the

untargeted LC-MS/MS data for *S. pratensis* culture extract detected neither CA nor Carb4550. Although the BS produced by *S. pratensis* showed a growth inhibitory activity when combined with penicillin G, our results demonstrated that it is different from CA or Carb4550, suggesting it is a novel compound requiring further characterization.

5.2. Introduction

The *Streptomyces* are widely distributed in natural environments and produce a diverse array of natural products. Over 13,000 of these are reported to have biological activity, and more than 100 microbial products are in use today as antibiotics, antitumor agents, and agrochemicals (Rigali et al., 2018; Singh & Pelaez, 2008). *Streptomyces pratensis* ATCC33331 (= IAF-45CD), which formerly was identified as *Streptomyces flavogriseus* strain IAF-45CD (= ATCC33331), was first isolated from compost in Laval, Canada (Ishaque & Kluepfel, 1980; Rong et al., 2013; Doroghazi & Buckley, 2014). *S. pratensis* has been reported to produce high levels of several extracellular enzymes, such as exoglucanase, proteases, cellulase, and xylanase (Ishaque & Kluepfel, 1980; MacKenzie et al., 1984). However, to date no production of specialized metabolites with antibacterial activities was reported in this species. Recently, the strain *S. pratensis* IIIM06, isolated from soils of the Himalayas, India, was found to produce actinomycin C1, C2, and C3, which exhibit potent antimicrobial activity against *Staphylococcus aureus* and *Mycobacterium tuberculosis* (Shah et al., 2017). Also, a marine strain *S. pratensis* NA-ZhouS1, isolated from sediment collected from the East China Sea, Zhoushan, produces two new angucycline antibiotics named stremycin A and B (Akhter et al., 2018). *S. pratensis* ATCC33331 is reported to possess a CA-like BGC with

similarity to the CA BGC of *S. clavuligerus*, but the production of CA has not been detected under the laboratory conditions tested (Álvarez-Álvarez et al. 2013). Transcriptional analysis showed that some of the late genes of the CA-like cluster are not expressed, indicating that the lack of CA production is due to the absence of some biosynthetic enzymes (Álvarez-Álvarez et al. 2013). Furthermore, a thienamycin-like (Thn-like) BGC, similar to the thienamycin BGC of *Streptomyces cattleya* NRRL8057, was also reported in the *S. pratensis* genome (Blanco, 2012). However, the production of Thn by *S. pratensis* has not been reported to date. Some *S. flavogriesus* strains such as MA4434 and MA4600 have been reported to produce epithienamycins (carbapenem related compounds), which showed *in vitro* antibacterial activity against a wide range of microorganisms (Stapley et al., 1981). The Thn-like BGC of *S. pratensis* is also highly similar to the BGC for producing another carbapenem metabolite called carbapenem MM4550 (Carb4550) in *Streptomyces argenteolus* (Li et al., 2014), and thus this cluster is herein referred to as the Carb4550-like BGC rather than the Thn-like BGC. Carb4550 is a member of the olivanic acids metabolites, which shows a broad spectrum of antibacterial activity against Gram-positive and Gram-negative bacteria (Brown et al., 1976; Butterworth et al., 1979). Carb4550 is also a potent inhibitor for several β-lactamases including those produced by *Staphylococcus aureus, Escherichia coli, Klebsiella, Citrobacter, Proteus* and *Pseudomonas* (Brown et al., 1976).

The availability of microbial genome sequences reveals the existence on average of ~30 BGCS of specialized metabolites in *Streptomyces* species; many of these are silent or poorly expressed (Xia et al., 2020). The awakening of these silent clusters is a challenge for researchers and requires an understanding of the molecular basis for the

lack of expression. To activate such gene clusters, some approaches have been developed for this purpose, and these have been extensively reviewed (Baltz, 2016; Onaka, 2017; Ren et al., 2017; Zhang et al., 2017; Kong et al., 2019; Xu and Wright, 2019). One of them is the one strain many compounds (OSMAC) approach, a simple and powerful tool that can activate many silent BGCs in single strains of bacteria (Zerikly & Challis, 2009). The OSMAC method includes changing medium composition and cultivation status, co-cultivation with other strains, adding enzyme inhibitors and metabolite precursors to allow production of more natural products (Bode et al., 2002; Pan et al., 2019).

5.3. Objectives

S. pratensis ATCC33331 possess 27 BGCs, as we will address in this chapter, but it has not been reported to produce any specialized metabolites. The expression of specialized metabolite BGCs is often conditional and dependent on culture conditions. Therefore, in this study, I first aimed to investigate the ability of *S. pratensis* ATCC33331 to produce a bioactive substance (BS) by following the OSMAC approach, through which various kinds of media were examined. Then I aimed to characterize the detected BS by testing different parameters. Also, I attempted to identify the genetic basis of production of the BS by genetic deletion and overexpression experiments. Finally, untargeted metabolomics analysis was conducted on *S. pratensis* ATCC33331 extracts from SM cultures using GNPS and MolNetEnhancer. My goal was to identify the classes of molecules/compounds produced by *S. pratensis* and to investigate the bioactive metabolites produced by this strain.

5.4. Results and discussion

5.4.1. Testing *S. pratensis* for the production of bioactive substances using different types of media

To investigate the ability of *S. pratensis* ATCC33331 to produce specialized metabolites, the OSMAC approach was followed by culturing *S. pratensis* in different types of liquid broths and solid media. Seeding cultures of *S. pratensis* were started from glycerol stocks in 5 ml of TSB and were used to inoculate different kinds of broth media, MEY, MS, R5A, SA, SM, TBO, and TSB (see 'Materials and Methods" for details). The supernatant samples were collected at 48, 96, and 120-h time points. Since *S. pratensis* ATCC33331 was reported to have CA-like and Carb4550-like BGCs, the supernatant samples were tested for the presence of β-lactamase inhibitors using disc-diffusion bioassays against *K. pnuemoniae* (*Kp*) (β-lactam resistant) with and without PenG (60 μg/ml). Also, the samples were tested against *E. coli* ESS (β-lactam sensitive strain) to detect the presence of β-lactam compounds. Unfortunately, no bioactivities were detected, neither against *K. pnuemoniae* nor *E. coli* ESS, indicating that *S. pratensis* does not produce any β-lactamase inhibitor or β-lactam metabolites when grown in the tested broth media. These results are consistent with those from Álvarez-Álvarez et al. (2013), who tested for the production of CA by fermenting *S. pratensis* in nine different liquid media without any success, and suggests that the CA-like BGC is not active in the tested conditions.

Next, I tried to investigate the production of specialized metabolites by culturing *S. pratensis* on different solid media plates. The bacteria were streaked onto SM, beef extract-starch (BES), ISP-4, SA, and TSA agar plates. Agar plug bioassays were

conducted against *K. pnuemoniae* (with and without PenG) and *E. coli* ESS to test for the production of β-lactamase inhibitor and/or β-lactam antibiotic compounds. Interestingly, zones of growth inhibition for *K. pnuemoniae* (with PenG) were detected around *S. pratensis* plugs taken from SM and BES plates but not from the other kinds of media. The results indicate that a BS is produced by *S. pratensis* when grown on SM and BES solid media. The activity of this BS appeared when combined with PenG, suggesting it has a β-lactamase inhibitory activity. No growth inhibitory activities were detected for the agar plugs against *E. coli* ESS, indicating that it is not sensitive to the BS produced by *S. pratensis*.

For further confirmation, agar-plot diffusion bioassays were performed against *K. pneumoniae* using cellophane membranes (Figure 5.1A) as described in Section 2.6.3 of the "Materials and Methods." The cellophane membranes, which allow the metabolites to diffuse through but not the bacterial cells, were used to exclude any possible physical contact between *S. pratensis* and *K. pneumoniae*. The agar plots were taken from *S. pratensis* cultures on SM, BES, ISP-4, and TSA plates and placed onto cellophane membranes on the top of TSA bioassay plates, and incubated for 48 h. A layer of TSA (0.8% agar) mixed with *K. pneumoniae* and PenG was poured after removing the agar-plot and the cellophane (Figure 5.1A). As controls, bioassay plates were prepared without adding PenG. As shown in Figure 5.1B, growth inhibition zones were detected in the spots of *S. pratensis* agar-plots taken from SM and BES plates but not from ISP-4 and TSA cultures, indicating that *S. pratensis* produced the BS when grown on SM and BES media. The BS inhibited *K. pnuemoniae* when combined with PenG antibiotic, while no

growth inhibition was detected in the plates without PenG (Figure 5.1B). For comparison, disc diffusion bioassays were carried out at the same time for *S. clavuligerus* (producer of CA) supernatant sample (Sc), CA solution, and *S. cattleya* (producer of thienamycin) supernatant sample (Scat) (Figure 5.1A). Zones of growth inhibition were noticed for Sc and CA samples but not for the Scat sample (Figure 5.1B). The results suggest that the BS produced by *S. pratensis* inhibited the β-lactamase enzymes in *K. pneumoniae* and render it sensitive to penicillin G.

The growth cycle of *S. pratensis* on SM and BES solid media takes around 11 – 14 days. To investigate the optimum time for the maximum production of this BS, agar plug bioassays were conducted over a time course using cultures of *S. pratensis* on SM and BES media. In SM, *S. pratensis* started showing bioactivity against *K. pnuemoniae* on day four of the culture and reached the maximum production (the largest diameter of growth inhibition) on day seven (Figure 5.1C). After that, the output of the BS started to decrease until it stopped on day 11 of the culture (Figure 5.1C). On BES media, the production of the BS occurred during a short period. The zones of growth inhibition were observed in days 6 – 8 of the culture and were smaller in size compared to those obtained using the SM plugs (Figure 5.1C). Based on these results, all the subsequent agar-plug/plot bioassays were conducted using day seven *S. pratensis* cultures.

5.4.2. Testing the activity range of the bioactive substance

To characterize the BS produced by *S. pratensis*, we tried to investigate the activity range of this BS by testing different parameters. Agar plot bioassays using *S. pratensis* SM cultures were conducted against *Enterobacter cloacae* (β-lactam resistant),

E. coli/pGEMT-amp^{R} (ampicillin-resistant), and *E. coli* ESS (β-lactam sensitive) and were also examined using penicillin G (60 μg/ml), ampicillin (100 μg/ml), or without antibiotics (Figure 5.2). Disc diffusion assays using Sc, CA, and Scat were combined with the bioassays for comparison. As shown in Figure 5.2, the BS produced by *S. pratensis* did not show any activity against *E. cloacae, E. coli*/pGEMT-amp^{R}, or *E. coli* ESS. Also, the BS did not show any activity when Amp antibiotic was used instead of PenG. *K. pnuemoniae* has a class A serine β-lactamase (TEM-1 and SHV-1) (Drawz & Bonomo, 2010; Tooke et al., 2019), but *E. cloacae* were reported to have class C β-lactamase (Joris et al., 1985; Parveen et al., 2010), suggesting that the BS of *S. pratensis* is effective against class A β-lactamase rather than class C β-lactamase. The use of Amp instead of PenG did not grant any advantage to the BS, which suggests that this BS works in synergy with PenG. The BS did not show any inhibitory activity against *E. coli* ESS, while the zone of growth inhibition was noticed around the Scat sample (producer of thienamycin) (Figure 5.2), which indicates that the *E. coli* ESS is not susceptible to this BS, and it is unlikely to be thienamycin. The Carb4550 produced by *S. argenteolus* was reported to have bioactivity against both *E. coli ESS* and *K. pneumoniae* (Li et al., 2014); however, the activity noticed for the BS was only against *K. pneumoniae* but not *E. coli* ESS (Figure 5.1B and 5.2). Moreover, Carb4550 combined with ampicillin (5 μg/ml) inhibits the growth of some β-lactamase-producing bacteria (Brown et al., 1976), an activity that was not observed for the BS even with an ampicillin concentration of 100 μg/ml (Figure 5.2). These results suggest that the BS produced by *S. pratensis* is not a Carb4550 metabolite. Altogether, *S. pratensis* produces the BS when cultured on SM and

BES solid media. The BS has a growth inhibitory effect against *K. pneumoniae* when combined with PenG only, and it exhibits no inhibitory effects against *E. cloacae* and *E. coli* ESS. The behaviour of this BS suggests that it is not thienamycin, but likely a β-lactamase inhibitor.

5.4.3. Examination for the interspecies interactions to produce the bioactive substance.

Streptomyces co-exist with hundreds or thousands of different microorganisms in their natural environment, allowing them to interact directly or indirectly by exchanging metabolites with other microbial community members (Curtis et al., 2002). These interactions significantly impact growth and specialized metabolite production (van Wezel & McDowall, 2011). *S. pratensis* did not show any production of the BS when cultivated in SM broth. The bioactivity of this substance appeared when agar-plug/plot bioassays were performed, in which the *S. pratensis* agar-plug/plots were in contact with the TSA containing *K. pneumoniae*, suggesting that direct physical contact between the two species might awaken a silent BGC in *S. pratensis* and start producing this BS. We used cellophane membranes in all agar-plot bioassays (as shown in Figure 5.1). The membranes are penetrable by metabolites but not by bacterial cells, excluding the probability of direct physical contact between the two species of bacteria. To further confirm that the direct contact between *S. pratensis* and *K. pnuemoniae* is not required for the production of the BS, a SM plate seven-day culture of *S. pratensis* was cut into two halves and placed in an upside-down position in a large sterile plate (Figure 5.3). PenG

was then spread over one-half of the two (See Materials and methods Section 2.8.3), and *K. pnuemoniae* was inoculated at three spots on each (+PenG and -PenG) SM half (Figure 5.3A). As a control, the same procedure was carried out for a fresh SM plate with no *S. pratensis*. After overnight incubation, no growth of *K. pnuemoniae* was noticed on the SM half with *S. pratensis* and PenG (Figure 5.3B), while *K. pnuemoniae* grew well on the half without PenG (-PenG). On the control plate (no *S. pratensis*), *K. pneumoniae* grew well on both halves with and without PenG (Figure 5.3B). The results demonstrate that direct interaction between *S. pratensis* and *K. pnuemoniae* is not required for the production of the BS.

5.4.4. The bacteri(cidal/ostatic) effect of the bioactive substance

To characterize the BS produced by *S. pratensis*, we tried to investigate the potency of the BS against *K. pnuemoniae*. Agar-plot bioassays were conducted for *S. pratensis* cultures against *K. pnuemoniae,* combined with discs diffusion bioassays for *S. clavuligerus* supernatant and CA-solution as controls (Figure 5.4A). After incubation, agar plugs from the zones of growth inhibition produced by the *S. pratensis* plot (Sp-plug), *S. clavuligerus* supernatant (Sc-plug), and CA solution (CA-plug) were removed and placed onto a TSA plate (Figure 5.4A). Plugs from the zone of *K. pneumoniae* growth (Kp-plug) were used as controls (Figure 5.4A). The plate was incubated for five days at 37 °C to monitor the re-growth of *K. pneumoniae*. As shown in Figure 5.4A, *K. pnuemoniae* started to regrow on the Sp-plug after 48 h of incubation. In comparison, no re-growth of *K. pnuemoniae* was detected on Sc-plug and CA-plug even after 172 h of incubation. The results indicate that the BS produced by *S. pratensis* has a bacteriostatic

effect that inhibited the growth of *K. pnuemoniae* for not more than 72 h before the bacteria regrew. In comparison, CA demonstrates bactericidal effects as shown by the Sc-plug and CA-plug, where *K. pneumoniae* did not grow after five days of incubation (Figure 5.4A). To ensure that the re-growth of *K. pnuemoniae* was not a result of a mutation in the genome of *K. pnuemoniae* that made it resistant to the BS, a bacterial suspension from the regrown *K. pnuemoniae* was prepared and used to perform a subsequent agar-plot bioassay to test its susceptibility to the BS, Sc supernatant and CA solution. As shown in Figure 5.4B, the zones of growth inhibition were again detected around the Sp-plot, Sc supernatant, and CA solution, indicating that *K. pnuemoniae* is still sensitive to the BS and CA. Thus, the *K. pnuemoniae* re-growth on the Sp-plug (Figure 5.4A) was due to the bacteriostatic effects of the BS.

To compare the activity of the BS with that of known bactericidal and bacteriostatic antibiotics, disc diffusion assays were performed against *K. pnuemoniae* using antibiotic discs for two bactericidal antibiotics, Gentamycin (GM) and Streptomycin (S), and two bacteriostatic antibiotics, Tetracycline (T) and Chloramphenicol (C). After overnight incubation, agar plugs were taken from the halo zone of inhibition and placed on a fresh plate. The plugs were incubated for five days to monitor the re-growth of *K. pnuemoniae*. As shown in Figure 5.5, *K. pnuemoniae* regrew on the plugs for bacteriostatic antibiotics tetracycline and chloramphenicol after 48 h of incubation but did not regrow on the plugs for bactericidal antibiotics gentamycin and streptomycin. The results obtained here for the bacteriostatic antibiotics are similar to those for the BS of *S. pratensis,* which supports our conclusion that this BS, when

combined with PenG, exerts its growth inhibitory effect on *K. pnuemoniae* as a bacteriostatic, not bactericidal, compound.

5.4.5. Genomic analysis of specialized metabolite BGCs in *S. pratensis*

To examine the *S. pratensis* genome for specialized metabolite BGCs, we employed the genome mining tool antiSMASH (antibiotics and Secondary Metabolite Analysis Shell) 4.0, a pipeline capable of identifying biosynthetic loci covering a whole range of known secondary metabolite compound classes (Blin et al., 2017). AntiSMASH predicted 27 putative specialized metabolite BGCs in the *S. pratensis* genome (Table 5.1), including five terpenes, four bacteriocins, and three BGCs each for polyketides (PKs), nonribosomal peptides (NRPs), and hybrid clusters. Two BGCs each for butyrolactone, siderophore, ectoine metabolites, and one BGC each for β-lactam, lantipeptide, and melanine were also detected (Table 5.1). While eight of the predicted BGCs displayed high levels of similarity (≥70%) to BGCs in the MIBiG database, two of them displayed moderate similarity (30 - 70%), and ten showed low similarity (<30%), the remaining seven clusters did not match any reference in the MIBiG database (Table 5.1). The combined length of the predicted BGCs is ~ 924 kb, accounting for ~12.6% of the *S. pratensis* genome (~7.3 Mbp). To date, none of the metabolites of the BGCs predicted in Table 5.1 has been reported to be produced by *S. pratensis*.

The CA-like and the Carb4550-like (= Thn-like) BGCs are located in region 5 and region 1 of the *S. pratensis* genome, respectively (Table 5.1), and they are the only predicted BGCs with β-lactamase inhibitory activity. The CA-like BGC of *S. pratensis* is highly similar to the CA BGC of *S. clavuligerus* (Figure 5.6A) (Jensen 2012; Álvarez-

Álvarez et al., 2013). The conserved genes of the CA-like BGC are organized in groups like those of *S. clavuligerus* but assembled in a different organization (Figure 5.6A). Chapter four of this thesis discussed some of the similarities and differences between the CA-like BGC of *S. pratensis* and the CA BGC of *S. clavuligerus*. Furthermore, a transcriptional study showed that the "early genes" of the CA-like cluster and the regulatory genes *ccaR* and *claR* are transcribed when *S. pratensis* is cultivated in broth media, whereas the "late genes" are not expressed under the same conditions (Alvarez-Alvarez et al., 2013), suggesting that the lack of CA production in *S. pratensis* is due to the lack of essential enzymes required for the late steps of CA biosynthesis (Alvarez-Alvarez et al., 2013).

The Carb4550-like BGC shows 65% and 50% similarities to the carbapenem MM4550 BGC of *S. argenteolus* and thienamycin BGC of *S. cattleya*, respectively (Table 5.1), due to which in this study it is called Carb4550-like BGC. Both Carb4550 and Thn (Figure 5.6C) belong to the carbapenem group of the β-lactam metabolites (see section 1.2.1 in the Introduction chapter). A genetic comparative analysis was previously described for the Thn BGC of *S. cattleya* and the Carb4550-like BGC of *S. pratensis* (formerly described as a thienamycin-like cluster in *S. flavogriseus*) (Blanco, 2012), and also for the Thn cluster of *S. cattleya* with Carb4550 cluster of *S. argenteolus* (Li et al. 2014) (Figure 5.6B). The Carb4550-like BGC of *S. pratensis* contains the *carI* gene (encoding regulator protein) and 17 genes (*carE – T* and *carSU*) that are homologous to those in the Carb4550 cluster of *S. argenteolus* (Figure 5.6B) and organized in the same order and orientation. However, *cmm22* and *cmm23* encoding a two-component system in

S. argenteolus (Li et al., 2014), are absent in the Carb4550-like cluster of *S. pratensis* (Figure 5.6B).

5.4.6. Transcriptional analysis of the CA-like and Carb4550-like gene clusters.

To determine if the CA-like and Carb4550-like BGCs are active in *S. pratensis*, we set about examining the expression of the essential genes for both clusters using RT-PCR. RNA was isolated from a 7-day culture of *S. pratensis* on SM agar. The plate was cut into two halves and used to perform bioassays against *K. pneumoniae* with and without PenG (Figure 5.7A and B) (see Section 2.8.6. in the Materials and Methods). After confirming the production of the BS by detecting the zone of inhibition around the *S. pratensis* culture (Figure 5.7 B), the spores/mycelia were immediately collected to isolate RNA and generate cDNA by reverse transcription (RT). An RNA sample was also isolated from a wt *S. pratensis* culture that was not in contact with *K. pneumoniae* or PenG. PCR was then carried out using specific primer pairs (Supplementary Table S2.1) for the genes of interest. As shown in Figure 5.7C, the expression of Sp-*cas2* and Sp-*ceaS2* of the CA-like BGC was found in all RNA samples. The two homologous genes *cas2* and *ceaS2* in *S. clavuligerus* are essential for the production of CA and are involved in the early stages of CA biosynthesis. The expression of Sp-*cpe*, one of the late steps genes, was not detected (Figure 5.7C), suggesting that the late stage of CA biosynthesis was blocked in *S. pratensis*. The results agreed with Alvarez-Alvarez et al. (2013), who showed that some of the early steps' genes (*ceaS2, oat2* and *oppA1*) are expressed when *S. pratensis* grows in broth media, while the late genes (*cyp, orf12, orf13, orf14*, and *oppA2*) are silent. In addition, our bioinformatic search for a paralogue of the Sp-*cpe* gene

did not find any match in the *S. pratensis* genome that could compensate the function of *cpe*. Taken together, these results suggest that the BS produced by *S. pratensis* is not CA.

For the Carb4550-like BGC, we examined the expression of Sp-*carE, carM, carP,* and *carI* (Figure 5.7D). Sp-*carE* and Sp-*carM* are genes homologoue to *cmmE* and *cmmM* in *S. argenteolus* and *thnE* and *thnM* in *S. cattleya,* respectively. These genes catalyze the first two steps of the carbapenem biosynthetic pathways to make the carbapenem bicyclic core (Sleeman and Schofield, 2004; Li et al. 2014). However, disruption of *cmmE* and *cmmM* eliminated the production of Carb4550 in *S. argenteolus* (Li et al., 2014). Sp-*carP* is a homologue of *cmmP* and *thnP* in *S. argenteolus* and *S. cattleya,* respectively. The deletion of *cmmP* affected the production of Carb4550 in *S. argenteolus* (Li et al., 2014). *carI* is similar to *cmmI* in *S. argenteolus*, which encodes LysR-family transcription regulator that positively controls the expression of Carb4550 gene cluster. A transcription analysis showed that the genes Sp-*carE, carM, carP,* and *carI* are not expressed in *S. pratensis,* indicating that the gene cluster is silent, and that the BS produced by *S. pratensis* is not Carb4550 (Figure 5.7D).

5.4.7. Heterologous expression of *S. clavuligerus cpe* (*orf12*) and *orf14* in *S. pratensis*.

Since the *cpe* gene is not expressed, and the transcription of *orf14* is reported to be undetectable in *S. pratensis* (Alvarez-Alvarez et al., 2013), we decided to study the heterologous expression of *S. clavuligerus cpe* and *orf14* in *S. pratensis*. The constructs pSET152/Sc-*cpe* and pHM11a/Sc-*orf14* were prepared by cloning the *cpe* and *orf14* from *S. clavuligerus* into pSET152 (Apr^R) and pHM11a (Hyg^R) plasmids, respectively, (Section 2.4.5. in the Materials and Methods). The constructs were introduced by

conjugation into *S. pratensis* separately or together to give *S. pratensis*/Sc-*cpe*, *S. pratensis*/Sc-*orf14* and *S. pratensis*/Sc-*cpe*-*orf14*. The heterologous expression and wt strains were fermented in seven broth media (MEY, MS, R5A, SA, SM, TBO, and TSB) and were tested by disc diffusion bioassays for the production of the BS, which was not detected in any of the samples.

The strains *S. pratensis*/Sc-*cpe*, *S. pratensis*/Sc-*orf14*, *S. pratensis*/Sc-*cpe*-*orf14* and the wt *S. pratensis* were also streaked onto solid SM and incubated for seven days before examining them for the production of the BS by agar plug bioassays. The zones of growth inhibition were detected around the agar plugs with no significant differences between the wt and the other heterologous expression strains (Figure 5.8), which indicates that the heterologous expression of Sc-*cpe* and Sc-*orf14* does not have any positive effect on the production of the BS. The results suggest that the BS is not CA, or, in case it is a low level of CA, the heterologous expression of *cpe* and *orf14* were not enough to improve the production since other genes in the late steps (*cyp*, *orf13*, and *oppA2*) are weakly expressed or undetectable (Alvarez-Alvarez et al., 2013).

5.4.8. The effect of Sp-*cas2* insertional inactivation on the production of the bioactive substance.

To further narrow down the BS being secreted by *S. pratensis*, an insertional inactivation approach was conducted to disrupt the Sp-*cas2* and *carE* genes in CA-like and Carb4550-like BGCs, respectively (Figure 5.9 and supplementary Figure S5.1). Two regions in the gene Sp-*cas2* were amplified and cloned in pIJ773 plasmid to construct pIJ773/Sp-*cas2*-KO-1 and pIJ773/Sp-*cas2*-KO-2, respectively (Figure 5.9 and Figure

5.10A; see Section 2.6.4 in the Materials and Methods). The two constructs were confirmed by sequencing (Figure 5.9D) and moved separately into wt *S. pratensis* by conjugation to achieve single crossover insertional inactivation in the Sp-*cas2* gene (Figure 5.10A). Three exconjugants (samples 1 – 3 in Figure 5.10) for each *S. pratensis*/pIJ773-*cas2*-KO-1 (Sp-Δ*cas2*-1) and *S. pratensis*/pIJ773-*cas2*-KO-2 (Sp-Δ*cas2*-2) were streaked on SM plates and incubated for seven days. Agar plug bioassays were carried out to examine the production of the BS, and PCR was conducted simultaneously to confirm the disruption of the target gene by the insertion of our constructs. As shown in Figure 5.10B, the *cas2* disrupted *S. pratensis* strains Sp-Δ*cas2*-1 and Sp-Δ*cas2*-2 inhibited the growth of *K. pneumoniae* around the plugs, indicating that the strains were still producing the BS and the insertion inactivation in Sp-*cas2* did not affect the production of this BS. The PCR results showed (Figure 5.10C) Sp-*cas2* was still disrupted in all samples of both mutants Sp-Δ*cas2*-1 and Sp-Δ*cas2*-2, as the pIJ773/Sp-*cas2*-KO-1 and pIJ773/Sp-*cas2*-KO-2, respectively, were still integrated. In addition, *S. pratensis* does not have a paralogue gene for Sp-*cas2* to compensate for the loss of Sp-*cas2*, in contrast to *S. clavuligerus,* which has *cas1* and *cas2* paralogues. Therefore, these results suggest that the CA-like BGC is not responsible for producing the BS by *S. pratensis*. To study the effect of the disruption of the *carE* gene in Carb4550-like BGC on the production of the BS, we followed the same single crossover insertional inactivation approach to knock out the gene (Supplementary Figure S5.1). Unfortunately, despite many trial times, I was unsuccessful in obtaining a *carE* disrupted mutant of *S. pratensis*.

5.4.9. Metabolomics analysis of *S. pratensis* (GNPS molecular networking)

To characterize the specialized metabolites produced by *S. pratensis*, the bacterium was cultured on SM plates for seven days, and the agar was extracted with methanol, ethyl acetate, or phosphate buffer saline (PBS). The extracts were analyzed by untargeted LC-MS/MS in both positive and negative ionization mode to detect as many compounds as possible. The LC-MS/MS analysis was performed by Dr. Kapil Tahlan in the Dr. Dorrestein lab at the UCSD. We set about identifying the classes of molecules found in the extracts of *S. pratensis*; the MS/MS spectral data were analyzed and annotated by the Network annotation propagation (NAP) in GNPS before being further analyzed using the MolNetEnhancer to identify the classes of the molecules/compounds found in our samples. The analysis results were visualized and interpreted as colored networks in Cytoscape 3.7 (Figure 5.11). More than 3000 spectral features of molecules were detected in the analysis, and ~50% of them were found to belong to 13 superclasses of compounds listed in Figure 5.11. The most abundant groups were the "organoheterocyclic compounds" with a total of 506 detected spectral features, followed by "organic acids and derivatives" and "lipids and lipid-like molecules" with 404 and 379 detected spectral features, respectively (Figure 5.11B). The detected spectral features for the remaining compounds range between 1 – 142 (Figure 5.11B).

To identify the specialized metabolites produced by *S. pratensis* growing on SM, the MS/MS data obtained from both positive and negative ionization modes were used to build a molecular network. Specialized metabolites were annotated by matching spectra against public libraries in GNPS and visualized in Cytoscape 3.7 (Figure 5.12). The analysis showed >6000 molecular nodes in the GNPS analysis (Figure 5.12), but ~4%

could be annotated by matching spectra with available libraries. Among the 18 predicted BGCs (Table 5.1), only the metabolites for two of them, ectoine and desferrioxamine, were detected in the GNPS network. The ions corresponding to desferrioxamine D (desmethylenylnocardamine, $[M + NH_4]^+$, *m/z* 604.7) and ectoine ($[M + H]^+$, *m/z* 143.082) (Figure 5.12) were found in the positive ionization mode with 0.8 and 0.77 cosine values, respectively. Since desferrioxamine and ectoine are produced by many Actinomycetes and are involved in general cellular growth/survival processes (Challis, 2005; Czech et al., 2018), finding them was expected. However, none of the CA, carbapenem MM4550, or thienamycin metabolites was detected in the network, suggesting that the BS is not any one of them, and the bioactivity shown by *S. pratensis* is due to other unknown metabolites.

5.5. Conclusion

Variation in the growing environment can have significant impacts on the quantity and diversity of bacterial specialized metabolites. Here, the OSMAC approach was followed to examine the effect of different media types, including seven liquid and five solid media, on specialized metabolite production in *S. pratensis*, which was found to produce BS when grown on SM and BES solid media. This BS has a bacteriostatic effect on *K. pneumoniae* but does not affect *E. cloacae* or *E. coli* ESS growth. The various bioassays conducted in this study indicated this BS is not thienamycin. It is instead a β-lactamase inhibitor or another metabolite with antimicrobial activity that works in synergy with the PenG to inhibit the growth of *K. pneumoniae*. Carb4550 and CA are the two candidates that have a β-lactamase inhibitory activity. However, the Carb4550

compound is reported to have growth inhibitory activity against both *E. coli ESS* and *K. pneumoniae* (Li et al., 2014); whereas the activity observed for the BS was only against *K. pneumoniae* but not *E. coli* ESS (Figure 5.1B and 5.2). In addition, Carb4550 is known to have bioactivity against a wide range of β-lactamase-producing bacteria when combined with ampicillin (5 μg/ml) (Brown et al., 1976). This action was not found for the BS produced by *S. pratensis* (Figure 5.2).

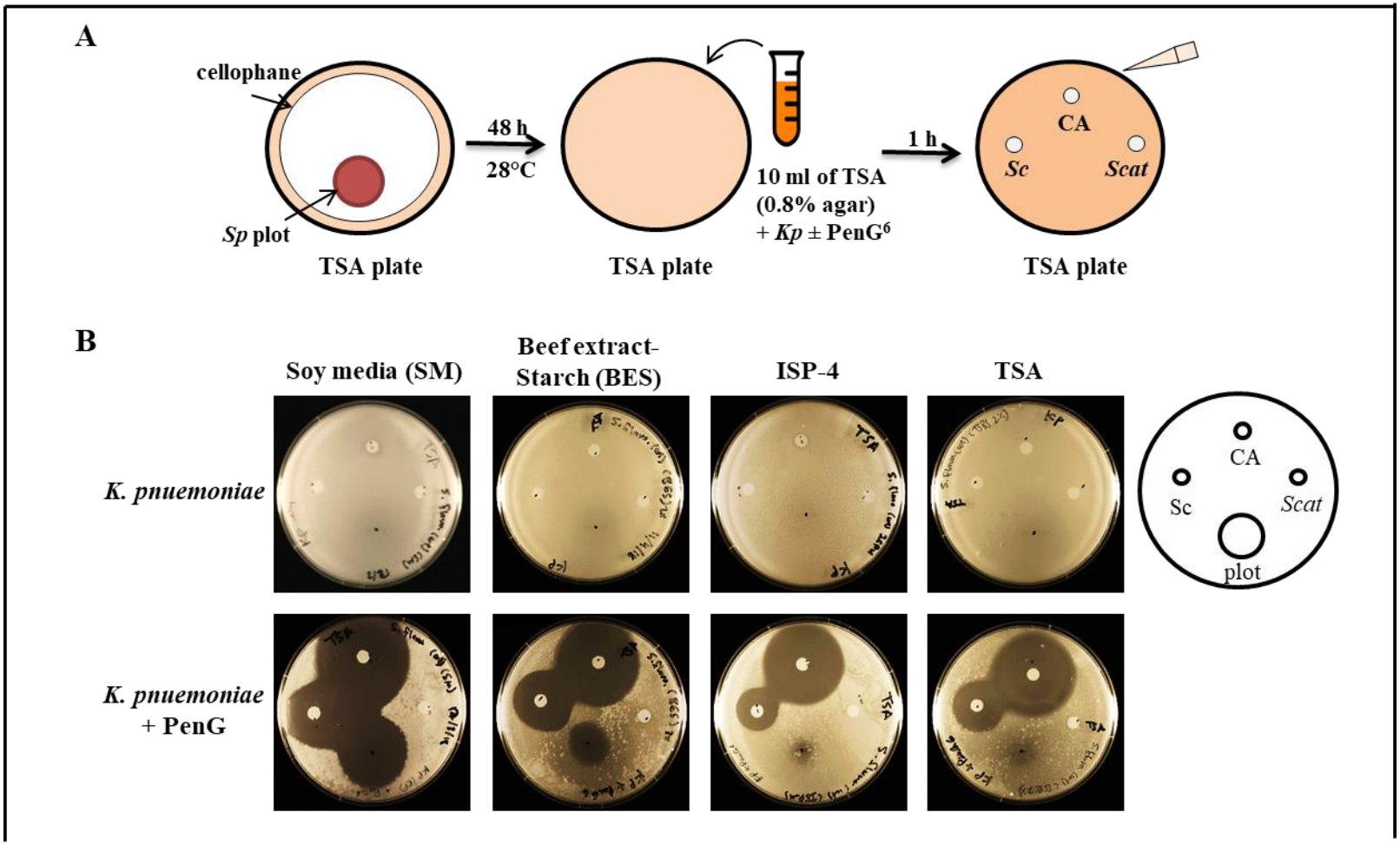
Further, the transcriptional analysis conducted in this study showed that the genes required for Carb4550 biosynthesis are not expressed. Having these results together leads to the conclusion that the BS produced by *S. pratensis* is not Carb4550. In addition, the ineffective deletion of Sp-*cas2* in CA-like BGC and the silence of the genes in the late steps of CA biosynthesis in *S. pratensis,* along with the non-detection of the CA in the GNPS network, support the conclusion that this BS is not CA. Therefore, I propose that the BS produced by *S. pratensis* may be a novel β-lactamase inhibitor or may be a metabolite that has antimicrobial activity and works synergistically with penicillin G.

It was previously reported that some antibiotics can work synergistically with penicillin and exert an inhibitory effect against β-lactam resistant bacteria. For instance, penicillin BRL1437 with cloxacillin was effective against both the plasmid-mediated TEM β-lactamase and the chromosomally mediated β-lactamase produced by *Klebsiella* species. Moreover, methicillin/nafcillin with isoxazolyl penicillin show a β-lactamase inhibition activity against some β-lactam resistant bacteria (Sutherland and Batchelor, 1964; Rolinson, 1991). In the genome of *S. pratensis*, one of the predicted BGCs has a high similarity (~88%) to the sceliphrolactam BGC of *Streptomyces* sp. SD85 (Table 5.1). Sceliphrolactam is a polyene macrocyclic lactam with antifungal biological activity

that inhibits the growth of amphotericin B-resistant *Candida albicans* (Oh et al., 2011). The same BGC also showed 61% similarity to the macrotermycins BGC of *Amycolatopsis* sp. M39 (Table 5.1). The product of this BGC (macrotermycin) demonstrated antibacterial activity against human-pathogenic *Staphylococcus aureus* (Beemelmanns et al., 2017). Neither of the metabolites was reported to be produced by *S. pratensis*. Whether they have a synergic effect with penicillin G, in case they are produced, on the growth of *K. pneumoniae* is a question that requires further study and investigation.

5.6. Figures and tables

5.6.1. Figures



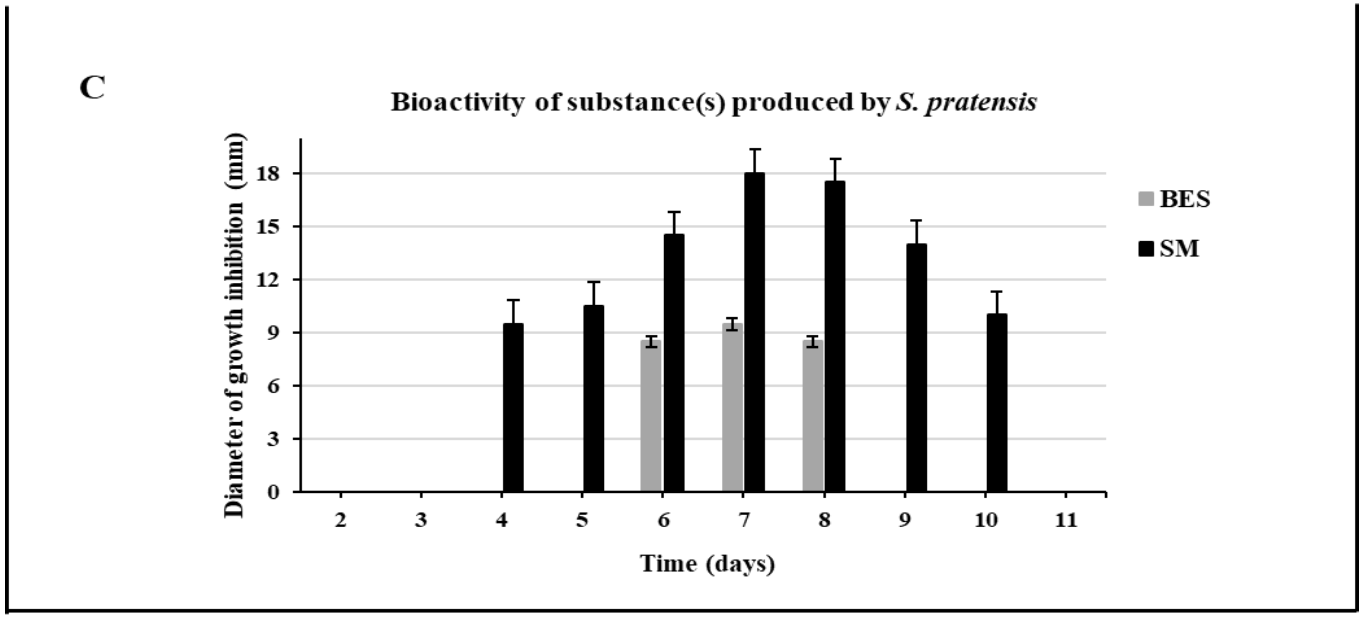


Figure 5.1. The production of bioactive substance by *S. pratensis* grown on different types of media. (A) Agar plots from *S. pratensis* culture were placed over the cellophane membrane on a TSA plate for 48 h. After removing the cellophane and the agar plots, 10 ml of TSA (0.8% agar) containing *K. pneumoniae* with or without $PenG^{60}$ were poured. After solidifying the agar, disc diffusion assays were conducted for supernatants from *S. clavuligerus* and *S. cattleya* cultures and CA solution. **(B)** Agar plots bioassay results for seven days *S. pratensis* culture on different types of media, soy (SM), beef extract-starch (BES), ISP-4, and TSA-S, with the bioassays performed on TSA plates with or without PenG against *K. pneumoniae* as indicator microorganism. **(C)** Column chart showing the production of the bioactive substance(s) by *S. pratensis* cultures on SM and BSE media for 11 days time course. Sc, *S. clavuligerus* supernatant; CA, clavulanic acid solution; Scat, *S. cattleya* supernatant; Kp, *K. pneumoniae*; SM, soy media.

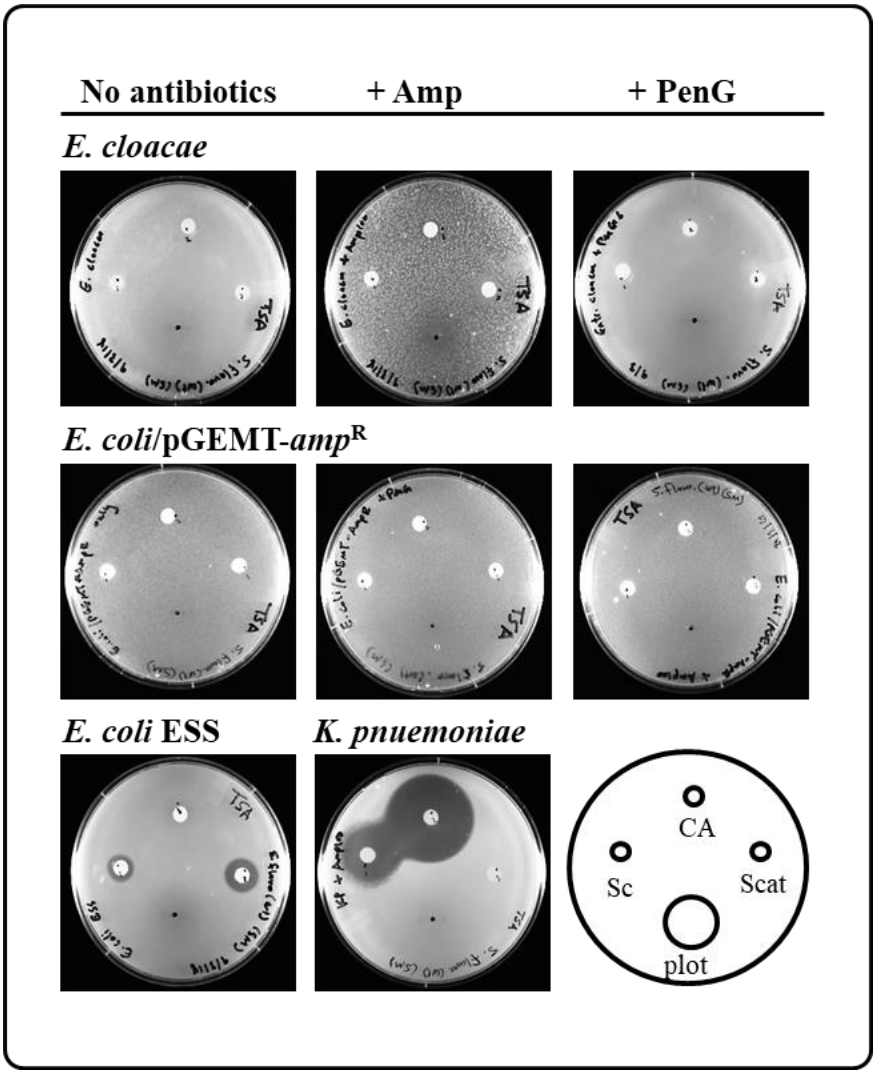


Figure 5.2. The activity range of *S. pratensis* bioactive substance. Agar plot bioassays were conducted against different indicator microorganisms; *E. cloacae* (β-lactam resistant), *E. coli*/pGEMT-amp^{R}, and *E. coli* ESS (β-lactam sensitive). The bioactivity was also tested with Amp^{100} or PenG^{60} to added to the TSA plates or without adding antibiotics. *S. clavuligerus* supernatant (Sc), clavulanic acid solution (CA), and *S. cattleya* supernatant (Scat) were used as controls.

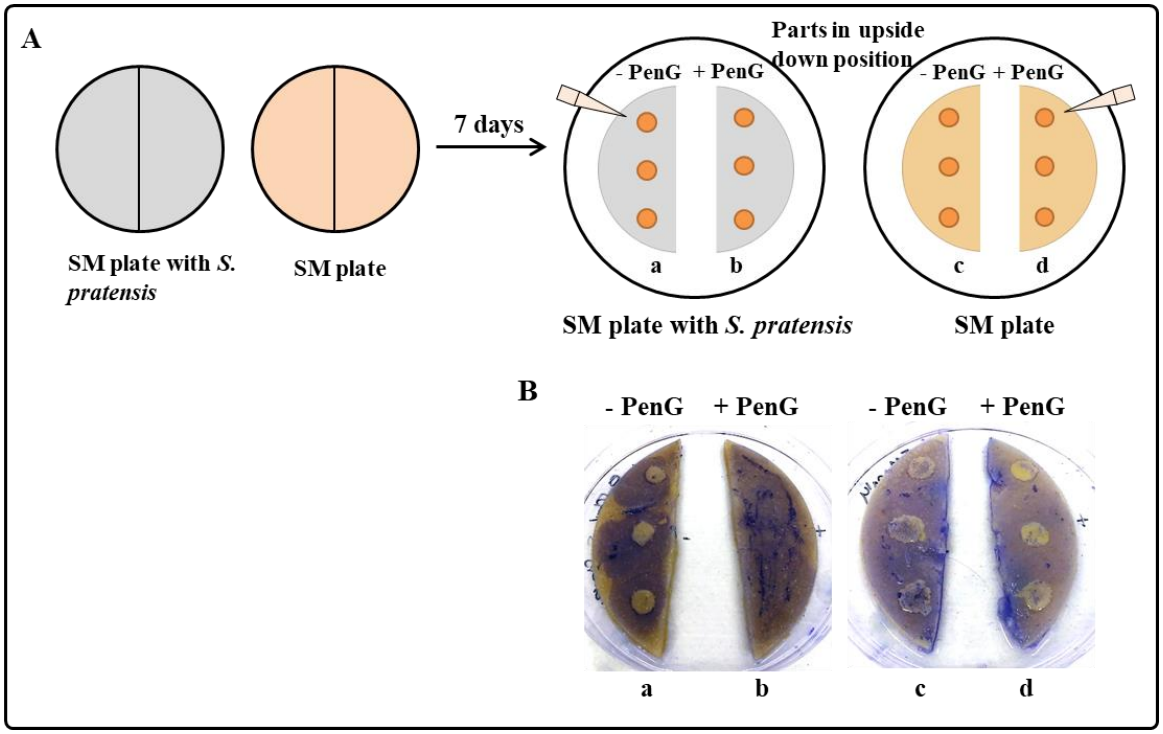


Figure 5.3. Physical interaction with *K. pneumoniae* is not necessary for bioactive substance production by *S. pratensis.* (A) The schematic diagram for seven days growth of *S. pratensis* on SM plate (left plate), the medium was cut into two halves and placed in upside-down positions. Penicillin G was spread over one side (+PenG), *K. pneumoniae* was inoculated in both halves at three spots and incubated overnight at 37° C. As a control, a blank SM plate without *S. pratensis* (right plate) was used. **(B)** Pictures show no growth of *K. pneumoniae* on the half (b) of SM, which has *S. pratensis* and PenG, compared to the one without PenG (a). *K. pneumoniae* grew well in both halves (c and d) of the blank SM plate. Coomassie blue stain was added to the media and contrasted with *K. pneumoniae* colonies to take the pictures.

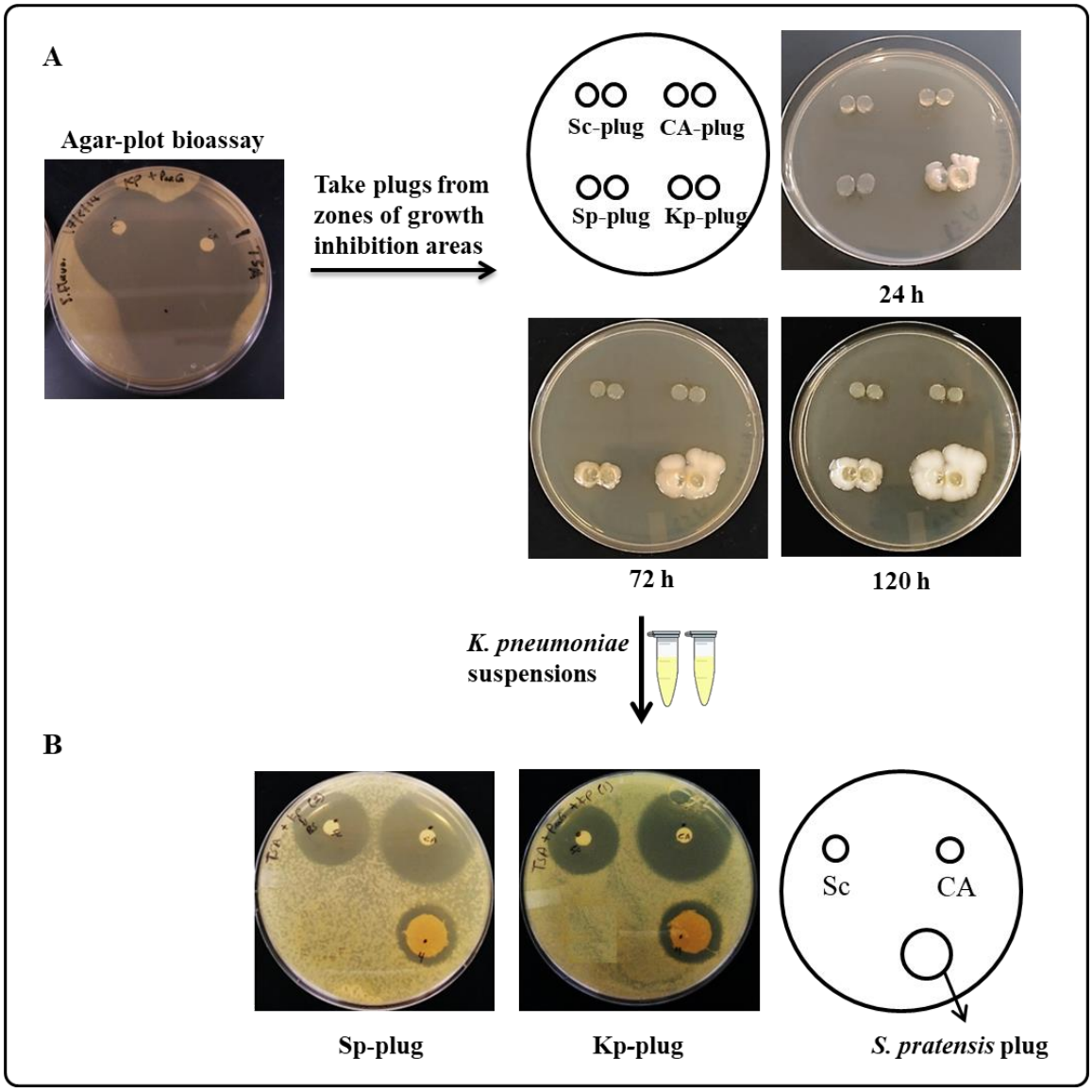


Figure 5.4. The bacteriostatic effect of the *S. pratensis* bioactive substance. (A) Media plugs were taken from the zone of growth inhibition (ZOI) resulting from the *S. pratensis* plot (Sp-plug), ZOI resulting from *S. clavuligerus* supernatant (Sc-plug), and ZOI resulting from CA solution (CA-plug). The picked plugs were placed in a TSA plate and incubated for five days at 37° C to monitor the re-growth of *K. pneumoniae*. Plugs from the zone of *K. pneumoniae* growth (Kp-plug) were used as control. **(B)** Suspensions of *K. pneumoniae* were prepared from the re-grown bacteria on Sp-plug and Kp-plug. Agar plug bioassays were conducted for the second time using suspensions of *K. pneumoniae* from (A), Sc: *S. clavuligerus* supernatant, CA: clavulanic acid solution.

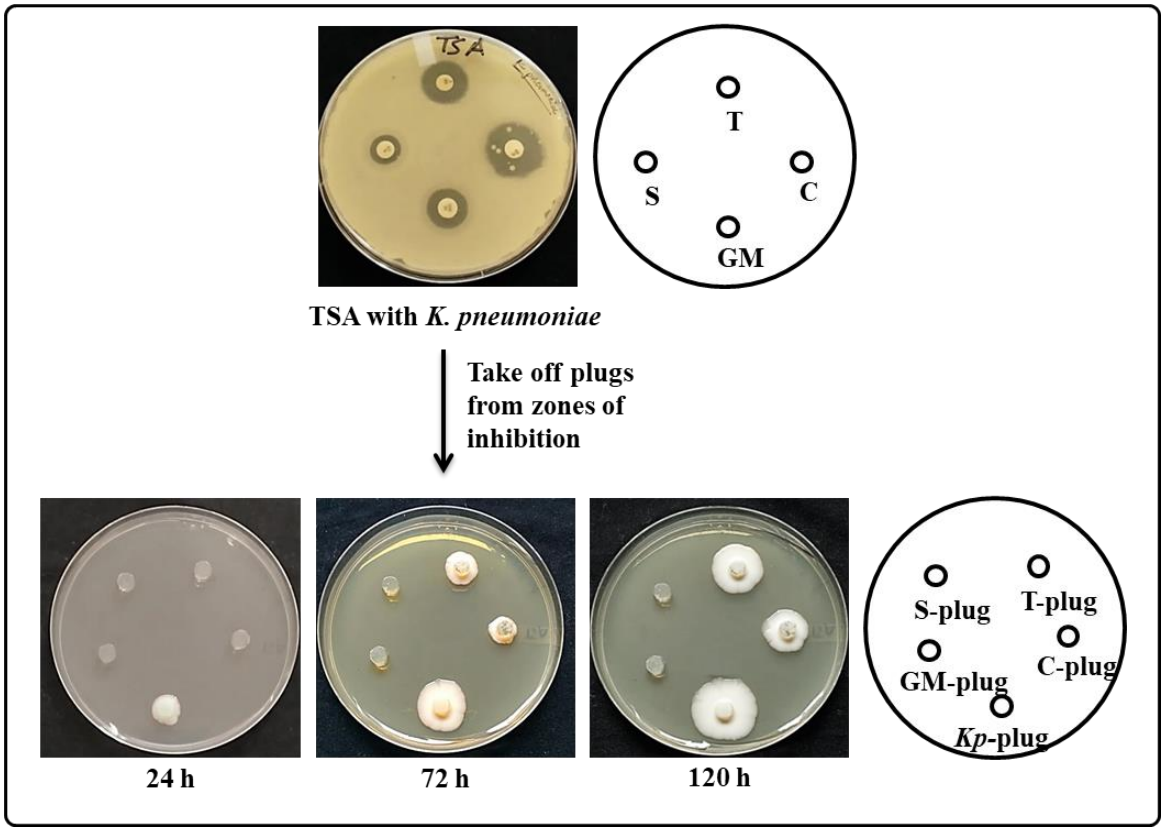
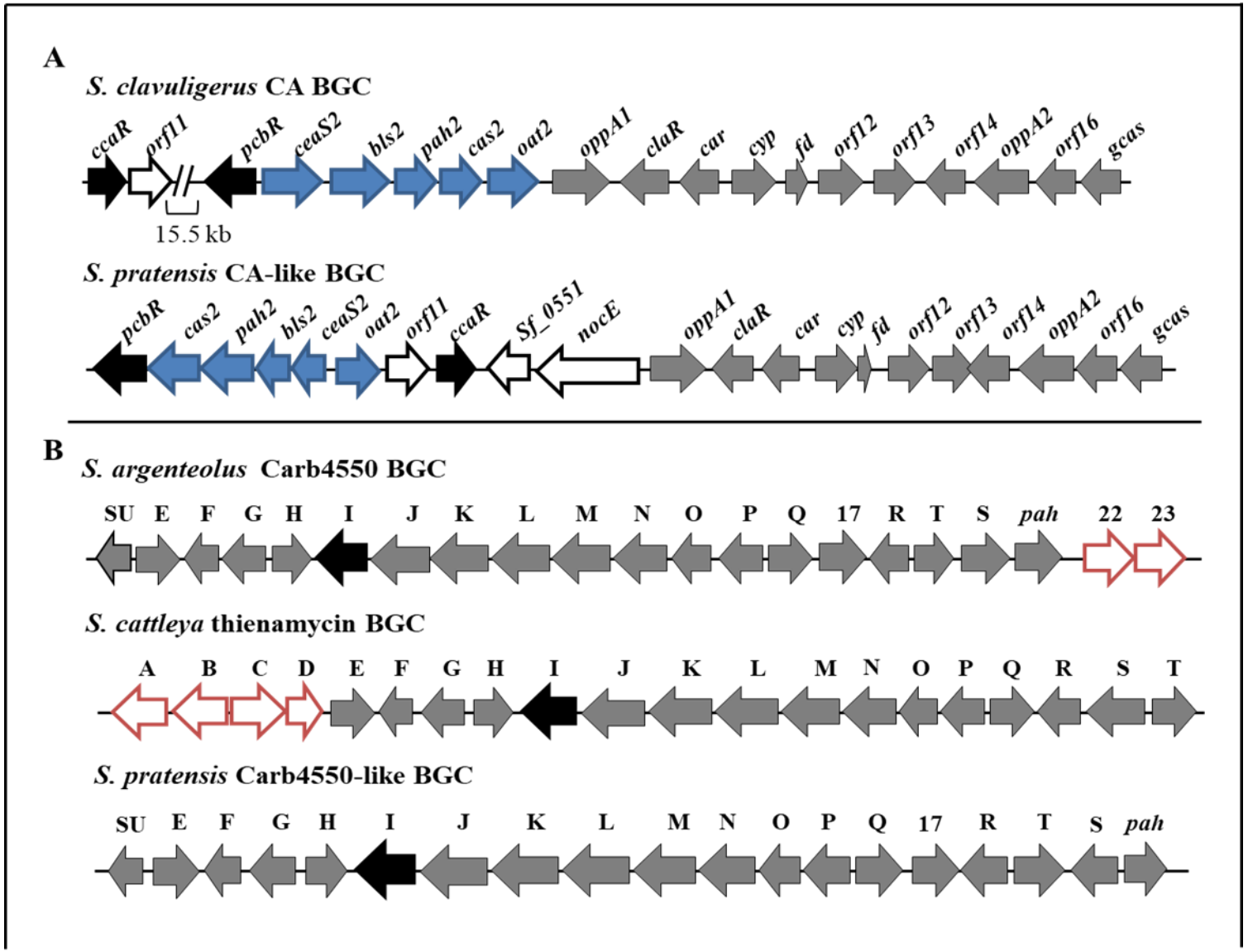


Figure 5.5. Antimicrobial susceptibility test for bacteriostatic and bactericidal antibiotics. Disk diffusion bioassays were performed against *K. pneumoniae* using bactericidal antibiotics gentamycin (GM) and streptomycin (S), and bacteriostatic antibiotics tetracycline (T) and chloramphenicol (C). Agar plugs were picked up from the area of no growth, placed on a TSA plate, and incubated for five days. Plug with *K. pneumoniae* (Kp-plug) was used as control. Re-growth of *K. pneumoniae* was noticed on the plugs of bacteriostatic antibiotics (T-plug) and (C-plug) but not in bactericidal antibiotics (S-plug) and (GM-plug).


A
S. clavuligerus CA BGC
ccaR
orf11
pcbR
ceaS2
bls2
pah2
cas2
oat2
oppA1
claR
car
cyp
fd
orf12
orf13
orf14
oppA2
orf16
gcas
15.5 kb
S. pratensis CA-like BGC
pcbR
cas2
pah2
bls2
ceaS2
oat2
orf11
ccaR
Sf_0551
nocE
oppA1
claR
car
cyp
fd
orf12
orf13
orf14
oppA2
orf16
gcas
B
S. argenteolus Carb4550 BGC
SU E F G H I J K L M N O P Q 17 R T S pah 22 23
S. cattleya thienamycin BGC
A B C D E F G H I J K L M N O P Q R S T
S. pratensis Carb4550-like BGC
SU E F G H I J K L M N O P Q 17 R T S pah

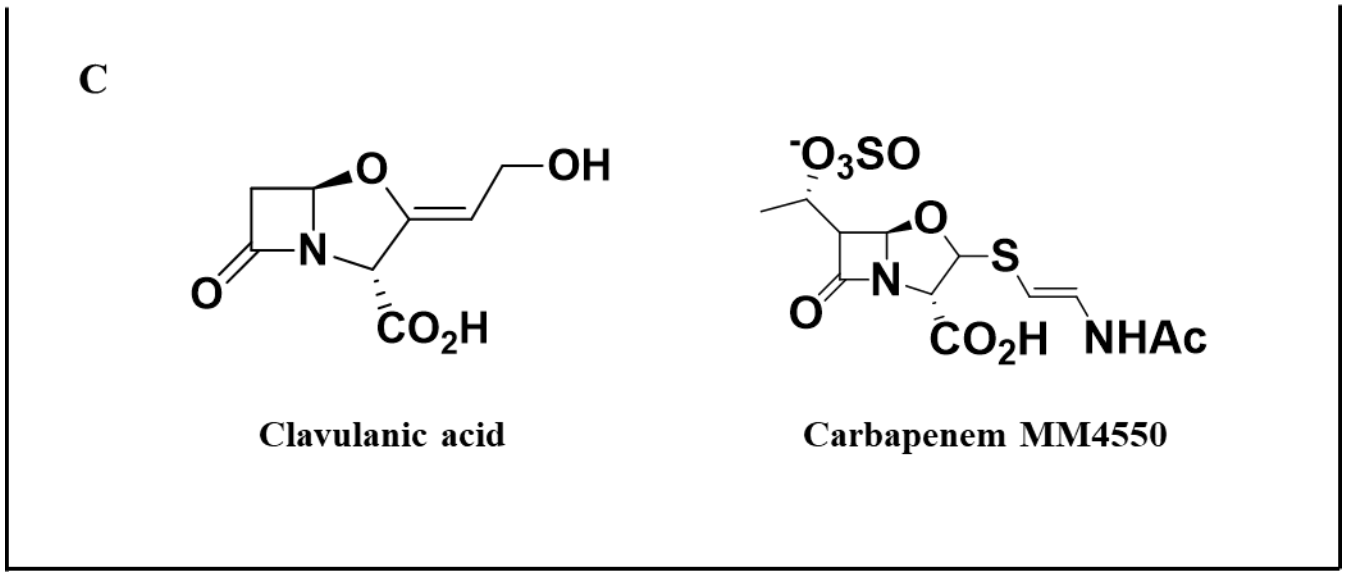


Figure 5.6. Organization and comparison of clavulanic acid and carbapenem MM4550 BGCs in *Streptomyces*. (A) Physical map and organization of the CA BGC of *S. clavuligerus* and CA-like BGC of *S. pratensis*, showing the early genes (blue) and the late genes (grey) of CA biosynthesis. Black arrows represent genes located or similar to genes in the Ceph-C of *S. clavuligerus*. The white arrows represent genes with unknown functions. **(B)** Physical map and organization of the carbapenem MM4550 BGC of *S. argenteolus*, thienamycin BGC of *S. cattleya*, and carbapenem MM4550-like BGC of *S. pratensis*. White arrows represent genes not conserved in the three clusters. Black arrows represent regulatory genes. **(C)** The chemical structure of clavulanic acid and carbapenem MM4550.

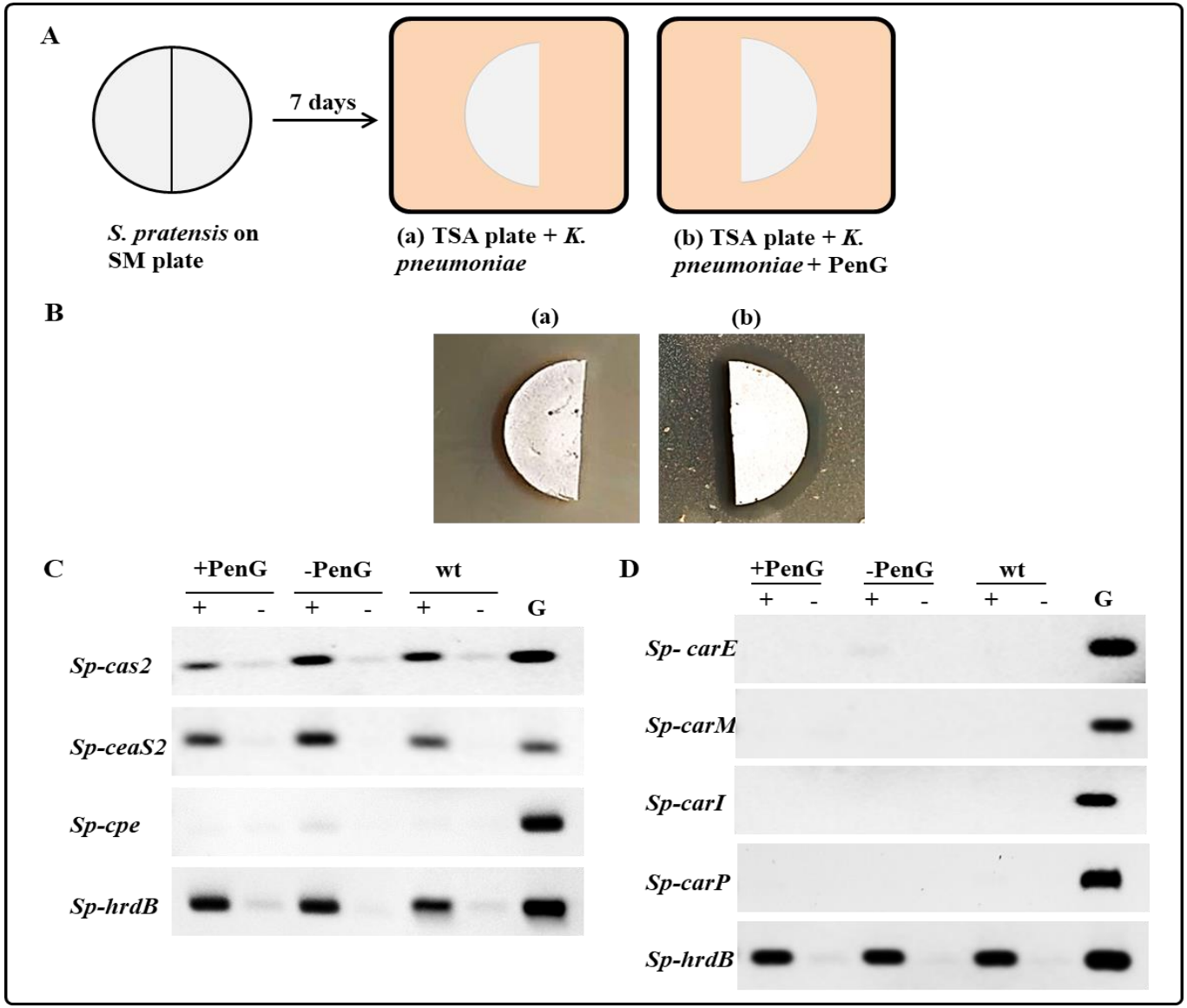

A
7 days
S. pratensis on SM plate
(a) TSA plate + K. pneumoniae
(b) TSA plate + K. pneumoniae + PenG
B
(a)
(b)
C
+PenG
-PenG
wt
+ - + - + - G
Sp-cas2
Sp-ceaS2
Sp-cpe
Sp-hrdB
D
+PenG
-PenG
wt
+ - + - + - G
Sp- carE
Sp-carM
Sp-carI
Sp-carP
Sp-hrdB

Figure 5.7. Transcriptional analysis of genes from the CA-like and Carb4550-like BGCs. (A) The schematic diagram for *S. pratensis* culture on SM plate from which RNA samples were extracted. On day 7 of culture, the *S. pratensis* plate was divided into two halves and used to perform the bioassay against *K. pneumoniae* on TSA plates with +PenG (right, b) or without (left, a). After confirming the bioactive substance production **(B)**, RNA isolation was conducted for the two halves shown on **(B)**. As a control, RNA isolation was also performed for seven days culture of wt *S. pratensis* on an SM plate that did not touch the bioassay TSA plates. **(C)** RT-PCR results for selected genes from the CA-like BGC; Sp-*cas2* and Sp-*ceaS2* from the early steps of CA biosynthesis, and Sp-*cpe* from the late stages of CA biosynthesis. **(D)** RT-PCR for selected essential genes from the Carb4550-like BGC. As controls for RT-PCR, RNA samples were used directly in PCR without RT or cDNA synthesis (-). The constitutively expressed *hrdB* gene expression was used as an internal control to normalize expression levels between different samples. G, PCR conducted from *S. pratensis* genomic DNA.

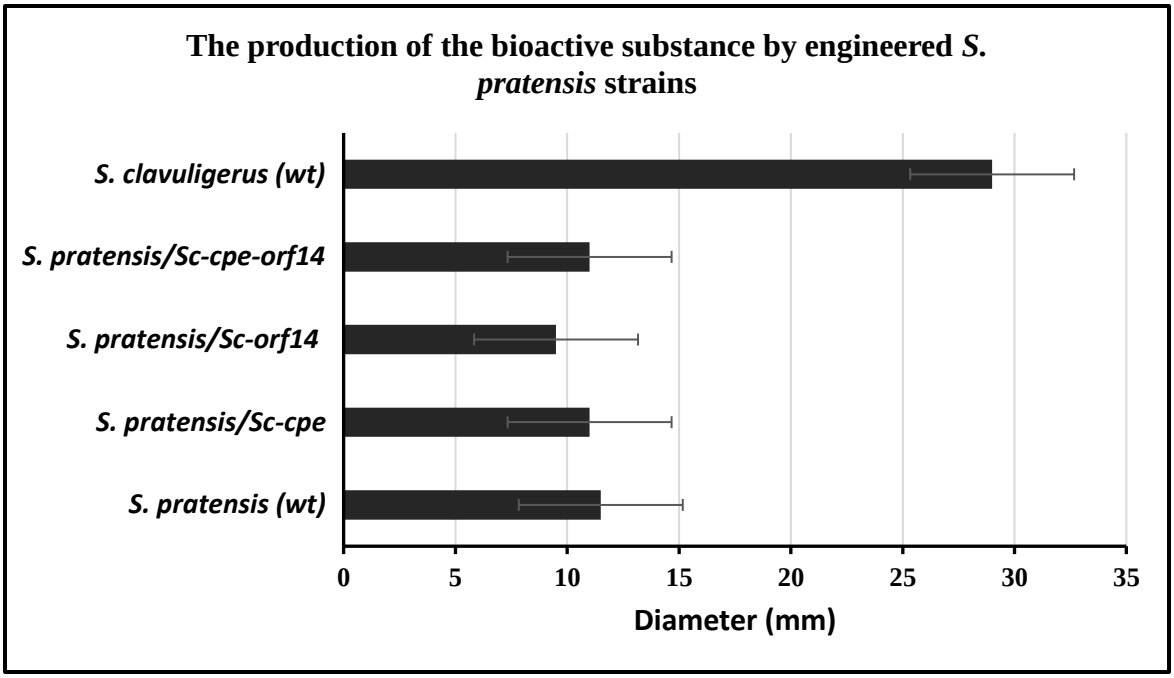


Figure 5.8. The production of a bioactive substance by *S. pratensis* strains expression two genes (*cpe* and *orf14*) from *S. clavuligerus*. The agar plug bioassays were conducted against *K. pneumoniae*. The zones of growth inhibition were measured in diameter (mm) for duplicate samples. Sample for wt *S. clavuligerus* was used as control.

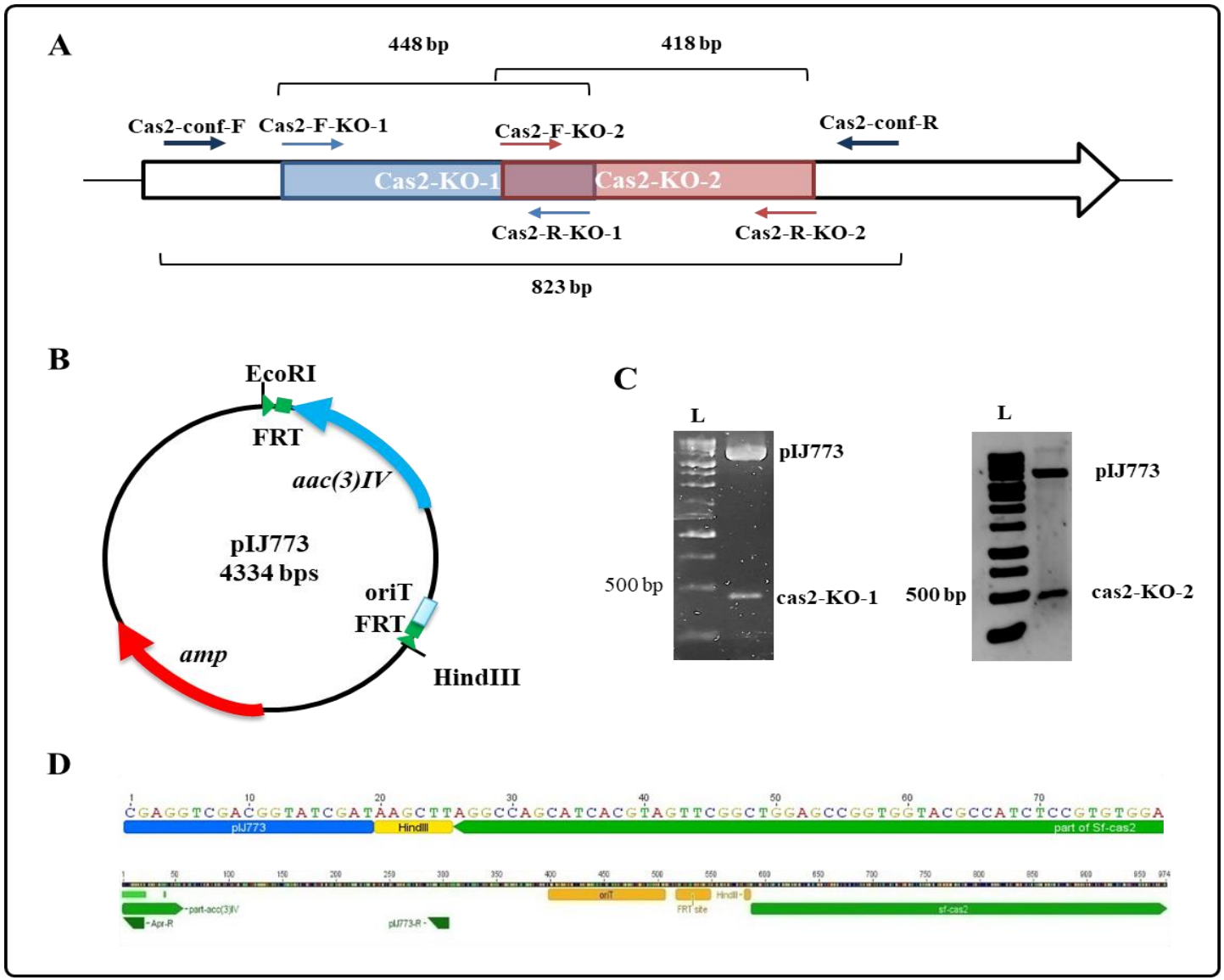
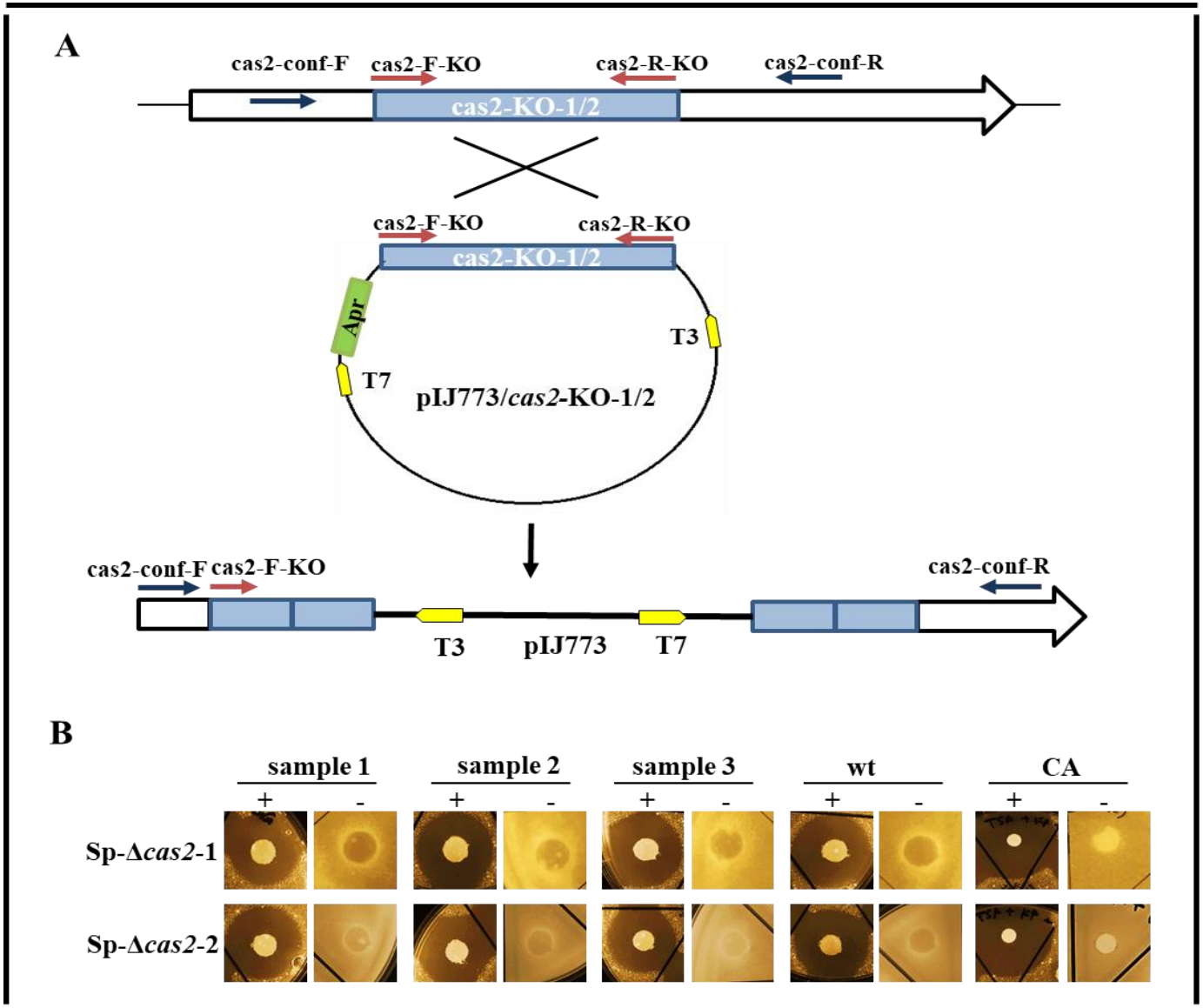

A
448 bp
418 bp
Cas2-conf-F
Cas2-F-KO-1
Cas2-F-KO-2
Cas2-conf-R
Cas2-KO-1
Cas2-KO-2
Cas2-R-KO-1
Cas2-R-KO-2
823 bp
B
EcoRI
FRT
aac(3)IV
pIJ773
4334 bps
oriT
FRT
amp
HindIII
C
L
pIJ773
500 bp
cas2-KO-1
L
pIJ773
500 bp
cas2-KO-2
D
CGAGGTCGACGGTATCGATAAGCTTAGGCCAGCATCACGTAGTTCGGCTGGAGCCGGTGGTACGCCATCTCCGTGTGGA
pIJ773
HindIII
part of Sf-cas2
oriT
HindIII
FRT site
sf-cas2
part-acc(3)IV
Apr-R
pIJ773-R

Figure 5.9. *Sp-cas2* cloning into pIJ773 for insertional inactivation. (A) Schematic graph for *S. pratensis cas2* gene showing the two regions *cas2*-KO-1 and *cas2*-KO-2 that were amplified and cloned into pIJ773 plasmid. **(B)** Schematic map for the pIJ773 plasmid that was used to clone *cas2*-KO regions at the HindIII site. **(C)** Gel electrophoresis results for pIJ773-*cas2*-KO-1 (left) and pIJ773-*cas2*-KO-2 (right) constructs after digestion with HindIII. **(D)** Sequencing analysis for pIJ773-*cas2*-KO-1 and pIJ773-*cas2*-KO-2.


A
cas2-conf-F
cas2-F-KO
cas2-R-KO
cas2-conf-R
cas2-KO-1/2
cas2-F-KO
cas2-R-KO
cas2-KO-1/2
Apr
T3
T7
pIJ773/cas2-KO-1/2
cas2-conf-F
cas2-F-KO
cas2-conf-R
T3
pIJ773
T7
B
sample 1
sample 2
sample 3
wt
CA
+
-
Sp-Δcas2-1
Sp-Δcas2-2

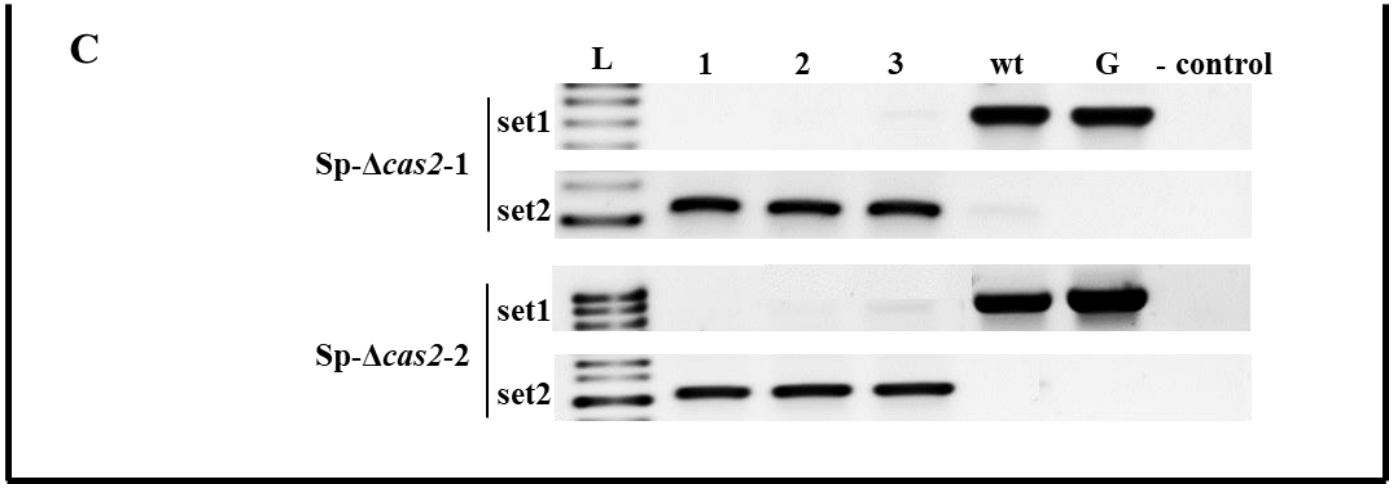
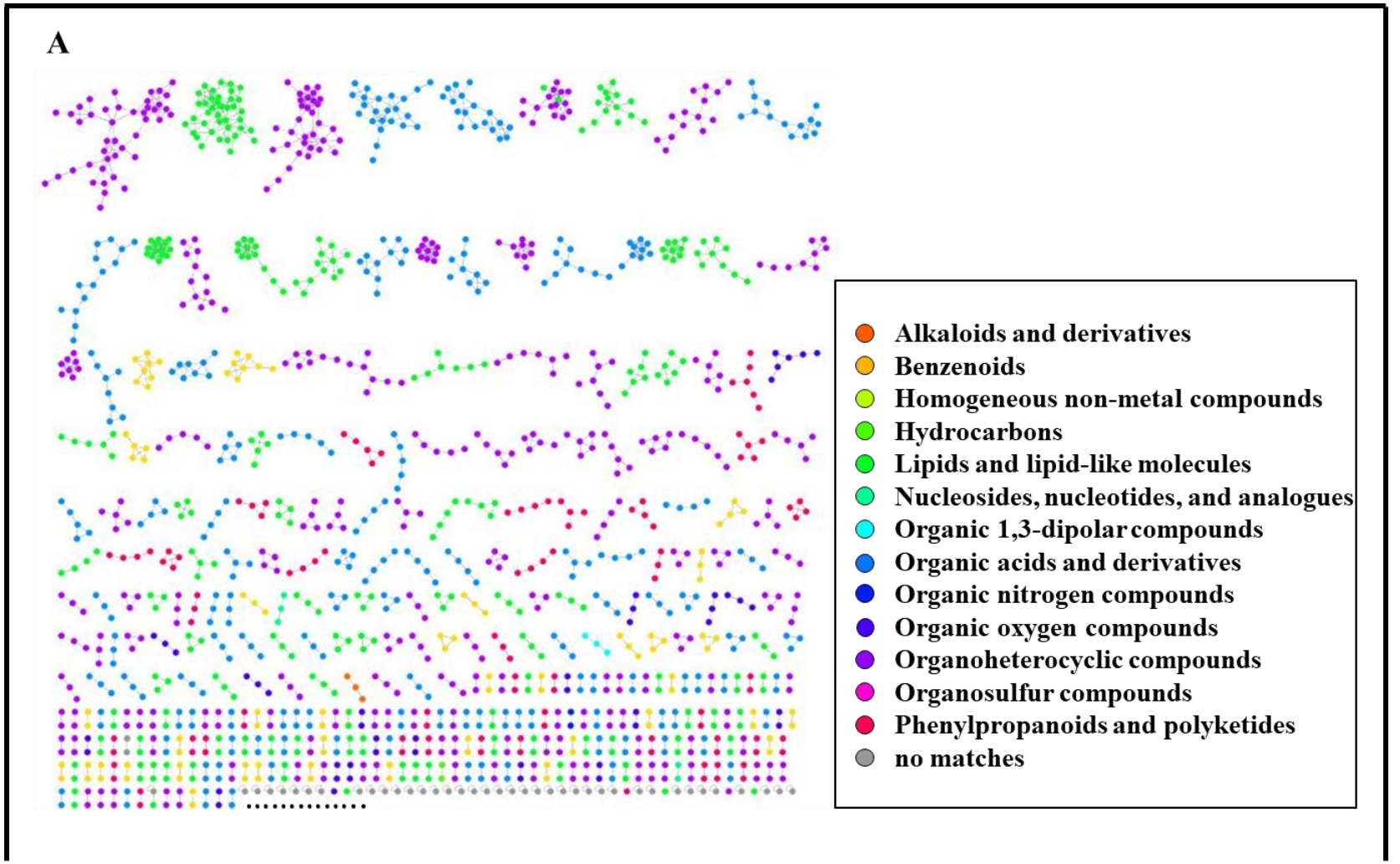


Figure 5.10. *S. pratensis cas2* insertional inactivation. (A) Schematic graph for the insertional inactivation for *cas2*-KO regions via single cross-over with Sp-*cas2* gene. The primer sets, used to confirm the insertional inactivation of Sp-*cas2*, are shown. The purpose of set1 (cas2-conf-F and cas2-conf-R) was to ensure the deletion of Sp-*cas2*, and Set2 (cas2-KO-F and T3) to confirm the cross-over between pIJ773/*cas2*-KO and the Sp-*cas2* gene itself. **(B)** Agar plug bioassays for the successful *cas2* insertional inactivation using both constructs pIJ773-cas2-KO-1 (upper row) and pIJ773-cas2-KO-2 (lower row). The bioassays were conducted using 7-day cultures of *S. pratensis* strains against *K. pneumoniae* on TSA plates with PenG (+) or without (-). The bioassays were performed in triplicates. Agar plugs from wt *S. pratensis* (wt), and CA solution (CA) were used as controls. **(C)** PCR results for the *S. pratensis* strains shown on (B) using Set1 and Set2 primers to confirm the insertional inactivation of Sp-*cas2*. As controls, PCR was also performed for wt *S. pratensis* grown on SM plate (wt) and *S. pratensis* extracted genome (G).


A
Alkaloids and derivatives
Benzenoids
Homogeneous non-metal compounds
Hydrocarbons
Lipids and lipid-like molecules
Nucleosides, nucleotides, and analogues
Organic 1,3-dipolar compounds
Organic acids and derivatives
Organic nitrogen compounds
Organic oxygen compounds
Organoheterocyclic compounds
Organosulfur compounds
Phenylpropanoids and polyketides
no matches

B

Superclass	Number of spectral features		
	Positive mode	Negative mode	Total
Alkaloids and derivatives	4	0	4
Benzenoids	49	65	114
Hydrocarbons	4	0	4
Lipids and lipid-like molecules	188	191	379
Nucleosides, nucleotides, and analogues	6	1	7
Organic 1,3-dipolar compounds	0	4	4
Organic acids and derivatives	256	148	404
Organic nitrogen compounds	3	0	3
Organic oxygen compounds	32	32	64
Organoheterocyclic compounds	280	226	506
Organosulfur compounds	1	0	1
Phenylpropanoids and polyketides	65	77	142
no matches	866	706	1572
Total	1754	1450	3204

Figure 5.11. Metabolomics analysis for the main groups of molecules in wt *S. pratensis.* (A) The molecular networking for spectral features detected by untargeted LC-MS and GNPS analysis and visualized by Cytoscape 3.8 program. The networks show >3000 nodes represent molecules/compounds of 13 superclasses/groups; their names are in the figure legend. **(B)** The table shows the number of spectral features detected for each superclass/group of molecules in the positive and negative ionization modes after *in silico* annotation using NAP in GNPS.

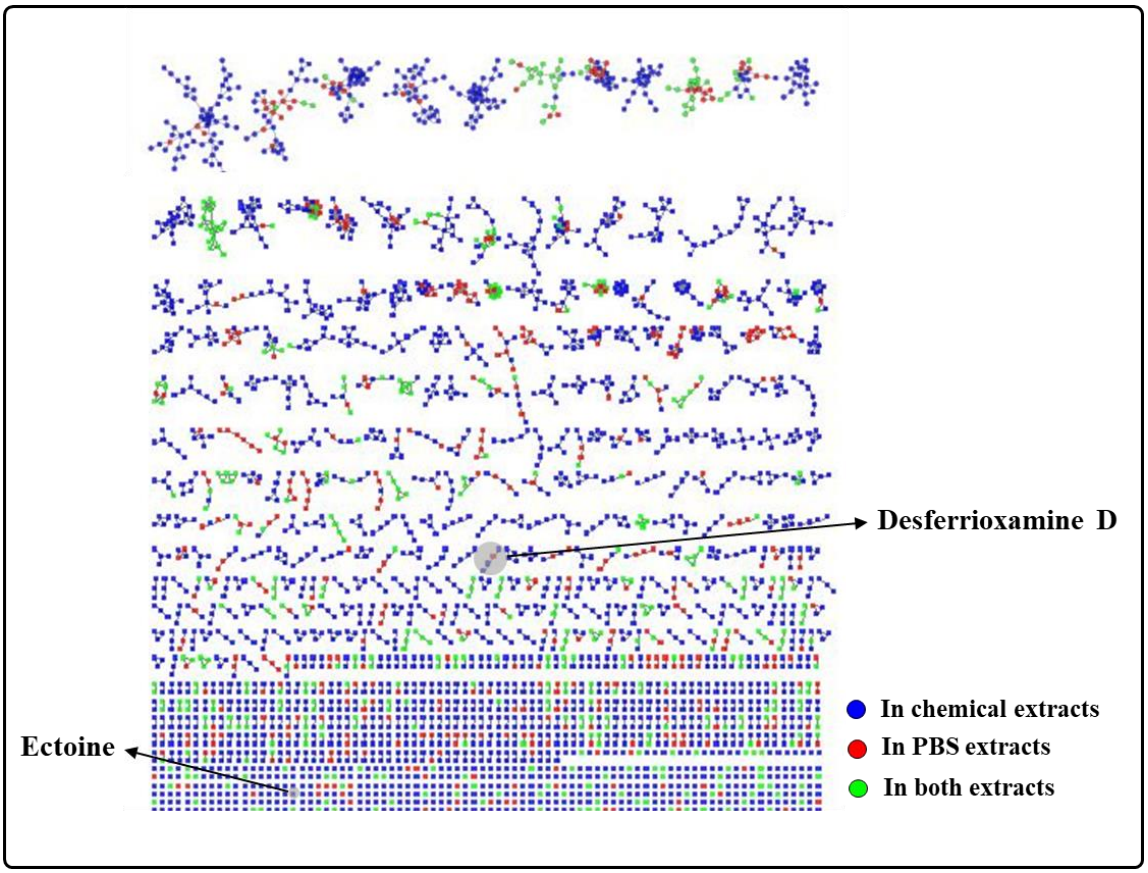


Figure 5.12. Metabolic network constructed using wt *S. pratensis* culture extracts. The nodes are color-coded according to the extraction protocol (bottom right legend), where each node depicts a mass spectrum and edges represent the relationship between different nodes. The clusters of nodes related to metabolites (Desferrioxamine D and Ectoine) are indicated by black arrows and their names.

5.6.2. Tables

Table 5.1: Specialized metabolite (SM) biosynthetic gene clusters predicted in the genome of *S. pratensis* ATCC33331 using antiSMASH 4.0.

Region [a]	Type	Most similar known cluster	Similarity	MIBiG [b]	Reference [c]
1	NRPS, T1PKs, β-lactam	Carbapenem MM4550	65%	BGC00842	*S. argenteolus*
		Thienamycin	50%	BGC00847	*S. cattleya*
		SGR PTMs	100%	BGC01043	*S. griseus* NBRC 13350
2	NRPs	Coelichelin	90%	BGC00325	*S. coelicolor* A3(2)
3	Terpene	Isorenieratene	100%	BGC00664	*S. griseus* NBRC 13350
4	Bacteriocin	NA [d]	NA	NA	NA
5	β-lactam	Clavulanic acid	20%	BGC00845	*S. clavuligerus*
6	Terpene	Hopene	69%	BGC00663	*S. coelicolor* A3(2)
7	T1PKS	Sceliphrolactam	88%	BGC01770	*Streptomyces* sp. SD85
		Macrotermycins	61%	BGC01658	*Amycolatopsis* sp. M39
		Vicenistatin	60%	BGC00167	*Streptomyces halstedii*
8	Bacteriocin	NA	NA	NA	NA
9	NRPS	Skyllamycin	4%	BGC00429	*Streptomyces* sp. Acta 289
10	Siderophore	Ficellomycin	3%	BGC01593	*Streptomyces ficellus*
11	Terpene	NA	NA	NA	NA
12	Bacteriocin	NA	NA	NA	NA
13	Butyrolactone	Lactonamycin	3%	BGC00238	*Streptomyces rishiriensis*
14	NRPS,T1PKS	Istamycin	11%	BGC00700	*Streptomyces tenjimariensis*
15	siderophore	Desferrioxamine	83%	BGC00940	*S. coelicolor* A3(2)
16	Lanthipeptide	Azalomycin F	8%	BGC01523	*Streptomyces* sp. 211726
17	Terpene	NA	NA	NA	NA

18	Ectoine	Ectoine	100%	BGC00853	*Streptomyces anulatus*
19	T2PKs, PKs-like	Cinerubin B	28%	BGC00212	*Streptomyces* sp. SPB074
20	Terpene	Steffimycin	16%	BGC00273	*Streptomyces steffisburgensis*
21	Terpene, Ectoine	Ectoine	100%	BGC00853	*Streptomyces anulatus*
22	Bacteriocin	NA	NA	NA	NA
23	T3PKs	Tetronasin	11%	BGC00163	*Streptomyces longisporoflavus*
24	Melanin	Melanin	100%	BGC00911	*S. griseus* NBRC 13350
25	T2PKs, Terpene	Spore pigment	83%	BGC00271	*Streptomyces avermitilis*
26	NRPS	Rimosamide	21%	BGC01760	*Streptomyces rimosus* ATCC 10970
27	Butyrolactone	NA	NA	NA	NA

[a] The regions as appear in *S. pratensis* genome.

[b] MIBiG BGC-ID: Minimal information about biosynthetic gene cluster-identification number.

[c] The name of bacteria that possess the reference BGC as appear in the MIBiG data base.

[d] NA: not available.

5.7. Supplementary materials

5.7.1. Supplementary Figures

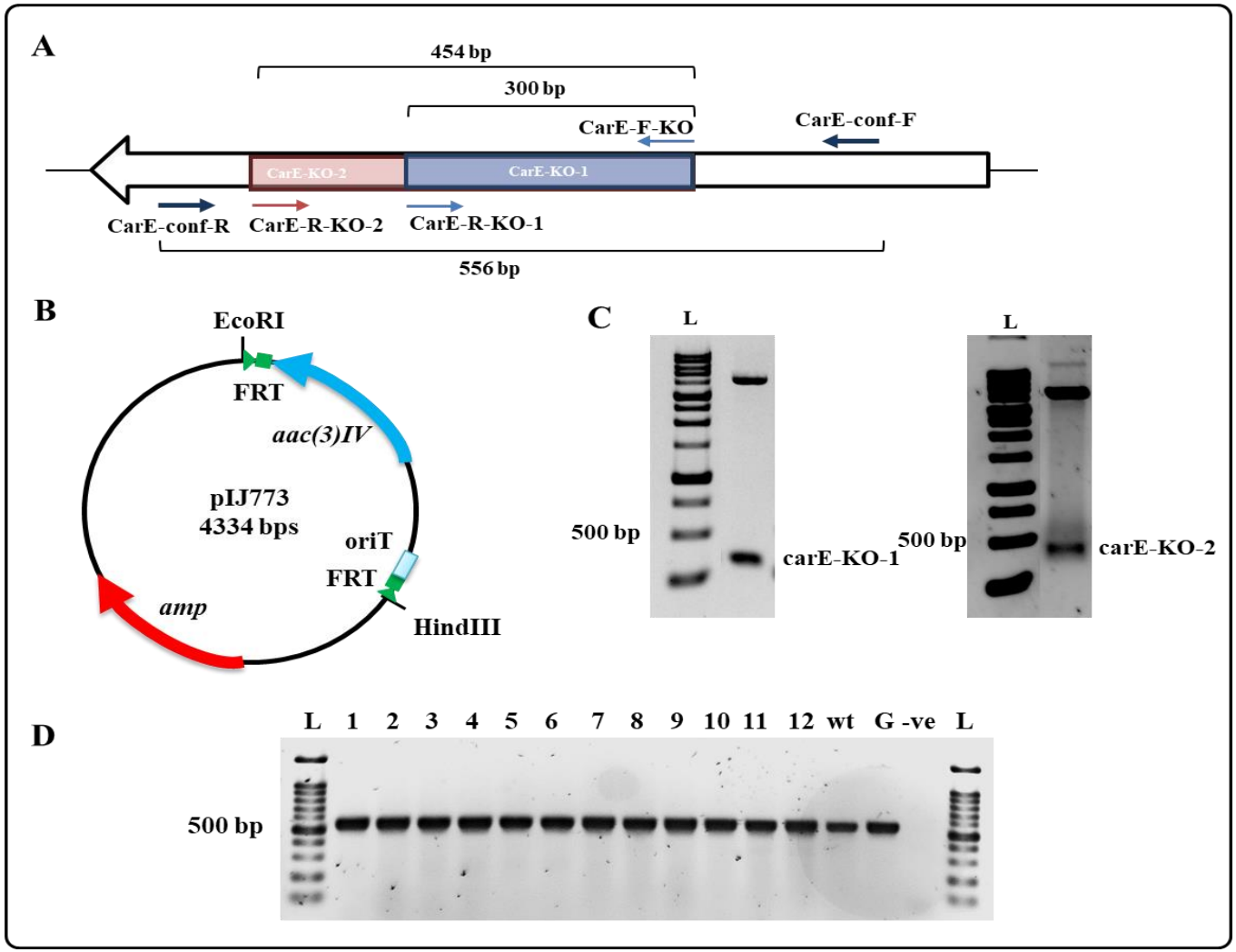


Figure S5.1. Sp-*carE* cloning into pIJ773 for insertional inactivation. (A) Schematic graph for *S. pratensis carE* gene showing the two regions *carE*-KO-1 and *carE*-KO-2 that were amplified and cloned into pIJ773 plasmid. **(B)** Schematic map for the pIJ773 plasmid that was used to clone *carE*-KO regions at the HindIII site. **(C)** Gel electrophoresis results for pIJ773-*carE*-KO-1 (left) and pIJ773-*carE*-KO-2 (right) constructs after digestion with HindIII. **(D)** Gel electrophoresis for PCR results for *S. pratensis*/*carE*-KO exconjugants showing the amplification of carE gene, indicating that the *carE* cut-off by insertional inactivation was not achieved.

CHAPTER VI

6.1. Summary and Perspectives

In Chapter III, the work conducted on *cpe* (*orf12*) and *orf14* provided a greater understanding of the roles of these two genes in CA and 5*S* clavams biosynthesis in *S. clavuligerus*. The two genes are thought to be involved in the late steps of CA biosynthesis, during which the conversion of clavaminic acid into clavaldehyde takes place (Figure 1.6) (Iqbal et al., 2010b; Srivastava et al., 2019). In addition, the deletion of *cpe* or *orf14* almost eliminates CA biosynthesis without affecting the 5*S* clavams production (Jensen et al., 2004a; Srivastava et al. 2019). The first objective of this study was to investigate if CPE has a regulatory function on the transcription of the CA BGC, but the deletion of *cpe* did not affect the expression of CA biosynthetic genes.

Interestingly, the overexpression of *cpe* induces the production of 2HMC when *S. clavuligerus* is grown in SA medium, a situation by which the 5*S* clavams are not produced by the wt strain. Also, the transcription analysis for some genes (*cas1, cvm1, cvm2, cvm5, cvm6p, cvm7p, orfA, and orfB*) that are essential for 5*S* clavam biosynthesis showed elevation in their expression in the *cpe*-overexpressed strain in comparison to wt *S. clavuligerus*, suggesting that CPE might have an indirect positive regulation effect on the production of 2HMC (Figure 3.9). These results are the first to link one of the "late genes" of the CA BGC with 5*S* clavam biosynthesis, and this opens new avenues of investigation on the effects of the late genes of the CA BGC (*orf10 – orf23*) on 5*S* clavam biosynthesis, which in turn will lead to a better understanding of the 5*S* clavam biosynthesis pathway. The 5*S* clavam biosynthetic pathway shares intermediate compounds with the CA biosynthesis pathway (Figure 1.5) and influences the overall

energetics and primary metabolic pools available for CA biosynthesis. Therefore, understanding the 5*S* clavam pathway and thus eliminating the competing steps for the CA pathway would be of benefit for constructing CA high-yielding strains of *S. clavuligerus* (Paradkar et al. 2001).

A hydrolysis step is required to convert 2FMC into 2HMC (Figure 1.7), but the gene product that accomplished this step is still unknown. In SA medium the 5*S* clavams genes of *S. clavuligerus* are silent. Therefore, it is hypothesized that CPE (when it is overexpressed) performs this hydrolysis step to produce the 2HMC. However, this activity for CPE is likely not its principal function, but instead it reflects a side, or "moonlighting" activity, as the previously reported esterase function for this protein (Valegård et al., 2013). The role of CPE in the CA biosynthetic pathway requires further studies and investigations.

The work presented in Chapter IV focused on the *nocE* gene, which is present within the CA-like BGC in CA non-producers and is outside of the CA BGC in *S. clavuligerus*. The presence of *nocE* combined with a β-lactam BGC in the genome of *Streptomyces* species was noticeable. Although the deletion of *nocE* or its constitutive expression did not affect the production of the β-lactam metabolites in *S. clavuligerus*, it did affect the growth and the general metabolism in *S. clavuligerus*.

The NocE protein belongs to the SGNH/GDSL-hydrolases family of proteins with esterase or lipase activity. Bielen et al. (2009) have hypothesized that lipases and other hydrolytic enzymes might be necessary to produce the specialized metabolites in *Streptomyces* (Horinouchi, 2002). The results presented here disagree with this

hypothesis; rather they show that the NocE protein is more involved in nutrient and general metabolism in *S. clavuligerus*.

Our searching in the data bases discovered the presence of more actinomycetes species (11 species found in this study) with β-lactam metabolite BGCs (CA, Ceph-C, and/or Carb4550). These species are not reported for their ability to produce specialized metabolites. However, investigating their abilities to produce the respective metabolites is a promising field of research.

In Chapter V of this thesis, the biochemical and molecular work presented provide a better picture of *S. pratensis* ATCC33331 and its capability to produce antimicrobial agents. This strain, isolated from compost in Laval, Canada (Ishaque & Kluepfel, 1980), has not previously been reported to produce specialized metabolites with antibacterial activity. Following the OSMAC approach to awaken silent BGCs, *S. pratensis* ATCC33331 was found to produce a BS that inhibits the growth of *K. pneumoniae* when combined with penicillin G, an action that resembles β-lactamase inhibitors. *S. pratensis* possesses two BGCs, one CA-like and one Carb4550-like, making them the first targets to investigate for identifying this BS. Using untargeted LC-MS/MS and GNPS, the mass spectra for the two metabolites, CA and Carb4550, could not be detected in bioactive culture extracts. Moreover, transcriptional analysis of the BGCs and the gene deletion experiment indicate that the BS produced by *S. pratensis* is neither CA nor Carb4550, but instead is likely a novel specialized metabolite. The genomic analysis of the specialized metabolites BGCs in *S. pratensis* predicted the presence of a BGC with high similarity to sceliphrolactam and macrotermycin BGCs (Table 5.1). Both of the metabolites showed

antimicrobial activities (Beemelmanns et al., 2017; Oh et al., 2011). Whether they are produced by *S. pratensis* or not and whether they can exert a bacteriostatic effect against *K. pneumoniae* requires further study and investigation. A targeted LC-MS could be useful to examine the production of sceliphrolactam or macrotermycin by *S. pratensis*. Moreover, A transcriptomic analysis for *S. pratensis* would be a valuable approach to figure out the BGC responsible for the production of the discovered BS.

References

Abe, I. (2018). Biosynthetic studies on teleocidins in *Streptomyces*. *Journal of Antibiotics*, *71*(9), 763–768. doi.org/10.1038/s41429-018-0069-4

Abraham, E. P., & Chain, E. (1940). An Enzyme from Bacteria able to Destroy Penicillin. *Nature*, *146*, 837.

AbuSara, N. F., Piercey, B. M., Moore, M. A., Shaikh, A. A., Nothias, L., Srivastava, S. K., Cruz-Morales, P., Tahlan, K. (2019). Comparative Genomics and Metabolomics Analyses of Clavulanic Acid-Producing *Streptomyces* Species Provides Insight Into Specialized Metabolism. *Frontiers in Microbiology*, *10*(November), 1–17. doi.org/10.3389/fmicb.2019.02550

Adusumilli, R., & Mallick, P. (2017). Data Conversion with ProteoWizard msConvert. In L. Comai, J. Katz, & P. Mallick (Eds.), *Proteomics. Methods in Molecular Biology* (Vol. 1550, pp. 508–510). New York: Humana Press, New York. doi.org/10.1007/978-1-4939-6747-6

Aidoo, K., Wong, A., Alexander, D., Rittammer, R., & Jensen, S. (1994). Cloning, sequencing and disruption of a gene from *Streptomyces clavuligerus* involved in clavulanic acid biosynthesis. *Gene*, *147*(1), 41–46.

Akhter, N., Liu, Y., Auckloo, B. N., Shi, Y., Wang, K., Chen, J., … Wu, B. (2018). Stress-driven discovery of new angucycline-type antibiotics from a marine *Streptomyces pratensis* NA-ZhouS1. *Marine Drugs*, *16*(9), 1–16. doi.org/10.3390/md16090331

Akoh, C. C., Lee, G. C., Liaw, Y. C., Huang, T. H., & Shaw, J. F. (2004). GDSL family of serine esterases/lipases. *Progress in Lipid Research*, *43*(6), 534–552. doi.org/10.1016/j.plipres.2004.09.002

Al-Bassam, M. M., Bibb, M. J., Bush, M. J., Chandra, G., & Buttner, M. J. (2014). Response Regulator Heterodimer Formation Controls a Key Stage in *Streptomyces* Development. *PLoS Genetics*, *10*(8). doi.org/10.1371/journal.pgen.1004554

Alexander, D. C., & Jensen, S. E. (1998). Investigation of the *Streptomyces clavuligerus* Cephamycin C Gene Cluster and Its Regulation by the CcaR Protein. *Journal of Bacteriology*, *180*(16), 4068–4079.

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., … Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology*, *37*(4), 420–423. hdoi.org/10.1038/s41587-019-0036-z

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. doi.org/10.1016/S0022-2836(05)80360-2

Álvarez-Álvarez, R., Martínez-Burgo, Y., Pérez-Redondo, R., Braña, A. F., Martín, J. F., & Liras, P. (2013). Expression of the endogenous and heterologous clavulanic acid cluster in *Streptomyces flavogriseus*: Why a silent cluster is sleeping. *Applied Microbiology and Biotechnology*, *97*(21), 9451–9463. doi.org/10.1007/s00253-013-5148-7

Álvarez-álvarez, R., Rodríguez-García, A., Martínez-Burgo, Y., Martín, J. F., & Liras, P. (2018). Transcriptional studies on a *Streptomyces clavuligerus oppA2* deletion mutant: Nacetylglycyl- clavaminic acid is an intermediate of clavulanic acid biosynthesis. *Applied and Environmental Microbiology*, *84*(22), 1–17. doi.org/10.1128/AEM.01701-18

Álvarez-Álvarez, Rubén, Botas, A., Albillos, S. M., Rumbero, A., Martín, J. F., & Liras, P. (2015). Molecular genetics of naringenin biosynthesis, a typical plant secondary metabolite produced by *Streptomyces clavuligerus*. *Microbial Cell Factories, 14*(1), 178. doi.org/10.1186/s12934-015-0373-7

Álvarez-Álvarez, Rubén, Martínez-Burgo, Y., Rodríguez-García, A., & Liras, P. (2017). Discovering the potential of *S. clavuligerus* for bioactive compound production: Crosstalk between the chromosome and the pSCL4 megaplasmid. *BMC Genomics*, *18*(1), 1–13. doi.org/10.1186/s12864-017-4289-y

Ambler, R. P. (1980). The structure of β-lactamases. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *289*(1036), 321–331. doi.org/https://doi.org/10.1098/rstb.1980.0049

Antunes, N. T., & Fisher, J. F. (2014). Acquired class D β-Lactamases. *Antibiotics*, *3*(3), 398–434. doi.org/10.3390/antibiotics3030398

Aoki, H., Sakai, H. I., Kohsaka, M., Konomi, T., Hosoda, J., Iguchi, E., … Kubochi, Y. (1976). Nocardicin A, a new monocyclic β-lactam antibiotic I. Discovery, isolation and characterization. *The Journal of Antibiotics*, *29*(5), 492–500.

doi.org/10.7164/antibiotics.29.492

Arinaitwe, E., Sandison, T. G., Wanzira, H., Kakuru, A., Homsy, J., Kalamya, J., … Dorsey, G. (2009). Artemether-lumefantrine versus dihydroartemisinin-piperaquine for falciparum malaria: A longitudinal, randomized trial in young Ugandan children. *Clinical Infectious Diseases*, *49*(11), 1629–1637. doi.org/10.1086/647946

Arruda, P., & Barreto, P. (2020). Lysine Catabolism Through the Saccharopine Pathway: Enzymes and Intermediates Involved in Plant Responses to Abiotic and Biotic Stress. *Frontiers in Plant Science*, *11*(May), 1–10. doi.org/10.3389/fpls.2020.00587

Arulanantham, H., Kershaw, N. J., Hewitson, K. S., Hughes, C. E., Thirkettle, J. E., & Schofield, C. J. (2006). ORF17 from the clavulanic acid biosynthesis gene cluster catalyzes the ATP-dependent formation of N-glycyl-clavaminic acid. *Journal of Biological Chemistry*, *281*(1), 279–287. doi.org/10.1074/jbc.M507711200

Ashihara, H., Yokota, T., & Crozier, A. (2013). Biosynthesis and catabolism of purine alkaloids. *Advances in Botanical Research* (1st ed., Vol. 68). Elsevier Inc. doi.org/10.1016/B978-0-12-408061-4.00004-3

Bachmann, B. O., Li, R., & Townsend, C. a. (1998). β-Lactam synthetase: A new biosynthetic enzyme. *Proceedings of the National Academy of Sciences*, *95*(16), 9082–9086. doi.org/10.1073/pnas.95.16.9082

Baggaley, K. H., Brownb, A. G., Schofield, C. J., Dyson, T., & Sciences, M. (1997). Chemistry and biosynthesis of clavulanic acid and other clavams. *Natural Product*

Reports, *14*, 309–333. https://doi.org/doi:10.1039/NP9971400309

Baltz, R. H. (1980). Genetic recombination by protoplast fusion in *Streptomyces*. *Developments in Inductrial Microbiology*, *21*, 43–54.

Baltz, R. H. (2008). Renaissance in antibacterial discovery from actinomycetes. *Current Opinion in Pharmacology*, *8*(5), 557–563. doi.org/10.1016/j.coph.2008.04.008

Baltz, R. H. (2016). Genetic manipulation of secondary metabolite biosynthesis for improved production in *Streptomyces* and other actinomycetes. *Journal of Industrial Microbiology and Biotechnology*, *43*(2–3), 343–370. doi.org/10.1007/s10295-015-1682-x

Baltz, R. H. (2019). Natural product drug discovery in the genomic era: realities, conjectures, misconceptions, and opportunities. *Journal of Industrial Microbiology and Biotechnology*, *46*(3–4), 281–299. doi.org/10.1007/s10295-018-2115-4

Barona-Gómez, F., Wong, U., Giannakopulos, A. E., Derrick, P. J., & Challis, G. L. (2004). Identification of a cluster of genes that directs desferrioxamine biosynthesis in *Streptomyces coelicolor* M145. *Journal of the American Chemical Society*, *126*(50), 16282–16283. doi.org/10.1021/ja045774k

Bartoňková, I., & Dvořák, Z. (2018). Essential oils of culinary herbs and spices display agonist and antagonist activities at human aryl hydrocarbon receptor AhR. *Food and Chemical Toxicology*, *111*(August 2017), 374–384. doi.org/10.1016/j.fct.2017.11.049

Baumgartner, J. T., Habeeb Mohammad, T. S., Czub, M. P., Majorek, K. A., Arolli, X., Variot, C., … Kuhn, M. L. (2021). Gcn5-Related N-Acetyltransferases (GNATs) With a Catalytic Serine Residue Can Play Ping-Pong Too. *Frontiers in Molecular Biosciences*, *8*(April). doi.org/10.3389/fmolb.2021.646046

Beemelmanns, C., Ramadhar, T. R., Kim, K. H., Klassen, J. L., Cao, S., Wyche, T. P., … Clardy, J. (2017). Macrotermycins A-D, Glycosylated Macrolactams from a Termite-Associated *Amycolatopsis* sp. M39. *Organic Letters*, *19*(5), 1000–1003. doi.org/10.1021/acs.orglett.6b03831

Bibb, M. J. (2005). Regulation of secondary metabolism in *Streptomycetes*. *Current Opinion in Microbiology*, *8*(114), 208–215. doi.org/10.1146/annurev.phyto.43.040204.140214

Bibb, M. J., Janssen, G. R., & Ward, J. M. (1985). Cloning and analysis of the promoter region of the erythromycin resistance gene (*ermE*) of *Streptomyces erythraeus*. *Gene*, *38*(1–3), 215–226. doi.org/10.1016/0378-1119(85)90220-3

Bielen, A., Ćetković, H., Long, P. F., Schwab, H., Abramić, M., & Vujaklija, D. (2009). The SGNH-hydrolase of *Streptomyces coelicolor* has (aryl)esterase and a true lipase activity. *Biochimie*, *91*(3), 390–400. doi.org/10.1016/j.biochi.2008.10.018

Bierman, M., Logan, R., O'Brien, K., Seno, E. T., Rao, R. N., & Schoner, B. E. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. *Gene*, *116*(1), 43–49. doi.org/10.1016/0378-1119(92)90627-2

Bignell, D. R. D., Jensen, S. E., & Leskiw, B. K. (2005). Expression of *ccaR*, Encoding the Positive Activator of Cephamycin C and Clavulanic Acid Production in *Streptomyces clavuligerus*, Is Dependent on bldG. *Antimicrobial Agents and Chemotherapy*, *49*(4), 1529–1541. doi.org/10.1128/AAC.49.4.1529

Bird, A. E., Bellis, J. M., & Gasson, B. C. (1982). Spectrophotometric assay of clavulanic acid by reaction with imidazole. *The Analyst*, *107*(1279), 1241–1245. doi.org/10.1039/AN9820701241

Blanco, G. (2012). Comparative analysis of a cryptic thienamycin-like gene cluster identified in *Streptomyces flavogriseus* by genome mining. *Archives of Microbiology*, *194*(6), 549–555. doi.org/10.1007/s00203-011-0781-y

Blin, K., Wolf, T., Chevrette, M. G., Lu, X., Schwalen, C. J., Kautsar, S. A., … Medema, M. H. (2017). AntiSMASH 4.0 - improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Research*, *45*(W1), W36–W41. doi.org/10.1093/nar/gkx319

Bode, H. B., Bethe, B., Höfs, R., & Zeeck, A. (2002). Big effects from small changes: Possible ways to explore nature's chemical diversity. *ChemBioChem*, *3*(7), 619–627. doi.org/10.1002/1439-7633(20020703)3:7<619::AID-CBIC619>3.0.CO;2-9

Bonnet, R. (2004). Growing Group of Extended-Spectrum β-Lactamases: The CTX-M Enzymes. *Antimicrobial Agents and Chemotherapy*, *48*(1), 1–14. doi.org/10.1128/AAC.48.1.1-14.2004

Boraston, A. B., Bolam, D. N., Gilbert, H. J., & Davies, G. J. (2004). Carbohydrate-binding modules: Fine-tuning polysaccharide recognition. *Biochemical Journal*, *382*(3), 769–781. doi.org/10.1042/BJ20040892

Bouwmeester, H. J., Gershenzon, J., Konings, M. C. J. M., & Croteau, R. (1998). Biosynthesis of the monoterpenes limonene and carvone in the fruit of caraway: I. Demonstration of enzyme activities and their changes with development. *Plant Physiology*, *117*(3), 901–912. doi.org/10.1104/pp.117.3.901

Bowler, L. D., Zhang, Q. Y., Riou, J. Y., & Spratt, B. G. (1994). Interspecies recombination between the *penA* genes of *Neisseria meningitidis* and commensal *Neisseria* species during the emergence of penicillin resistance in *N. meningitidis*: Natural events and laboratory simulation. *Journal of Bacteriology*, *176*(2), 333–337. doi.org/10.1128/jb.176.2.333-337.1994

Breilh, D., Texier-Maugein, J., Allaouchiche, B., Saux, M.-C., & Boselli, E. (2013). Carbapenems. *Journal of Chemotherapy*, *25*(1), 1–17. doi.org/10.1179/1973947812Y.0000000032

Brown, A. G., Butterworth, D., Cole, M., Hanscomb, G., Hood, J. D., & Reading, C. (1976). Naturally-occurring β-lactamase inhibitors with antibacterial activity. *The Journal of Antibiotics*, *XXIX*(6), 668–669.

Brown, D., Evans, J. R., & Fletton, R. A. (1979). Structures of Three Novel B-Lactams isolated from *Streptomyces clavuligerus*. *Journal of Chemical Society, Chemical Communications*, (6) 282–283. doi.org/https://doi.org/10.1039/C39790000282

Burton, K. (1968). Determination of DNA Concentration with Diphenylamine. *Methods in Enzymology*, *12*(B), 163–166.

Bush, K, Jacoby, G. A., & Medeiros, A. A. (1995). A functional classification scheme for β-lactamases and its correlation with molecular structure. *Antimicrobial Agents and Chemotherapy*, *39*(6), 1211–1233. doi.org/10.1128/AAC.39.6.1211

Bush, K, Macalintal, C., Rasmussen, B. A., Lee, V. J., & Yang, Y. (1993). Kinetic interactions of tazobactam with β-lactamases from all major structural classes. *Antimicrobial Agents and Chemotherapy*, *37*(4), 851–858. doi.org/10.1128/AAC.37.4.851

Bush, Karen, & Bradford, P. A. (2016). β -Lactams and β -Lactamase Inhibitors: An Overview. *Cold Spring Harbor Perspectives in Medicine*, *6*(a025247). doi.org/10.1101/cshperspect.a025247

Bush, Karen, & Jacoby, G. A. (2010). Updated functional classification of β-lactamases. *Antimicrobial Agents and Chemotherapy*, *54*(3), 969–976. doi.org/10.1128/AAC.01009-09

Butterworth, D., Cole, M., Handscomb, G., & Rolinson, G. N. (1979). Olivanic acids, a family of β-lactam antibiotics with β-lactamase inhibitory properties produced by *Streptomyces* species. I. Detection, properties and fermentation studies. *The Journal of Antibiotics*, *32*(4), 287–294. doi.org/10.7164/antibiotics.32.287

Challis, G. L. (2005). A widely distributed bacterial pathway for siderophore biosynthesis

independent of nonribosomal peptide synthetases. *ChemBioChem*, *6*(4), 601–611. doi.org/10.1002/cbic.200400283

Challis, G. L., & Hopwood, D. a. (2003). Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proceedings of the National Academy of Sciences of the United States of America*, *100 Suppl*, 14555–14561. doi.org/10.1073/pnas.1934677100

Chater, K. F. (1993). Genetics of Differentiation in Streptomyces. *Annual Review of Microbiology*, *47*, 685–713.

Chater, K. F. (2006). *Streptomyces* inside-out: A new perspective on the bacteria that provide us with antibiotics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1469), 761–768. doi.org/10.1098/rstb.2005.1758

Chater, K. F. (2016). Recent advances in understanding *Streptomyces*. *F1000Research*, *5*(0), 2795. doi.org/10.12688/f1000research.9534.1

Chong, J., Soufan, O., Li, C., Caraus, I., Li, S., Bourque, G., … Xia, J. (2018). MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. *Nucleic Acids Research*, *46*(W1), W486–W494. doi.org/10.1093/nar/gky310

Chowdhury, S. M., Shi, L., Yoon, H., Ansong, C., M, L., Norbeck, A. D., … Smith, R. D. (2009). A method for investigating protein-protein interactions related to *Salmonella typhimurium* pathogenesis. *Proteome*, *8*(3), 1504–1514.

doi.org/10.1021/pr800865d.A

Circosta, C., De Pasquale, R., Samperi, S., Pino, A., & Occhiuto, F. (2007). Biological and analytical characterization of two extracts from *Valeriana officinalis*. *Journal of Ethnopharmacology*, *112*(2), 361–367. https://doi.org/10.1016/j.jep.2007.03.021

Curtis, T. P., Sloan, W. T., & Scannell, J. W. (2002). Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(16), 10494–10499. https://doi.org/10.1073/pnas.142680199

Czech, L., Hermann, L., Stöveken, N., Richter, A. A., Höppner, A., Smits, S. H. J., … Bremer, E. (2018). Role of the extremolytes ectoine and hydroxyectoine as stress protectants and nutrients: Genetics, phylogenomics, biochemistry, and structural analysis. *Genes*, *9*(4), 1–58. https://doi.org/10.3390/genes9040177

Davidsen, J. M., & Townsend, C. A. (2009). Identification and characterization of NocR as a positive transcriptional regulator of the β-Lactam Nocardicin A in *Nocardia uniformis*. *Journal of Bacteriology*, *191*(3), 1066–1077. https://doi.org/10.1128/JB.01833-07

De Kraker, J. W., Franssen, M. C. R., Joerink, M., De Groot, A., & Bouwmeester, H. J. (2002). Biosynthesis of costunolide, dihydrocostunolide, and leucodin. Demonstration of cytochrome P450-catalyzed formation of the lactone ring present in sesquiterpene lactones of chicory. *Plant Physiology*, *129*(1), 257–268. https://doi.org/10.1104/pp.010957

de la Fuente, A., Martín, J., Rodríguez-García, A., & Liras, P. (2004). Two Proteins with Ornithine Acetyltransferase Activity Show Different Functions in *Streptomyces clavuligerus*: Oat2 Modulates Clavulanic Acid Biosynthesis in Response to Arginine. *Journal of Bacteriology*, *186*(19), 6501–6507. https://doi.org/10.1128/JB.186.19.6501

de Lima Procópio, R. E., da Silva, I. R., Martins, M. K., de Azevedo, J. L., & de Araújo, J. M. (2012). Antibiotics produced by Streptomyces. *Brazilian Journal of Infectious Diseases*, *16*(5), 466–471. https://doi.org/10.1016/j.bjid.2012.08.014

Demain, A L, & Elander, R. P. (1999). The β-lactam antibiotics: past, present, and future. *Antonie van Leeuwenhoek*, *75*(1–2), 5–19. http://www.ncbi.nlm.nih.gov/pubmed/10422578

Dillman, R. O. (2004). Pentostatin (Nipent®) in the treatment of chronic lymphocyte leukemia and hairy cell leukemia. *Expert Review of Anticancer Therapy*, *4*(1), 27–36. https://doi.org/10.1586/14737140.4.1.27

Docquier, J. D., & Mangani, S. (2018). An update on β-lactamase inhibitor discovery and development. *Drug Resistance Updates*, *36*(June 2017), 13–29. https://doi.org/10.1016/j.drup.2017.11.002

Doroghazi, J. R., & Buckley, D. H. (2014). Intraspecies comparison of *Streptomyces pratensis* genomes reveals high levels of recombination and gene conservation between strains of disparate geographic origin. *BMC Genomics*, *15*, 1–14.

Doumith, M., Ellington, M. J., Livermore, D. M., & Woodford, N. (2009). Molecular mechanisms disrupting porin expression in ertapenem-resistant *Klebsiella* and *Enterobacter* spp. clinical isolates from the UK. *Journal of Antimicrobial Chemotherapy*, *63*(4), 659–667. https://doi.org/10.1093/jac/dkp029

Drawz, S. M., & Bonomo, R. A. (2010). Three decades of β-lactamase inhibitors. *Clinical Microbiology Reviews*, *23*(1), 160–201. https://doi.org/10.1128/CMR.00037-09

Egan, L. a., Busby, R. W., Iwata-Reuyl, D., & Townsend, C. a. (1997). Probable role of clavaminic acid as the terminal intermediate in the common pathway to clavulanic acid and the antipodal clavam metabolites. *Journal of the American Chemical Society*, *119*(10), 2348–2355. https://doi.org/10.1021/ja963107o

Ehmann, D. E., Jahić, H., Ross, P. L., Gu, R. F., Hu, J., Durand-Réville, T. F., … Fisher, S. L. (2013). Kinetics of avibactam inhibition against class A, C, and D β-lactamases. *Journal of Biological Chemistry*, *288*(39), 27960–27971. https://doi.org/10.1074/jbc.M113.485979

Eiamphungporn, W., Schaduangrat, N., Malik, A. A., & Nantasenamat, C. (2018). Tackling the antibiotic resistance caused by class a β-lactamases through the use of β-lactamase inhibitory protein. *International Journal of Molecular Sciences*, *19*(8). https://doi.org/10.3390/ijms19082222

English, A. R., Retsema, J. A., Girard, A. E., Lynch, J. E., & Barth, W. E. (1978). CP-45,899, a β-lactamase inhibitor that extends the antibacterial spectrum of β-lactams: Initial bacteriological characterization. *Antimicrobial Agents and Chemotherapy*,

14(3), 414–419. https://doi.org/10.1128/AAC.14.3.414

Ferguson, N. L., Peña, L., Moore, M. A., Bignell, D. R. D., & Tahlan, K. (2016). Proteomics analysis of global regulatory cascades involved in clavulanic acid production and morphological development in *Streptomyces clavuligerus*. *Journal of Industrial Microbiology & Biotechnology*, *43*(4), 537–555. https://doi.org/10.1007/s10295-016-1733-y

Fernández-Martínez, L. T., & Bibb, M. J. (2014). Use of the Meganuclease I-SceI of *Saccharomyces cerevisiae* to select for gene deletions in actinomycetes. *Scientific Reports*, *4*, 7100. https://doi.org/10.1038/srep07100

Fisher, J., Belasco, J. G., Charnas, R. L., Khosla, S., & Knowles, J. R. (1980). β-Lactamase inactivation by mechanism-based reagents. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *289*, 309–319.

Flärdh, K., & Buttner, M. J. (2009). *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. *Nature Reviews Microbiology*, *7*(1), 36–49. https://doi.org/10.1038/nrmicro1968

Fu, J., Qin, R., Zong, G., Liu, C., Kang, N., Zhong, C., & Cao, G. (2019). The CagRS Two-Component System Regulates Clavulanic Acid Metabolism via Multiple Pathways in *Streptomyces clavuligerus* F613-1. *Frontiers in Microbiology*, *10*(February), 1–17. https://doi.org/10.3389/fmicb.2019.00244

Fumimoto, R., Sakai, E., Yamaguchi, Y., Sakamoto, H., Fukuma, Y., Nishishita, K., …

Tsukuba, T. (2012). The coffee diterpene kahweol prevents osteoclastogenesis via impairment of NFATc1 expression and blocking of Erk phosphorylation. *Journal of Pharmacological Sciences*, *118*(4), 479–486. https://doi.org/10.1254/jphs.11212FP

Galleni, M., Lamotte-Brasseur, J., Raquet, X., Dubus, A., Monnaie, D., Knox, J. R., & Frère, J. M. (1995). The enigmatic catalytic mechanism of active-site serine β-lactamases. *Biochemical Pharmacology*, *49*(9), 1171–1178. https://doi.org/10.1016/0006-2952(94)00502-D

Gao, Y., Jiang, W., Dong, C., Li, C., Fu, X., Min, L., … Shen, J. (2012). Anti-inflammatory effects of sophocarpine in LPS-induced RAW 264.7 cells via NF-κB and MAPKs signaling pathways. *Toxicology in Vitro*, *26*(1), 1–6. https://doi.org/10.1016/j.tiv.2011.09.019

Geddes, A. M., Klugman, K. P., & Rolinson, G. N. (2007). Introduction: historical perspective and development of amoxicillin/clavulanate. *International Journal of Antimicrobial Agents*, *30*(SUPPL. 2), 109–112. https://doi.org/10.1016/j.ijantimicag.2007.07.015

Ghelardini, C., Galeotti, N., Di Cesare Mannelli, L., Mazzanti, G., & Bartolini, A. (2001). Local anaesthetic activity of β-caryophyllene. *Farmaco*, *56*(5–7), 387–389. https://doi.org/10.1016/S0014-827X(01)01092-8

Gunsior, M., Breazeale, S. D., Lind, A. J., Ravel, J., Janc, J. W., & Townsend, C. A. (2004). The Biosynthetic Gene Cluster for a Monocyclic Beta-Lactam Antibiotic, Nocardicin A. *Chemistry and Biology*, *11*(7), 927–938.

https://doi.org/http://dx.doi.org/10.1016/j.chembiol.2004.04.012

Gust, B., Challis, G. L., Fowler, K., Kieser, T., & Chater, K. F. (2003). PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(4), 1541–1546. https://doi.org/10.1073/pnas.0337542100

Hamed, R. B., Gomez-Castellanos, J. R., Henry, L., Ducho, C., McDonough, M. A., & Schofield, C. J. (2013). The enzymes of β-lactam biosynthesis. *Natural Product Reports*, *30*(1), 21–107. https://doi.org/10.1039/c2np20065a

Higgens, C. E., Hamill, R. L., Sands, T. H., Hoehn, M. M., Davis, N. E., Nagarajan, R., & Boeck, L. D. (1974). The occurrence of deacetoxycephalosporin C in fungi and streptomycetes. *The Journal of Antibiotics*, *27*(4), 298–300. https://doi.org/10.7164/antibiotics.27.298

Higgens, C. E., & Kastner, R. E. (1971). Streptomyces clavuligerus sp. nov., a β-lactam antibiotic producer. *International Journal of Systematic Bacteriology*, *21*(4), 326–331.

Higgins, P. G., Wisplinghoff, H., Stefanik, D., & Seifert, H. (2004). In Vitro Activities of the β-Lactamase Inhibitors Clavulanic Acid, Sulbactam, and Tazobactam Alone or in Combination with β-Lactams against Epidemiologically Characterized Multidrug-Resistant *Acinetobacter baumannii* Strains. *Antimicrobial Agents and Chemotherapy*, *48*(5), 1586–1592. https://doi.org/10.1128/AAC.48.5.1586–

1592.2004

Hopkins, J. M., & Towner, K. J. (1990). Enhanced resistance to cefotaxime and imipenem associated with outer membrane protein alterations in enterobacter aerogenes. *Journal of Antimicrobial Chemotherapy*, *25*(1), 49–55. https://doi.org/10.1093/jac/25.1.49

Hopwood, D. A. (1999). Forty years of genetics with Streptomyces: From *in vivo* through *in vitro* to *in silico*. *Microbiology*, *145*(9), 2183–2202. https://doi.org/10.1099/00221287-145-9-2183

Horinouchi, S. (2002). A Micorobial Hormone, A-Factor, As A Master Switch For Morphological Differentiation and Secondary Metabolism In *Streptomyces griesues*. *Frontiers in Bioscience, 7*, 2045–2057.

Hou, Z., Raymond, K. N., O'Sullivan, B., Esker, T. W., & Nishio, T. (1998). A Preorganized Siderophore: Thermodynamic and Structural Characterization of Alcaligin and Bisucaberin, Microbial Macrocyclic Dihydroxamate Chelating Agents. *Inorganic Chemistry*, *37*(26), 6630–6637. https://doi.org/10.1021/ic9810182

Hudzicki, J. (2009). Kirby-Bauer Disk Diffusion Susceptibility Test Protocol. *American Society For Microbiology*, 1–23. Retrieved from https://www.asm.org/Protocols/Kirby-Bauer-Disk-Diffusion-Susceptibility-Test-Pro

Humeniuk, C., Arlet, G., Gautier, V., Grimont, P., Labia, R., & Philippon, A. (2002). β-lactamases of *Kluyvera ascorbata*, probable progenitors of some plasmid-encoded

CTX-M types. *Antimicrobial Agents and Chemotherapy*, *46*(9), 3045–3049. https://doi.org/10.1128/AAC.46.9.3045-3049.2002

Hwang, S., Lee, N., Jeong, Y., Lee, Y., Kim, W., Cho, S., … Cho, B. K. (2019). Primary transcriptome and translatome analysis determines transcriptional and translational regulatory elements encoded in the *Streptomyces clavuligerus* genome. *Nucleic Acids Research*, *47*(12), 6114–6129. https://doi.org/10.1093/nar/gkz471

Hyun, S. K., Yong, J. L., Chang, K. L., Sun, U. C., Yeo, S. H., Yong, I. H., … Nihira, T. (2004). Cloning and characterization of a gene encoding the γ-butyrolactone autoregulator receptor from *Streptomyces clavuligerus*. *Archives of Microbiology*, *182*(1), 44–50. https://doi.org/10.1007/s00203-004-0697-x

Inoue, H., Nojima, H., & Okayama, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene*, *96*(1), 23–28. https://doi.org/10.1016/0378-1119(90)90336-P

Iqbal, A., Arunlanantham, H., Brown, T., Chowdhury, R., Clifton, I. J., Kershaw, N. J., … Schofield, C. J. (2010). Crystallographic and mass spectrometric analyses of a tandem GNAT protein from the clavulanic acid biosynthesis pathway. *Proteins: Structure, Function and Bioinformatics*, *78*(6), 1398–1407. https://doi.org/10.1002/prot.22653

Ishaque, M., & Kluepfel, D. (1980). Cellulase complex of a mesophilic *Streptomyces* strain. *Canadian Journal of Microbiology*, *26*(2), 183–189. http://www.ncbi.nlm.nih.gov/pubmed/6773643

Jacoby, G. A., Mills, D. M., & Chow, N. (2004). Role of β-lactamases and porins in resistance to ertapenem and other β-lactams in *Klebsiella pneumoniae*. *Antimicrobial Agents and Chemotherapy*, *48*(8), 3203–3206. https://doi.org/10.1128/AAC.48.8.3203-3206.2004

Jacqueline, C., Howland, K., & Chesnel, L. (2017). *In vitro* activity of ceftolozane/tazobactam in combination with other classes of antibacterial agents. *Journal of Global Antimicrobial Resistance*, *10*, 326–329. https://doi.org/10.1016/j.jgar.2017.04.003

Jakimowicz, D., & Van Wezel, G. P. (2012). Cell division and DNA segregation in *Streptomyces*: How to build a septum in the middle of nowhere? *Molecular Microbiology*, *85*(3), 393–404. https://doi.org/10.1111/j.1365-2958.2012.08107.x

Jensen, S., Elder, K. J., Aidoo, K. a, & Paradkar, A. S. (2000). Enzymes Catalyzing the Early Steps of Clavulanic Acid Biosynthesis Are Encoded by Two Sets of Paralogous Genes in *Streptomyces clavuligerus. Antimicrobial Agents and Chemotherapy*, *44*(3), 720–726. https://doi.org/10.1128/AAC.44.3.720-726.2000.Updated

Jensen, S E, & Paradkar, A. S. (1999). Biosynthesis and molecular genetics of clavulanic acid. *Antonie van Leeuwenhoek*, *75*(1–2), 125–133. https://doi.org/10.1023/A:1001755724055

Jensen, S E, Paradkar, A. S., Mosher, R. H., Anders, C., Beatty, P. H., Brumlik, M. J., … Barton, B. (2004a). Five Additional Genes Are Involved in Clavulanic Acid

Biosynthesis in *Streptomyces clavuligerus*. *Antimicrobial Agents and Chemotherapy*, *48*(1), 192–202. doi.org/10.1128/AAC.48.1.192

Jensen, Susan E., Wong, A., Griffin, A., & Barton, B. (2004b). *Streptomyces clavuligerus* Has a Second Copy of the Proclavaminate Amidinohydrolase Gene. *Antimicrobial Agents and Chemotherapy*, *48*(2), 514–520. doi.org/10.1128/AAC.48.2.514-520.2004

Jensen, Susan E. (2012). Biosynthesis of clavam metabolites. *Journal of Industrial Microbiology and Biotechnology*, *39*(10), 1407–1419. doi.org/10.1007/s10295-012-1191-0

Jensen, Susan E, Elder, K. J., Aidoo, K. a, & Paradkar, A. S. (2000). Enzymes Catalyzing the Early Steps of Clavulanic Acid Biosynthesis Are Encoded by Two Sets of Paralogous Genes in *Streptomyces clavuligerus*. *Antimicrobial Agents and Chemotherapy*, *44*(720–726). doi.org/10.1128/AAC.44.3.720-726.2000.

Jnawali, H. N., Oh, T. J., Liou, K., & Park, B. C. (2008). A two-component regulatory system involved in clavulanic acid production. *The Journal of Antibiotics*, *61*(11), 651–659. doi.org/10.1038/ja.2008.92

Jones, S. E., & Elliot, M. A. (2018). 'Exploring' the regulation of *Streptomyces* growth and development. *Current Opinion in Microbiology*, *42*, 25–30. doi.org/10.1016/j.mib.2017.09.009

Jones, S. E., Ho, L., Rees, C. A., Hill, J. E., Nodwell, J. R., & Elliot, M. A. (2017).

Streptomyces exploration is triggered by fungal interactions and volatile signals. *ELife*, *6*, 1–21. doi.org/10.7554/eLife.21738

Joris, B., Meester, F. D. E., Galleni, M., Reckinger, G., Coyette, J., & Frere, J. (1985). The β-lactamase of *Enterobacter cloacae* P99. *Biochemical Journal*, *228*, 241–248.

Kahan, J. S., Kahan, F. M., Goegelman, R., Currie, S. A., Jackson, M., Stapley, E. O., … Hernandez, S. (1979). Thienamycin, a new β-lactam antibiotic i. discovery, taxonomy, isolation and physical properties. *The Journal of Antibiotics*, *32*(1), 1–12. doi.org/10.7164/antibiotics.32.1

Kalantari, K., Moniri, M., Moghaddam, A. B., Rahim, R. A., Ariff, A. Bin, Izadiyan, Z., & Mohamad, R. (2017). A Review of the biomedical applications of zerumbone and the techniques for its extraction from ginger rhizomes. *Molecules*, *22*(10). doi.org/10.3390/molecules22101645

Kalinovskaya, N. I., Romanenko, L. A., Irisawa, T., Ermakova, S. P., & Kalinovsky, A. I. (2011). Marine isolate *Citricoccus* sp. KMM 3890 as a source of a cyclic siderophore nocardamine with antitumor activity. *Microbiological Research*, *166*(8), 654–661. doi.org/10.1016/j.micres.2011.01.004

Kaltenpoth, M. (2009). Actinobacteria as mutualists: general healthcare for insects? *Trends in Microbiology*, *17*(12), 529–535. https://doi.org/10.1016/j.tim.2009.09.006

Kanno, S. I., Tomizawa, A., Hiura, T., Osanai, Y., Shouji, A., Ujibe, M., … Ishikawa, M. (2005). Inhibitory effects of naringenin on tumor growth in human cancer cell lines

and sarcoma S-180-implanted mice. *Biological and Pharmaceutical Bulletin*, *28*(3), 527–530. doi.org/10.1248/bpb.28.527

Kelley, L. A., Mezulis, S., Yates, C., Wass, M., & Sternberg, M. (2015). The Phyre2 web portal for protein modelling, prediction, and analysis. *Nature Protocols*, *10*(6), 845–858. doi.org/10.1038/nprot.2015-053

Kenig, M., & Reading, C. (1979). Holomycin and An Antibiotic (MM 19290) Related To Tunicamycin, Mmetabolites of *Streptomyces clavuligerus*. *The Journal of Antibiotics*, *33*(6), 549–554. doi.org/10.1017/CBO9781107415324.004

Kershaw, N. J., McNaughton, H. J., Hewitson, K. S., Hernández, H., Griffin, J., Hughes, C., … Schofield, C. J. (2002). ORF6 from the clavulanic acid gene cluster of *Streptomyces clavuligerus* has ornithine acetyltransferase activity. *European Journal of Biochemistry*, *269*(8), 2052–2059. doi.org/10.1046/j.1432-1033.2002.02853.x

Khaleeli, N., Li, R., & Townsend, C. A. (1999). Origin of the β-lactam carbons in clavulanic acid from an unusual thiamine pyrophosphate-mediated reaction. *Journal of the American Chemical Society*, *121*(39), 9223–9224. doi.org/10.1021/ja9923134

Khan, M., Ding, C., Rasul, A., Yi, F., Li, T., Gao, H., … Ma, T. (2012). Isoalantolactone induces reactive oxygen species mediated apoptosis in pancreatic carcinoma PANC-1 cells. *International Journal of Biological Sciences*, *8*(4), 533–547. doi.org/10.7150/ijbs.3753

Khan, S. T., Komaki, H., Motohashi, K., Kozone, I., Mukai, A., Takagi, M., & Shin-Ya,

K. (2011). *Streptomyces* associated with a marine sponge *Haliclona* sp.; biosynthetic genes for secondary metabolites and products. *Environmental Microbiology*, *13*(2), 391–403. https://doi.org/10.1111/j.1462-2920.2010.02337.x

Kieser, T., Bibb, M., Buttner, M., Chater, K., & Hopwood, D. (2000). *Practical Streptomyces genetics*. Norwich, UK: John Innes Foundation.

Kim, H. S., & Park, Y. I. (2008). Isolation and identification of a novel microorganism producing the immunosuppressant tacrolimus. *Journal of Bioscience and Bioengineering*, *105*(4), 418–421. https://doi.org/10.1263/jbb.105.418

Kinoshita, H., Tsuji, T., Ipposhi, H., Nihira, T., & Yamada, Y. (1999). Characterization of binding sequences for butyrolactone autoregulator receptors in *Streptomycetes*. *Journal of Bacteriology*, *181*(16), 5075–5080. doi.org/10.1128/jb.181.16.5075-5080.1999

Kois-Ostrowska, A., Strzałka, A., Lipietta, N., Tilley, E., Zakrzewska-Czerwińska, J., Herron, P., & Jakimowicz, D. (2016). Unique Function of the Bacterial Chromosome Segregation Machinery in Apically Growing *Streptomyces* - Targeting the Chromosome to New Hyphal Tubes and its Anchorage at the Tips. *PLoS Genetics*, *12*(12), 1–25. doi.org/10.1371/journal.pgen.1006488

Kong, D., Wang, X., Nie, J., & Niu, G. (2019). Regulation of Antibiotic Production by Signaling Molecules in *Streptomyces*. *Frontiers in Microbiology*, *10*(December), 1–11. doi.org/10.3389/fmicb.2019.02927

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. doi.org/10.1093/molbev/msw054

Kwong, T., Zelyas, N. J., Cai, H., Tahlan, K., Wong, A., & Jensen, S. E. (2012). 5*S* clavam biosynthesis is controlled by an atypical two-component regulatory system in *Streptomyces clavuligerus*. *Antimicrobial Agents and Chemotherapy*, *56*(9), 4845–4855. https://doi.org/10.1128/AAC.01090-12

Labeda, D. P., Goodfellow, M., Brown, R., Ward, A. C., Lanoot, B., Vanncanneyt, M., … Hatano, K. (2012). Phylogenetic study of the species within the family Streptomycetaceae. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *101*(1), 73–104. doi.org/10.1007/s10482-011-9656-0

Labeda, David P., Dunlap, C. A., Rong, X., Huang, Y., Doroghazi, J. R., Ju, K. S., & Metcalf, W. W. (2017). Phylogenetic relationships in the family Streptomycetaceae using multi-locus sequence analysis. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *110*(4), 563–583. doi.org/10.1007/s10482-016-0824-0

Lan, X., Wang, W., Li, Q., & Wang, J. (2016). The Natural Flavonoid Pinocembrin: Molecular Targets and Potential Therapeutic Applications. *Molecular Neurobiology*, *53*(3), 1794–1801. doi.org/10.1007/s12035-015-9125-2

Lawlor, E. J., Baylis, H. A., & Chater, K. F. (1987). Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product

in *Streptomyces coelicolor* A3(2). *Genes & Development*, *1*(10), 1305–1310. doi.org/10.1101/gad.1.10.1305

Lee, H., Shin, H. J., Jang, K. H., Kim, T. S., Oh, K., & Shin, J. (2005). Cyclic Peptides of the Nocardamine Class from a Marine-Derived Bacterium of the Genus *Streptomyces*, 623–625.

Lewis, K. A., Tzilivakis, J., Warner, D. J., & Green, A. (2016). An international database for pesticide risk assessments and management. *Human and Ecological Risk Assessment*, *22*(4), 1050–1064. doi.org/10.1080/10807039.2015.1133242

Li, B., & Walsh, C. T. (2010). Identification of the gene cluster for the dithiolopyrrolone antibiotic holomycin in *Streptomyces clavuligerus*. *Proc Natl Acad Sci U S A*, *107*(46), 19731–19735. https://doi.org/10.1073/pnas.1014140107

Li, R., Khaleeli, N., & Townsend, C. A. (2000). Expansion of the clavulanic acid gene cluster: Identification and *in vivo* functional analysis of three new genes required for biosynthesis of clavulanic acid by *Streptomyces clavuligerus*. *Journal of Bacteriology*, *182*(14), 4087–4095. doi.org/10.1128/JB.182.14.4087-4095.2000

Li, R., Lloyd, E. P., Moshos, K. A., & Townsend, C. A. (2014). Identification and characterization of the carbapenem MM 4550 and its gene cluster in *Streptomyces argenteolus* ATCC 11009. *ChemBioChem*, *15*(2), 320–331. doi.org/10.1002/cbic.201300319

Li, X. Z., Zhang, L., & Poole, K. (2000). Interplay between the MexA-MexB-OprM

multidrug efflux system and the outer membrane barrier in the multiple antibiotic resistance of *Pseudomonas aeruginosa*. *Journal of Antimicrobial Chemotherapy*, *45*(4), 433–436. doi.org/10.1093/jac/45.4.433

Liu, C. L., Wang, J. M., Chu, C. Y., Cheng, M. T., & Tseng, T. H. (2002). *In vivo* protective effect of protocatechuic acid on tert-butyl hydroperoxide-induced rat hepatotoxicity. *Food and Chemical Toxicology*, *40*(5), 635–641. doi.org/10.1016/S0278-6915(02)00002-9

Liu, J. B., Chen, D., Bao, T. T., Fan, F. T., & Yu, C. (2020). The anticancer effects of atractylenolide III associate with the downregulation of Jak3/Stat3-dependent IDO expression. *Frontiers in Pharmacology*, *10*(January), 1–12. doi.org/10.3389/fphar.2019.01505

Liu, W.-B., Shi, Y., Yao, L.-L., Zhou, Y., & Ye, B.-C. (2013). Prediction and Characterization of Small Non-Coding RNAs Related to Secondary Metabolites in *Saccharopolyspora erythraea*. *PLoS ONE*, *8*(11), e80676. doi.org/10.1371/journal.pone.0080676

Long, J., Ding, Y. H., Wang, P. P., Zhang, Q., & Chen, Y. (2013). Protection-group-free semisyntheses of parthenolide and its cyclopropyl analogue. *Journal of Organic Chemistry*, *78*(20), 10512–10518. doi.org/10.1021/jo401606q

López-García, M. T., Santamarta, I., & Liras, P. (2010). Morphological differentiation and clavulanic acid formation are affected in a *Streptomyces clavuligerus adpA*-deleted mutant. *Microbiology*, *156*(8), 2354–2365. doi.org/10.1099/mic.0.035956-0

Lorenzana, L. M., Pérez-Redondo, R., Santamarta, I., Martín, J. F., & Liras, P. (2004). Two oligopeptide-permease-encoding genes in the clavulanic acid cluster of *Streptomyces clavuligerus* are essential for production of the β-lactamase inhibitor. *Journal of Bacteriology*, *186*(11), 3431–3438. doi.org/10.1128/JB.186.11.3431-3438.2004

Lougheed, K. E. A., Bennett, M. H., & Williams, H. D. (2014). An *in vivo* crosslinking system for identifying mycobacterial protein-protein interactions. *Journal of Microbiological Methods*, *105*, 67–71. doi.org/10.1016/j.mimet.2014.07.012

Lucas, X., Senger, C., Erxleben, A., Grüning, B. A., Döring, K., Mosch, J., … Günther, S. (2013). StreptomeDB: A resource for natural compounds isolated from *Streptomyces* species. *Nucleic Acids Research*, *41*(D1), 1130–1136. doi.org/10.1093/nar/gks1253

MacKenzie, A. K., Kershaw, N. J., Hernandez, H., Robinson, C. V., Schofield, C. J., & Andersson, I. (2007). Clavulanic acid dehydrogenase: Structural and biochemical analysis of the final step in the biosynthesis of the β-lactamase inhibitor clavulanic acid. *Biochemistry*, *46*(6), 1523–1533. doi.org/10.1021/bi061978x

MacKenzie, A. K., Valegård, K., Iqbal, A., Caines, M. E. C., Kershaw, N. J., Jensen, S. E., … Andersson, I. (2010). Crystal structures of an oligopeptide-binding protein from the biosynthetic pathway of the β-lactamase inhibitor clavulanic acid. *Journal of Molecular Biology*, *396*(2), 332–344. doi.org/10.1016/j.jmb.2009.11.045

MacKenzie, C. R., Bilous, D., & Johnson, K. G. (1984). Purification and characterization of an exoglucanase from *Streptomyces flavogriseus*. *Canadian Journal of*

Microbiology, *30*(9), 1171–1178. doi.org/10.1139/m84-183

MacNeil, D. J., Gewain, K. M., Ruby, C. L., Dezeny, G., Gibbons, P. H., & MacNeil, T. (1992). Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene*, *111*(1), 61–68. doi.org/10.1016/0378-1119(92)90603-M

Marquez-Martin, A., Puerta, R. D. La, Fernandez-Arche, A., Ruiz-Gutierrez, V., & Yaqoob, P. (2006). Modulation of cytokine secretion by pentacyclic triterpenes from olive pomace oil in human mononuclear cells. *Cytokine*, *36*(5–6), 211–217. doi.org/10.1016/j.cyto.2006.12.007

Martínez-Burgo, Y., Álvarez-Álvarez, R., Rodríguez-García, A., & Liras, P. (2015). The 'pathway specific' regulator ClaR of *Streptomyces clavuligerus* has a global effect on the expression of genes for secondary metabolism and differentiation. *Applied and Environmental Microbiology*, *81*(19), 6637-6648. doi.org/10.1128/AEM.00916-15

Martínez-Burgo, Y., Santos-Aberturas, J., Rodríguez-García, A., Barreales, E. G., Tormo, J. R., Truman, A. W., … Liras, P. (2019). Activation of secondary metabolite gene clusters in *Streptomyces clavuligerus* by the pimm regulator of streptomyces natalensis. *Frontiers in Microbiology*, *10*(MAR), 1–14. doi.org/10.3389/fmicb.2019.00580

Maveyraud, L., Pratt, R. F., & Samama, J. P. (1998). Crystal structure of an acylation transition-state analog of the TEM-1 β-lactamase. Mechanistic implications for class

A β-lactamases. *Biochemistry*, *37*(8), 2622–2628. doi.org/10.1021/bi972501b

McCartney, L., Blake, A. W., Flint, J., Bolam, D. N., Boraston, A. B., Gilbert, H. J., & Knox, J. P. (2006). Differential recognition of plant cell walls by microbial xylan-specific carbohydrate-binding modules. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(12), 4765–4770. doi.org/10.1073/pnas.0508887103

Medema, M. H., Trefzer, A., Kovalchuk, A., Van Den Berg, M., Müller, U., Heijne, W., … Takano, E. (2010). The sequence of a 1.8-Mb bacterial linear plasmid reveals a rich evolutionary reservoir of secondary metabolic pathways. *Genome Biology and Evolution*, *2*(1), 212–224. doi.org/10.1093/gbe/evq013

Mellado, E., Lorenzana, L. M., Rodríguez-Sáiz, M., Díez, B., Liras, P., & Barredo, J. L. (2002). The clavulanic acid biosynthetic cluster of *Streptomyces clavuligerus*: Genetic organization of the region upstream of the car gene. *Microbiology*, *148*(5), 1427–1438.

Melland, A., & McLaren, D. (1998). Efficacy of herbicides against serrated tussock (*Nassella trichotoma*) in a pot trial. *Plant Protection Quarterly*, *13*(2), 3052.

Miller, J. M., Baker, C. N., & Thornsberry, C. (1978). Inhibition of β-lactamase in *Neisseria gonorrhoeae* by sodium clavulanate. *Antimicrobial Agents and Chemotherapy*, *14*(5), 794–796. doi.org/10.1128/AAC.14.5.794

Mølgaard, A., Kauppinen, S., & Larsen, S. (2000). Rhamnogalacturonan acetylesterase

elucidates the structure and function of a new family of hydrolases. *Structure, 8*(4), 373–383. doi.org/10.1016/S0969-2126(00)00118-0

Mosher, R. H., Paradkar, A. S., Anders, C., Barton, B., & Jensen, S. E. (1999). Genes specific for the biosynthesis of clavam metabolites antipodal to clavulanic acid are clustered with the gene for clavaminate synthase 1 in *Streptomyces clavuligerus*. *Antimicrobial Agents and Chemotherapy*, *43*(5), 1215–1224.

Motamedi, H., Shafiee, A., & Cai, S. (1995). Integrative vectors for heterologous gene expression in. *GENE*, *160*, 25–31.

Mushtaq, S., Vickers, A., Woodford, N., Haldimann, A., & Livermore, D. M. (2019). Activity of nacubactam (RG6080/OP0595) combinations against MBL-producing enterobacteriaceae. *Journal of Antimicrobial Chemotherapy*, *74*(4), 953–960. doi.org/10.1093/jac/dky522

Mussi, M. A., Limansky, A. S., & Viale, A. M. (2005). Acquisition of resistance to carbapenems in multidrug-resistant clinical strains of *Acinetobacter baumannii*: Natural insertional inactivation of a gene encoding a member of a novel family of β-barrel outer membrane proteins. *Antimicrobial Agents and Chemotherapy*, *49*(4), 1432–1440. doi.org/10.1128/AAC.49.4.1432-1440.2005

Naas, T., Oueslati, S., Bonnin, R. A., Dabos, M. L., Zavala, A., Dortet, L., … Iorga, B. I. (2017). Beta-lactamase database (BLDB)–structure and function. *Journal of Enzyme Inhibition and Medicinal Chemistry*, *32*(1), 917–919. doi.org/10.1080/14756366.2017.1344235

Naas, T., Poirel, L., & Nordmann, P. (2008). Minor extended-spectrum β-lactamases. *Clinical Microbiology and Infection*, *14*(SUPPL. 1), 42–52. doi.org/10.1111/j.1469-0691.2007.01861.x

Nagarajan, R., Boeck, L. D., Gorman, M., Hamill, R. L., Higgens, C. E., Hoehn, M. M., … Whitney, J. G. (1971). β-Lactam Antibiotics from *Streptomyces*. *Journal of the American Chemical Society*, *93*(9), 2308–2310. doi.org/10.1021/ja00738a035

Neil Marsh, E., Chang, M. D. T., & Townsend, C. A. (1992). Two Isozymes of Clavaminate Synthase Central to Clavulanic Acid Formation: Cloning and Sequencing of Both Genes from *Streptomyces clavuligerus*. *Biochemistry*, *31*(50), 12648–12657. doi.org/10.1021/bi00165a015

Netolitzky, D. J., Wu, X., Jensen, S. E., & Roy, K. L. (1995). Giant linear plasmids of β-lactam antibiotic producing Streptomyces. *FEMS Microbiology Letters*, *131*(1), 27–34. doi.org/10.1016/0378-1097(95)00230-3

Nett, M., Ikeda, H., & Moore, B. S. (2009). Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Natural Product Reports*, *26*(11), 1362–1384. doi.org/10.1039/b817069j

Newton, G. G. F., & Abraham, E. P. (1955). Cephalosporin C, a new antibiotic containing sulphur and D-α-aminoadipic acid. *Nature*, *175*(4456), 548. doi.org/10.1038/175548a0

Oh, D. C., Poulsen, M., Currie, C. R., & Clardy, J. (2011). Sceliphrolactam, a polyene

macrocyclic lactam from a wasp-associated *Streptomyces* sp. *Organic Letters*, *13*(4), 752–755. doi.org/10.1021/ol102991d

Onaka, H. (2017a). Novel antibiotic screening methods to awaken silent or cryptic secondary metabolic pathways in actinomycetes. *Journal of Antibiotics*, *70*(8), 865–870. doi.org/10.1038/ja.2017.51

Onaka, H. (2017b). Novel antibiotic screening methods to awaken silent or cryptic secondary metabolic pathways in actinomycetes. *Journal of Antibiotics*, *70*(8), 865–870. doi.org/10.1038/ja.2017.51

Page, M. G. (2007). Resistance mediated by penicllin-binding proteins. In R. A. Bonomo & M. E. Tolmasky (Eds.), *Enzyme-mediated resistance to antibiotics: mechanisms, dissemination, and prospects for Inhibition* (pp. 81–99). Washington, DC: ASM Press.

Paget, M. S. B., Chamberlin, L., Atrih, A., Foster, S. J., & Buttner, M. J. (1999). Evidence that the Extracytoplasmic Function Sigma Factor σ^E Is Required for Normal Cell Wall Structure in *Streptomyces coelicolor* A3 (2). *Journal of Bacteriology*, *181*(1), 204–211.

Palzkill, T. (2013). Metallo-β-lactamase structure and function. *Annals of the New York Academy of Sciences*, *1277*(1), 91–104. doi.org/10.1111/j.1749-6632.2012.06796.x

Pan, R., Bai, X., Chen, J., Zhang, H., & Wang, H. (2019). Exploring structural diversity of microbe secondary metabolites using OSMAC strategy: A literature review.

Frontiers in Microbiology, *10*(FEB), 1–20. doi.org/10.3389/fmicb.2019.00294

Papp-Wallace, K. M., Endimiani, A., Taracila, M. A., & Bonomo, R. A. (2011). Carbapenems: Past, present, and future. *Antimicrobial Agents and Chemotherapy*, *55*(11), 4943–4960. doi.org/10.1128/AAC.00296-11

Paradkar, A S, Griffin, A., Griffin, J., Hughes, C., Greaves, P., Barton, B., & Jensen, S. E. (2001). Applications of Gene Replacement Technology to *Streptomyces clavuligerus* Strain Development for Clavulanic Acid Production. *Applied and Environmental Microbiology*, *67*(5), 2292–2297. doi.org/10.1128/AEM.67.5.2292

Paradkar, AS, Aidoo, K. A., Wong, A., & Jensen, S. E. (1996). Molecular analysis of a β-lactam resistance gene encoded within the cephamycin gene cluster of Streptomyces clavuligerus. *Journal of Bacteriology*, *178*(21), 6266–6274.

Paradkar, Ashish. (2013). Clavulanic acid production by *Streptomyces clavuligerus*: Biogenesis, regulation and strain improvement. *Journal of Antibiotics*, *66*(7), 411–420. doi.org/10.1038/ja.2013.26

Paradkar, Ashish, & Jensen, S. E. (1995). Functional Analysis of the Gene Encoding the Clavaminate Synthase 2 Isoenzyme Involved in Clavulanic Acid Biosynthesis in *Streptomyces clavuligerus*. *Journal of Bacteriology*, *177*(5), 1307–1314.

Paradkar, Ashish S., Aidoo, K. A., & Jensen, S. E. (1998). A pathway-specific transcriptional activator regulates late steps of clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *Molecular Microbiology*, *27*(4), 831–843.

doi.org/10.1046/j.1365-2958.1998.00731.x

Paradkar, Ashish, Trefzer, A., Chakraburtty, R., & Stassi, D. (2003). *Streptomyces* genetics: a genomic perspective. *Critical Reviews in Biotechnology*, *23*(1), 1–27. doi.org/10.1080/713609296

Parveen, M., Harish, B. N., & Parija, S. C. (2010). AmpC beta lactamases among gram negative clinical isolates from a tertiary hospital, South India. *Brazilian Journal of Microbiology*, *41*(3), 596–602. doi.org/10.1590/s1517-83822010000300009

Peraud, O., Biggs, J. S., Hughen, R. W., Light, A. R., Concepcion, G. P., Olivera, B. M., & Schmidt, E. W. (2009). Microhabitats within venomous cone snails contain diverse actinobacteria. *Applied and Environmental Microbiology*, *75*(21), 6820–6826. doi.org/10.1128/AEM.01238-09

Perez-Llarena, F. J., Liras, P., Rodriguez-Garcia, A., & Martin, J. F. (1997). A regulatory gene (*ccaR*) required for cephamycin and clavulanic acid production in *Streptomyces clavuligerus*: amplification results in overproduction of both β-lactam compounds. *Journal of Bacteriology*, *179*(6), 2053–2059.

Pérez-Redondo, R., Rodríguez-García, A., Martín, J. F., & Liras, P. (1998). The *claR* gene of *Streptomyces clavuligerus*, encoding a LysR-type regulatory protein controlling clavulanic acid biosynthesis, is linked to the clavulanate-9-aldehyde reductase (*car*) gene. *Gene*, *211*(2), 311–321. doi.org/10.1016/S0378-1119(98)00106-1

Pérez-Redondo, R., Rodríguez-García, A., Marťin, J. F., & Liras, P. (1999). Deletion of the *pyc* gene blocks clavulanic acid biosynthesis except in glycerol-containing medium: Evidence for two different genes in formation of the C3 unit. *Journal of Bacteriology*, *181*(22), 6922–6928.

Plaskitt, K. A., & Chater, K. F. (1995). Influences of developmental genes on localized glycogen deposition in colonies of a mycelial prokaryote, *Streptomyces coelicolor* A3(2): a possible interface between metabolism and morphogenesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *347*, 105–121.

Pluskal, T., Castillo, S., Villar-Briones, A., & Orešič, M. (2010). MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics*, *11*(395). doi.org/10.1186/1471-2105-11-395

Podder, M. P., Rogers, L., Daley, P. K., Keefe, G. P., Whitney, H. G., & Tahlan, K. (2014). *Klebsiella* species associated with bovine mastitis in Newfoundland. *PLoS ONE*, *9*(9), 1–5. doi.org/10.1371/journal.pone.0106518

Poole, K. (2004). Efflux-mediated multiresistance in Gram-negative bacteria. *Clinical Microbiology and Infection*, *10*(1), 12–26. doi.org/10.1111/j.1469-0691.2004.00763.x

Popoola, O. K., Elbagory, A. M., Ameer, F., & Hussein, A. A. (2013). Marrubiin. *Molecules*, *18*(8), 9049–9060. https://doi.org/10.3390/molecules18089049

Pratt, R. F., & McLeish, M. J. (2010). Structural relationship between the active sites of β-Lactam- recognizing and amidase signature enzymes: Convergent evolution? *Biochemistry*, *49*(45), 9688–9697. doi.org/10.1021/bi1012222

Pruess, D. L., & Kellett, M. (1983). Ro 22-5417, a new clavam antibiotic from *Streptomyces clavuligerus*. I. Discovery and biological activity. *J Antibiot (Tokyo)*, *36*(3), 208–212.

Queenan, A. M., & Bush, K. (2007). Carbapenemases: The versatile β-lactamases. *Clinical Microbiology Reviews*, *20*(3), 440–458. doi.org/10.1128/CMR.00001-07

Quinn, G. A., Banat, A. M., Abdelhameed, A. M., & Banat, I. M. (2020). *Streptomyces* from traditional medicine: sources of new innovations in antibiotic discovery. *Journal of Medical Microbiology*, *69*(8), 1040–1048. doi.org/10.1099/jmm.0.001232

Ramirez-Malule, H. (2018). Bibliometric analysis of global research on clavulanic acid. *Antibiotics*, *7*(4). doi.org/10.3390/antibiotics7040102

Rauha, J.-P., Remes, S., Heinonen, M., Hopia, A., Kahkonen, M., Kujala, T., … Vuorela, P. (2000). Antimicrobial effects of Finnish plant extracts containing flavonoids and other phenolic compounds. *International Journal of Food Microbiology*, *56*, 3–12.

Reading, C., & Cole, M. (1977). Clavulanic acid: a β-lactamase-inhiting β-lactam from *Streptomyces clavuligerus*. *Antimicrobial Agents and Chemotherapy*, *11*(5), 852–857. doi.org/10.1128/AAC.11.5.852

Ren, H., Wang, B., & Zhao, H. (2017). Breaking the silence: new strategies for

discovering novel natural products. *Current Opinion in Biotechnology*, *48*, 21–27. doi.org/10.1016/j.copbio.2017.02.008

Rigali, S., Anderssen, S., Naômé, A., & van Wezel, G. P. (2018). Cracking the regulatory code of biosynthetic gene clusters as a strategy for natural product discovery. *Biochemical Pharmacology*, *153*, 24–34. doi.org/10.1016/j.bcp.2018.01.007

Rodriguez, M., Munez, L. E., Brana, A. F., Mendez, C., Salas, J. A., & Blanco, G. (2008). Identification of transcriptional activators for thienamycin and cephamycin C biosynthetic genes within the thienamycin gene cluster from *Streptomyces cattleya*. *Molecular Microbiology*, *69*(3), 633–645.

Rolinson, G. N. (1991). Evolution of β-Lactamase Inhibitors. *Reviews of Infectious Diseases*, *13*(9), S727–S732.

Romagnoli, S., & Tabita, F. R. (2007). Phosphotransfer reactions of the CbbRRS three-protein two-component system from Rhodopseudomonas palustris CGA010 appear to be controlled by an internal molecular switch on the sensor kinase. *Journal of Bacteriology*, *189*(2), 325–335. doi.org/10.1128/JB.01326-06

Romero, J., Liras, P., & Martin, J. F. (1984). Dissociation of cephamycin and clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *Applied Microbiology and Biotechnology*, *20*, 318–325. doi.org/10.1007/BF00270593

Rong, X., Doroghazi, J. R., Cheng, K., Zhang, L., Buckley, D. H., & Huang, Y. (2013). Classification of Streptomyces phylogroup pratensis (Doroghazi and Buckley, 2010)

based on genetic and phenotypic evidence, and proposal of *Streptomyces pratensis* sp. nov. *Systematic and Applied Microbiology*, *36*(6), 401–407. doi.org/10.1016/j.syapm.2013.03.010

Rouseff, R. L., Youtsey, C. O., & Martin, S. F. (1987). Quantitative Survey of Narirutin, Naringin, Hesperidin, and Neohesperidin in Citrus. *Journal of Agricultural and Food Chemistry*, *35*(6), 1027–1030. doi.org/10.1021/jf00078a040

Sadeghi, A., Soltani, B. M., Nekouei, M. K., Jouzani, G. S., Mirzaei, H. H., & Sadeghizadeh, M. (2014). Diversity of the ectoines biosynthesis genes in the salt tolerant Streptomyces and evidence for inductive effect of ectoines on their accumulation. *Microbiological Research*, *169*(9–10), 699–708. doi.org/10.1016/j.micres.2014.02.005

Salowe, S. P., Krol, W. J., Iwata-Reuyl, D., & Townsend, C. A. (1991). Elucidation of the Order of Oxidations and Identification of an Intermediate in the Multistep Clavaminate Synthase Reaction. *Biochemistry*, *30*(8), 2281–2292. doi.org/10.1021/bi00222a034

Sambrook, J., & Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.

Santamarta, I., López-García, M. T., Kurt, A., Nárdiz, N., Álvarez-Álvarez, R., Pérez-Redondo, R., … Liras, P. (2011). Characterization of DNA-binding sequences for CcaR in the cephamycin-clavulanic acid supercluster of *Streptomyces clavuligerus*. *Molecular Microbiology*, *81*(4), 968–981. doi.org/10.1111/j.1365-

2958.2011.07743.x

Santamarta, Irene, López-García, M. T., Pérez-Redondo, R., Koekman, B., Martín, J. F., & Liras, P. (2007). Connecting primary and secondary metabolism: AreB, an IclR-like protein, binds the AREccaR sequence of *S. clavuligerus* and modulates leucine biosynthesis and cephamycin C and clavulanic acid production. *Molecular Microbiology*, *66*(2), 511–524. doi.org/10.1111/j.1365-2958.2007.05937.x

Saudagar, P. S., Survase, S. A., & Singhal, R. S. (2008). Clavulanic acid: A review. *Biotechnology Advances*, *26*(4), 335–351. doi.org/10.1016/j.biotechadv.2008.03.002

Sauvage, E., Kerff, F., Terrak, M., Ayala, J. A., & Charlier, P. (2008). The penicillin-binding proteins: Structure and role in peptidoglycan biosynthesis. *FEMS Microbiology Reviews*, *32*(2), 234–258. doi.org/10.1111/j.1574-6976.2008.00105.x

Seipke, R. F., Barke, J., Brearley, C., Hill, L., Yu, D. W., Goss, R. J. M., & Hutchings, M. I. (2011). A single Streptomyces symbiont makes multiple antifungals to support the fungus farming ant acromyrmex octospinosus. *PLoS ONE*, *6*(8), 4–11. doi.org/10.1371/journal.pone.0022028

Seipke, R. F., Kaltenpoth, M., & Hutchings, M. I. (2012). *Streptomyces* as symbionts: An emerging and widespread theme? *FEMS Microbiology Reviews*, *36*(4), 862–876. https://doi.org/10.1111/j.1574-6976.2011.00313.x

Servín-González, L., Castro, C., Pérez, C., Rubio, M., & Valdez, F. (1997). bldA-dependent expression of the *Streptomyces exfoliatus* M11 lipase gene (lipA) is

mediated by the product of a contiguous gene, *lipR*, encoding a putative transcriptional activator. *Journal of Bacteriology*, *179*(24), 7816–7826. doi.org/10.1128/jb.179.24.7816-7826.1997

Shah, A. M., Shakeel-u-Rehman, Hussain, A., Mushtaq, S., Rather, M. A., Shah, A., … Hassan, Q. P. (2017). Antimicrobial investigation of selected soil actinomycetes isolated from unexplored regions of Kashmir Himalayas, India. *Microbial Pathogenesis*, *110*, 93–99. doi.org/10.1016/j.micpath.2017.06.017

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, R., … Ideker, T. (2003). Cytoscape: A Software Environment for Integrated Models. *Genome Research*, *13*, 2498–2504. doi.org/10.1101/gr.1239303.metabolite

Sharma, V., Qayum, A., Kaul, S., Singh, A., Kapoor, K. K., Mukherjee, D., … Dhar, M. K. (2019). Carbohydrate Modifications of Neoandrographolide for Improved Reactive Oxygen Species-Mediated Apoptosis through Mitochondrial Pathway in Colon Cancer. *ACS Omega*, *4*(24), 20435–20442. doi.org/10.1021/acsomega.9b01249

Shiono, Y., Hiramatsu, F., Murayama, T., Koseki, T., Funakoshi, T., Ueda, K., & Yasuda, H. (2007). Two drimane-type sesquiterpenes, strobilactones A and B, from the liquid culture of the edible mushroom *Strobilurus ohshimae*. *Zeitschrift Fur Naturforschung - Section B Journal of Chemical Sciences*, *62*(12), 1585–1589. doi.org/10.1515/znb-2007-1218

Singh, S. B., & Pelaez, F. (2008). Biodiversity, chemical diversity and drug discovery.

Progress in Drug Research, *65*, 143–174.

Sleeman, M. C., & Schofield, C. J. (2004). Carboxymethylproline Synthase (CarB), an Unusual Carbon-Carbon Bond-forming Enzyme of the Crotonase Superfamily Involved in Carbapenem Biosynthesis. *Journal of Biological Chemistry*, *279*(8), 6730–6736. doi.org/10.1074/jbc.M311824200

Snee, R. D. (1972). On the analysis of response curve data. *Technometrics*, *14*(1), 47–62. doi.org/10.1080/00401706.1972.10488882

Song, J. Y., Jensen, S. E., & Lee, K. J. (2010). Clavulanic acid biosynthesis and genetic manipulation for its overproduction. *Applied Microbiology and Biotechnology*, *88*(3), 659–669. doi.org/10.1007/s00253-010-2801-2

Song, J. Y., Kim, E. S., Kim, D. W., Jensen, S. E., & Lee, K. J. (2009). 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. *Journal of Industrial Microbiology and Biotechnology*, *36*(2), 301–311. doi.org/10.1007/s10295-008-0499-2

Spicer, R. A., Salek, R., & Steinbeck, C. (2017). Compliance with minimum information guidelines in public metabolomics repositories. *Scientific Data*, *4*(September), 1–8. doi.org/10.1038/sdata.2017.137

Srivastava, S. K., King, K. S., AbuSara, N. F., Malayny, C. J., Piercey, B. M., Wilson, J. A., & Tahlan, K. (2019). In vivo functional analysis of a class A β-lactamase-related

protein essential for clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *PLoS ONE*, *14*(4), 1–23. doi.org/10.1371/journal.pone.0215960

Stapley, E. O., Cassidy, P. J., Tunac, J., Monaghan, R. L., Jackson, M., Hernandez, S., … Hendlin, D. (1981). Epithienamycins - Novel β-lactams Related to Thienamycin, I. Production and Antibacterial Activity. *The Jornal of Antibiotics*, *XXIV*(6), 628–636. doi.org/10.1046/j.1365-2672.2002.01707.x

Taechowisan, T., Peberdy, J. F., & Lumyong, S. (2003). Isolation of endophytic actinomycetes from selected plants and their antifungal activity. *World Journal of Microbiology and Biotechnology*, *19*(4), 381–385. doi.org/10.1023/A:1023901107182

Tahlan, K., Anders, C., & Jensen, S. E. (2004). The paralogous pairs of genes involved in clavulanic acid and clavam metabolite biosynthesis are differently regulated in *Streptomyces clavuligerus*. *Journal of Bacteriology*, *186*(18), 6286–6297. doi.org/10.1128/JB.186.18.6286-6297.2004

Tahlan, K., Anders, C., Wong, A., Mosher, R. H., Beatty, P. H., Brumlik, M. J., … Jensen, S. E. (2007). 5*S* Clavam Biosynthetic Genes Are Located in Both the Clavam and Paralog Gene Clusters in *Streptomyces clavuligerus*. *Chemistry and Biology*, *14*(2), 131–142. doi.org/10.1016/j.chembiol.2006.11.012

Tahlan, K., & Jensen, S. E. (2013). Origins of the β-lactam rings in natural products. *Journal of Antibiotics*, *66*(7), 401–410. doi.org/10.1038/ja.2013.24

Tahlan, K., Moore, M. A., & Jensen, S. E. (2017). δ-(l-α-aminoadipyl)-l-cysteinyl-d-valine synthetase (ACVS): discovery and perspectives. *Journal of Industrial Microbiology and Biotechnology*, *44*(4–5), 517–524. doi.org/10.1007/s10295-016-1850-7

Tahlan, K., Park, H. U., & Jensen, S. E. (2004). Three unlinked gene clusters are involved in clavam metabolite biosynthesis in *Streptomyces clavuligerus*. *Canadian Journal of Microbiology*, *50*(10), 803–810. doi.org/10.1139/w04-070

Tahlan, K., Park, H. U., Wong, A., Perrin, H. B., & Jensen, S. E. (2004). Two Sets of Paralogous Genes Encode the Enzymes Involved in the Early Stages of Clavulanic Acid and Clavam Metabolite Biosynthesis in *Streptomyces clavuligerus*. *Antimicrobial Agents and Chemotherapy*, *48*(3), 930. doi.org/10.1128/AAC.48.3.930

Takano, E., Tao, M., Long, F., Bibb, M. J., Wang, L., Li, W., … Chater, K. F. (2003). A rare leucine codon in adpA is implicated in the morphological defect of bldA mutants of *Streptomyces coelicolor*. *Molecular Microbiology*, *50*(2), 475–486. https://doi.org/10.1046/j.1365-2958.2003.03728.x

Thomson, K. S. (2010). Extended-spectrum-β-lactamase, AmpC, and carbapenemase issues. *Journal of Clinical Microbiology*, *48*(4), 1019–1025. https://doi.org/10.1128/JCM.00219-10

Tooke, C. L., Hinchliffe, P., Bragginton, E. C., Colenso, C. K., Hirvonen, V. H. A., Takebayashi, Y., & Spencer, J. (2019). β-Lactamases and β-Lactamase Inhibitors in the 21st Century. *Journal of Molecular Biology*, *431*(18), 3472–3500.

doi.org/10.1016/j.jmb.2019.04.002

Toth, M., Antunes, N. T., Stewart, N. K., Frase, H., Bhattacharya, M., Smith, C. A., & Vakulenko, S. B. (2016). Class D β-lactamases do exist in Gram-positive bacteria. *Nature Chemical Biology*, *12*(1), 9–14. doi.org/10.1038/nchembio.1950

Toussaint, K. A., & Gallagher, J. C. (2015). β-Lactam/β-Lactamase Inhibitor Combinations: From Then to Now. *Annals of Pharmacotherapy*, *49*(1), 86–98. doi.org/10.1177/1060028014556652

Traxler, M. F., & Kolter, R. (2015). Natural products in soil microbe interactions and evolution. *Nat. Prod. Rep.*, *32*(7), 956–970. doi.org/10.1039/C5NP00013K

Trepanier, N. K., Jensen, S. E., Alexander, D. C., & Leskiw, B. K. (2002). The positive activator of cephamycin C and clavulanic acid production in *Streptomyces clavuligerus* is mistranslated in a bldA mutant. *Microbiology*, *148*(3), 643–656. doi.org/10.1099/00221287-148-3-643

Tyers, M., & Wright, G. D. (2019). Drug combinations: a strategy to extend the life of antibiotics in the 21st century. *Nature Reviews Microbiology*, *17*(3), 141–155. doi.org/10.1038/s41579-018-0141-x

Upton, C., & Buckley, J. T. (1995). A new family of lipolytic. *Trends Biochem*, 178–179. doi.org/10.1016/S0968-0004(00)89002-7

Valegård, K., Iqbal, A., Kershaw, N. J., Ivison, D., Généreux, C., Dubus, A., … Mcdonough, M. a. (2013). Structural and mechanistic studies of the *orf12* gene

product from the clavulanic acid biosynthesis pathway. *Acta Crystallographica Section D: Biological Crystallography*, *69*(8), 1567–1579. doi.org/10.1107/S0907444913011013

Van Bueren, A. L., Morland, C., Gilbert, H. J., & Boraston, A. B. (2005). Family 6 carbohydrate binding modules recognize the non-reducing end of β-1,3-linked glucans by presenting a unique ligand binding surface. *Journal of Biological Chemistry*, *280*(1), 530–537. doi.org/10.1074/jbc.M410113200

van der Meij, A., Worsley, S. F., Hutchings, M. I., & van Wezel, G. P. (2017). Chemical ecology of antibiotic production by actinomycetes. *FEMS Microbiology Reviews*, *41*(3), 392–416. doi.org/10.1093/femsre/fux005

van Wezel, G. P., & McDowall, K. J. (2011). The regulation of the secondary metabolism of *Streptomyces*: new links and experimental advances. *Natural Product Reports*, *28*(7), 1311. doi.org/10.1039/c1np00003a

Vandavasi, V. G., Weiss, K. L., Cooper, J. B., Erskine, P. T., Tomanicek, S. J., Ostermann, A., … Coates, L. (2016). Exploring the Mechanism of β-Lactam Ring Protonation in the Class A β-lactamase Acylation Mechanism Using Neutron and X-ray Crystallography. *Journal of Medicinal Chemistry*, *59*(1), 474–479. doi.org/10.1021/acs.jmedchem.5b01215

Vetting, M. W., Roderick, S. L., Yu, M., & Blanchard, J. S. (2003). Crystal structure of mycothiol synthase (Rv0819) from. *Protein Science*, *12*, 1954–1959. doi.org/10.1110/ps.03153703.codes

Viana-Marques, D. de A., Machado, S. E. F., Carvalho Santos Ebinuma, V., Duarte, C. de A. L., Converti, A., & Porto, A. L. F. (2018). Production of β-lactamase inhibitors by *Streptomyces* species. *Antibiotics*, *7*(3), 1–26. doi.org/10.3390/antibiotics7030061

Vujaklija, D., Schröder, W., Abramić, M., Zou, P., Leščić, I., Franke, P., & Pigac, J. (2002). A novelStreptomycete lipase: Cloning, sequencing and high-level expression of the *Streptomyces rimosus* GDS(L)-lipase gene. *Archives of Microbiology*, *178*(2), 124–130. doi.org/10.1007/s00203-002-0430-6

Wagner, U. G., Petersen, E. I., Schwab, H., & Kratky, C. (2009). EstB from *Burkholderia gladioli*: A novel esterase with a β-lactamase fold reveals steric factors to discriminate between esterolytic and β-lactam cleaving activity. *Protein Science*, *11*(3), 467–478. doi.org/10.1110/ps.33002

Walsh, T. R., Toleman, M. A., Poirel, L., & Nordmann, P. (2005). Metallo-β-lactamases: The quiet before the storm? *Clinical Microbiology Reviews*, *18*(2), 306–325. doi.org/10.1128/CMR.18.2.306-325.2005

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., … Bandeira, N. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology*, *34*(8), 828–837. doi.org/10.1038/nbt.3597

Wang, X. J., Yan, Y. J., Zhang, B., An, J., Wang, J. J., Tian, J., … Xiang, W. S. (2010). Genome sequence of the milbemycin-producing bacterium *Streptomyces bingchenggensis*. *Journal of Bacteriology*, *192*(17), 4526–4527.

doi.org/10.1128/JB.00596-10

Ward, J. M., & Hodgson, J. E. (1993). The biosynthetic genes for clavulanic acid and cephamycin production occur as a “super-cluster” in three *Streptomyces*. *FEMS Microbiology Letters, 110*(2), 239–242. doi.org/10.1111/j.1574-6968.1993.tb06326.x

Watve, M. G., Tickoo, R., Jog, M. M., & Bhole, B. D. (2001). How many antibiotics are produced by the genus *Streptomyces*? *Archives of Microbiology*, *176*(5), 386–390. doi.org/10.1007/s002030100345

Wei, Y., Schottel, J. L., Derewenda, U., Swenson, L., Patkar, S., & Derewenda, Z. S. (1995). A novel variant of the catalytic triad in the *Streptomyces scabies* esterase. *Nature Structural Biology*, *2*(3), 218–223. doi.org/10.1038/nsb0395-218

Wijeratne, E. M. K., Turbyville, T. J., Zhang, Z., Bigelow, D., Pierson, L. S., VanEtten, H. D., … Gunatilaka, A. A. L. (2003). Cytotoxic Constituents of *Aspergillus terreus* from the Rhizosphere of *Opuntia versicolor* of the Sonoran Desert. *Journal of Natural Products*, *66*(12), 1567–1573. doi.org/10.1021/np030266u

Williams, J. D. (1997). β-Lactamase Inhibition and in Vitro Activity of Sulbactam and Sulbactam/Cefoperazone. *Clinical Infectious Diseases*, *24*(3), 494–497. doi.org/10.1093/clinids/24.3.494

Wong, J., d’Espaux, L., Dev, I., van der Horst, C., & Keasling, J. (2018). De novo synthesis of the sedative valerenic acid in *Saccharomyces cerevisiae*. *Metabolic*

Engineering, *47*(March), 94–101. doi.org/10.1016/j.ymben.2018.03.005

Xia, H., Zhan, X., Mao, X. M., & Li, Y. Q. (2020). The regulatory cascades of antibiotic production in *Streptomyces*. *World Journal of Microbiology and Biotechnology*, *36*(1), 1–9. doi.org/10.1007/s11274-019-2789-4

Xu, M., & Wright, G. D. (2019). Heterologous expression-facilitated natural products' discovery in actinomycetes. *Journal of Industrial Microbiology and Biotechnology*, *46*(3–4), 415–431. doi.org/10.1007/s10295-018-2097-2

Zelyas, N. J., Cai, H., Kwong, T., & Jensen, S. E. (2008). Alanylclavam biosynthetic genes are clustered together with one group of clavulanic acid biosynthetic genes in *Streptomyces clavuligerus*. *Journal of Bacteriology*, *190*(24), 7957–7965. doi.org/10.1128/JB.00698-08

Zerikly, M., & Challis, G. L. (2009). Strategies for the discovery of new natural products by genome mining. *ChemBioChem*, *10*(4), 625–633. doi.org/10.1002/cbic.200800389

Zhang, M. M., Wong, F. T., Wang, Y., Luo, S., Lim, Y. H., Heng, E., … Zhao, H. (2017). CRISPR-Cas9 strategy for activation of silent Streptomyces biosynthetic gene clusters. *Nature Chemical Biology*, *13*(6), 607–609. doi.org/10.1038/nchembio.2341

Zhao, Y., Xiang, S., Dai, X., & Yang, K. (2013). A simplified diphenylamine colorimetric method for growth quantification. *Applied Microbiology and Biotechnology*, *97*(11), 5069–5077. doi.org/10.1007/s00253-013-4893-y

frontiers in Microbiology

ORIGINAL RESEARCH
published: 08 November 2019
doi: 10.3389/fmicb.2019.02550

Comparative Genomics and Metabolomics Analyses of Clavulanic Acid-Producing *Streptomyces* Species Provides Insight Into Specialized Metabolism

*Nader F. AbuSara[1†], Brandon M. Piercey[1†], Marcus A. Moore[1†], Arshad Ali Shaikh[1†], Louis-Félix Nothias[2], Santosh K. Srivastava[1], Pablo Cruz-Morales[3], Pieter C. Dorrestein[2], Francisco Barona-Gómez[3] and Kapil Tahlan[1]**

[1] Department of Biology, Memorial University of Newfoundland, St. John's, NL, Canada, [2] Collaborative Mass Spectrometry Innovation Center, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA, United States, [3] Evolution of Metabolic Diversity Laboratory, Unidad de Genómica Avanzada (Langebio), Cinvestav-IPN, Irapuato, Mexico

OPEN ACCESS

Edited by:
Haike Antelmann,
Freie Universität Berlin, Germany

Reviewed by:
Jörn Kalinowski,
Bielefeld University, Germany
Juan F. Martin,
Universidad de León, Spain

***Correspondence:**
Kapil Tahlan
ktahlan@mun.ca

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

Received: *19 July 2019*
Accepted: *22 October 2019*
Published: *08 November 2019*

Citation:
AbuSara NF, Piercey BM, Moore MA, Shaikh AA, Nothias L-F, Srivastava SK, Cruz-Morales P, Dorrestein PC, Barona-Gómez F and Tahlan K (2019) Comparative Genomics and Metabolomics Analyses of Clavulanic Acid-Producing Streptomyces Species Provides Insight Into Specialized Metabolism. Front. Microbiol. 10:2550. doi: 10.3389/fmicb.2019.02550

Clavulanic acid is a bacterial specialized metabolite, which inhibits certain serine β-lactamases, enzymes that inactivate β-lactam antibiotics to confer resistance. Due to this activity, clavulanic acid is widely used in combination with penicillin and cephalosporin (β-lactam) antibiotics to treat infections caused by β-lactamase-producing bacteria. Clavulanic acid is industrially produced by fermenting *Streptomyces clavuligerus*, as large-scale chemical synthesis is not commercially feasible. Other than *S. clavuligerus*, *Streptomyces jumonjinensis* and *Streptomyces katsurahamanus* also produce clavulanic acid along with cephamycin C, but information regarding their genome sequences is not available. In addition, the *Streptomyces* contain many biosynthetic gene clusters thought to be "cryptic," as the specialized metabolites produced by them are not known. Therefore, we sequenced the genomes of *S. jumonjinensis* and *S. katsurahamanus*, and examined their metabolomes using untargeted mass spectrometry along with *S. clavuligerus* for comparison. We analyzed the biosynthetic gene cluster content of the three species to correlate their biosynthetic capacities, by matching them with the specialized metabolites detected in the current study. It was recently reported that *S. clavuligerus* can produce the plant-associated metabolite naringenin, and we describe more examples of such specialized metabolites in extracts from the three *Streptomyces* species. Detailed comparisons of the biosynthetic gene clusters involved in clavulanic acid (and cephamycin C) production were also performed, and based on our analyses, we propose the core set of genes responsible for producing this medicinally important metabolite.

Keywords: *Streptomyces*, specialized metabolism, metabolomics, genomics, gene clusters, β-lactams, clavulanic acid

INTRODUCTION

Bacteria from the genus *Streptomyces* produce numerous and diverse specialized (or secondary) metabolites (SMs), many of which have medicinal applications (Baltz, 2008). Some of these SMs are also used as antibiotic adjuvants, agents administered in conjunction with antibiotics to potentiate or restore their activities against resistant bacteria (Tyers and Wright, 2019). Clavulanic acid (CA, a 5*R* clavam SM, **Figure 1**) is an irreversible inhibitor of certain class A and D serine β-lactamases, which are enzymes that hydrolyze β-lactam antibiotics such as the penicillins and cephalosporins to confer resistance (Drawz and Bonomo, 2010). Therefore, CA is widely used in human and veterinary medicine in combination with β-lactam antibiotics to treat otherwise resistant infections caused by β-lactamase-producing bacteria (Brown, 1986).

Clavulanic acid is industrially produced by fermenting the bacterium *Streptomyces clavuligerus* (Jensen and Paradkar, 1999; Townsend, 2002; Saudagar et al., 2008), which was first identified during screens for microorganisms capable of producing β-lactam antibiotics such as cephamycin C (Ceph-C) (Brown et al., 1976). Apart from *S. clavuligerus*, *Streptomyces jumonjinensis* and *Streptomyces katsurahamanus* are the only other species known to produce CA along with Ceph-C (Ward and Hodgson, 1993; Jensen, 2012). In addition, CA production in *S. clavuligerus* generally occurs in conjunction with Ceph-C (Romero et al., 1984; Jensen and Paradkar, 1999), even though both metabolites are products of distinct biosynthetic pathways (Hamed et al., 2013). As in the case of other Actinobacterial SMs (van der Heul et al., 2018), the regulation of CA production in *S. clavuligerus* is complex and involves cluster-situated regulators, global mechanisms, and signaling cascades (Liras et al., 2008; Song et al., 2010a; Paradkar, 2013; Ferguson et al., 2016; Álvarez-Álvarez et al., 2017). *S. clavuligerus* is also unique among the CA producers described so far due to its ability to produce the structurally related 5*S* clavams (Brown et al., 1979; Pruess and Kellett, 1983), which partially share a common biosynthetic pathway with CA (**Figure 1**; Egan et al., 1997; Jensen, 2012). The 5*S* clavams have the opposite stereochemistry as compared to CA and are therefore not inhibitory toward β-lactamases, but instead some of them display weak antibacterial, antifungal, or antimetabolite activities (Jensen, 2012). In comparison, some *Streptomyces* species only synthesize the 5*S* clavams but not CA, suggesting that the ability to produce clavams with both stereochemistries (5*R* and 5*S*, **Figure 1**) might be unique to *S. clavuligerus* (Jensen and Paradkar, 1999; Challis and Hopwood, 2003).

It is now recognized that the *Streptomyces* contain many biosynthetic gene clusters (BGCs) thought to be "cryptic or silent," as the SMs produced by them are not known (Katz and Baltz, 2016). On average, each *Streptomyces* species contains ∼30 BGCs but only produces 3–5 SMs under laboratory conditions. Additionally, recent reports have shown that *S. clavuligerus* produced some SMs only thought to originate from plants (Álvarez-Álvarez et al., 2015), highlighting the need for thoroughly cataloging specialized metabolism, even from well-studied organisms. Due to the small number of *Streptomyces* species known to produce CA, it is of interest to determine if *S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus* also share other metabolic capabilities. Therefore, we sequenced the genomes of *S. jumonjinensis* and *S. katsurahamanus*, conducted comparative metabolomics analysis on the three CA producers to identify SMs, and correlated their biosynthesis with predicted BGCs wherever possible.

The described analyses also provide information on BGC content from *S. jumonjinensis* and *S. katsurahamanus*, insight that was not available previously. In *S. clavuligerus,* three separate BGCs are involved in clavam metabolite biosynthesis (Tahlan et al., 2004a). The clavulanic acid BGC is primarily associated with CA production (Jensen et al., 2000, 2004a; Li et al., 2000; Mellado et al., 2002), whereas the clavam and paralog BGCs are involved in the biosynthesis of the 5*S* clavams (**Figure 1**; Tahlan et al., 2007; Zelyas et al., 2008). Because of the common biosynthetic origins of CA and the 5*S* clavams, it has been suggested that there is sharing of intermediates between the pathways (**Figure 1**). Therefore, many gene products from the CA, clavam, and paralog BGCs contribute to the early part of the biosynthetic pathway involved in both CA and the 5S clavam production (**Figure 1**; Jensen, 2012; Hamed et al., 2013; Álvarez-Álvarez et al., 2018). Previous genetic mapping studies have shown that the BGCs for CA and Ceph-C are clustered together on the chromosomes of all CA producers to form "β-lactam superclusters" (Ward and Hodgson, 1993), but details about their sequences from *S. jumonjinensis* and *S. katsurahamanus* were lacking. It has been hypothesized that CA biosynthesis evolved in an ancestral 5*S* clavam producer, after it acquired the ability to produce Ceph-C by horizontal gene transfer (Challis and Hopwood, 2003). Such an arrangement leads to the coordinated biosynthesis of Ceph-C and CA, or the production of a β-lactam antibiotic and a synergistically acting β-lactamase inhibitor, respectively. The complete biosynthetic pathway leading to Ceph-C has been elucidated (Liras, 1999), but some late steps required for CA production remain unknown (Jensen, 2012; Hamed et al., 2013). Additionally, not all genes from the proposed *S. clavuligerus* CA BGC are required for CA production (**Supplementary Table S1**), and the exact function of many gene products remains to be deciphered (Jensen et al., 2004a; Valegård et al., 2013; Álvarez-Álvarez et al., 2018; Srivastava et al., 2019). Recently available genome sequences have revealed that CA-like BGCs (without any associated Ceph-C BGCs) are also present in other organisms such as *Streptomyces pratensis* ATCC 33331 (formerly called *Streptomyces flavogriseus*) and *Saccharomonospora viridis* DSM 43017, but neither have been shown to produce CA to date (Jensen, 2012; Álvarez-Álvarez et al., 2013). Therefore, it is still not clear as to what defines the boundaries of a functional (or minimal) CA BGC, a question that we also address in the current study.

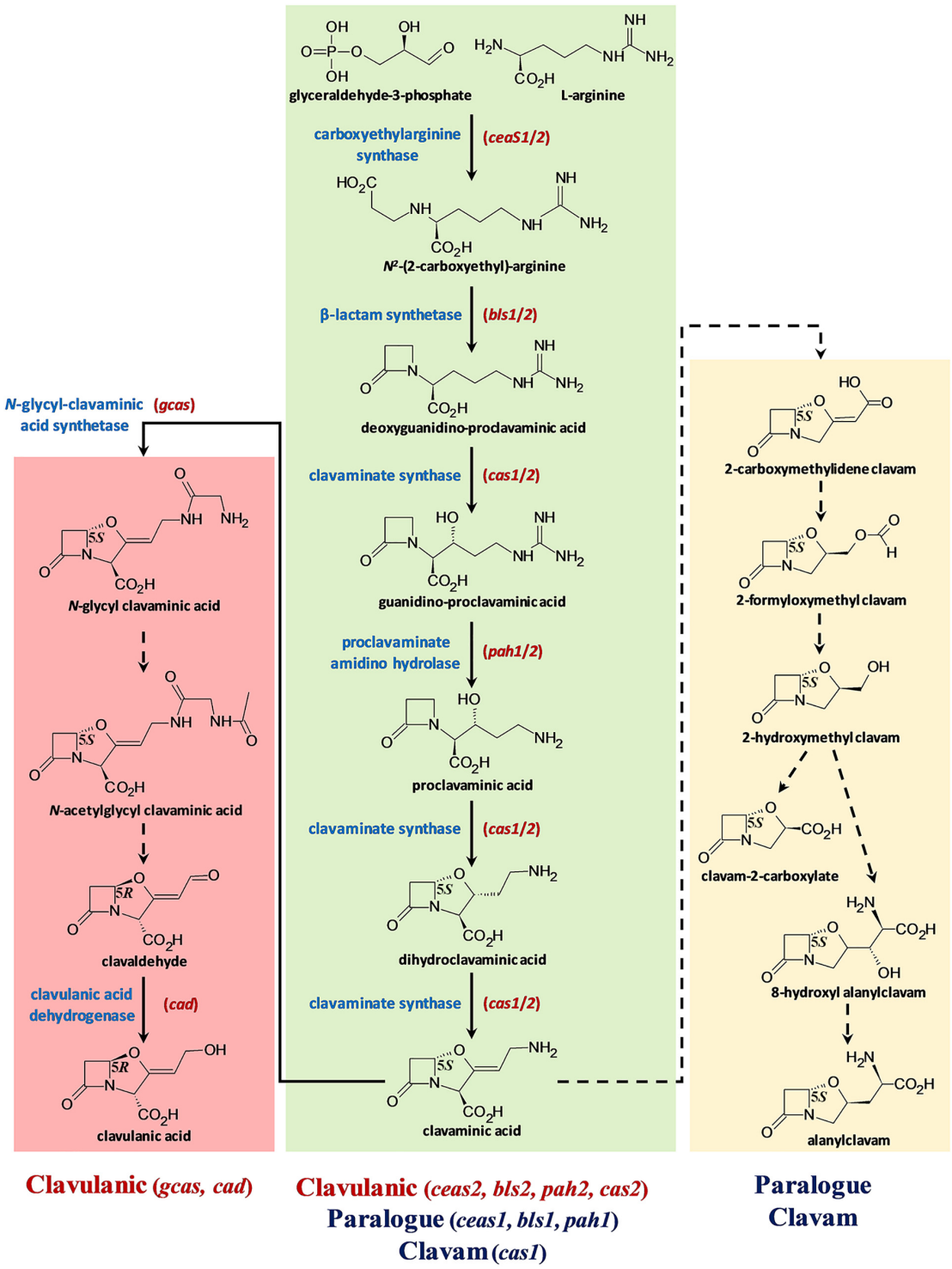


FIGURE 1 | The *S. clavuligerus* clavulanic acid and 5*S* clavam biosynthetic pathways. The pathway is depicted in three parts: the central green box represents the steps shared between CA and 5*S* clavam biosynthesis, whereas the pink (left) and yellow (right) boxes indicate the late steps specific for CA or 5*S* clavam production, respectively. The solid arrows represent known reactions and broken arrows indicate uncharacterized steps, which could potentially involve more than one unknown gene product/enzyme. The names of core biosynthetic enzymes (blue) catalyzing known reactions and the respective gene(s) encoding them (red) are included where applicable. The stereochemistries (*R*/*S*) of the intermediates/products are also included along with their names. The identities of the gene clusters involved in each stage of biosynthesis is indicated at the bottom of the figure. Note that the shared part of the pathway (green) involves substitutable isozymes (CeaS, Bls, Cas, and Pah), which are encoded by two sets of genes (1 and 2) residing in three separate gene clusters. Additional genes from the respective clusters for which exact biosynthetic functions have not been assigned are not shown to simply interpretation.

MATERIALS AND METHODS

Bacterial Strains, Plasmids, Media/Culture Conditions, and Molecular Methods

Bacterial strains and plasmids used in the current study are described in **Table 1**. All media/reagents were purchased from Fisher Scientific or VWR International (Canada). For routine analysis, *Streptomyces* cultures were maintained on International *Streptomyces* Project (ISP) medium 4 agar or were grown in Trypticase Soy Broth supplemented with 1% (w/v) soluble starch (TSB-S). Cultures for metabolite analysis were grown using glycerol, sucrose, proline, and glutamic acid (GSPG); starch asparagine (SA); soy; or TSB-S media (Romero et al., 1997; Tahlan et al., 2004b). All *Streptomyces* cultures were incubated at 28°C and liquid cultures were agitated at 250 rpm. Plasmid-bearing *Streptomyces* cultures were supplemented with appropriate antibiotics when required (Tahlan et al., 2004b), whereas *Escherichia coli* was grown and maintained as described previously (Sambrook, 2001). Standard methods were used for isolating and manipulating DNA from *E. coli* (Sambrook, 2001) and *Streptomyces* (Kieser et al., 2000; Tahlan et al., 2004b). Total RNA was isolated from *S. clavuligerus* grown on SA medium as described previously (Srivastava et al., 2019), and RT was performed using the Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Scientific, United States). PCRs were carried out using the Phusion or Taq DNA polymerase kits (ThermoFisher, United States). When required, PCR products were cloned into the pGEM-T Easy vector (Promega, United States) according to the manufacturer's instruction and were sequenced at the Centre for Applied Genomics, University of Toronto (Canada). All DNA oligonucleotide primers used in the study (**Supplementary Table S2**) were obtained from Integrated DNA Technologies (United States).

TABLE 1 | Bacterial strains and plasmids used in this study.

Strain/plasmid	Description[a]	Source/ Reference
Bacterial strain		
E. coli ESS	Indicator strain for Ceph-C bioassays	Wang et al., 2004
E. coli ET12567/ pUZ8002	Non-methylating conjugation host carrying helper plasmid pUZ8002 (Cml^R and Kan^R)	Kieser et al., 2000
E. coli DH5α	General laboratory cloning host	Promega
K. pneumoniae ATCC 15380	Indicator strain for CA bioassays (Pen^R)	ATCC
S. clavuligerus ATCC27064	Wild-type CA producer	ATCC
S. clavuligerus Δ*nocE*	*nocE* null mutant	This study
S. clavuligerus pIJ8668-*ermE*p*-*nocE*	Strain constitutively expressing *nocE*	This study
S. jumonjinensis NRRL 5741	Wild-type CA producer	Jensen and Paradkar, 1999
S. katsurahamanus T272	Wild type CA producer	Jensen and Paradkar, 1999
Plasmids		
pGEMT-Easy	Plasmid for cloning PCR products	Promega
pIJ8668-*ermE*p*	Conjugative *Streptomyces* suicide vector containing *ermE*p* for chromosomal promoter insertion (Apr^R)	(Tahlan et al., 2017)
pIJ8668-*ermE*p*-*nocE*	pIJ8668-*ermE*p* containing a portion of the 5′ end of *nocE* from *S. clavuligerus* (Apr^R)	This study
pIJ12738	Conjugative *Streptomyces* suicide vector containing an I-SceI site for gene targeting (Apr^R)	Fernández-Martínez and Bibb, 2014
pIJ12738-*nocE*-UP-DN	pIJ12738 containing regions upstream and downstream of *nocE* from *S. clavuligerus* (Apr^R)	This study
pIJ12742	Plasmid expressing the Meganuclease I-SceI in *Streptomyces* for gene disruption (Apr^R and Tsr^R)	Fernández-Martínez and Bibb, 2014

[a] *Apr^R, apramycin resistance; Cml^R, chloramphenicol resistance; Kan^R, kanamycin resistance; Pen^R, penicillin G resistance; and Tsr^R, thiostrepton resistance.*

Genome Sequencing, Gene Cluster Identification, and Bioinformatics Analyses

The *S. jumonjinensis* and *S. katsurahamanus* genomes were sequenced using Illumina MiSEQ in paired-end format with read lengths of 300 bp. A chromosomal DNA library was prepared for each organism using the PCR-based method adjusted for high GC DNA according to the manufacturer's instructions (Illumina, United States). Raw reads were filtered with trimmomatic (Bolger et al., 2014) with a cutoff of 26 bp and a minimum length of 150 bp. The remaining reads were assembled using Velvet (Zerbino and Birney, 2008). *k*-mers from 30 to 170 were tested for selecting optimal contig length and the assembled genomes (31–46 × coverage, **Supplementary Table S3**) were submitted to NCBI (accession numbers: *S. jumonjinensis* NRRL 5741, VCLA00000000 and *S. katsurahamanus* T-272, VDEQ00000000). Genome completeness was calculated (**Supplementary Table S3**) using BUSCO (Simao et al., 2015) and QUAST (Gurevich et al., 2013). Annotations were carried out using RAST (Overbeek et al., 2014) and also manually in Artemis (Rutherford et al., 2000). Specialized metabolite (SM) biosynthetic gene clusters (BGCs) were identified using antiSMASH 4.0 (Blin et al., 2017) and polyketide synthases/nonribosomal peptide synthetase genes were predicted using PRISM 3 (Skinnider et al., 2017). The DNA sequences of *S. jumonjinensis* and *S. katsurahamanus* BGCs were manually examined for possible frame shifts and other ambiguities. In some cases, PCR amplification was performed using custom primers (**Supplementary Table S2**) followed by Sanger sequencing of products to verify results. The genome sequences of *S. clavuligerus* ATCC 27064 (NZ_CM000913.1, NZ_CM000914.1), *S. pratensis* ATCC 33331 (NC_016114), *S. viridis* DSM 43017 (CP001683.1), *Streptomyces* sp. M41(2017) (NZ_MWFK00000000), *Streptomyces* sp. PAMC26508 (NC_021055), *Streptomyces* sp. NRRL S-325 (NZ_JOIW00000000), *Streptomyces* sp. NRRL B-24051

(NZ_JOAE00000000), *Streptomyces flavovirens* NRRL B-2182 (NZ_JOAB00000000), *Streptomyces fulvoviridis* NRRL ISP-5210 (NZ_JNXH00000000), and *Streptomyces olivaceus* NRRL B-3009 (NZ_JOFH00000000) were included for comparison as the latter harbor CA-like BGCs containing homologs of all genes currently known to be involved in CA production in *S. clavuligerus* (Jensen, 2012). In addition, the sequences of the Ceph-C BGCs from *Streptomyces cattleya* 8057 (NC_017586.1) and *Nocardia lactamdurans* (also known as *Amycolatopsis lactamdurans*) (Z13971.1-Z13974.1, Z21681.1-Z21686.1 and X57310.1) were also included in the analysis. Geneious 8.1.9 (Biomatters Ltd., New Zealand) was used for sequence comparisons and constructing phylogenetic trees. Protein homologs were identified using NCBI BLAST and secretory signals were predicted using the SignalP-5.0 Server (Almagro Armenteros et al., 2019).

Preparation of the *S. clavuligerus* ΔnocE and ermEp*-nocE Strains

The *S. clavuligerus nocE* gene mutant was prepared using the meganuclease I-SceI marker-less gene deletion system (Fernández-Martínez and Bibb, 2014). DNA fragments (1–1.2 kb each) containing regions immediately upstream and downstream of *nocE* from the *S. clavuligerus* chromosome were amplified using PCR along with engineered primers (**Supplementary Table S2**) and were separately cloned into the pGEM-T Easy vector (**Table 1**). The upstream fragment was released from pGEM-T Easy by digestion with *Hin*dIII and *Eco*RI and was introduced into the same sites of pIJ12738 to give pIJ12738/*nocE*-UP. The downstream fragment was then introduced into the *Eco*RI and *Xba*I sites of pIJ12738-*nocE*-UP to give pIJ12738/*nocE*-UP-DN, which functioned as the *nocE* disruption construct (**Table 1**). pIJ12738-*nocE*-UP-DN was conjugated into *S. clavuligerus* to obtain the apramycin-resistant single crossover strain, which was confirmed using genomic DNA PCR (**Supplementary Table S2**). The plasmid pIJ12742 expressing the I-SceI meganuclease (**Table 1**) was then conjugated into *S. clavuligerus* pIJ12738-*nocE*-UP-DN to obtain apramycin and thiostrepton resistant exconjugants, which were made to undergo sporulation at 28°C without any selection to facilitate double homologous recombination and loss of pIJ12738 from the chromosome. Spore stocks were prepared and re-streaked onto ISP-4 plates without selection and incubated for 5 days at 37°C to promote the loss of temperature-sensitive pIJ12742. This led to the isolation of the apramycin and thiostrepton-sensitive *S. clavuligerus* Δ*nocE* mutant, which was verified using genomic DNA PCR (**Supplementary Table S2**).

To prepare an *S. clavuligerus* strain constitutively expressing *nocE* (**Table 1**), the *ermE*p* promoter (Bibb et al., 1985) was inserted upstream of the gene in the *S. clavuligerus* chromosome. A 1.1-kb DNA fragment from the 5′ end of the gene was amplified by PCR (**Supplementary Table S2**) and was cloned into pGEM-T Easy. The insert was re-isolated as an *Nde*I and *Eco*RI fragment and was ligated with similarly digested pIJ8668-*ermE*p* to give pIJ8668-*ermE*p*-*nocE* (**Table 1**), which was introduced into wt *S. clavuligerus* by conjugation. This resulted in the *S. clavuligerus ermE*p*-*nocE* strain, which was confirmed using genomic DNA PCR (**Supplementary Table S2**) and was used to examine the effect of constitutively expressing *nocE* in *S. clavuligerus*.

Bioassays and Bacterial Growth Measurement

The production of CA and Ceph-C in culture supernatants was detected (and quantified in the case of Ceph-C) using bioassays employing indicator organisms (**Table 1**), as described previously (Paradkar and Jensen, 1995; Wang et al., 2004). Growth in liquid cultures was determined using a modified diphenylamine colorimetric method to measure DNA content (Zhao et al., 2013) and statistical analysis (ANOVA repeated measure) was performed using R 3.4.3. To assess for growth characteristics on solid media, 10-fold dilutions of a spore stock (4×10^4 spores/μl) were prepared, and 5 μl of which were spotted onto two different agar media (SA and TSB-S with 1.5% agar). The plates were then incubated at 28°C and visually scored for growth over a 7-day period.

Liquid Chromatography–Mass Spectrometry (LC-MS and LC-MS/MS) Analysis

The production of clavam metabolites in 96-h broth cultures was analyzed by targeted LC-MS after imidazole derivatization using an XTerra column (2.1 × 150 mm, 3.5 μm, 125 Å; Waters Scientific, United States) as described previously (Srivastava et al., 2019). Untargeted metabolomics was conducted using bacteria grown on solid media. One hundred microliters of a standardized spore stock (4×10^4 spores/μl) of each species was used to inoculate agar plates in duplicate, and each plate was extracted using 15 ml of methanol or ethyl acetate. Two milliliters of each extract was dried, resuspended in 130 μl of 70% methanol containing 0.2 μM of amitriptyline (internal standard), and transferred to a 96-well plate, which was centrifuged at 2000 rpm for 15 min at 4°C. One hundred microliters of each sample was then transferred to a new 96-well plate for LC-MS/MS analysis. Samples were analyzed using a Vanquish UHPLC System coupled Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Scientific, United States). Chromatographic separation was performed in mixed mode (allowing weak anion/cation exchange) on a Scherzo SM-C18 column (2 × 250 mm, 3 μm, 130 Å; Imtakt, United States) maintained at 40°C. Ten microliters of each sample was injected for analysis and the mobile phase consisted of (A) 0.1% formic acid in water and (B) 0.1% formic acid in acetonitrile. Chromatography was performed at a flow rate of 0.5 ml/min using the following program: 0–5 min, 98% A; 5–8 min, gradient of 98–50% A (or 50% B); 8–13 min, gradient 50–100% B; 13–14.00 min, 100% B; 14–14.10 min, 100–2% B; 14.10–18 min, 2% B.

Mass spectrometry was performed using a heated electrospray ionization source (heater temperature, 370°C and capillary temperature, 350°C) in either positive or negative ionization mode (± 3000.0 V; S-lens RF, 55; sheath gas flow rate, 55; and auxiliary gas flow rate, 20). MS^1 and MS^2 scans (at 200 *m*/*z*) were acquired from 0.48 to 16.0 min at a resolution of

35,000 and 17,500, respectively, for the 100–1500 *m*/*z* range. The automatic gain control (AGC) target value and maximum injection time were set at 5×10^5 and 150 ms. Up to four MS^2 scans in data-dependent mode were acquired for most abundant ions per duty cycle, with a starting value of 70 *m*/z, and exclusion parameter of 10 s. Higher-energy collision-induced dissociation was performed with a normalized collision energy of 20, 35, and 50 eV. The apex trigger mode was used (2–7 s) and the isotopes were excluded. Inclusion lists of ions for molecules observed in *Streptomyces* extracts were generated from the Dictionary of Natural Products[1] and the StreptomeDB (Lucas et al., 2013), and were used for prioritizing the acquisition of their MS^2 when observed. The raw LC-MS/MS data files were converted to .mzXML format using ProteoWizard (Adusumilli and Mallick, 2017). All metabolomics MS data have been deposited on the MassIVE public repository[2] under the accession number MSV000083835.

MS Data Annotation and Analysis

Molecular networks were generated using positive and negative ionization mode data in GNPS (Wang et al., 2016). The resulting networks were visualized in Cytoscape (Shannon et al., 2003), allowing nodes associated with uninoculated media controls to be removed. Annotations were first obtained by matching spectra in public libraries (Wang et al., 2016), including NIST17[3]. Library annotations were manually validated using mirror plots (maximum ion mass accuracy = 5 ppm) corresponding to level 2 annotation based on the Minimum Standard Initiative (Spicer et al., 2017). The data were deposited to the GNPS library (CCMSLIB00005435954-CCMSLIB00000531493), which enabled the annotation of putative tunicamycin derivatives (CCMSLIB00005435941-42) and lyngbyatoxin (CCMSLIB00005435954-55) using molecular networks. In some cases, Sirius 4.0.1 was used to confirm the molecular formulas of certain predicted metabolites (Böcker et al., 2009).

To generate a heat map using the *S. clavuligerus* wt, Δ*nocE*, and *ermEp*-nocE* strains, feature-based detection and alignment of positive mode ionization data were performed (parameters: MS^1 noise level of 25000, MS^2 noise level of 1000) using the MZmine 2 toolbox (v2.39) (Pluskal et al., 2010). Chromatograms were built using the ADAP module (parameters: min group size in # of scans = 4, group intensity threshold = 700,000, min highest intensity = 100,000, max *m/z* tolerance = 10 ppm), which were then deconvoluted (parameters: S/N threshold = 10.0, min feature height = 7000000, coefficient/area threshold = 60.0, peak duration range = 0.01–0.5 min, RT wavelet range = 0.01–0.1 s). Fragmentation spectra were paired with deconvoluted peaks using 0.02 Da and 0.2 min windows, and LC-MS features were annotated using the Peak-Grouping module (parameters: deisotope = true, remove features without isotope pattern = false, minimal intensity for interval selection = 0.1, minimal intensity overlap = 0.7, minimal correlation = 0.7). Features were aligned in the JoinAligner module (parameters: ppm tolerance = 7, weight for *m/z* = 75.0, retention time tolerance = 0.5 min, weight for RT = 25.0; require same charge state = false, require same ID = false, compare isotope pattern = false). The aligned peaklist was filtered with the row filter module to keep only features with at least two isotopic ions, two occurrences, and at least one MS^2 spectrum before gap filling (parameters: intensity = 5%, ppm window = 5, retention time tolerance = 0.15). The aligned peaklist containing 3149 features was exported as a .CSV file, and the spectral data as .MGF files using the GNPSExport module for further processing. The signal intensities of the features (.CSV) were normalized to that of an internal standard (*m/z* 278.189; retention time, 9.2 min) and only 1684 features with an intensity 3-fold higher than in experimental controls (uncultivated media) were retained. MetaboAnalyst4.0 (Chong et al., 2018) was used to perform the hierarchal clustering, which was visualized as a heat map.

RESULTS AND DISCUSSION

Three *Streptomyces* species are known to produce CA, but details about the involved BGCs are only available for the genome sequenced industrial producer, *S. clavuligerus* (Medema et al., 2010; Song et al., 2010b; Cao et al., 2016). Therefore, we sequenced the genomes of the other two CA producers, *S. jumonjinensis* and *S. katsurahamanus* (**Table 1**), for comparative studies. The published genome sequence of *S. pratensis* ATCC 33331 was also included during some of the analyses (**Figure 2A**), as it contains a CA-like BGC (**Figures 3A,C**), and has been shown not to produce the metabolite under tested conditions (Álvarez-Álvarez et al., 2013). Examination of the *S. jumonjinensis* and *S. katsurahamanus* genomes revealed that they each contain 49 and 44 known or predicted SM BGCs (**Table 2** and **Supplementary Table S4**), respectively, which is much higher than the average number found in many *Streptomyces* species. Additionally, *S. clavuligerus* contains 43 SM BGCs, although re-sequencing of its genome suggests that it may contain many more (Hwang et al., 2019). This prompted us to further investigate the specialized metabolic capabilities of the three CA producers to determine similarities or differences between these microorganisms.

SM-BGCs and Metabolism in *S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus*

Detailed analysis of the *S. jumonjinensis* and *S. katsurahamanus* genome sequences using antiSMASH 4.0 (Blin et al., 2017) and manual curation showed that both organisms contain numerous BGCs for diverse SMs (**Figure 2A** and **Supplementary Table S4**). Therefore, *S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus* were grown on SA, GSPG, and TSB-S media for assessing CA/Ceph-C production (**Figure 2B** and **Supplementary Figure S1**) and for preparing methanol/ethyl acetate extracts for liquid chromatography–tandem mass spectrometry (LC-MS/MS)-based metabolomics. The MS/MS data obtained from both positive and negative ionization mode were used to build a molecular network (**Figure 2C**), and

[1] http://dnp.chemnetbase.com

[2] massive.ucsd.edu

[3] www.nist.gov

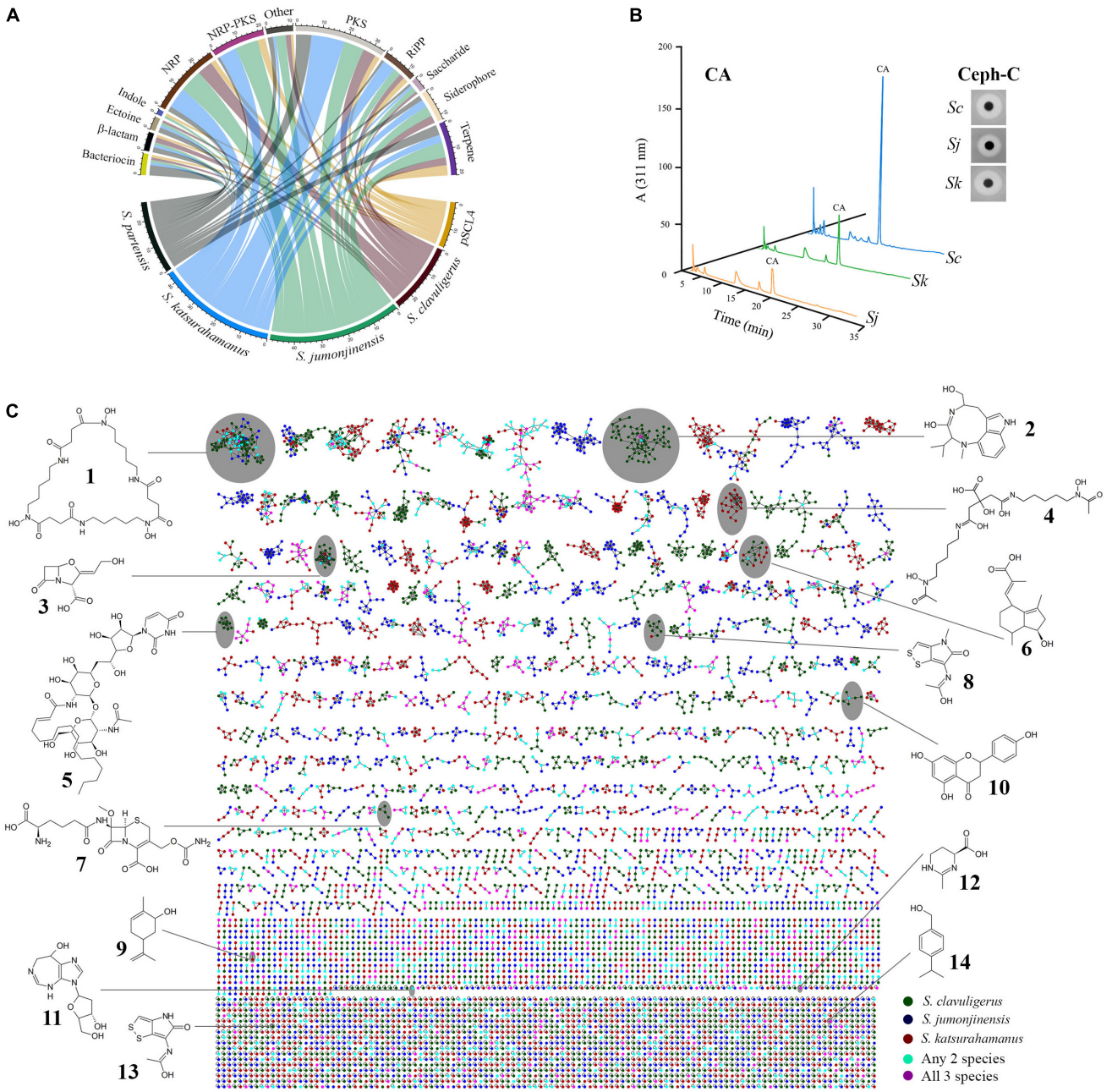


FIGURE 2 | Biosynthetic gene cluster (BGC) content and metabolomics analysis of clavulanic acid (CA)-producing *Streptomyces* species (*S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus*). **(A)** Circular chord diagram representing all predicted BGC classes present in the respective *Streptomyces* species. *S. pratensis* was included for comparison as the bacterium contains a CA-like BGC, but does not produce the metabolite. The sequence of the giant linear plasmid pSCL4 from *S. clavuligerus* was also included separately due to the presence of multiple BGCs on it. The lower arc represents genomes/plasmid, while the upper arc represents different classes of BGCs and the color-coded ribbons connecting them indicate the presence of a BGC in the specific species. **(B)** Detection of CA and cephamycin C (Ceph-C) in 96-h SA culture supernatants of *S. clavuligerus* (*Sc*), *S. jumonjinensis* (*Sj*), and *S. katsurahamanus* (*Sk*) using LC-MS (after imidazole derivatization) and bioassays (inset), respectively. The peak corresponding to CA in HPLC chromatograms is noted and the zones of inhibition in the inset panel demonstrate relative amounts of Ceph-C production. **(C)** Metabolic network constructed using *S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus* culture extracts (culture conditions and details are described in the section "Materials and Methods"). The network is color-coded according to source organism (bottom right legend), where each node depicts a mass spectrum and edges represent the relationship between different nodes. Structures of natural products detected in the extracts at high confidence in the three species are shown and the clade in the network containing the node corresponding to the respective metabolite is also indicated. 1, desferrioxamine E; 2, (−)-indolactam V; 3, clavulanic acid; 4, arthrobactin; 5, tunicamycin C2; 6, hydroxyvalerenic acid; 7, cephamycin C; 8, thiolutin; 9, (−)-carveol; 10, naringenin; 11, pentostatin, 12, ectoine; 13, holomycin; and 14, cuminyl alcohol.

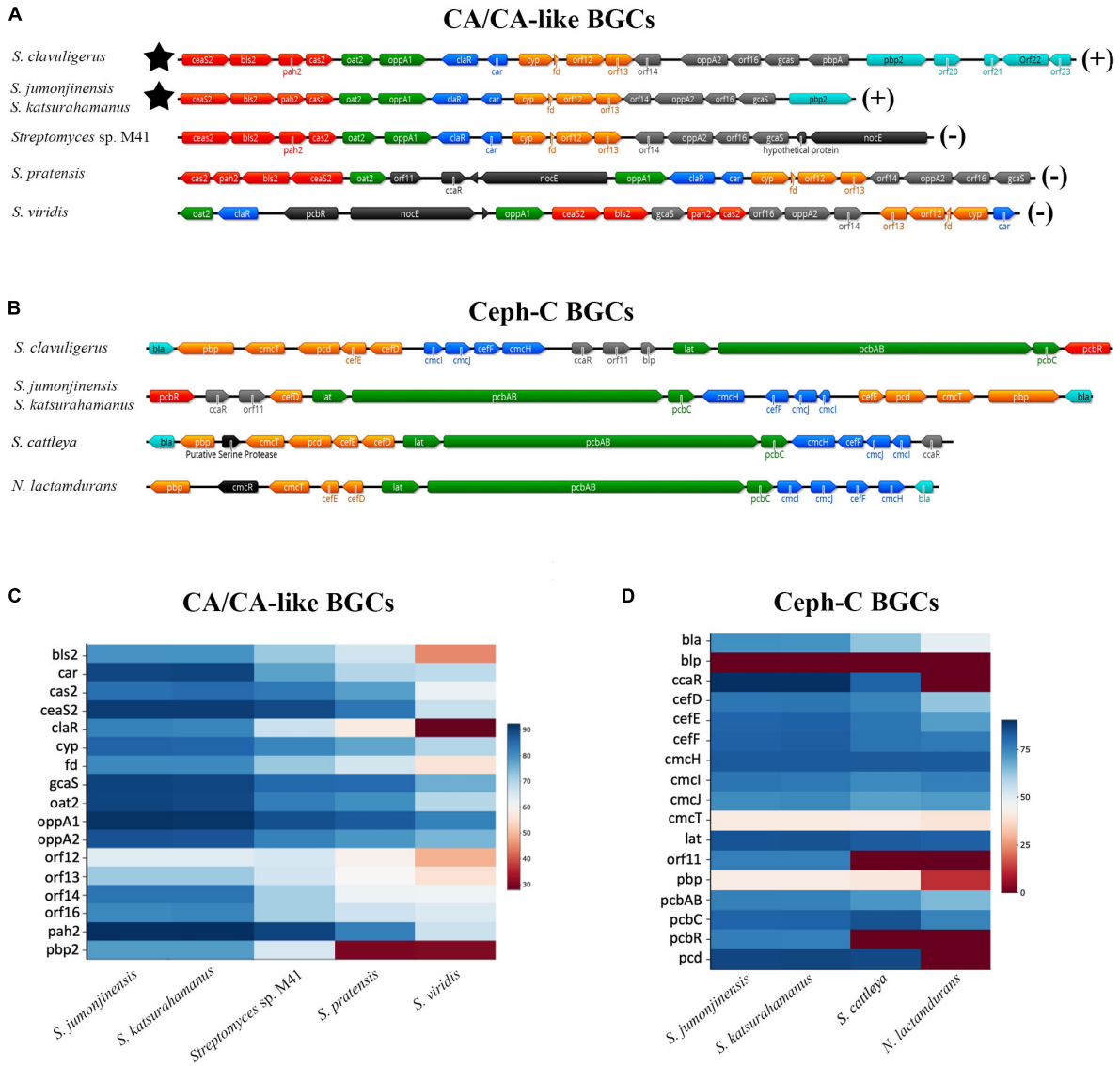


FIGURE 3 | Organization of CA/CA-like **(A)** and Ceph-C **(B)** BGCs in select CA-producing/non-producing *Streptomyces* species. The CA non-producers *Saccharomonospora viridis* **(A,C)** and *Nocardia* (or *Amycolatopsis*) *lactamdurans* **(B,D)** were also included for comparison as both are phylogenetically distinct from the *Streptomyces*. **(A,B)** The architecture of respective BGCs from the described organisms showing their gene content and relative organization. Genes are color-coded based on known or predicted transcriptional units. **(A)** The star symbol represents the location of the Ceph-C BGC (if present) and the CA production status (±) of each organism is indicated on the right. **(B)** All species included are Ceph-C and CA producers except for *S. cattleya* and *N. lactamdurans*, which only produce the former. **(C,D)** Relative identities of protein products from the CA/CA-like **(C)** and Ceph-C **(D)** BGCs of described organisms as compared to corresponding homologs from *S. clavuligerus*. The legend on the right shows colors indicating percent identities between respective gene products.

TABLE 2 | Genome features relevant to specialized metabolism in *S. jumonjinensis* and *S. katsurahamanus* as compared to *S. clavuligerus* (CA producer) and *S. pratensis* (CA non-producer).

Feature	*S. jumonjinensis* NRRL 5741	*S. katsurahamanus* T272	*S. clavuligerus* ATCC27064	*S. pratensis* ATCC 33331
Genome size (Mbp)	8.47[a]	7.25[a]	8.56	7.34
Coding sequences	7423	6123	7281	6537
SM BGCs[b] (PKS/NRPS)[c]	49 (8/18)	44 (9/9)	43 (10/9)	27 (5/9)

[a] *Estimated sizes based on sequence analysis from the current study.* [b] *Specialized metabolite (SM) biosynthetic gene clusters (BGCs) were predicted using antiSMASH 4.0.* [c] *Polyketide synthase (PKS) and nonribosomal peptide synthetase (NRPS) genes were predicted using PRISM 3.*

metabolites were annotated by matching spectra against public libraries corresponding to level 2 annotation based on the Metabolomics Standard Initiative (Wang et al., 2016). During the analysis, ions corresponding to CA ($[M\text{-}H]^-$, *m/z* 198.039), Ceph-C ($[M\text{-}H]^-$, *m/z* 445.104) and numerous other SMs were also detected in extracts from one or more *Streptomyces* species (**Figure 2C**, **Supplementary Tables S5, S6**), some of which are discussed below.

The desferrioxamines (**Figure 2C**) comprise a group of nonpeptide hydroxamate siderophores produced by many bacteria (Barona-Gómez et al., 2004), including *S. clavuligerus* (Álvarez-Álvarez et al., 2017). In the current study, ions corresponding to desferrioxamine E (Nocardamine, $[M + H]^+$, *m/z* 601.356) and desferrioxamine B (Desferal, $[M + H]^+$, *m/z* 561.361) were detected in extracts from *S. clavuligerus*/*S. jumonjinensis* and *S. clavuligerus*/*S. katsurahamanus*, respectively (**Supplementary Table S5**). Desferrioxamine E exhibits antitumor activity (Kalinovskaya et al., 2011), while desferrioxamine B is used in therapy for secondary iron overload disease (Olivieri and Brittenham, 1997). We also identified BGCs in *S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus* (**Supplementary Table S4**) that have high degrees of similarity (80–100%) with BGCs from known desferrioxamine producers such as *Streptomyces griseus* (Yamanaka et al., 2005; Ohnishi et al., 2008) and *Streptomyces coelicolor* A3(2) (Bentley et al., 2002; Barona-Gómez et al., 2004). The siderophore arthrobactin (**Figure 2C**) was also detected in *S. katsurahamanus* extracts ($[M + H]^+$, *m/z* 477.256) (**Supplementary Table S5**), but since the genes responsible for arthrobactin production are not known (Burrell et al., 2012), we were unable to identify an associated BGC in this organism. However, our analysis showed that *S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus* each contain additional siderophore-like BGCs of unknown function (**Supplementary Table S4**), which could potentially be involved in the production of such metabolites. Ectoine is another commonly produced metabolite that helps bacteria survive extreme osmotic stress (Sadeghi et al., 2014), and it was detected ($[M + H]^+$, *m/z* 143.082) in extracts from all three CA-producing species (**Figure 2C** and **Supplementary Table S5**). In addition, *S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus* contain BGCs that are similar to the known ectoine BGC from *Streptomyces anulatus* (previously called *Streptomyces chrysomallus*) (Prabhu et al., 2004). Since the desferrioxamines and ectoine are produced by many Actinomycetes and are involved in general cellular growth/survival processes (Challis, 2005; Czech et al., 2018), finding them in culture extracts from the three CA producers in the current study was not surprising.

Streptomyces clavuligerus is a known producer of the dithiolopyrrolone antibiotic holomycin (Kenig and Reading, 1979) and the associated BGC has been identified in this organism (Li and Walsh, 2010). In the current study, holomycin ($[M + H]^+$, *m/z* 214.994) and thiolutin (another dithiolopyrrolone, $[M + H]^+$, *m/z* 229.010) were detected in extracts from *S. clavuligerus*, but not in those from *S. jumonjinensis* or *S. katsurahamanus* (**Figure 2C** and **Supplementary Table S5**). Recently, a dithiolopyrrolone with the same molecular weight as thiolutin (predicted to be *N*-propionylholothin) was also detected in extracts from *S. clavuligerus* strains lacking the giant linear plasmid pSCL4 (Álvarez-Álvarez et al., 2017). Since holomycin and thiolutin (**Figure 2C**), and the respective BGCs involved in their biosynthesis (from *S. clavuligerus* and *Saccharothrix algeriensis* NRRL B-24137, respectively), are very similar (**Supplementary Table S4**), it is possible that a single pathway in *S. clavuligerus* produces both metabolites. It has also been reported that there is some sort of cross regulation between CA and holomycin production in *S. clavuligerus* (de la Fuente et al., 2002; Álvarez-Álvarez et al., 2017). Our results showed that *S. jumonjinensis* and *S. katsurahamanus* lack dithiolopyrrolone BGCs (**Supplementary Table S4**) and therefore do not have a similar link between holomycin and CA production as observed in *S. clavuligerus*.

We also detected certain nucleoside SMs during the current analysis (**Figure 2C**). For example, the purine nucleoside pentostatin, which is also used as an anticancer agent (Dillman, 2004), was identified ($[M + 2H]^{2+}$, *m/z* 135.066) in *S. clavuligerus* extracts (**Figure 2C**, **Supplementary Table S5**). A putative pentostatin-like BGC was recently shown to be present in *S. clavuligerus* (Wu et al., 2017), but production of the metabolite has not been reported in this organism previously. Therefore, our results suggest that the *S. clavuligerus* pentostatin BGC can be activated under laboratory conditions. The tunicamycins also comprise a mixture of related nucleoside antibiotics, some of which (A, B, C, and I) were detected in extracts from *S. clavuligerus* (**Figure 2C** and **Supplementary Table S5**), but not in those from *S. jumonjinensis* or *S. katsurahamanus*. *S. clavuligerus* is a known producer of tunicamycin and the BGC involved in its production has been identified (Kenig and Reading, 1979; Chen et al., 2010). In addition, certain derivatives of tunicamycin I with different acyl chains were detected in *S. clavuligerus* extracts recently (Martínez-Burgo et al., 2019), which were also present in our samples (**Supplementary Table S5**). Our results demonstrated that *S. jumonjinensis* and *S. katsurahamanus* do not possess tunicamycin BGCs (**Supplementary Table S4**), further distinguishing *S. clavuligerus* from the other CA producers due to its ability to produce such nucleoside SMs.

Metabolomics analysis also revealed the presence of certain plant-associated SMs in the *Streptomyces* extracts. It was recently shown that *S. clavuligerus* produces the citrus flavonoid naringenin and the genes involved in the production of this metabolite were also identified (Álvarez-Álvarez et al., 2015). Naringenin exhibits antibacterial, antifungal, and anticancer activities (Rauha et al., 2000; Kanno et al., 2005), and its production by a bacterium was unexpected since it was previously isolated from plants only (Álvarez-Álvarez et al., 2015). We detected naringenin (**Figure 2C**, $[M\text{-}H]^-$, *m/z* 271.062) in extracts from *S. clavuligerus* and *S. jumonjinensis*, but not from *S. katsurahamanus* (**Supplementary Table S5**). In addition, the genes involved in naringenin production were also found in both *S. jumonjinensis* and *S. katsurahamanus* (**Supplementary Table S4**), suggesting that the metabolite might be produced at undetectable levels in *S. katsurahamanus* or that the genes

are not expressed in this species under the conditions tested. Also detected in all three *Streptomyces* extracts were the plant-associated monoterpenes, carveol ($[M\text{-}H_2O + H]^+$, *m/z* 135.117), and cuminyl alcohol ($[M\text{-}H_2O + H]^+$, *m/z* 133.101), whereas hydroxyvalerenic acid (another plant terpene, $[2M\text{-}H]^-$, *m/z* 499.307) was found in *S. clavuligerus* extracts only (**Figure 2C** and **Supplementary Table S5**). The pathways involved in the production of the latter three metabolites are not fully known (Wong et al., 2018), however, *S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus* possess many terpene-like BGCs of unknown function, which could potentially be involved in their biosynthesis (**Supplementary Table S4**). Therefore, our results suggest that certain *Streptomyces* also harbor the capacity to produce carveol, cuminyl alcohol, and hydroxyvalerenic acid along with naringenin, a finding that can be potentially exploited for further development.

The indole alkaloid, (−)-indolactam V is a protein kinase C activator (Heikkila and Akerman, 1989) and functions as an intermediate during the biosynthesis of other SMs in certain Actinomycetes (Abe, 2018). We detected (−)-indolactam V (**Figure 2C**, $[M\text{-}CO + H]^+$, *m/z* 274.191) and some of its alkylated derivatives in extracts from *S. clavuligerus*, but not in those from *S. jumonjinensis* or *S. katsurahamanus* (**Supplementary Table S5**). The genes normally associated with (−)-indolactam V biosynthesis could not be identified in the current study, warranting further investigation into its production in *S. clavuligerus*. Other metabolites were also detected during the analysis (**Supplementary Table S6**), but we were unable to find details about their biosynthesis in bacteria or predict associated BGCs, and therefore we did not include them in the discussion. In addition, *S. jumonjinensis* and *S. katsurahamanus* contain several BGCs related to known pathways for which products could not be detected (**Supplementary Table S4**). For example, there is an NRPS-containing BGC in *S. jumonjinensis* that is 100% similar to the BGC in *Streptomyces* sp. DSM 11171, which produces the antiviral metabolite feglymycin (**Supplementary Table S4**; Gonsior et al., 2015). We also identified indole-associated BGCs in *S. clavuligerus* and *S. jumonjinensis* (**Supplementary Table S4**), which are similar to the one from *Streptomyces* sp. TP-A0274 responsible for producing the anticancer agent staurosporine (Onaka et al., 2002). Similarly, BGCs for polycyclic tetramate macrolactams (PTMs, NRP/PKs) are present in both *S. jumonjinensis* and *S. katsurahamanus*, which are 100% similar to a SGR-PTM BGC from the known producer *S. griseus* (**Supplementary Table S4**; Luo et al., 2013). PTMs possess antifungal and antioxidant properties, and cryptic PTM-like BGCs are commonly found in *Streptomyces* genomes (Zhang et al., 2016). Moreover, BGCs for many other classes of SMs including enediynes (Rudolf et al., 2016) and the ribosomally synthesized and post-translationally modified peptides (RiPPs) (Hetrick and van der Donk, 2017) were also identified in *S. jumonjinensis* and *S. katsurahamanus* (**Supplementary Table S4**), but further work is required to detect their production in these organisms. In the current study, >14,000 molecular nodes were obtained using MS-based metabolomics and GNPS analysis (**Figure 2C**), but only 10% could be annotated by matching spectra with available libraries. Therefore, many of the unannotated nodes could represent products of so-called "cryptic" BGCs, a situation that should change over time as databases are populated with more spectra from authentic samples.

Comparative Sequence Analysis of CA-BGCs From *Streptomyces* Species

In addition to analyzing the overall SM production capabilities of CA producers, we were also interested in specifically examining the BGCs involved in β-lactam biosynthesis from *S. jumonjinensis* and *S. katsurahamanus* for comparison with *S. clavuligerus* (**Figure 3**). The genome sequences of *S. jumonjinensis* and *S. katsurahamanus* revealed that they both contain identical CA and Ceph-C BGCs (**Figure 3**), but lack the clavam and paralog gene clusters (**Supplementary Table S4**). This would explain why they do not produce the 5*S* clavams as compared to *S. clavuligerus* (Jensen, 2012). The results further confirm that intact 5*S* clavam and paralog BGCs are not essential for CA production (**Figure 1**), since both *S. jumonjinensis* and *S. katsurahamanus* can produce the metabolite (**Figure 2B** and **Supplementary Figure S1**). The paralog gene cluster from *S. clavuligerus* contains second copies of certain genes (*ceaS1, bls1*, and *pah1*) from the CA BGC (Jensen et al., 2004b; Tahlan et al., 2004b), which encode enzymes involved in the early shared stages of CA and 5*S* clavam biosynthesis (**Figure 1**). It has also been shown that the remaining un-duplicated genes from the paralog gene cluster and almost all genes from the clavam gene cluster (except one; *cas1*) are exclusively involved in 5*S* clavam production (Mosher et al., 1999; Tahlan et al., 2007; Zelyas et al., 2008). Therefore, our results provide additional support for the hypothesis that the clavam and paralog gene clusters are associated with 5*S* clavam biosynthesis, and that some gene products from the two clusters augment CA production in *S. clavuligerus* by contributing to a common pool of precursors (**Figure 1**; Jensen, 2012; Hamed et al., 2013). Although, it should be noted that in *S. clavuligerus*, there is some cross regulation between the chromosomal CA and plasmid-borne paralog gene clusters (Kwong et al., 2013; Álvarez-Álvarez et al., 2017), which is again not expected to occur in the other two CA producers since they only contain the CA BGC. This also highlights the complexity of the regulatory pathways controlling CA and 5*S* clavam production in *S. clavuligerus* (Liras et al., 2008). For this reason, we focused our analysis and discussion on the comparison of biosynthetic genes (and BGCs), instead of regulation. In the current study, CA production levels in *S. jumonjinensis* and *S. katsurahamanus* could never match those observed in wt *S. clavuligerus*, whereas all three species produced Ceph-C at comparable levels (**Figure 2B**). It has been previously suggested that higher CA yields in *S. clavuligerus* might be explained in part by increased precursor supply for biosynthesis due to the presence of the paralog and clavam gene clusters in this species (**Figure 1**). In addition, enhanced levels of biosynthetic gene expression could be another reason why *S. clavuligerus* is currently the preferred industrial producer and was first identified in screens for β-lactamase inhibitors, as higher CA yields would make it easier to detect during assays (Jensen, 2012).

Closer examination of the CA BGCs from *S. jumonjinensis* and *S. katsurahamanus* showed that they each contain most of the genes from the corresponding *S. clavuligerus* BGC in the same order, except that *orf18* (*pbpA*), *orf20*, *orf21*, *orf22*, and *orf23* are absent (**Figure 3A**). *pbpA* is predicted to encode a high-molecular-weight penicillin-binding protein (PBP), but its role in CA production remains unknown (Jensen et al., 2004a). Previous studies have also shown that disruption of *orf19* (*pbp2*) (Jensen et al., 2004a), *orf20* (cytochrome P-450) (Song et al., 2009), *orf21* (putative sigma factor), *orf22* (sensor kinase), or *orf23* (response regulator) (Fu et al., 2019a) in *S. clavuligerus* does not abolish CA or Ceph-C production (Song et al., 2009; **Supplementary Table S1**). Since the respective genes are not present in *S. jumonjinensis* and *S. katsurahamanus* (**Figure 3A**), it is apparent that they are not part of the core BGC required for biosynthesis, but instead have accessory roles in *S. clavuligerus*. In a previous study, it was also shown that the expression of *orf18–21* was not significantly affected in a *S. clavuligerus* mutant defective in ClaR, the cluster-situated regulator responsible for controlling CA biosynthesis (Martínez-Burgo et al., 2015). Therefore, we propose that the core CA BGC comprises *ceaS2* (encoding carboxyethylarginine synthase), *gcas* (encoding *N*-glycyl-clavaminic acid synthetase), and the intervening genes (**Figure 3A**, and **Supplementary Table S1**).

The CA and Ceph-C BGCs in *S. jumonjinensis* and *S. katsurahamanus* also form "β-lactam superclusters" as observed in *S. clavuligerus*, which agrees with previous restriction mapping studies (Ward and Hodgson, 1993). The linkage of the Ceph-C and CA BGCs in *S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus*, and the coordinated production of the two metabolites in *S. clavuligerus* (Pérez-Llarena et al., 1997), provides further evidence for the simultaneous acquisition of the two BGCs by producing species. It has been proposed that the CA BGC might have evolved by the duplication of an ancestral 5*S* clavam BGC and the acquisition of the ability to produce Ceph-C in the same organism (Challis and Hopwood, 2003). Such a situation led to the selection for the ability to produce a β-lactamase inhibitor, resulting in the assembly of the currently known CA BGC, and the formation of the β-lactam supercluster (Challis and Hopwood, 2003). Our results showed that the Ceph-C BGCs from *S. jumonjinensis* and *S. katsurahamanus* are identical to each other, but differ slightly from those present in *S. clavuligerus* and other Ceph-C-producing Actinobacteria (**Figures 3B,D**). The positions of genes forming individual operons (or transcriptional units) in all three CA producers is very similar (except for the location of *cefD*), but the relative arrangement of operons is different in *S. jumonjinensis* and *S. katsurahamanus* as compared to *S. clavuligerus* (**Figure 3B**). In addition, the Ceph-C BGCs of *S. jumonjinensis*, *S. katsurahamanus* and other previously reported Ceph-C producers (other than *S. clavuligerus*) (Liras et al., 1998) do not contain *blp* (**Figure 3B**), which encodes a product resembling β-lactamase inhibitory proteins (Blip), but has been shown to lack any such activity (Gretes et al., 2009). Previous studies have shown that disruption of *blp* does not affect Ceph-C or CA production in *S. clavuligerus* (Alexander and Jensen, 1998; Thai et al., 2001). Therefore, *blp* does not seem to a part of the core Ceph-C BGC since *S. jumonjinensis*, *S. katsurahamanus*, and other species shown in **Figure 3B** can still produce the metabolite in its absence. Another noticeable feature of Ceph-C BGCs from the three CA producers is the presence of *pcbR*, which is missing from the homologous BGCs of species that only produce Ceph-C, but not CA (**Figure 3B**). PcbR resembles PBPs (Paradkar et al., 1996), but it is not essential for Ceph-C biosynthesis since it is not present in the BGCs of all organisms capable of producing the metabolite (**Figure 3B**, more details below).

Overall, the "β-lactam superclusters" from *S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus* are very similar to each other (**Figures 3C,D**). In comparison, CA-like BGCs from non-producers are markedly different, and do not form "β-lactam superclusters" as they lack Ceph-C BGCs (Jensen, 2012). The non-producers (including some *Streptomyces*) are also phylogenetically distinct from CA-producing species (**Supplementary Figure S2**), and their CA-like BGCs show three distinct patterns in terms of gene content and arrangement (**Figure 3A**). Many organisms in the database contain CA-like BGCs identical to the one found in *S. pratensis*, whereas we could only find one example each of the types present in *Streptomyces* sp. M41 and *S. viridis*, respectively (**Figure 3A**). In addition, CA-like BGCs from *S. pratensis* and *S. viridis* contain the *pcbR*, *orf11*, and *nocE* genes (Álvarez-Álvarez et al., 2013), which are not present in the CA BGCs of *S. clavuligerus*, *S. jumonjinensis*, or *S. katsurahamanus* (**Figure 3A**). Interestingly, *pcbR* and *orf11* are included in the Ceph-C BGCs of CA producers, whereas *nocE* is located elsewhere on the chromosome in the three *Streptomyces* species (**Figure 3B**). As mentioned earlier, *pcbR* encodes a PBP involved in β-lactam resistance (Paradkar et al., 1996), whereas *orf11* encodes a predicted protein of unknown function. Previous reports have shown that disruption of neither *pcbR* nor *orf11* in *S. clavuligerus* affected Ceph-C or CA production (Paradkar et al., 1996; Alexander and Jensen, 1998), suggesting that they are not required for the biosynthesis of the respective metabolites.

The presence of *nocE* homologs in CA producers and in the CA-like BGCs of all non-producers is intriguing (**Figure 3A**), as they are similar to a gene from the nocardicin A monobactam BGC of *Nocardia uniformis* (Gunsior et al., 2004). The *nocE* genes are predicted to encode proteins containing C-terminal SGNH/GDSL hydrolase family domains, which are normally associated with esterases or lipases (Upton and Buckley, 1995), but their function during β-lactam metabolite biosynthesis is not obvious. The disruption of *nocE* in *N. uniformis* does not affect nocardicin A production (Davidsen and Townsend, 2009), but the role of the gene in β-lactam-producing *Streptomyces* has not been examined to date.

Examination of the Function of nocE in *S. clavuligerus*

In previous studies, every gene from the proposed CA BGC of *S. clavuligerus* (**Figure 3A**) was systematically disrupted

(**Supplementary Table S1**), to determine if it had any effect on CA or Ceph-C production. It has been suggested that *nocE* might have some role during CA biosynthesis in *S. clavuligerus*, but since the gene is not part of the CA BGC, a mutant has not been prepared and analyzed to date (Jensen, 2012). Therefore, the function of *nocE* was examined in the model CA producer, *S. clavuligerus*. RT-PCR analysis of RNA isolated from wt *S. clavuligerus* grown in SA medium demonstrated that *nocE* is temporally expressed along with *ceaS2* and *cas2* (**Figure 4A**), genes that are essential for CA biosynthesis

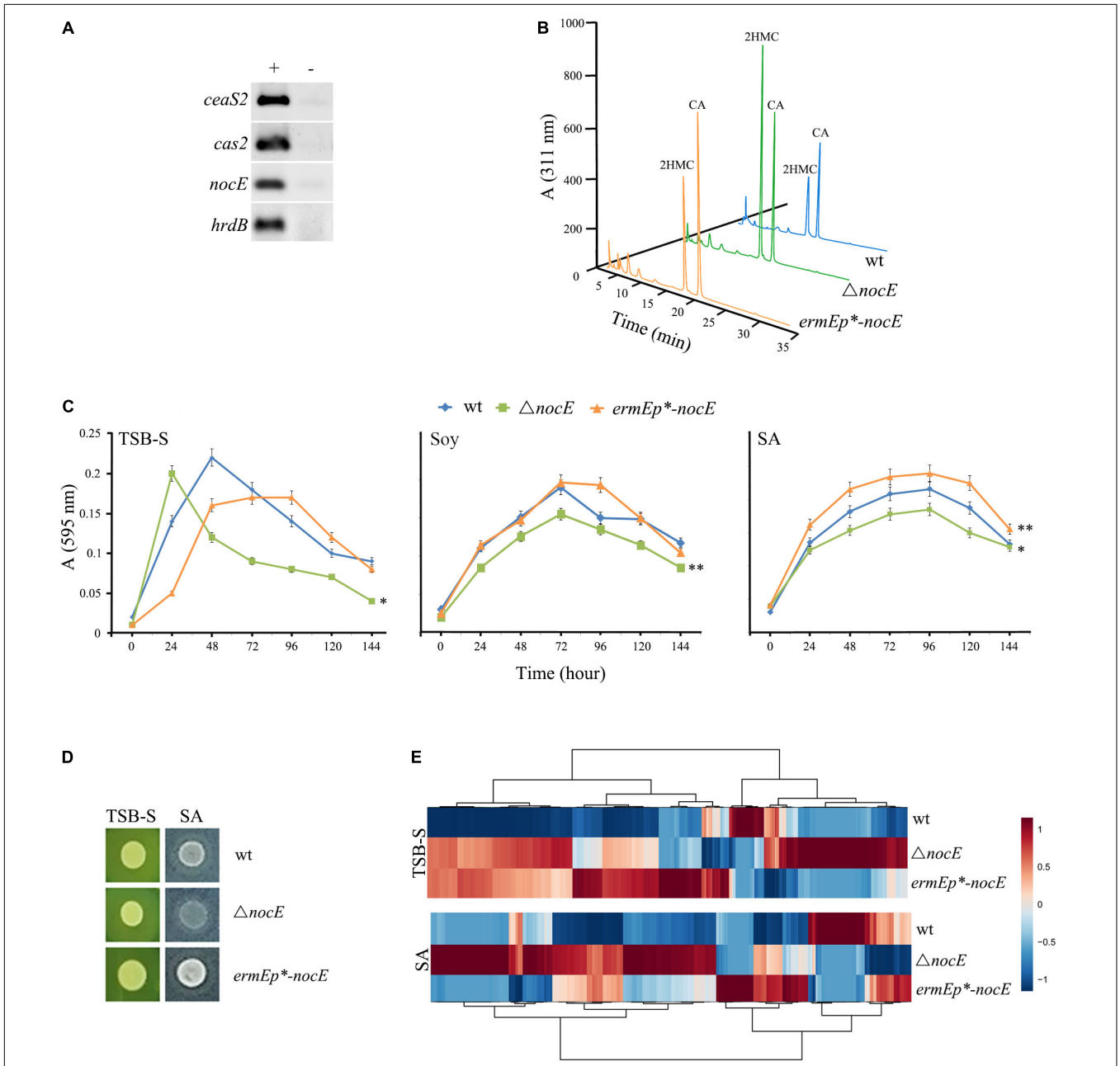


FIGURE 4 | Examination of the function of *nocE* in *S. clavuligerus*. **(A)** RT-PCR analysis (+) of RNA isolated from 96-h *S. clavuligerus* SA cultures showing the expression of *nocE* during CA production. Transcription of *ceaS2* and *cas2* was used as a reporter for CA-BGC expression, whereas that of the constitutively expressed *hrdB* was used as a control. Negative controls (–) consisted of RNA samples subjected to PCR without undergoing RT. **(B)** LC-MS analysis of imidazole derivatized 96-h soy culture (different media from **Figure 2B**) supernatants form the *S. clavuligerus* wt, Δ*nocE*, and *ermEp*-nocE* (constitutive expression) strains to assess CA and 5*S* clavam metabolite production. **(C,D)** Growth characteristics of the *S. clavuligerus* wt, Δ*nocE*, and *ermEp*-nocE* strains in broth **(D)** or agar **(D)** cultures under different nutritional conditions, where (*) and (**) indicate *p* values of less than 0.05 and 0.001, respectively. **(E)** Comparative metabolomics of the *S. clavuligerus* wt, Δ*nocE* and *ermEp*-nocE* strains grown on two different media as shown in panel **(D)**. The heat map was constructed by hierarchical clustering of ∼1000 statically significant features to show overall differences between the three strains.

(**Figure 1**). However, when *S. clavuligerus* strains were prepared in which *nocE* was either deleted (Δ*nocE*) or constitutively expressed (*ermE*p*-*nocE*) (**Table 1**), the production of CA, 5*S* clavams, or Ceph-C was found to be unaffected (**Figure 4B** and **Supplementary Figure S3**), demonstrating that the gene is not required for β-lactam metabolite production in *S. clavuligerus*. The predicted lipase/esterase-like domain present in NocE is also found in hydrolytic enzymes from other *Streptomyces* species, some of which are known to be secreted (Wei et al., 1995; Vujaklija et al., 2002). Closer examination of the predicted NocE amino acid sequence from *S. clavuligerus* suggested that it is also a secreted protein, as it contains a highly conserved N-terminal Sec-signal sequence ($p > 0.9$) (Almagro Armenteros et al., 2019). These findings further ruled out the direct involvement of NocE in CA production, which occurs in the cytoplasm, and suggested that NocE might have some other exocellular hydrolytic function instead. Therefore, the *S. clavuligerus* wt, Δ*nocE*, and *ermE*p*-*nocE* strains were assessed for growth under different nutritional condition using TSB-S (rich), soy (complex fermentation), or SA (defined fermentation) media (**Figure 4C**). It was observed that the growth of the *S. clavuligerus*Δ*nocE* mutant was significantly reduced in each medium tested, whereas that of the *ermE*p*-*nocE* strain was enhanced in SA medium only, when compared to the wt strain (**Figure 4C**). The growth of the three strains was also assessed on TSB-S and SA agar, which again showed that the *S. clavuligerus* Δ*nocE* mutant did not grow as well as the other strains in the latter medium (**Figure 4D**). To examine the influence of *nocE* on primary metabolism in *S. clavuligerus*, the wt, Δ*nocE*, and *ermE*p*-*nocE* strains were grown on TSB-S and SA agar for metabolomics analysis, which showed marked differences in overall metabolite levels between the respective strains (**Figure 4E**). Furthermore, metabolomics analysis showed that SM production in *S. clavuligerus* was unaffected in the Δ*nocE* mutant as compared to the wt strain. Therefore, based on all evidence collected so far, it seems plausible that NocE could have some extracellular role in nutrient acquisition in *S. clavuligerus,* but like *pcbR* and *orf11*, it is not required for CA or Ceph-C production under the tested conditions.

CONCLUSION

To summarize, we have shown that *S. clavuligerus*, *S. jumonjinensis*, and *S. katsurahamanus* contain numerous BGCs and that they synthesize many SMs, including the plant-associated metabolites, naringenin, and valerenic acid. It is possible that genes encoding enzymes for the synthesis of plant-associated metabolites are present in *Streptomyces* genomes, but they are not easily identified due to their organization, since some of them do not form BGCs (Álvarez-Álvarez et al., 2015; Nybo et al., 2017). In addition, plants normally produce metabolites like valerenic acid in low amounts, and for this reason, their heterologous production has been recently attempted in *Saccharomyces* and *Escherichia coli* (Nybo et al., 2017; Wong et al., 2018). The finding that certain *Streptomyces* species can synthesize these metabolites naturally could provide future avenues for their overproduction in a native host. Our results also show similarities and differences in the overall specialized metabolic capabilities of CA-producing *Streptomyces* species under different nutritional conditions, which, to the best of our knowledge, is the first report on the subject. Although the current study did not examine or address regulation, we would like to point out that many of the genes known to control Ceph-C and CA production in *S. clavuligerus* are also conserved in the two other producers (Liras et al., 2008; Ferguson et al., 2016; Fu et al., 2019b). It has been noted that deciphering the complete CA biosynthetic pathway in *S. clavuligerus* is challenging due to the presence of the 5*S* clavam biosynthetic pathway. The current report provides a framework for future studies on CA biosynthesis using *S. jumonjinensis* or *S. katsurahamanus* as models due to the absence of such competing or overlapping pathways in these organisms. Our analyses have also allowed us to propose the core group of genes involved in CA biosynthesis and have helped us to rule out the involvement of *nocE* and other genes in the production of this important metabolite.

DATA AVAILABILITY STATEMENT

The datasets generated and/or analyzed during this study can be found in the NCBI sequence database (ncbi.nlm.nih.gov/genome) and the MassIVE public repository (massive.ucsd.edu). All accession numbers are provided in the Materials and Methods section.

AUTHOR CONTRIBUTIONS

KT contributed conception, resources, and supervision. FB-G and PD provided reagents, resources, and supervision for genomics and metabolomics analysis, respectively. MM and PC-M performed the genome sequencing and annotation. NA and BP conducted the described comparative genomics analysis. NA prepared and analyzed the *S. clavuligerus nocE* mutant and overexpression strains. NA and SS prepared extracts for LC-MS/MS analysis, which was performed by L-FN. AS and L-FN carried out the metabolomics analysis and compound annotation. NA and MM wrote the first draft of the manuscript, whereas BP, AS, and L-FN wrote specific sections. NA, BP, AS, L-FN, FB-G, and KT contributed to manuscript revision.

FUNDING

This work was supported by operating grants from the Natural Sciences and Engineering Research Council of Canada (NSERC: 386417–2010 and 2018–05949) to KT. NA and MM were the recipients of NSERC graduate student awards, including a Michael Smith Foreign Study Supplement to MM. Memorial

University of Newfoundland also provided graduate student support to NA, BP, MM, and AS.

ACKNOWLEDGMENTS

We thank Zhenglong Cheng (Memorial University of Newfoundland) for technical assistance. We would also like to express our deep gratitude to Dr. Susan E. Jensen (University of Alberta) for sharing the *Streptomyces* species used in the current study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.02550/full#supplementary-material

REFERENCES

Abe, I. (2018). Biosynthetic studies on teleocidins in *Streptomyces*. *J. Antibiot.* 71, 763–768. doi: 10.1038/s41429-018-0069-4

Adusumilli, R., and Mallick, P. (2017). Data conversion with proteowizard msconvert. *Methods Mol. Biol.* 1550, 339–368. doi: 10.1007/978-1-4939-6747-6_23

Alexander, D. C., and Jensen, S. E. (1998). Investigation of the *Streptomyces clavuligerus* cephamycin C gene cluster and its regulation by the CcaR protein. *J. Bacteriol.* 180, 4068–4079.

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423. doi: 10.1038/s41587-019-0036-z

Álvarez-Álvarez, R., Botas, A., Albillos, S. M., Rumbero, A., Martín, J. F., and Liras, P. (2015). Molecular genetics of naringenin biosynthesis, a typical plant secondary metabolite produced by *Streptomyces clavuligerus*. *Microb. Cell Fact.* 14:178. doi: 10.1186/s12934-015-0373-7

Álvarez-Álvarez, R., Martínez-Burgo, Y., Pérez-Redondo, R., Braña, A. F., Martín, J. F., and Liras, P. (2013). Expression of the endogenous and heterologous clavulanic acid cluster in *Streptomyces flavogriseus*: why a silent cluster is sleeping. *Appl. Microbiol. Biotechnol.* 97, 9451–9463. doi: 10.1007/s00253-013-5148-7

Álvarez-Álvarez, R., Martínez-Burgo, Y., Rodríguez-García, A., and Liras, P. (2017). Discovering the potential of *S. clavuligerus* for bioactive compound production: cross-talk between the chromosome and the pSCL4 megaplasmid. *BMC Genomics* 18:907. doi: 10.1186/s12864-017-4289-y

Álvarez-Álvarez, R., Rodríguez-García, A., Martínez-Burgo, Y., Martín, J. F., and Liras, P. (2018). Transcriptional studies on a *Streptomyces clavuligerus* oppA2 deletion mutant: n-acetylglycyl-clavaminic acid is an intermediate of clavulanic acid biosynthesis. *Appl. Environ. Microbiol.* 84:e1701-18. doi: 10.1128/AEM.01701-18

Baltz, R. H. (2008). Renaissance in antibacterial discovery from actinomycetes. *Curr. Opin. Pharmacol.* 8, 557–563. doi: 10.1016/j.coph.2008.04.008

Barona-Gómez, F., Wong, U., Giannakopulos, A. E., Derrick, P. J., and Challis, G. L. (2004). Identification of a cluster of genes that directs desferrioxamine biosynthesis in *Streptomyces coelicolor* M145. *J. Am. Chem. Soc.* 126, 16282–16283. doi: 10.1021/ja045774k

Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., et al. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417, 141–147. doi: 10.1038/417141a

Bibb, M. J., Janssen, G. R., and Ward, J. M. (1985). Cloning and analysis of the promoter region of the erythromycin resistance gene (ermE) of *Streptomyces erythraeus*. *Gene* 38, 215–226. doi: 10.1016/0378-1119(85)90220-3

Blin, K., Wolf, T., Chevrette, M. G., Lu, X., Schwalen, C. J., Kautsar, S. A., et al. (2017). antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* 45, W36–W41. doi: 10.1093/nar/gkx319

Böcker, S., Letzel, M. C., Lipták, Z., and Pervukhin, A. (2009). SIRIUS: decomposing isotope patterns for metabolite identification. *Bioinformatics* 25, 218–224. doi: 10.1093/bioinformatics/btn603

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Brown, A. G. (1986). Clavulanic acid, a novel beta-lactamase inhibitor-a case study in drug discovery and development. *Drug Des. Deliv.* 1, 1–21.

Brown, A. G., Butterworth, D., Cole, M., Hanscomb, G., Hood, J. D., Reading, C., et al. (1976). Naturally-occurring beta-lactamase inhibitors with antibacterial activity. *J. Antibiot.* 29, 668–669. doi: 10.7164/antibiotics.29.668

Brown, D., Evans, J. R., and Fletton, R. A. (1979). Structures of three novel β-lactams isolated from *Streptomyces clavuligerus*. *J. Chem. Soc. Chem. Commun.* 6, 282–283. doi: 10.1039/c39790000282

Burrell, M., Hanfrey, C. C., Kinch, L. N., Elliott, K. A., and Michael, A. J. (2012). Evolution of a novel lysine decarboxylase in siderophore biosynthesis. *Mol. Microbiol.* 86, 485–499. doi: 10.1111/j.1365-2958.2012.08208.x

Cao, G., Zhong, C., Zong, G., Fu, J., Liu, Z., Zhang, G., et al. (2016). Complete genome sequence of *Streptomyces clavuligerus* F613-1, an industrial producer of clavulanic acid. *Genome Announc.* 4:e1020-16. doi: 10.1128/genomeA.01020-16

Challis, G. L. (2005). A widely distributed bacterial pathway for siderophore biosynthesis independent of nonribosomal peptide synthetases. *Chembiochem* 6, 601–611. doi: 10.1002/cbic.200400283

Challis, G. L., and Hopwood, D. A. (2003). Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proc. Natl. Acad. Sci. U.S.A.* 100(Suppl. 2), 14555–14561. doi: 10.1073/pnas.1934677100

Chen, W., Qu, D., Zhai, L., Tao, M., Wang, Y., Lin, S., et al. (2010). Characterization of the tunicamycin gene cluster unveiling unique steps involved in its biosynthesis. *Protein Cell* 1, 1093–1105. doi: 10.1007/s13238-010-0127-6

Chong, J., Soufan, O., Li, C., Caraus, I., Li, S., Bourque, G., et al. (2018). MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res.* 46, W486–W494. doi: 10.1093/nar/gky310

Czech, L., Hermann, L., Stöveken, N., Richter, A. A., Hoppner, A., Smits, S. H. J., et al. (2018). Role of the extremolytes ectoine and hydroxyectoine as stress protectants and nutrients: genetics, phylogenomics, biochemistry, and structural analysis. *Genes* 9:E177. doi: 10.3390/genes9040177

Davidsen, J. M., and Townsend, C. A. (2009). Identification and characterization of NocR as a positive transcriptional regulator of the beta-lactam nocardicin A in *Nocardia uniformis*. *J. Bacteriol.* 191, 1066–1077. doi: 10.1128/JB.01833-07

de la Fuente, A., Lorenzana, L. M., Martín, J. F., and Liras, P. (2002). Mutants of *Streptomyces clavuligerus* with disruptions in different genes for clavulanic acid biosynthesis produce large amounts of holomycin: possible cross-regulation of two unrelated secondary metabolic pathways. *J. Bacteriol.* 184, 6559–6565. doi: 10.1128/jb.184.23.6559-6565.2002

Dillman, R. O. (2004). Pentostatin (Nipent) in the treatment of chronic lymphocyte leukemia and hairy cell leukemia. *Expert Rev. Anticancer Ther.* 4, 27–36. doi: 10.1586/14737140.4.1.27

Drawz, S. M., and Bonomo, R. A. (2010). Three decades of beta-lactamase inhibitors. *Clin. Microbiol. Rev.* 23, 160–201. doi: 10.1128/CMR.00037-09

Egan, L. A., Busby, R. W., Iwata-Reuyl, D., and Townsend, C. A. (1997). Probable role of clavaminic acid as the terminal intermediate in the common pathway to clavulanic acid and the antipodal clavam metabolites. *J. Am. Chem. Soc.* 119, 2348–2355. doi: 10.1021/ja963107o

Ferguson, N. L., Peña-Castillo, L., Moore, M. A., Bignell, D. R., and Tahlan, K. (2016). Proteomics analysis of global regulatory cascades involved in clavulanic acid production and morphological development in *Streptomyces clavuligerus*. *J. Ind. Microbiol. Biotechnol.* 43, 537–555. doi: 10.1007/s10295-016-1733-y

Fernández-Martínez, L. T., and Bibb, M. J. (2014). Use of the meganuclease I-SceI of *Saccharomyces cerevisiae* to select for gene deletions in actinomycetes. *Sci. Rep.* 4:7100. doi: 10.1038/srep07100

Fu, J., Qin, R., Zong, G., Liu, C., Kang, N., Zhong, C., et al. (2019a). The CagRS two-component system regulates clavulanic acid metabolism via multiple pathways in *Streptomyces clavuligerus* F613-1. *Front. Microbiol.* 10:244. doi: 10.3389/fmicb.2019.00244

Fu, J., Qin, R., Zong, G., Zhong, C., Zhang, P., Kang, N., et al. (2019b). The two-component system CepRS regulates the cephamycin C biosynthesis in *Streptomyces clavuligerus* F613-1. *AMB Express* 9:118. doi: 10.1186/s13568-019-0844-z

Gonsior, M., Mühlenweg, A., Tietzmann, M., Rausch, S., Poch, A., and Süssmuth, R. D. (2015). Biosynthesis of the peptide antibiotic feglymycin by a linear nonribosomal peptide synthetase mechanism. *Chembiochem* 16, 2610–2614. doi: 10.1002/cbic.201500432

Gretes, M., Lim, D. C., de Castro, L., Jensen, S. E., Kang, S. G., Lee, K. J., et al. (2009). Insights into positive and negative requirements for protein-protein interactions by crystallographic analysis of the beta-lactamase inhibitory proteins BLIP, BLIP-I, and BLP. *J. Mol. Biol.* 389, 289–305. doi: 10.1016/j.jmb.2009.03.058

Gunsior, M., Breazeale, S. D., Lind, A. J., Ravel, J., Janc, J. W., and Townsend, C. A. (2004). The biosynthetic gene cluster for a monocyclic beta-lactam antibiotic, nocardicin A. *Chem. Biol.* 11, 927–938. doi: 10.1016/j.chembiol.2004.04.012

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086

Hamed, R. B., Gomez-Castellanos, J. R., Henry, L., Ducho, C., McDonough, M. A., and Schofield, C. J. (2013). The enzymes of beta-lactam biosynthesis. *Nat. Prod. Rep.* 30, 21–107. doi: 10.1039/c2np20065a

Heikkila, J., and Akerman, K. E. (1989). (-)-Indolactam V activates protein kinase C and induces changes in muscarinic receptor functions in SH-SY5Y human neuroblastoma cells. *Biochem. Biophys. Res. Commun.* 162, 1207–1213. doi: 10.1016/0006-291x(89)90802-4

Hetrick, K. J., and van der Donk, W. A. (2017). Ribosomally synthesized and post-translationally modified peptide natural product discovery in the genomic era. *Curr. Opin. Chem. Biol.* 38, 36–44. doi: 10.1016/j.cbpa.2017.02.005

Hwang, S., Lee, N., Jeong, Y., Lee, Y., Kim, W., Cho, S., et al. (2019). Primary transcriptome and translatome analysis determines transcriptional and translational regulatory elements encoded in the *Streptomyces clavuligerus* genome. *Nucleic Acids Res.* 47, 6114–6129. doi: 10.1093/nar/gkz471

Jensen, S. E. (2012). Biosynthesis of clavam metabolites. *J. Ind. Microbiol. Biotechnol.* 39, 1407–1419. doi: 10.1007/s10295-012-1191-0

Jensen, S. E., Elder, K. J., Aidoo, K. A., and Paradkar, A. S. (2000). Enzymes catalyzing the early steps of clavulanic acid biosynthesis are encoded by two sets of paralogous genes in *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.* 44, 720–726. doi: 10.1128/aac.44.3.720-726.2000

Jensen, S. E., and Paradkar, A. S. (1999). Biosynthesis and molecular genetics of clavulanic acid. *Antonie Van Leeuwenhoek* 75, 125–133.

Jensen, S. E., Paradkar, A. S., Mosher, R. H., Anders, C., Beatty, P. H., Brumlik, M. J., et al. (2004a). Five additional genes are involved in clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.* 48, 192–202. doi: 10.1128/aac.48.1.192-202.2004

Jensen, S. E., Wong, A., Griffin, A., and Barton, B. (2004b). *Streptomyces clavuligerus* has a second copy of the proclavaminate amidinohydrolase gene. *Antimicrob. Agents Chemother.* 48, 514–520. doi: 10.1128/aac.48.2.514-520.2004

Kalinovskaya, N. I., Romanenko, L. A., Irisawa, T., Ermakova, S. P., and Kalinovsky, A. I. (2011). Marine isolate Citricoccus sp. KMM 3890 as a source of a cyclic siderophore nocardamine with antitumor activity. *Microbiol. Res.* 166, 654–661. doi: 10.1016/j.micres.2011.01.004

Kanno, S., Tomizawa, A., Hiura, T., Osanai, Y., Shouji, A., Ujibe, M., et al. (2005). Inhibitory effects of naringenin on tumor growth in human cancer cell lines and sarcoma S-180-implanted mice. *Biol. Pharm. Bull.* 28, 527–530. doi: 10.1248/bpb.28.527

Katz, L., and Baltz, R. H. (2016). Natural product discovery: past, present, and future. *J. Ind. Microbiol. Biotechnol.* 43, 155–176. doi: 10.1007/s10295-015-1723-5

Kenig, M., and Reading, C. (1979). Holomycin and an antibiotic (MM 19290) related to tunicamycin, metabolites of *Streptomyces clavuligerus*. *J. Antibiot.* 32, 549–554. doi: 10.7164/antibiotics.32.549

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000). *Practical Streptomyces Genetics*. Norwich: The John Innes Foundation.

Kwong, T., Tahlan, K., Anders, C. L., and Jensen, S. E. (2013). Carboxyethylarginine synthase genes show complex cross-regulation in *Streptomyces clavuligerus*. *Appl. Environ. Microbiol.* 79, 240–249. doi: 10.1128/AEM.02600-12

Li, B., and Walsh, C. T. (2010). Identification of the gene cluster for the dithiolopyrrolone antibiotic holomycin in *Streptomyces clavuligerus*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19731–19735. doi: 10.1073/pnas.1014140107

Li, R., Khaleeli, N., and Townsend, C. A. (2000). Expansion of the clavulanic acid gene cluster: identification and *in vivo* functional analysis of three new genes required for biosynthesis of clavulanic acid by *Streptomyces clavuligerus*. *J. Bacteriol.* 182, 4087–4095. doi: 10.1128/jb.182.14.4087-4095.2000

Liras, P. (1999). Biosynthesis and molecular genetics of cephamycins. Cephamycins produced by actinomycetes. *Antonie Van Leeuwenhoek* 75, 109–124.

Liras, P., Gomez-Escribano, J. P., and Santamarta, I. (2008). Regulatory mechanisms controlling antibiotic production in *Streptomyces clavuligerus*. *J. Ind. Microbiol. Biotechnol.* 35, 667–676. doi: 10.1007/s10295-008-0351-8

Liras, P., Rodríguez-García, A., and Martín, J. F. (1998). Evolution of the clusters of genes for beta-lactam antibiotics: a model for evolutive combinatorial assembly of new beta-lactams. *Int. Microbiol.* 1, 271–278.

Lucas, X., Senger, C., Erxleben, A., Gruning, B. A., Doring, K., Mosch, J., et al. (2013). StreptomeDB: a resource for natural compounds isolated from *Streptomyces* species. *Nucleic Acids Res.* 41, D1130–D1136. doi: 10.1093/nar/gks1253

Luo, Y., Huang, H., Liang, J., Wang, M., Lu, L., Shao, Z., et al. (2013). Activation and characterization of a cryptic polycyclic tetramate macrolactam biosynthetic gene cluster. *Nat. Commun.* 4:2894. doi: 10.1038/ncomms3894

Martínez-Burgo, Y., Álvarez-Álvarez, R., Rodríguez-García, A., and Liras, P. (2015). The pathway-specific regulator ClaR of *Streptomyces clavuligerus* has a global effect on the expression of genes for secondary metabolism and differentiation. *Appl. Environ. Microbiol.* 81, 6637–6648. doi: 10.1128/AEM.00916-15

Martínez-Burgo, Y., Santos-Aberturas, J., Rodríguez-García, A., Barreales, E. G., Tormo, J. R., Truman, A. W., et al. (2019). Activation of secondary metabolite gene clusters in *Streptomyces clavuligerus* by the pimm regulator of *Streptomyces natalensis*. *Front. Microbiol.* 10:580. doi: 10.3389/fmicb.2019.00580

Medema, M. H., Trefzer, A., Kovalchuk, A., van den Berg, M., Muller, U., Heijne, W., et al. (2010). The sequence of a 1.8-mb bacterial linear plasmid reveals a rich evolutionary reservoir of secondary metabolic pathways. *Genome Biol. Evol.* 2, 212–224. doi: 10.1093/gbe/evq013

Mellado, E., Lorenzana, L. M., Rodriguez-Saiz, M., Diez, B., Liras, P., and Barredo, J. L. (2002). The clavulanic acid biosynthetic cluster of *Streptomyces clavuligerus*: genetic organization of the region upstream of the car gene. *Microbiology* 148(Pt 5), 1427–1438. doi: 10.1099/00221287-148-5-1427

Mosher, R. H., Paradkar, A. S., Anders, C., Barton, B., and Jensen, S. E. (1999). Genes specific for the biosynthesis of clavam metabolites antipodal to clavulanic acid are clustered with the gene for clavaminate synthase 1 in *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.* 43, 1215–1224. doi: 10.1128/aac.43.5.1215

Nybo, S. E., Saunders, J., and McCormick, S. P. (2017). Metabolic engineering of *Escherichia coli* for production of valerenadiene. *J. Biotechnol.* 262, 60–66. doi: 10.1016/j.jbiotec.2017.10.004

Ohnishi, Y., Ishikawa, J., Hara, H., Suzuki, H., Ikenoya, M., Ikeda, H., et al. (2008). Genome sequence of the streptomycin-producing microorganism *Streptomyces griseus* IFO 13350. *J. Bacteriol.* 190, 4050–4060. doi: 10.1128/JB.00204-08

Olivieri, N. F., and Brittenham, G. M. (1997). Iron-chelating therapy and the treatment of thalassemia. *Blood* 89, 739–761. doi: 10.1182/blood.v89.3.739

Onaka, H., Taniguchi, S., Igarashi, Y., and Furumai, T. (2002). Cloning of the staurosporine biosynthetic gene cluster from *Streptomyces* sp. TP-A0274 and its heterologous expression in *Streptomyces lividans*. *J. Antibiot.* 55, 1063–1071. doi: 10.7164/antibiotics.55.1063

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., et al. (2014). The seed and the rapid annotation of microbial genomes using subsystems

technology (RAST). *Nucleic Acids Res.* 42, D206–D214. doi: 10.1093/nar/gkt1226

Paradkar, A. (2013). Clavulanic acid production by *Streptomyces clavuligerus*: biogenesis, regulation and strain improvement. *J. Antibiot.* 66, 411–420. doi: 10.1038/ja.2013.26

Paradkar, A. S., Aidoo, K. A., Wong, A., and Jensen, S. E. (1996). Molecular analysis of a beta-lactam resistance gene encoded within the cephamycin gene cluster of *Streptomyces clavuligerus*. *J. Bacteriol.* 178, 6266–6274. doi: 10.1128/jb.178.21.6266-6274.1996

Paradkar, A. S., and Jensen, S. E. (1995). Functional analysis of the gene encoding the clavaminate synthase 2 isoenzyme involved in clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *J. Bacteriol.* 177, 1307–1314. doi: 10.1128/jb.177.5.1307-1314.1995

Pérez-Llarena, F. J., Liras, P., Rodríguez-García, A., and Martín, J. F. (1997). A regulatory gene (ccaR) required for cephamycin and clavulanic acid production in *Streptomyces clavuligerus*: amplification results in overproduction of both beta-lactam compounds. *J. Bacteriol.* 179, 2053–2059. doi: 10.1128/jb.179.6.2053-2059.1997

Pluskal, T., Castillo, S., Villar-Briones, A., and Oresic, M. (2010). MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 11:395. doi: 10.1186/1471-2105-11-395

Prabhu, J., Schauwecker, F., Grammel, N., Keller, U., and Bernhard, M. (2004). Functional expression of the ectoine hydroxylase gene (thpD) from *Streptomyces chrysomallus* in *Halomonas elongata*. *Appl. Environ. Microbiol.* 70, 3130–3132. doi: 10.1128/aem.70.5.3130-3132.2004

Pruess, D. L., and Kellett, M. (1983). Ro 22-5417, a new clavam antibiotic from *Streptomyces clavuligerus*. *I. Discovery and biological activity*. *J. Antibiot.* 36, 208–212. doi: 10.7164/antibiotics.36.208

Rauha, J. P., Remes, S., Heinonen, M., Hopia, A., Kahkonen, M., Kujala, T., et al. (2000). Antimicrobial effects of Finnish plant extracts containing flavonoids and other phenolic compounds. *Int. J. Food Microbiol.* 56, 3–12.

Romero, J., Liras, P., and Martín, J. F. (1984). Dissociation of cephamycin and clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *Appl. Microbiol. Biotechnol.* 20, 318–325. doi: 10.1093/femsle/fnv215

Romero, J., Martín, J. F., Liras, P., Demain, A. L., and Rius, N. (1997). Partial purification, characterization and nitrogen regulation of the lysine epsilon-aminotransferase of *Streptomyces clavuligerus*. *J. Ind Microbiol. Biotechnol.* 18, 241–246.

Rudolf, J. D., Yan, X., and Shen, B. (2016). Genome neighborhood network reveals insights into enediyne biosynthesis and facilitates prediction and prioritization for discovery. *J. Ind. Microbiol. Biotechnol.* 43, 261–276. doi: 10.1007/s10295-015-1671-0

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., et al. (2000). Artemis: sequence visualization and annotation. *Bioinformatics* 16, 944–945. doi: 10.1093/bioinformatics/16.10.944

Sadeghi, A., Soltani, B. M., Nekouei, M. K., Jouzani, G. S., Mirzaei, H. H., and Sadeghizadeh, M. (2014). Diversity of the ectoines biosynthesis genes in the salt tolerant *Streptomyces* and evidence for inductive effect of ectoines on their accumulation. *Microbiol. Res.* 169, 699–708. doi: 10.1016/j.micres.2014.02.005

Sambrook, J. R. D. (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Saudagar, P. S., Survase, S. A., and Singhal, R. S. (2008). Clavulanic acid: a review. *Biotechnol. Adv.* 26, 335–351. doi: 10.1016/j.biotechadv.2008.03.002

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351

Skinnider, M. A., Merwin, N. J., Johnston, C. W., and Magarvey, N. A. (2017). PRISM 3: expanded prediction of natural product chemical structures from microbial genomes. *Nucleic Acids Res.* 45, W49–W54. doi: 10.1093/nar/gkx320

Song, J. Y., Jensen, S. E., and Lee, K. J. (2010a). Clavulanic acid biosynthesis and genetic manipulation for its overproduction. *Appl. Microbiol. Biotechnol.* 88, 659–669. doi: 10.1007/s00253-010-2801-2

Song, J. Y., Jeong, H., Yu, D. S., Fischbach, M. A., Park, H. S., Kim, J. J., et al. (2010b). Draft genome sequence of *Streptomyces clavuligerus* NRRL 3585, a producer of diverse secondary metabolites. *J. Bacteriol.* 192, 6317–6318. doi: 10.1128/JB.00859-10

Song, J. Y., Kim, E. S., Kim, D. W., Jensen, S. E., and Lee, K. J. (2009). 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. doi: 10.1007/s10295-008-0499-2

Spicer, R. A., Salek, R., and Steinbeck, C. (2017). Compliance with minimum information guidelines in public metabolomics repositories. *Sci. Data* 4:170137. doi: 10.1038/sdata.2017.137

Srivastava, S. K., King, K. S., AbuSara, N. F., Malayny, C. J., Piercey, B. M., Wilson, J. A., et al. (2019). In vivo functional analysis of a class A beta-lactamase-related protein essential for clavulanic acid biosynthesis in *Streptomyces clavuligerus*. *PLoS One* 14:e0215960. doi: 10.1371/journal.pone.0215960

Tahlan, K., Anders, C., Wong, A., Mosher, R. H., Beatty, P. H., Brumlik, M. J., et al. (2007). 5S clavam biosynthetic genes are located in both the clavam and paralog gene clusters in *Streptomyces clavuligerus*. *Chem. Biol.* 14, 131–142. doi: 10.1016/j.chembiol.2006.11.012

Tahlan, K., Moore, M. A., and Jensen, S. E. (2017). delta-(L-alpha-aminoadipyl)-L-cysteinyl-D-valine synthetase (ACVS): discovery and perspectives. *J. Ind. Microbiol. Biotechnol.* 44, 517–524. doi: 10.1007/s10295-016-1850-7

Tahlan, K., Park, H. U., and Jensen, S. E. (2004a). Three unlinked gene clusters are involved in clavam metabolite biosynthesis in *Streptomyces clavuligerus*. *Can. J. Microbiol.* 50, 803–810. doi: 10.1139/w04-070

Tahlan, K., Park, H. U., Wong, A., Beatty, P. H., and Jensen, S. E. (2004b). Two sets of paralogous genes encode the enzymes involved in the early stages of clavulanic acid and clavam metabolite biosynthesis in *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.* 48, 930–939. doi: 10.1128/aac.48.3.930-939.2004

Thai, W., Paradkar, A. S., and Jensen, S. E. (2001). Construction and analysis of ss-lactamase-inhibitory protein (BLIP) non-producer mutants of *Streptomyces clavuligerus*. *Microbiology* 147(Pt 2), 325–335. doi: 10.1099/00221287-147-2-325

Townsend, C. A. (2002). New reactions in clavulanic acid biosynthesis. *Curr. Opin. Chem. Biol.* 6, 583–589.

Tyers, M., and Wright, G. D. (2019). Drug combinations: a strategy to extend the life of antibiotics in the 21st century. *Nat. Rev. Microbiol.* 17, 141–155. doi: 10.1038/s41579-018-0141-x

Upton, C., and Buckley, J. T. (1995). A new family of lipolytic enzymes? *Trends Biochem. Sci.* 20, 178–179.

Valegård, K., Iqbal, A., Kershaw, N. J., Ivison, D., Généreux, C., Dubus, A., et al. (2013). Structural and mechanistic studies of the orf12 gene product from the clavulanic acid biosynthesis pathway. *Acta Crystallogr. D Biol. Crystallogr.* 69(Pt 8), 1567–1579. doi: 10.1107/S0907444913011013

van der Heul, H. U., Bilyk, B. L., McDowall, K. J., Seipke, R. F., and van Wezel, G. P. (2018). Regulation of antibiotic production in Actinobacteria: new perspectives from the post-genomic era. *Nat. Prod. Rep.* 35, 575–604. doi: 10.1039/c8np00012c

Vujaklija, D., Schröder, W., Abramiæ, M., Zou, P., Leščič, I., Franke, P., et al. (2002). A novel streptomycete lipase: cloning, sequencing and high-level expression of the *Streptomyces rimosus* GDS(L)-lipase gene. *Arch. Microbiol.* 178, 124–130. doi: 10.1007/s00203-002-0430-6

Wang, L., Tahlan, K., Kaziuk, T. L., Alexander, D. C., and Jensen, S. E. (2004). Transcriptional and translational analysis of the ccaR gene from *Streptomyces clavuligerus*. *Microbiology* 150(Pt 12), 4137–4145. doi: 10.1099/mic.0.27245-0

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., et al. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* 34, 828–837. doi: 10.1038/nbt.3597

Ward, J. M., and Hodgson, J. E. (1993). The biosynthetic genes for clavulanic acid and cephamycin production occur as a 'super-cluster' in three *Streptomyces*. *FEMS Microbiol. Lett.* 110, 239–242. doi: 10.1111/j.1574-6968.1993.tb06326.x

Wei, Y., Schottel, J. L., Derewenda, U., Swenson, L., Patkar, S., and Derewenda, Z. S. (1995). A novel variant of the catalytic triad in the *Streptomyces scabies* esterase. *Nat. Struct. Biol.* 2, 218–223.

Wong, J., d'Espaux, L., Dev, I., van der Horst, C., and Keasling, J. (2018). De novo synthesis of the sedative valerenic acid in *Saccharomyces cerevisiae*. *Metab. Eng.* 47, 94–101. doi: 10.1016/j.ymben.2018.03.005

Wu, P., Wan, D., Xu, G., Wang, G., Ma, H., Wang, T., et al. (2017). An unusual protector-protege strategy for the biosynthesis of purine nucleoside antibiotics. *Cell Chem. Biol.* 24, 171–181. doi: 10.1016/j.chembiol.2016.12.012

Yamanaka, K., Oikawa, H., Ogawa, H. O., Hosono, K., Shinmachi, F., Takano, H., et al. (2005). Desferrioxamine E produced by *Streptomyces griseus* stimulates growth and development of *Streptomyces tanashiensis*. *Microbiology* 151(Pt 9), 2899–2905. doi: 10.1099/mic.0.28139-0

Zelyas, N. J., Cai, H., Kwong, T., and Jensen, S. E. (2008). Alanylclavam biosynthetic genes are clustered together with one group of clavulanic acid biosynthetic genes in *Streptomyces clavuligerus*. *J. Bacteriol.* 190, 7957–7965. doi: 10.1128/JB.00698-08

Zerbino, D. R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi: 10.1101/gr.074492.107

Zhang, G., Zhang, W., Saha, S., and Zhang, C. (2016). Recent advances in discovery, biosynthesis and genome mining of medicinally relevant polycyclic tetramate macrolactams. *Curr. Top. Med. Chem.* 16, 1727–1739.

Zhao, Y., Xiang, S., Dai, X., and Yang, K. (2013). A simplified diphenylamine colorimetric method for growth quantification. *Appl. Microbiol. Biotechnol.* 97, 5069–5077. doi: 10.1007/s00253-013-4893-y

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

PLOS ONE

RESEARCH ARTICLE

In vivo functional analysis of a class A β-lactamase-related protein essential for clavulanic acid biosynthesis in *Streptomyces clavuligerus*

Santosh K. Srivastava☯, Kelcey S. King☯, Nader F. AbuSara, Chelsea J. Malayny, Brandon M. Piercey, Jaime A. Wilson, Kapil Tahlan*

Department of Biology, Memorial University of Newfoundland, St. John's, NL, Canada

☯ These authors contributed equally to this work.
* ktahlan@mun.ca

Citation: Srivastava SK, King KS, AbuSara NF, Malayny CJ, Piercey BM, Wilson JA, et al. (2019) *In vivo* functional analysis of a class A β-lactamase-related protein essential for clavulanic acid biosynthesis in *Streptomyces clavuligerus*. PLoS ONE 14(4): e0215960. https://doi.org/10.1371/journal.pone.0215960

Editor: Pradeep Kumar, Rutgers New Jersey Medical School, UNITED STATES

Received: January 21, 2019

Accepted: April 12, 2019

Published: April 23, 2019

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: The described work was funded by an operating grant from the Natural Science and Engineering Research Council of Canada (NSERC: 386417-2010 and 2018-05949) to KT. KSK, NFA and CJM were also the recipients of NSERC graduate/undergraduate student awards. In addition, we acknowledge Memorial University of

Abstract

In *Streptomyces clavuligerus*, the gene cluster involved in the biosynthesis of the clinically used β-lactamase inhibitor clavulanic acid contains a gene (*orf12* or *cpe*) encoding a protein with a C-terminal class A β-lactamase-like domain. The *cpe* gene is essential for clavulanic acid production, and the recent crystal structure of its product (Cpe) was shown to also contain an N-terminal isomerase/cyclase-like domain, but the function of the protein remains unknown. In the current study, we show that Cpe is a cytoplasmic protein and that both its N- and C-terminal domains are required for *in vivo* clavulanic acid production in *S. clavuligerus*. Our results along with those from previous studies allude towards a biosynthetic role for Cpe during the later stages of clavulanic acid production in *S. clavuligerus*. Amino acids from Cpe essential for biosynthesis were also identified, including one (Lys_{89}) from the recently described N-terminal isomerase-like domain of unknown function. Homologues of Cpe from other clavulanic acid-producing *Streptomyces* spp. were shown to be functionally equivalent to the *S. clavuligerus* protein, whereas those from non-producers containing clavulanic acid-like gene clusters were not. The suggested *in vivo* involvement of an isomerase-like domain recruited by an ancestral β-lactamase related protein, supports a previous hypothesis that Cpe could be involved in a step requiring the opening and modification of the clavulanic acid core during its biosynthesis from 5*S* precursors.

Introduction

The β-lactam class of antibiotics have broad-spectrum activity and include some of the most commonly prescribed agents used for treating bacterial infections [1–3]. They have a long history of use in medicine, but as with other antibiotics, the emergence of resistance is a major problem [3–5]. There are several mechanisms responsible for β-lactam resistance, which include the production of secreted β-lactamases, enzymes that hydrolyze and inactivate certain members of this antibiotic class [6, 7]. Combinations of β-lactamase inhibitors such as

Newfoundland for providing graduate student support to KSK, NFA and BMP. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

clavulanic acid along with β-lactam antibiotics are often used as a strategy for treating some infections caused by β-lactamase-producing antibiotic resistant bacteria [8, 9]. Clavulanic acid belongs to the clavam family of specialized metabolites and it irreversibly inhibits class A β-lactamases, thereby restoring the activity of β-lactam antibiotics against target organisms in such combinations [10, 11]. The activity of clavulanic acid is attributed in part to its 3*R*,5*R* stereochemistry, as other naturally occurring clavams have a 5*S* configuration (collectively referred to as the 5*S* clavams) and do not inhibit β-lactamases [8, 12]. Commercial production of clavulanic acid is achieved by fermenting *Streptomyces clavuligerus*, and a cluster of ~18 genes referred to as the clavulanic acid biosynthetic gene cluster (CA-BGC) encodes components of the core biosynthetic pathway [13]. It has previously been reported that *Streptomyces jumonjinensis* and *Streptomyces katsurahamanus* also produce clavulanic acid, but the sequences of their respective CA-BGCs are not available [12, 14]. On the other hand, the genome sequences of organisms such as *Streptomyces flavogriseus* (ATCC 33331, also known as *S. pratensis*) and *Saccharomonospora viridis* (DSM 43017) contain gene clusters closely resembling the *S. clavuligerus* CA-BGC, but neither has been shown to produce the metabolite to date [13, 15]. In addition, *S. clavuligerus* is somewhat unique among clavulanic acid producers as it also produces certain 5*S* clavams as products of a pathway related to clavulanic acid [13, 16]. Clavulanic acid and the 5*S* clavams have common biosynthetic origins and the pathway involved in their production can be roughly divided into two parts in *S. clavuligerus* (Fig 1). The "early" steps leading up to the intermediate clavaminic acid are shared during the production of both types of metabolites, with all intermediates possessing 5*S* configuration [17]. Beyond clavaminic acid (also a 5*S* clavam) the pathway diverges into specific "late" steps leading to either the 5*S* clavams or to clavulanic acid (Fig 1) [18].

The early shared portion of the pathway has been well characterized along with the genes involved in the process [19], but specific reactions involved in the production of each type of metabolite are yet to be elucidated [13]. It is currently hypothesized that during clavulanic acid production, the intermediate clavaminic acid undergoes oxidative deamination and ring inversion leading to clavaldehyde (Fig 1), which has 5*R* stereochemistry and is the immediate precursor of clavulanic acid [20]. The enzymes responsible for clavaldehyde formation are not known, but the products of *orf10-17* from the CA-BGC are thought to play a role in the process [13, 17]. Previous reports have shown that the disruption of individual genes from the *orf10-17* region abolishes or reduces clavulanic acid production without affecting 5*S* clavam levels [21–23]. Under certain conditions, the concomitant accumulation of acylated clavaminic acid derivatives was also observed in the *orf15-16* mutants [23, 24], suggesting that the respective metabolites are intermediates from the clavulanic acid arm of the biosynthetic pathway (Fig 1). Because of the clinical applications of clavulanic acid, there is considerable interest in understanding how the metabolite is produced in *S. clavuligerus*.

Of particular relevance to the current study is the product of *orf12* (SCLAV_4187) from the CA-BGC of *S. clavuligerus*, which resembles class A β-lactamases and also contains similar SXXK, SDN and KAG amino acid motifs [25]. *orf12* is co-transcribed with *orf13*, which encodes a putative membrane transport protein (Fig 2A), and their relative arrangement also suggests possible translational coupling [23]. Due to the bioactivities of specialized metabolites (especially when the product is an antibacterial), producer organisms often employ self-resistance strategies for protection [26, 27]. Intrinsic resistance in β-lactam-producing organisms is often attributed to the presence of altered penicillin-binding proteins (PBPs, the targets of β-lactam antibiotics) with reduced binding affinities for endogenously-produced antibiotics [28], but BGCs from such organisms also contain genes encoding β-lactamases and efflux transporters [29, 30]. Studies have shown that *orf12* is required for clavulanic acid, but not 5*S* clavam production [23] and that the encoded protein lacks any detectible β-lactamase activity.

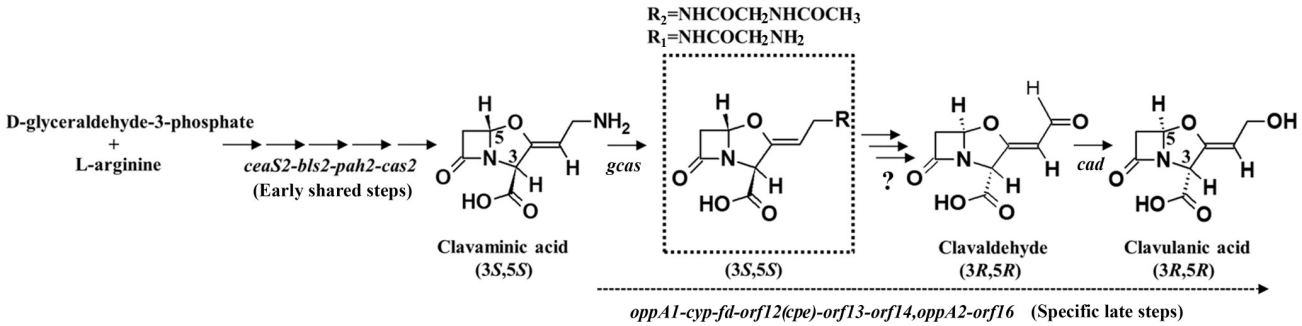


Fig 1. Diagrammatic representation of the partial clavulanic acid biosynthetic pathway from *Streptomyces clavuligerus*. Genes encoding enzymes known to be involved in the "early" shared stages of 5*S* clavam and clavulanic acid production, and those predicted to encode proteins involved exclusively in the biosynthesis of clavulanic acid ("late" steps) are indicated. In addition, genes encoding enzymes with known biosynthetic functions are shown next to arrows representing the respective reactions catalyzed by them, and the question mark indicates the unknown protein(s) responsible for the 5*S* to 5*R* epimerization and side chain modification of clavam intermediates during clavulanic acid biosynthesis. The two 5*S* clavam intermediates related to clavaminic acid (R_1 = *N*-glycyl and R_2 = *N*-acetyl-glycyl, respectively), which accumulate in the *orf15* and *orf16* gene mutants are shown in the dashed box.

https://doi.org/10.1371/journal.pone.0215960.g001

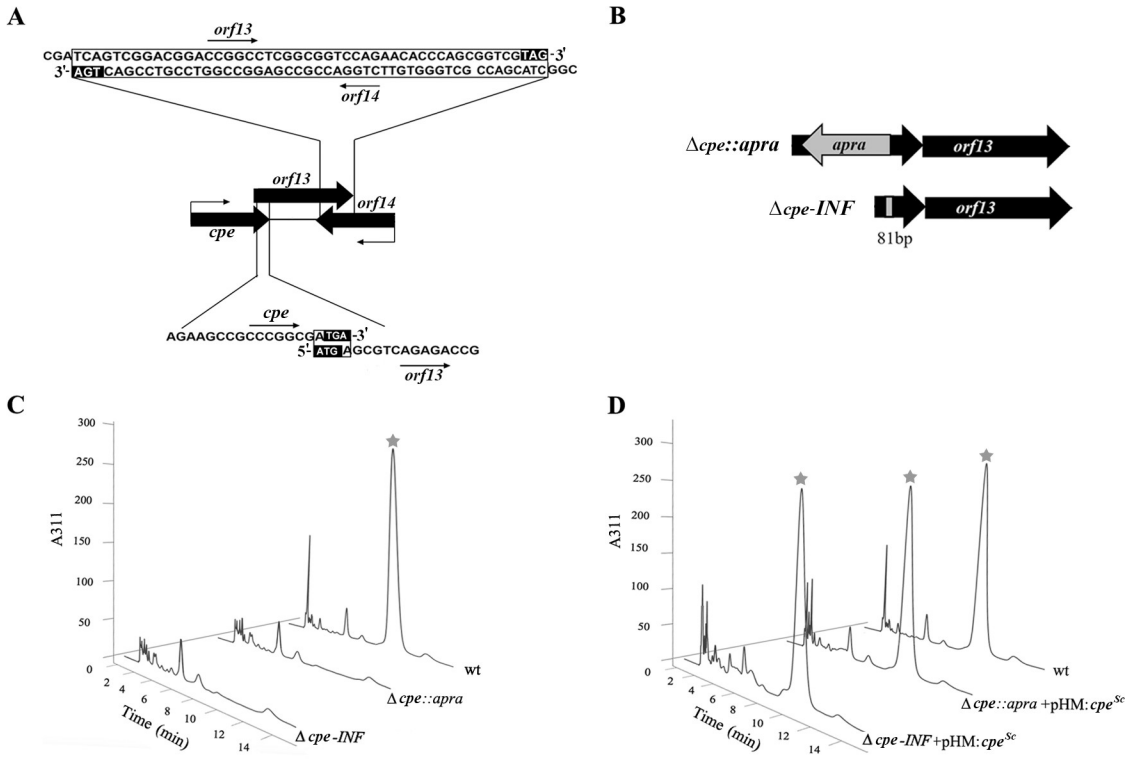


Fig 2. Preparation and analysis of *S. clavuligerus cpe* deletion mutants. (A and B) The thick arrows depict genes with arrowheads indicating the direction of transcription. (A) The relative arrangement of genes from the chromosomal locus surrounding *cpe* in *S. clavuligerus* is shown, and the bent arrows represent the known promoters for the different transcriptional units. The DNA sequences of the overlapping regions between *cpe-orf13* (bottom) and *orf13-orf14* (top) are indicated in open boxes, whereas the respective start and stop codons are shown in filled boxes. (B) Diagrammatic representation of the *Δcpe::apra* and *Δcpe-INF* mutants, which were prepared such that the 5' and 3' ends of *cpe* were retained and intervening DNA sequences were replaced by an apramycin resistance cassette or an in-frame 81-bp sequence in the respective mutants. (C and D) HPLC analysis of 96 hour wt *S. clavuligerus* and different *cpe* mutant SA culture supernatants for assessing clavulanic acid production using the phosphate buffer system [32]. Peaks corresponding to imidazole-derivatized clavulanic acid are indicated by the star symbol, which were detected at 311nm.

https://doi.org/10.1371/journal.pone.0215960.g002

Instead, heterologously expressed and purified Orf12 was shown to function as a cephalosporin esterase under *in vitro* conditions [25], due to which it is henceforth referred to as Cpe (for cephalosporin esterase).

The recently solved crystal structure of Cpe showed that in addition to a β-lactamase-like domain located in its C-terminus (residues 128–458), the protein also contains a previously unrecognized N-terminal domain (residues 1–127) resembling those found in steroid isomerases and polyketide cyclases [25]. In addition, two molecules of clavulanic acid were found to be bound non-covalently to Cpe when crystals of the protein were soaked in a solution of the metabolite during structural studies [25]. The first molecule (CA-1) was positioned in an active site pocket lined by residues (His88, Ser173, Thr209, Ser234, Ser278,Met383, Phe374, Ala376 and Phe385) from both the N- and the C-terminal domains, whereas the second molecule (CA-2) bound to a mostly hydrophobic cleft at the interface of the two domains *via* weak ionic interactions [25]. It was also shown that apart from Ser_{173}, Ser_{234} and Ser_{378}, other residues from Cpe or its N-terminal domain are not essential for its *in vitro* esterase activity. The ability of Cpe to bind clavulanic acid non-covalently under *in vitro* conditions is intriguing [25], as bona fide class A β-lactamases form irreversible covalent suicide adducts with the inhibitor [9]. In addition, β-lactamases are secreted out of the cell to inactive their antibiotic substrates [7], but the cellular location of Cpe in *S. clavuligerus* is not known. It is also not clear if Cpe undergoes post-translational processing in *S. clavuligerus*, or if both of its N- and C-terminal domains and associated amino acid residues are required for clavulanic acid production in the native host. Therefore, questions regarding the actual *in vivo* role of the Cpe gene product still remain unanswered, many of which are examined in the current study.

Materials and methods

Bacterial strains, plasmids and culture conditions

Dehydrated media components and reagents were purchased from VWR International, Fisher Scientific or Sigma-Aldrich (Canada). Details of bacterial strains and plasmids used in the current study are described in Tables 1 and 2, respectively. *Escherichia coli* and *S. clavuligerus* cultures were grown and manipulated as described previously [31, 32]. Other *Streptomyces* species were cultured using tryptic soy broth (TSB) or ISP4 media, whereas *S. viridis* was grown in nutrient broth (BD, Canada). Unless otherwise specified, all *E. coli*, *Streptomyces* and *S. viridis* cultures were grown at 37, 28 and 42˚C, respectively. Appropriate antibiotics were included in the media when required [31, 33], and liquid cultures were agitated at 200 rpm. For assessment of metabolite production, *S. clavuligerus* strains were grown in duplicate in starch asparagine (SA) or soy fermentation media as described previously [19]. All production phenotypes were verified using at least two independent fermentations.

DNA isolation, manipulation and analysis

All oligonucleotide primers used in the current study were purchased from Integrated DNA Technologies (USA) and are listed in S1, S2 and S3 Tables. Standard techniques were used to introduce, isolate, manipulate and analyze plasmid DNA from *E. coli* (35). Restriction enzymes used in the study were purchased from New England Biolabs Ltd. (Canada). Chromosomal DNA was isolated from *Streptomyces* and *S. viridis* cultures using the QIAamp DNA Mini Kit (QIAGEN, Canada) and a SpeedMill PLUS Bead Homogenizer (Analytik Jena, Germany), which was also used in all subsequent bead-beating purposes. PCR was performed using either the Fisher BioReagents *Taq* DNA polymerase or the Phusion High-Fidelity DNA Polymerase kits (Fisher Scientific, Canada) according to the manufacturer's recommendations, except that 5% DMSO was included in problematic reactions. DNA fragments were purified after

Table 1. Bacterial strains used in the current study.

Bacterial strain	Antibiotic resistance marker (s)[a]	Description	Source/ Reference[b]
***Escherichia coli* strains**			
E. coli NEB5α	NA	DH5α derived cloning host	NEB
E. coli BL21(DE3)	NA	Host for protein expression	NEB
E. coli ET12567(pUZ8002)	Cam^R, Kan^R	DNA methylation deficient conjugation host containing the plasmid pUZ8002	[33]
E. coli BW25118 (pIJ790)	Cam^R, Kan^R	Host containing the plasmid pIJ790 for λ RED mediated ReDirect PCR targeting of genes	[34]
E. coli DH5α (BT340)	Amp^R, Cam^R	Strain containing plasmid BT340 used for expressing the FLP recombinase	[35]
***Streptomyces* and other strains**			
Streptomyces clavuligerus NRRL 3585	NA	Wild type; cephamycin and clavulanic acid producer	NRRL
Streptomyces clavuligerus Δcpe::apra	Apr^R	*cpe* deletion mutant; gene replaced by disruption cassette from plasmid pIJ773	This study
Streptomyces clavuligerus Δcpe-INF	NA	*cpe* deletion mutant; gene replaced by 81bp marker less in-frame scar sequence	This study
Streptomyces flavogriseus ATCC 33331	NA	Wild type; clavulanic acid non-producer	ATCC
Saccharomonospora viridis ATCC 15386	NA	Wild type; clavulanic acid non-producer	ATCC
Streptomyces katsurahamanus	NA	Wild type; cephamycin and clavulanic acid producer	[12]
Streptomyces jumonjinensis	NA	Wild type; cephamycin and clavulanic acid producer	[12]
Klebsiella pneumoniae ATCC 15380	NA	Indicator organism for clavulanic acid bioassays	[32]

[a] Amp^R, ampicillin resistance; Apr^R, apramycin resistance; Cam^R, chloramphenicol resistance; Kan^R, kanamycin resistance; NA, Not applicable.

[b] ATCC, American Type Culture Collection; NEB, New England Biolabs; NRRL, Northern Regional Research Laboratory.

https://doi.org/10.1371/journal.pone.0215960.t001

standard TBE agarose gel electrophoresis using the EZ-10 Spin Column DNA Gel Extraction Kit according to the manufacturer's instructions (Bio Basic Canada Inc.). Unless otherwise specified, all PCR products were cloned into the pGEM-T Easy vector (Promega, USA) and the DNA sequences of all inserts were determined at the Centre for Applied Genomics, University of Toronto, Canada. Plasmid and cosmid constructs were introduced into *S. clavuligerus* through intergeneric conjugation using *E. coli* ET12567/pUZ8002 as described previously [19, 33].

Preparation of the *S. clavuligerus Δcpe::apra* and *Δcpe-INF* mutants

The pWE15 vector based cosmid clone 12B8 (Table 2) containing the entire clavulanic acid gene cluster was used to prepare the *S. clavuligerus Δcpe* mutants according to the previously described ReDirect PCR-Targeting method [19, 34]. Specific oligonucleotide primers (S1 Table) along with pIJ773 as template were used to amplify a PCR product containing the apramycin resistance cassette (*apra*) to target *cpe* in 12B8. This led to the replacement of an internal fragment of *cpe* by the *apra* disruption cassette to give the mutant cosmid 12B8-*Δcpe*::*apra*. In addition, the *apra* cassette comprising the *aac3(IV)* gene and RK2 *oriT* flanked by FLP recombinase target sites (FRT), was inserted in the direction opposite to *cpe* transcription in the mutant cosmid. 12B8-*Δcpe*::*apra* was then introduced into wt *S. clavuligerus* for double homologous recombination and isolation of the apramycin resistant, *Δcpe*::*apra* mutant.

In order to prepare the in-frame (*INF*) marker-less *Δcpe-INF* mutant, cosmid 12B8-*Δcpe*::*apra* from above was introduced in *E. coli* DH5α/BT340, which expresses the FLP recombinase

Table 2. Plasmids and cosmids used in the current study.

Plasmid/cosmid	Antibiotic resistance marker(s)[a]	Description	Source/ Reference
pGEMT-Easy	Amp^R	General *E. coli* cloning vector	Promega
pET30b	Kan^R	*E. coli* protein expression vector	Novagen
pHM11a	Hyg^R	Integrative *Streptomyces* expression vector containing the constitutive *ermE*p*	[36]
pSET152	Apr^R	Integrative *Streptomyces* cloning vector	[37]
pIJ773	Apr^R	Template plasmid for preparing the ReDirect *apra* disruption cassette	[34]
pIJ10700	Hyg^R	Template plasmid for preparing the ReDirect *hyg* disruption cassette	[34]
pET30b-*cpe*^{Sc}	Kan^R	Plasmid vector used to express C-terminal 6×His-tagged Cpe in *E. coli* for purification	This study
12B8	Amp^R, Kan^R	Cosmid clone containing the clavulanic acid biosynthetic gene cluster from *S. clavuligerus*	[19]
12B8-*Δcpe*::*apra*	Apr^R, Amp^R, Kan^R	Mutant cosmid 12B8 in which *cpe* has been replaced by the disruption cassette from plasmid pIJ773 using the ReDirect system	This study
12B8-*Δcpe-INF*	Amp^R, Kan^R	Mutant cosmid 12B8 in which *cpe* has been replaced by the 81-bp in-frame "scar" sequence using the ReDirect system	This study
12B8-*Δcpe-INF-Δamp*::*hyg*	Hyg^R, Kan^R	Cosmid 12B8-*Δcpe-INF* in which ampicillin resistance gene replaced by the *hyg* cassette from plasmid pIJ10700 using the ReDirect system	This study
pHM:*cpe*^{Sc}	Hyg^R	Expression plasmid pHM11a containing *cpe* from *S. clavuligerus*	This study
pHM:*cpe*^{Sf}	Hyg^R	Expression plasmid pHM11a containing *cpe* from *S. flavogriseus*	This study
pHM:*cpe*^{Sv}	Hyg^R	Expression plasmid pHM11a containing *cpe* from *S. viridis*	This study
pHM:*cpe*^{Sj}	Hyg^R	Expression plasmid pHM11a containing *cpe* from *S. jumonjinensis*	This study
pHM:*cpe*^{Sk}	Hyg^R	Expression plasmid pHM11a containing *cpe* from *S. katsurahamanus*	This study
pHM:*blip*^{FLAG}	Hyg^R	Expression plasmid pHM11a containing *blip* from *S. clavuligerus* with a C-terminal FLAG tag	This study
pHM:*ccaR*^{FLAG}	Hyg^R	Expression plasmid pHM11a containing *ccaR* from *S. clavuligerus* with a C-terminal FLAG tag	This study
pHM:*cpe*$^{Sc\text{-}FLAG}$	Hyg^R	Expression plasmid pHM11a containing *cpe* from *S. clavuligerus* with a C-terminal FLAG tag	This study
pHM:*cpe*$^{Sc\text{-}6\times his}$	Hyg^R	Expression plasmid pHM11a containing *cpe* from *S. clavuligerus* with a C-terminal 6×His tag	This study
pHM:*cpe*^{Ct}	Hyg^R	Expression plasmid pHM11a containing the C-terminal domain of *cpe* from *S. clavuligerus*	This study
pHM:*cpe*^{Nt}	Hyg^R	Expression plasmid pHM11a containing the N-terminal domain of *cpe* from *S. clavuligerus*	This study
pHM-*cpe*$^{Ct+Nt}$	Hyg^R	Expression plasmid pHM11a containing the N-terminal and C- terminal domains of *cpe* from *S. clavuligerus*, each expressed independently under the control of the *ermE*p*	This study
pSET:*cpe*^{Sc}	Apr^R	Plasmid pSET152 containing the *S. clavuligerus cpe* gene along with *ermE*p* from pHM11a was used as template to prepare all described *cpe*^{Sc} site directed mutants	This study

[a]Amp^R, ampicillin resistance; Apr^R, apramycin resistance; Kan^R, kanamycin resistance, Hyg^R, hygromycin resistance.

https://doi.org/10.1371/journal.pone.0215960.t002

[35]. FLP caused the excision of the FRT-flanked *apra* cassette in 12B8-*Δcpe*::*apra*, leaving an 81-bp in-frame DNA sequence ("scar") in its place in the mutant cosmid 12B8-*Δcpe-INF* (Table 2). Since *oriT* is part of the *apra* cassette, it was also lost, and 12B8-*Δcpe-INF* could not be transferred to *S. clavuligerus via* conjugation. Therefore, an *oriT* was introduced into 12B8-*Δcpe-INF* using a second round of ReDirect PCR-Targeting [34]. Specified primers (S1 Table) were used along with pIJ10700 as a template to amplify a PCR product containing the hygromycin resistance cassette (*hyg*) to target the ampicillin resistance gene present on the pWE15 vector backbone of 12B8-*Δcpe-INF*. The resulting cosmid 12B8-*Δcpe-INF-Δamp*::*hyg* (Table 2), containing the *hyg* cassette (which in turn contains an *oriT*) in place of the ampicillin resistance gene, was transferred to the *S. clavuligerus Δcpe*::*apra* mutant by conjugation with *E. coli*. Hygromycin-resistant colonies that arose were then made to undergo sporulation without any antibiotic selection to isolate the apramycin and hygromycin sensitive *S. clavuligerus Δcpe-INF* mutant. The replacement of the wt *cpe* gene with *Δcpe*::*apra* and *Δcpe-INF* in the respective *S. clavuligerus* mutants was confirmed by genomic DNA PCR and sequencing of products using specific primers (S1 Table).

Preparation of *cpe* complementation plasmids

Specific oligonucleotide primers (S1 Table) with engineered NdeI and HindIII/BamHI restriction sites were used to PCR amplify DNA fragments containing the *cpe* genes from *S. clavuligerus* (^{Sc}), *S. jumonjinensis* (^{Sj}), *S. katsurahamanus* (^{Sk}), *S. flavogriseus* (^{Sf}) and *S. viridis* (^{Sv}) for complementation studies. Since the sequences of *cpe* from *S. jumonjinensis* and *S. katsurahamanus* were not know, degenerate oligonucleotide primers with engineered restriction sites were designed based on known *cpe* DNA sequences from the three other species. After PCR amplification, the DNA fragments were directly cloned into the NdeI and HindIII/BamHI sites of the *Streptomyces* expression plasmid pHM11a [36] to give pHM:*cpe*^{Sj}, pHM:*cpe*^{Sk}, pHM:*cpe*^{Sf} and pHM:*cpe*^{Sv} (Table 2). The DNA sequences of all inserts were also verified/determined for comparison using custom primers (S1 Table).

To examine the *in vivo* roles of the N- and C-terminal domains of Cpe^{Sc}, custom oligonucleotide primers were used to amplify DNA fragments containing each domain separately (S1 Table). The respective PCR fragments were cloned into pHM11a at NdeI and BamHI after their sequences had been verified to give pHM:*cpe*^{Nt} and pHM:*cpe*^{Ct}, which functioned as the Cpe^{Sc} N- and C-terminal domain expression constructs, respectively (Table 2). To prepare a construct that could express the two domains separately at the same time from a single plasmid, the insert from pHM:*cpe*^{Ct} was released as a BglII-BamHI fragment and ligated to BamHI-digested pHM:*cpe*^{Nt}. This led to the plasmid pHM:*cpe*$^{Nt+Ct}$, in which the expression of each domain (not as part of the same protein) was driven independently by *ermE*p* (Table 2). Plasmid constructs were introduced into either the *S. clavuligerus Δcpe*::*apra* and/or *Δcpe-INF* mutants for complementation studies.

Detection and localization of Cpe^{Sc} in *S. clavuligerus*

Engineered oligonucleotide primers were used to add C-terminal FLAG tags onto Cpe, CcaR and Blip (S1 Table). PCR fragments containing the three respective genes (*cpe*$^{Sc\text{-}FLAG}$, *ccaR*^{FLAG} and *blip*^{FLAG}) were cloned into pHM11a and introduced into wt *S. clavuligerus* for localization studies (Table 2). One hundred milliliters *S. clavuligerus* SA cultures expressing each protein were separately grown for 48 hours, after which the cultures were subjected to centrifugation and the mycelial pellets were separated from the supernatants. Cell pellets were resuspended in 5 ml of lysis buffer (150 mM HEPES and 150 mM NaCl) and were sonicated on ice using a 5/64-inch probe (VWR International, Canada). The lysates were centrifuged at high-speed (27,000 × *g*) for 15 minutes to clarify the cytoplasmic fraction contained in the supernatants for subsequent use. Approximately 87 ml of culture supernatant (separated from the above mycelial pellet in the first step) was centrifuged at 27,000 × *g* for 15 minutes and was then filtered through 0.2 μm vacuum membranes (VWR International, Canada) to remove any residual particulate or insoluble material. To precipitate secreted proteins, 44.9 g of ammonium sulfate was added gradually to 500 ml of the filtered supernatant (final volume is made up by using lysis buffer) with constant stirring at 4˚C to give 80% saturation. Precipitated protein fractions were collected by high-speed centrifugation as described above, after which the supernatant was discarded, and the protein pellet was left to air dry for 10 minutes. The pellet was then resuspended in 500 μl of 1M phosphate buffer (sodium phosphate, pH-7.0) for future analysis.

C-terminal 6×His tagged protein (Cpe$^{Sc\text{-}6\times His}$) was also expressed in *S. clavuligerus* and *E. coli*. Engineered oligonucleotide primers were used to introduce a C-terminal 6×His tag during the amplification of *cpe*^{Sc} (S1 Table), which was cloned into pHM11a for expression in *S. clavuligerus*. For expressing Cpe$^{Sc\text{-}6\times His}$ in *E. coli*, the gene was PCR amplified using primers listed in S1 Table, was cloned into pET30b for expression at 15˚C for 24 hours. Cpe$^{Sc\text{-}6\times His}$

protein was purified using Ni-NTA resin as per the manufacturer's instructions (Qiagen, USA) and was stored in 20 mM Tris-HCl, 150 mM NaCl (pH 7.6) + 20% (v/v) glycerol.

For western analysis, 20–50 μg of cell-free extract or 0.5–1 μg of purified $Cpe^{Sc\text{-}6\times His}$ was subjected to standard 12% SDS-PAGE before being transferred to Immobilon-P PVDF membranes according to the manufacturer's recommendations (Millipore, Canada). Membranes were washed with TBS-T buffer (50 mM Tris-HCl pH 7.6, 150 mM NaCl, and 0.5% v/v Tween-20) and were blocked overnight at 4˚C in blocking buffer (TBS-T with 10% w/v non-fat milk). The membranes were probed using anti-FLAG or anti-6×His antibodies (Thermo Scientific Pierce, USA) at 1:500 final dilutions before being washed several times with TBS-T buffer. The secondary antibody (Thermo Scientific Pierce, USA) was added at 1:400 dilution in TBS-T buffer and the membranes were processed using the ECL Western Blot Substrate (Promega, USA) for imaging using a GE ImageQuant LAS 4000 Digital Imaging System (GE Healthcare, USA).

RNA isolation and RT-PCR analysis

S. clavuligerus wt and *Δcpe-INF* strains were used to isolate RNA after 48 hours of growth in SA medium using the innuSPEED Bacteria/Fungi RNA Kit and a bead beater as per the manufacturer's instructions (Analytik Jena, Germany). The cDNA was synthesized using 500 ng of DNaseI-treated RNA using random hexameric primers provided with the SuperScript II reverse transcriptase (RT) kit as per the manufacturer's recommendations (Invitrogen, USA). PCR was performed using 2.5μl of the RT product from above in a final volume of 20μl using the GoTaq DNA Polymerase (Promega, Canada). Thirty cycle PCR was performed to detect *ceaS2*, *oat2*, *oppA1*, *claR*, *car*, *cyp*, *cpe (orf12)*, *orf13*, *orf14*, *oppA2*, *orf16*, *gcas*, *pbpA*, and *hrdB* cDNA using gene-specific primers (S2 Table). Control reactions contained DNaseI-treated RNA preparations without reverse transcription for each reaction.

Site-directed mutagenesis of Cpe^{Sc}

The cpe^{Sc} gene along with the *ermE*p* from pHM11a was isolated as a BglII/BamHI fragment and inserted into the BamHI site of pSET152 (45) to prepare a smaller expression plasmid (pSET:cpe^{Sc}), which would be more amenable for site-directed mutagenesis (Table 2). The QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies, USA) along with mutagenic oligonucleotide primers (S3 Table) and pSET:cpe^{Sc} as template was to prepare selected single amino acid variants of Cpe^{Sc} according to the manufacturer's instructions. All introduced mutations were verified by DNA sequencing, and plasmids expressing Cpe^{Sc} variants (Table 2) were introduced into the *S. clavuligerus Δcpe-INF* mutant for complementation studies.

Metabolite detection and analysis

S. clavuligerus strains were grown for fermentation studies and culture supernatants were assessed for clavulanic acid production using bioassays as described previously [32]. High-performance liquid chromatography analysis of imidazole-derivatized culture supernatants was performed using a 1260 Infinity system (Agilent Technologies, USA) and a Bondclone C18 (100×8mm, 10μm, 148Å) column (Phenomenex, USA) [23]. Selected supernatants were also analyzed by liquid chromatography-mass spectrometry on an LC-MS-Trap system (1100 LC-MS Agilent Technologies, USA) as previously described [23, 38], with the exception that an Xterra (2.1×150 mm, 3.5μm, 125Å) column (Waters Scientific, USA) was used in the analysis.

Results

Preparation and complementation of the *S. clavuligerus Δcpe*::*apra* and *Δcpe-INF* deletion mutants

In *S. clavuligerus*, *cpe* (*orf12*) and *orf13* are transcribed together as a polycistronic mRNA and the stop codon of *cpe* also overlaps with the start codon of *orf13* [23]. In addition, there is a 48 bp overlap between the 3' ends of *orf13* and *orf14*, which are encoded on opposite DNA strands (Fig 2A). Therefore, there is potential for polar effects on the expression (transcription and/or translation) of *orf13* in a *cpe* gene mutant, depending on how it was prepared. To test this hypothesis, two different *cpe* mutants (Table 1) were prepared using the ReDirect two-step protocol [34]. In the first mutant, the apramycin (*apra*) cassette flanked by FLP recombinase target (FRT) sites from the plasmid pIJ773 was used to delete an internal region of *cpe* (39 bp from the 5' end to 39 bp from the 3' end), leading to the *S. clavuligerus Δcpe*::*apra* mutant (Fig 2B). The *apra* gene was inserted in the orientation opposite to *cpe* transcription to maximize the potential for polar effects on the expression of downstream genes. For preparing the second mutant, the *apra* cassette was excised from the *Δcpe*::*apra* mutant and replaced with an 81 bp scar sequence in the correct reading frame to give the *S. clavuligerus Δcpe-INF* (in frame deletion) mutant (Fig 2B), which has the least potential for producing polar effects on the expression of the downstream genes. The prepared mutants were verified by genomic DNA PCR and were complemented using the *cpe* gene from *S. clavuligerus* (cpe^{Sc}) expressed under the control of the constitutive $ermE^*$ promoter ($ermEp^*$) in the plasmid pHM11a (Table 2). Wild-type and *cpe* mutant strains of *S. clavuligerus* containing either pHM11a (control) or pHM:cpe^{Sc} were grown in SA medium for up to 120 hours to assess for clavulanic acid production. Bioassays and HPLC analysis of culture supernatants demonstrated that both the *S. clavuligerus Δcpe*::*apra* and *Δcpe-INF* mutants were completely blocked in clavulanic acid production when compared to the wt strain (Fig 2C). Introduction of pHM:cpe^{Sc} restored clavulanic acid production to 60%-70% of wt levels in both mutants (Fig 2D), suggesting that the *cpe* disruption(s) was not associated with any significant polarity. The marker-less *S. clavuligerus Δcpe-INF* mutant was chosen for further analysis in the current study.

Cellular localization of Cpe^{Sc} and its influence on the expression of other genes from the clavulanic acid biosynthetic gene cluster of *S. clavuligerus*

The Cpe^{Sc} protein shares many sequence and structural similarities with class A β-lactamases, but it has been shown to lack any detectable β-lactamase activity under *in vitro* conditions [25]. Most bona fide β-lactamases are secreted proteins that inactivate β-lactam antibiotics in the periplasm, the site of peptidoglycan biosynthesis and crosslinking [39]. However, the predicted amino acid sequence of Cpe^{Sc} does not contain any detectable secretion signals, warranting further investigation into its exact cellular location in *S. clavuligerus*. C-terminal FLAG ($Cpe^{Sc\text{-}FLAG}$) and 6×His ($Cpe^{Sc\text{-}6\times His}$) epitope-tagged copies of the protein were separately expressed in wt *S. clavuligerus* using the constitutive $ermEp^*$ from plasmid pHM11a (Table 2). As controls for protein localization studies, *S. clavuligerus* strains expressing C-terminal FLAG-tagged copies of known cytoplasmic ($CcaR^{FLAG}$) [40] and secreted ($Blip^{FLAG}$) [41] proteins were also prepared separately (Table 2). *S. clavuligerus* strains expressing FLAG-tagged copies of the respective proteins were grown in SA medium for 48 hours for isolating different cellular protein fractions. Mycelial pellets were used to obtain cytoplasmic and cell wall-associated fractions, whereas enriched secreted fractions were prepared by using salt to precipitate soluble proteins from culture supernatants. Western blot analysis of different cellular fractions using anti-FLAG polyclonal antibodies demonstrated that $Cpe^{Sc\text{-}FLAG}$ was only detected in the

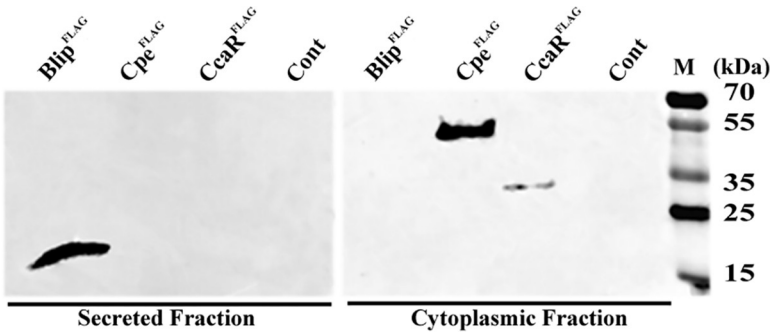


Fig 3. Cellular localization of Cpe in *S. clavuligerus*. C-terminal FLAG-tagged copies of Cpe, secreted Blip and cytoplasmic CcaR were expressed in wt *S. clavuligerus* separately for western blot analysis. Cultures were used for isolating secreted (left panel) and cell/cytoplasmic (right panel) fractions, which were probed using anti-FLAG antibodies. The analysis of protein fractions from *S. clavuligerus* strains containing plasmids pHM:*blip*^{FLAG} (expressing Blip^{FLAG}), pHM:*cpe*^{FLAG} (expressing Cpe^{FLAG}), pHM:*ccaR*^{FLAG} (expressing CcaR^{FLAG}) or pHM11a (Cont, empty vector) is shown. Lane M contains the PageRuler Plus Prestained Protein Ladder, which functioned as the molecular weight marker for resolving protein samples during 12% SDS-PAGE.

https://doi.org/10.1371/journal.pone.0215960.g003

cytoplasmic fraction (Fig 3). As expected, the CcarR^{FLAG} and Blip^{FLAG} controls were detected in cytoplasmic and secreted fractions, respectively (Fig 3). In addition, cultures of *S. clavuligerus* expressing Cpe$^{Sc\text{-}6\times His}$ were also used for isolating fractions for western analysis, which confirmed that Cpe^{Sc} is a cytoplasmic protein (S1 Fig). During the described western blot analysis, the size of epitope tagged Cpe^{Sc} was determined to be ~54 kDa based on the signal obtained using anti-FLAG and anti-6×His antibodies (S1 Fig). This corresponded to the size of 6×His-tagged Cpe^{Sc} heterologously expressed and purified from *E. coli*, which was used as a control (S1 Fig).

Clavulanic acid has been shown to bind non-covalently with Cpe^{Sc} under *in vitro* conditions [25], but the relevance of this interaction is still not clear as the protein did not catalyze any associated reaction. In addition, Cpe^{Sc} is located in the cytoplasm of *S. clavuligerus* (Fig 3), and *cpe* mutants are completely blocked in clavulanic acid production (Fig 2C). This raised the possibility that the protein could have a role in functioning as a cytoplasmic sensor/receptor for clavulanic acid or related metabolites to indirectly regulate production under *in vivo* conditions. To test this hypothesis, we analyzed the expression level of the first gene from each transcriptional unit (Fig 4A) from the clavulanic acid gene cluster of *S. clavuligerus* in the *Δcpe-INF* mutant and compared it with that from the wt strain (Fig 4B). RT-PCR analysis showed that only expression of the *cpe* gene was altered in the comparison, which was expected (Fig 4B). The analysis also demonstrated that the *Δcpe-INF* mutation is not associated with any transcriptional polarity as the expression of *orf13* was unaffected in the strain. Therefore, it is clear that the deletion of *cpe* does not in any way influence the expression of other genes from the clavulanic acid gene cluster in *S. clavuligerus*.

Assessing the requirement of the N- and C-terminal domains of Cpe^{Sc} for clavulanic acid production in *S. clavuligerus*

The crystal structure of heterologously expressed Cpe^{Sc} from *E. coli* demonstrated that it contains distinct N-terminal (residues 1–127) and C-terminal (residues 128–458) domains resembling ketosteroid isomerases/polyketide cyclases and β-lactamases, respectively [25]. It was also shown that the C-terminal domain was responsible for the observed *in vitro* cephalosporin esterase activity of Cpe^{Sc}, but a function or phenotype could not be assigned at the time for the

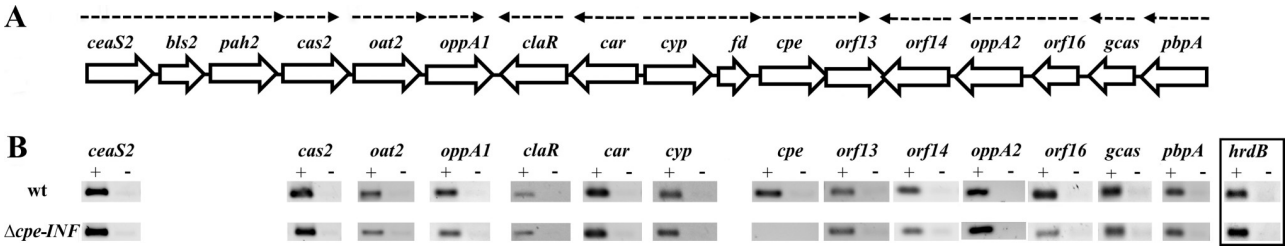


Fig 4. Transcriptional analysis of genes from the clavulanic acid biosynthetic gene cluster (BGC) in wt *S. clavuligerus* and the *Δcpe-INF* mutant. (A) The overall architecture of the BGC is shown with each hollow arrow representing a gene and the arrowhead its orientation. The known transcriptional units are also indicated, and the broken lines represent transcripts (B) The first gene from each transcriptional unit in (A) was selected for analysis to determine its comparative expression level in the two respective strains. RNA isolated from wt *S. clavuligerus* and the *Δcpe-INF* mutant after 48 hours of growth in SA medium was used for RT-PCR (+) analysis. As controls, treated RNA samples were used directly in PCR without RT or cDNA synthesis (-). The expression of the constitutively expressed *hrdB* gene (extreme right boxed panel) was used as internal control to normalize expression levels between different samples/strains.

https://doi.org/10.1371/journal.pone.0215960.g004

N-terminal domain based on the assays used [25]. Results from the western blot analysis described above indicate that Cpe^{Sc} does not undergo posttranslational processed in *S. clavuligerus*, but it is not known if the N-terminal domain is required for the activity of the protein in the native host. To investigate the *in vivo* roles of the N- and C-terminal domains during clavulanic acid biosynthesis, three additional Cpe^{Sc} expression constructs were prepared for analysis. The N- and C-terminal domains were expressed separately or together (as separate polypeptides, Table 2) in complementation studies using the *S. clavuligerus Δcpe-INF* mutant. Analysis of SA and soy culture supernatants showed that except for full-length Cpe^{Sc}, none of the other expression plasmids restored clavulanic acid production in the *Δcpe-INF* mutant, suggesting that both domains need to be part of a single polypeptide for biosynthesis to occur (Fig 5A and S2 Fig). Since the 5*S* clavams are not produced by wt *S. clavuligerus* when grown in SA medium [42], soy cultures were included in the analysis. Results showed that none of the strains accumulated any of the known intermediates from the clavulanic acid arm of the pathway (Fig 1 and S4 Table), and production of the 5*S* clavams was also unaffected in all of them when cultured in soy medium (S5 Table).

Examination of the ability of other *cpe* homologues to support clavulanic acid production in the *S. clavuligerus Δcpe-INF* mutant

The *S. flavogriseus* and *S. viridis* genome sequences revealed that they encode clavulanic-like BGCs [13], which are thought to be "silent or cryptic" as the two organisms are not known to produce any clavam metabolites [15]. Whereas other studies have shown that *S. jumonjinensis* and *S. katsurahamanus* can also produce clavulanic acid [12], but details regarding the sequences of their respective gene clusters are unavailable [14]. Therefore, we amplified the *cpe* homologues from the four organisms using genomic DNA as a template for complementation studies, and we also determined the complete sequence of the genes from *S. jumonjinensis* and *S. katsurahamanus* (S3 Fig). The predicted amino acid sequences of the Cpe proteins from *S. jumonjinensis* (Cpe^{Sj}) and *S. katsurahamanus* (Cpe^{Sk}) share ~68% identity with Cpe from *S. clavuligerus* (Cpe^{Sc}) (S4 Fig). In comparison, the predicted sequences of the proteins from *S. flavogriseus* (Cpe^{Sf}) and *S. viridis* (Cpe^{Sv}) showed 58.8% and 48.8% identity to Cpe^{Sc}, respectively. The predicted C-terminal domains of all four proteins contain the characteristic class A β-lactamase SXXK and SDN catalytic motifs, whereas the KTG motif was replaced by KGG in the non-producers (*S. flavogriseus* and *S. viridis*) and KAG in the producers (*S. clavuligerus*, *S. jumonjinensis* and *S. katsurahamanus*), respectively (S4 Fig). In addition, all proteins also contained an extra N-terminal domain resembling that of Cpe^{Sc} to different extents.

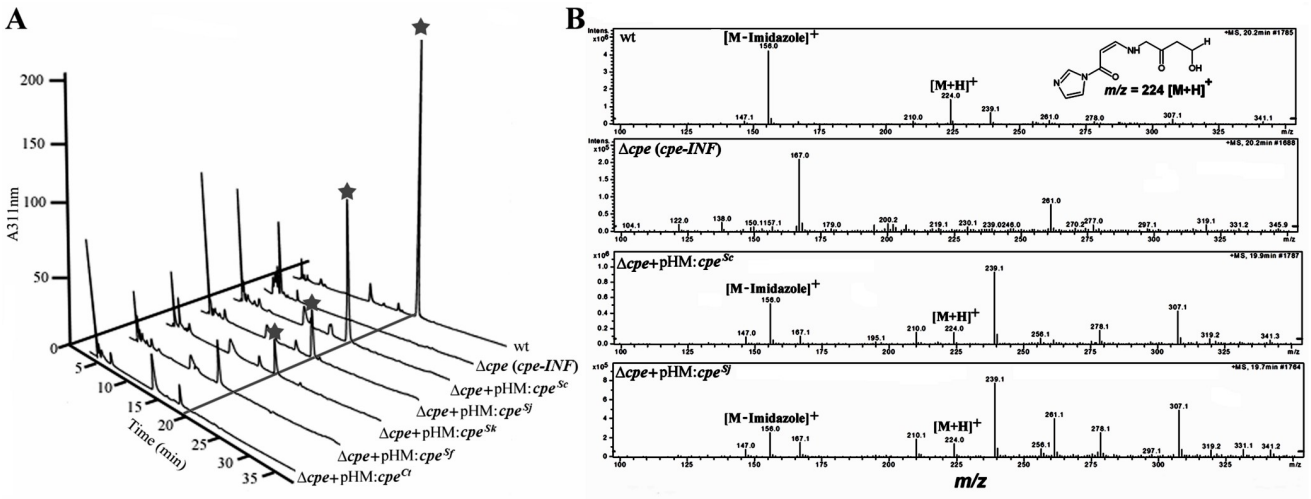


Fig 5. Functional analysis of different domains of Cpe^{Sc} and its homologues during clavulanic acid production in *S. clavuligerus*. (A and B) LC-MS analysis of 96 hour SA culture supernatants after imidazole derivatization using the ammonium bicarbonate buffer system [38]. Cultures of wt *S. clavuligerus* or the *Δcpe-INF* mutant expressing Cpe from *S. clavuligerus* (pHM:*cpe*^{Sc}), *S. flavogriseus* (pHM:*cpe*^{Sf}), *S. jumonjinensis*/*S. katsurahamanus* (pHM:*cpe*$^{Sj/Sk}$) or the C-terminal domain of Cpe^{Sc} (pHM:*cpe*^{Ct}) were used in the analysis. (A) Liquid chromatography profiles showing the elution of the peaks corresponding to imidazole-derivatized clavulanic acid (indicated by the star symbol). (B) Mass spectra of the major peaks corresponding imidazole-derivatized clavulanic acid $[M+H]^+$ (m/z = 224) and the fragmented product [M-imidazole]$^+$ (m/z = 156), which were only detected in supernatants from clavulanic acid producing strains shown in (A).

https://doi.org/10.1371/journal.pone.0215960.g005

Sequence analysis showed that the Cpe proteins from clavulanic acid producers are more closely related to each other as compared to those from the non-producers (S4 and S5 Figs). To determine the significance of this finding, the respective *cpe* genes from different sources were expressed under the control of *ermEp** in the *S. clavuligerus Δcpe-INF* mutant for complementation studies (Table 2). It was found that Cpe from *S. jumonjinensis* and *S. katsurahamanus* restored clavulanic acid production in the *S. clavuligerus Δcpe-INF* mutant to varying degrees (Fig 5A and 5B), whereas no complementation was observed in the case of Cpe^{Sv} and Cpe^{Sf}. Therefore, only the *cpe* genes from clavulanic acid producers (Cpe^{Sj} and Cpe^{Sk}) seem to be functionally equivalent to the known homologue from *S. clavuligerus* during clavulanic acid biosynthesis.

Identification of amino acid residues from Cpe^{Sc} required for clavulanic acid production in *S. clavuligerus*

The crystal structure of Cpe^{Sc} revealed that two molecules of clavulanic acid (CA-1 and CA-2, respectively) bind to the monomeric protein [25]. CA-1 binds to an active site pocket to form hydrogen bonds with Lys_{89}, Tyr_{359} and Arg_{418} and its C2 side chain carboxylate is positioned deep in the active site of Cpe^{Sc}, where it interacts with Lys_{375} [25]. Lys_{375} is part of the Cpe^{Sc} KTG motif, where the equivalent catalytic residues in PBPs/β-lactamases also interact with the analogous carboxylates from penicillin and cephalosporin substrates, respectively [43]. In comparison, binding of CA-2 occurs in a mostly hydrophobic cleft comprised of Trp_{91}, Leu_{362}, Leu_{415}, Arg_{418} and Ala_{422} at the interface of the N- and C-terminal domains [25]. The Trp_{91} and Arg_{418} residues are also highly conserved in the other predicted Cpe proteins (Cpe^{Sj}, Cpe^{Sk}, Cpe^{Sf}, Cpe^{Sv}), where Arg_{418} from Cpe^{Sc} is also involved in binding to CA-1 (S4 Fig). Therefore, Lys_{89}, Trp_{91} Tyr_{359}, Lys_{375} and Arg_{418} from Cpe^{Sc} were selected for mutagenesis studies to examine their *in vivo* contributions during the clavulanic acid production in *S. clavuligerus*. The *cpe*^{Sc} gene along with the *ermEp** was transferred from pHM:*cpe*^{Sc} to pSET152

and subjected to site-directed mutagenesis, and the prepared *cpe*^{Sc} variants were assessed for their ability to complement the *S. clavuligerus Δcpe-INF* mutant. Replacement of Lys_{89}, Tyr_{359}, Lys_{375} or Arg_{418} with Ala individually in Cpe^{Sc} led to a complete loss in clavulanic acid production (Table 3 and S6 Fig). However, when Lys_{375} was replaced with arginine (both being basic amino acids), clavulanic acid production was restored to 40% production levels in the *Δcpe-INF* mutant as compared to the wt strain (Table 3 and S6 Fig). As well, partial complementation was also observed in the case of the Cpe^{Sc} Trp_{91}Ala variant (Table 3).

Other amino acids from Cpe^{Sc} have also been shown to interact with clavulanic acid, some of which contributed to its *in vitro* cephalosporin esterase activity [25]. These include residues from the SXXK (Ser_{173}) and SDN (Ser_{234}) motifs comprising the catalytic tetrad (Ser_{173}/Lys_{176}/Ser_{234}/Lys_{375}), which is conserved in all four Cpe protein sequences described above (S4 Fig). Valegård, et al. reported that the Cpe^{Sc} Ser_{173}Ala mutant showed a 100-fold reduction in esterase activity, whereas the Ser_{234}Ala and Ser_{378}A mutants were not affected to the same extent [25]. Since the roles of the respective amino acids during clavulanic acid are not known, Ser_{173}, Lys_{176}, Ser_{234}, Ser_{378} were also individually substituted with Ala in Cpe^{Sc} for *in vivo* analysis.

Table 3. Clavulanic acid production in wild type (wt) *S. clavuligerus* and the *Δcpe-INF* mutant expressing different variants of Cpe^{Sc}.

S. clavuligerus strain[a]	Cpe^{Sc} protein variant[b] (*cpe*^{Sc} codon substitution)	Bioactivity[c]	
		SA	soy
wt	NA	++++	+++++
Δcpe-INF	NA	-	-
Δcpe-INF (pSET-152)	NA	-	-
Δcpe-INF (pSET:*cpe*^{Sc})	wt (none)	+++	++++
Δcpe-INF (pSET:*cpe*$^{Sc-\ Ser27Ala}$)	Ser_{27}Ala (TCC→GCC)	+++	++++
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Lys89Ala}$)	Lys_{89}Ala (AAG→GCG)	-	-
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Trp91Ala}$)	Trp_{91}Ala (TGG→GCG)	++	+++
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Arg115Ala}$)	Arg_{115}Ala (CGC→GCC)	+++	++++
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Ser173Ala}$)	Ser_{173}Ala (TCG→GCG)	-	-
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Lys176Ala}$)	Lys_{176}Ala (AAG→GCG)	-	-
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Ser206Ala}$)	Ser_{206}Ala (AGC→GCC)	+++	++++
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Ser234Ala}$)	Ser_{234}Ala (AGC→GCC)	-	-
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Arg311Ala}$)	Arg_{311}Ala (CGC→GCC)	+++	++++
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Gln321Ala}$)	Gln_{321}Ala (CAG→GCG)	+++	++++
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Trp326Ala}$)	Trp_{326}Ala (TGG→GCG)	+++	++++
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Arg346Ala}$)	Arg_{346}Ala (CGG→GCG)	+++	++++
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Tyr359Ala}$)	Tyr_{359}Ala (TAC→GCC)	-	-
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Lys375Ala}$)	Lys_{375}Ala (AAG→GCG)	-	-
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Lys375Arg}$)	Lys_{375}Arg (AAG→AGG)	++	+++
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Ser378Ala}$)	Ser_{378}Ala (TCC→GCC)	-	-
Δcpe-INF (pSET:*cpe*$^{Sc\text{-}Arg418Ala}$)	Arg_{418}Ala (CGC→GCC)	-	-

[a] Strains of *S. clavuligerus* were fermented in either SA or soy media for 96 hours and culture supernatants were used in bioassays for detecting clavulanic acid production.

[b] Single amino acid variants of Cpe^{Sc} used in the analysis are shown and the corresponding codon changes in *cpe*^{Sc} leading to the respective substitutions are indicated in parenthesis; NA, Not applicable.

[c] Zones of inhibition relative the wt strain grown in each media are indicated, where (+) indicates clavulanic acid production and (-) indicates the lack of production, respectively.

https://doi.org/10.1371/journal.pone.0215960.t003

All four variants were unable to complement the *Δcpe-INF* mutant, demonstrating that they essential for clavulanic acid production in *S. clavuligerus* (Table 3 and S6 Fig).

In the current analysis, amino acids were also identified that are either highly and/or partially conserved in all five Cpe proteins (S4 Fig). These include the Ser_{27}, Arg_{115}, Ser_{206}, Arg_{311}, Gln_{321}, Trp_{326} and Arg_{346} from Cpe^{Sc}, which are not part of any conserved motif and do not interact with clavulanic acid directly based on the reported crystal structure of the protein (29). When each of these residues was replaced with Ala, the respective Cpe^{Sc} variants restored clavulanic acid production in the *S. clavuligerus Δcpe-INF* mutant to varying degrees (Table 3), demonstrating that they are not essential for production. Overall, a detailed set of residues were identified in Cpe^{Sc}, some of which contribute to both the *in vitro* and *in vivo* activities of the protein, whereas others are only relevant during the latter process. These are important findings as they allude to the actual biochemical role/function of the protein, which occurs during *in vivo* clavulanic acid production in *S. clavuligerus*

Discussion

In the current study, we examined the function of *cpe* from the CA-BGC of *S. clavuligerus*, starting with the significance of the relative arrangement of neighboring genes located in its immediate vicinity. Polycistronic mRNAs often allow for the concerted expression of gene products involved in related biosynthetic pathways [44, 45], and gene knockout studies have implicated *cpe* as being essential for clavulanic acid production in *S. clavuligerus* [21, 23]. *cpe* is transcribed as part of a bicistronic operon along with *orf13*, the start codon of which also overlaps with the stop codon of *cpe* (Fig 2A), suggesting potential co-translation [46–48]. In addition, the 3′ ends of *orf13* and *orf14* overlap (Fig 2A), which is unusual in bacteria [49]. It can be challenging to decipher the precise roles of genes located within operons, particularly in cases where co-translation is involved [50–53]. The disruption of genes located in the 5′ regions of operons can influence the expression of downstream genes and also impact the relative stoichiometry of encoded gene products, thereby leading to polar effects [47, 51, 54]. Therefore, we prepared an in-frame *S. clavuligerus cpe* deletion mutant for use in the current study, while maintaining its stop codon and context with *orf13* to minimize the potential for polar effects. During the process, we also prepared the *S. clavuligerus Δcpe::apra* deletion mutant in which a disruption cassette was inserted in the opposite orientation to *cpe* transcription. It was noted that both the in-frame and the insertional mutant could be successfully complemented to restore clavulanic acid production using a plasmid-borne copy of *cpe*, demonstrating that polar effects were not associated with either of them. Therefore, it seems that despite their relative organization, alternate mechanisms exist to facilitate the translation of Orf13 in the *cpe* mutants, enabling us to use the in-frame mutant for more detailed *in vivo* studies.

The Cpe protein resembles class A β-lactamases [25], proteins which are secreted to inactivate β-lactam antibiotics before they can inhibit peptidoglycan crosslinking on the outer surface of the cytoplasmic membrane. In addition, when *cpe* was first sequenced it was reported to share some similarity with the LpqF lipoprotein from *Mycobacterium tuberculosis*, [25], the function of which is still unknown [55]. Most β-lactamases are secreted using the Sec pathway [39], however, some are translocated by the Tat system in mycobacteria [56]. In addition, certain proteins unrelated to β-lactamases have been reported to be secreted in the absence of any recognizable N-terminal signal sequences [57]. To narrow down its biological function, we investigated the cellular location of Cpe^{Sc} in its native host and showed that it is a cytoplasmic protein, with no evidence of any association with the cell wall or secreted fractions. Therefore, it can be inferred that unlike β-lactamases, the *in vivo* role of Cpe^{Sc} lies in the cytoplasm, which is also the site of clavulanic acid biosynthesis in *S. clavuligerus*.

During the biosynthesis of certain bioactive natural products, mechanisms exist to coordinate different stages involved in the production of the terminal metabolite [58]. The strategy is used to regulate the expression of specific genes, including those involved in export or self-resistance, when threshold concentrations of a specific intermediate(s) from the pathway is achieved in the cell [59, 60]. As well, feedback inhibition during natural product biosynthesis by end products is also well documented [61]. The cytoplasmic location of Cpe^{Sc} and its ability to bind to certain cephalosporins and clavulanic acid raises the possibility that it could function as an intracellular receptor for sensing such metabolites to elicit an associated response directly or indirectly. For example, the membrane-associated sensor kinase BlaR from *Staphylococcus aureus* also contains a PBP-like domain (related to β-lactamases) that binds to β-lactams and triggers the proteolysis of the cytoplasmic BlaI repressor to activate the expression of the BlaZ β-lactamase [62]. To investigate this hypothesis, the expression of key transcriptional units (comprising *ceaS2*, *oat2*, *oppA1*, *claR*, *car*, *cyp*, *cpe*, *orf13*, *orf14*, *orf16*, *gcas* and *pbpA*) from the CA-BGC was analyzed in the *S. clavuligerus Δcpe-INF* mutant for comparison with the wt strain. Except for the expression of *cpe* itself, the transcription of all other analyzed genes was unaffected in the mutant (Fig 4B). Results also clearly demonstrated that the transcription of *orf13* (and *orf14*) was not affected in the *Δcpe-INF* mutant, despite their complex transcriptional/transnational arrangement (Fig 2A). Therefore, based on the results, we can rule out any apparent sensory or regulatory role for Cpe^{Sc} during clavulanic acid biosynthesis in *S. clavuligerus*.

Homologues of cpe^{Sc} are also present in related BGCs from other clavulanic acid producing (*S. jumonjinensis* and *S. katsurahamanus*) and non-producing (*S. flavogriseus* and *S. viridis*) organisms (S4 Fig). It was found that only expression of Cpe from producer species (Cpe^{Sj}/Cpe^{Sk}) could complement the *S. clavuligerus Δcpe-INF* mutant. The probability of two proteins having a similar biological function increases proportionally with the relatedness of their respective amino acid sequences [63]. This might explain the complementation phenotypes as Cpe^{Sj}/Cpe^{Sk} are more closely related to Cpe^{Sc} than Cpe^{Sf}/Cpe^{Sv} (S4 and S5 Figs). The inability of the corresponding *S. flavogriseus* and *S. viridis cpe* homologues to complement the *S. clavuligerus* mutant suggests that portions of the respective clavulanic acid-like BGCs from the non-producers (including *cpe*) might also be defective in addition to being "silent" [15], a hypothesis that is currently being explored. Since the same plasmid(s) and constitutive promoter (*ermEp**) was used to drive the expression of all *cpe* genes independently in the current study, the lack of observed complementation in some cases is unlikely due to differences in expression levels.

A previous study showed that the C-terminal domain of recombinant Cpe^{Sc} displays *in vitro* O-acetyl cephalosporin esterase activity, but a function could not be assigned for its N-terminal domain [25]. It was also suggested that Cpe^{Sc} might undergo *in vivo* post-translational processing in *S. clavuligerus* to separate the two domains, which could not be addressed at the time since the protein was heterologous expressed and purified from *E. coli* [25]. Therefore, we examined the requirement of the two Cpe^{Sc} domains during *in vivo* clavulanic acid production in *S. clavuligerus* by using them to complement the *Δcpe-INF* mutant. Results suggest that both domains are required, and that they should be present to on a single peptide for production to take place (S2 Fig). It is also possible that the inclusion of the two domains on separate plasmid constructs could lead to reduced expression of the respective peptides or they could become unstable/misfolded, which might explain the lack of complementation. This seems unlikely, as the C-terminal domain of Cpe^{Sc} (completely lacking the N-terminus) was previously expressed and purified from *E. coli* for biochemical and structural studies [25]. Therefore the inability of the Cpe^{Sc} C-terminal domain to complement the *S. clavuligerus Δcpe-INF* mutant is most likely due to the missing region of the protein, which is the N-

terminus isomerase like domain. We also show that Cpe^{Sc} with specific amino acid substitutions (but not all) in either its N- or C-terminal domain is unable to complement the *Δcpe-INF* mutant and that only a single band corresponding to intact Cpe^{Sc} was observed in protein fractions from *S. clavuligerus* during western analysis (Fig 3 and S1 Fig). Therefore, results clearly demonstrate that Cpe^{Sc} does not undergo of *in vivo* proteolytic processing in *S. clavuligerus* and that its N-terminal isomerase-like domain is required for clavulanic acid production, which is the first direct evidence for its involvement in the process.

The crystal structure of Cpe^{Sc} showed that the protein binds to two molecules of clavulanic acid [25]. The first molecule (CA-1) forms hydrogen bonds with Tyr_{359}, Arg_{418} and Lys_{89} from the active site, which also contains Ser_{173} and Ser_{234} from the $S_{173}XXK_{176}$ and $S_{234}DN$ motifs, respectively. In addition, these 5 amino acids are conserved across all five Cpe homologues included in the current study (S4 Fig). Ser_{173}, Ser_{234} and Ser_{378} from Cpe^{Sc} are also important for the *in vitro* cephalosporin esterase activity of the protein [25], whereas in class A β-lactamases the equivalent residues (including Lys_{176}) are required for the binding and acylation of β-lactam substrates during catalysis [64–66]. It has been reported that certain esterases also exhibit β-lactamase activity and that some PBPs can conversely function as esterases [67, 68], which is not a true representation of their actual physiological function(s). This could also be the case for Cpe^{Sc}, where the protein can function as an esterase if given permissive substrates (25). In the current study, we demonstrated that Ser_{173}, Lys_{176}, Ser_{234} and Ser_{378} from Cpe^{Sc} are essential for *in vivo* clavulanic acid production in *S. clavuligerus*, reminiscent of their catalytic roles in class A β-lactamases. The K_{234} $(T/S)_{235}G$ catalytic motifs of serine β-lactamases also contain a conserved Ser/Thr residue, which interacts with the carboxylates of corresponding β-lactam substrates [69, 70]. Substitution of this Ser/Thr by amino acids with non-hydroxylated side chains (such as Ala) significantly reduces β-lactamase activity, especially against cephalosporin substrates [70, 71]. It is interesting to note that this conserved Ser/Thr residue is replaced by Gly (KGG) or Ala (KAG) in Cpe from clavulanic acid non-producers and producers, respectively (S4 Fig), which might explain why Cpe^{Sc} lacks any detectable *in vitro* β-lactamase activity. As well, substitution of Lys_{234} by Thr (but not Arg) in class A β-lactamases substantially reduces their catalytic activities and ability to bind to clavulanic acid for inhibition [72, 73]. We also show that the positively charged electrostatic feature at position 375 is essential for the *in vivo* functional activity of Cpe^{Sc}, as replacement of Lys ($K_{375}TG$) with Ala but not Arg in the protein completely abolished clavulanic acid production in *S. clavuligerus* (Table 3 and S6 Fig).

The *in vivo* role of amino acids from Cpe^{Sc} that interact with the second molecule of clavulanic acid (CA-2) *via* weak electrostatic interactions was also examined. These include the Trp_{91} or Arg_{418} residues [25], the substitution of which with Ala either reduced or abolish clavulanic acid production (Table 3). The blocked phenotype of the Arg_{418} mutant is consistent with the role of Arg_{418} as part of the CA-1 binding active site described above. In contrast, substitution of all other residues from the CA-2 binding cleft in Cpe^{Sc} did not significantly impact clavulanic acid production in *S. clavuligerus* (Table 3), suggesting that they do not contribute towards catalysis. Therefore, the role of the second clavulanic acid binding site in Cpe^{Sc} remains unclear. It is possible that the site occupied by CA-2 binds some other ligand and/or is an artifact of co-crystallizing purified Cpe^{Sc} with clavulanic acid during structural studies (86, 87), possibilities that warrant further examination. In addition, the substitution of other residues in Cpe^{Sc} that are conserved across all five homologues (S4 Fig), but which do not interact with clavulanic acid and/or are not part of any recognizable motif, did not affect the *in vivo* activity of the protein in *S. clavuligerus*.

To summarize, in the current study specific residues from Cpe^{Sc} and its two domains were shown to be essential for *in vivo* clavulanic acid production in *S. clavuligerus* (Table 3). These

are novel finding and allude towards a biosynthetic role for the protein during production. The described Cpe proteins share many similarities with class A serine β-lactamases, but some crucial differences are also apparent. For example, class A β-lactamases possess the characteristic Ω loop containing residue(s) involved in deacylation and subsequent release of hydrolyzed substrates [74], which are missing in Cpe^{Sc}. In comparison, class C and D serine β-lactamases also lack the Ω loop and are believed to use alternate mechanisms involving the SXXK and SDN motifs for deacylation instead [75]. The corresponding residues from Cpe^{Sc} (S_{173}XXK and S_{234}DN) bind to CA-1 in the crystal structure of the protein [25], but it is also possible that they might interact with some other intermediate(s) from the clavulanic acid biosynthetic pathway. Such precursors would not be detected in co-crystals reconstituted using heterologously expressed Cpe^{Sc} and purified clavulanic acid, as all other components of the biosynthetic pathway would be missing. Therefore, Ser_{173} and Ser_{234} could promote a nucleophilic attack on a still unknown substrate to form the primary Cpe-substrate complex, while other essential residues including some form the N-terminal domain (Lys_{89}, Lys_{375} and Arg_{418}) might be involved in stabilizing the intermediate followed by isomerization and product formation. The N-terminal domain of Cpe^{Sc} (residues 1–127) is structurally similar to a putative ketosteroid isomerase from *Shewanella frigidimarina*, which also contains the equivalent Lys_{89} from Cpe^{Sc} shown to interact with CA-1 [25]. Lys_{89} and some of its neighboring residues are conserved in all five Cpe homologues included in the current study (S4 Fig). For the first time, we show that the N-terminal domain of Cpe^{Sc} and specifically the Lys_{89} residue from it plays an essential role during clavulanic acid production in *S. clavuligerus* (S2 Fig and Table 3). We did not detect precursors or shunt products from the clavulanic acid pathway in culture supernatants from different *S. clavuligerus* mutants (S1 Table), suggesting that reaction intermediates remain tightly/covalently associated with Cpe^{Sc} during catalysis or are perhaps unstable [25]. Therefore, it is possible that Cpe^{Sc} is involved in an "altered/modified" β-lactamase-derived reaction required for clavulanic acid production by itself or in combination with another proteins(s), a hypothesis that is currently under investigation. It has been previously suggested that Cpe could be involved in the epimerization of 5*S* precursors to the 5*R* configuration during clavulanic acid biosynthesis [25], but this has not been demosntrated. Such a role for Cpe^{Sc} is conceivable based on the stereospecificity and reversible nature of enzyme-catalyzed reactions, but is not trivial to examine as the natural substrate(s) of Cpe are unknown [25]. It is intriguing that the biosynthetic pathway for a β-lactamase inhibitor (clavulanic acid) has recruited an enzyme for its production that is evolutionarily related to the very proteins that it inhibits. In the long term, deciphering the roles of different residues from Cpe^{Sc} involved in catalysis can enable us to engineer protein variants with the ability to accept altered substrates. Such a strategy would allow for the production of clavulanic acid analogues for future studies and possible applications.

Supporting information

S1 Fig. Western blot analysis of C-terminal 6×His tagged Cpe ($Cpe^{Sc\text{-}6\times His}$) expressed and purified from *E. coli* (using pET 30b-cpe^{Sc}) or expressed in *Streptomyces clavuligerus* (using pHM:$cpe^{Sc\text{-}6\times his}$). Purified protein (*E. coli*) or cell free lysates (*S. clavuligerus*) were used in the analysis along with anti-6×His antibodies for detecting epitope-tagged Cpe. The lane labeled as "Mock prep" contains *S. clavuligerus* pHM11a empty vector lysate as control to account for any non-specific antibody binding. The size of the band corresponding to $Cpe^{Sc\text{-}6\times His}$ in all lanes was approximately 50–55 kDa, and the prestained protein ladder (Marker) was used as a reference for estimating molecular weights during 12% SDS-PAGE.
(PDF)

S2 Fig. HPLC analysis of 96 hour SA culture supernatants from the *S. clavuligerus Δcpe-INF* strain expressing full-length Cpe^{Sc} (pHM:*cpe*^{Sc}), its N-terminus (pHM:*cpe*^{Nt}), C-terminus (pHM:*cpe*^{Ct}) or both the N- and C-terminal domains at the same time as separate peptides (pHM:*cpe*$^{Ct+Nt}$). The peak corresponding to imidazole-derivatized clavulanic acid (CA) is indicated and was only observed when the full-length protein was used in the analysis.
(PDF)

S3 Fig. DNA sequences of *cpe* homologues from *S. jumonjinensis and S. katsurahamanus*. The complete gene sequences starting from initiation (ATG) to the stop (TGA) codon for each gene were determined as part of the current study and are reported.
(PDF)

S4 Fig. Multiple sequence alignment of different Cpe proteins described in the current study. Analysis was performed with Clustal Omega (ver 1.2.1) using translated Cpe amino acid sequences from *S. clavuligerus* (WP_003952519.1), *S. flavogrisius* (WP_014152684.1), *S. viridis* (WP_015787620), *S. jumonjinensis* and *S. katsurahamanus*. The DNA sequences of *cpe* from the latter two producers (*S. jumonjinensis* and *S. katsurahamanus*) were determined as part of the current study and are reported in S3 Fig. The boxes in black highlight the conserved SXXK, SDN and KTG/KAG motifs present in class A β-lactamses and the respective Cpe proteins, whereas the box in blue represents the N-terminus domain indentifed in Cpe^{Sc}. The arrows indicate amino acids from Cpe^{Sc} that were selected for mutagenesis and the ones highlighted in red were shown to be essential for *in vivo* clavulanic acid production in *S. clavuligerus*.
(PDF)

S5 Fig. Phylogenetic relationship between select class A β-lactamases and Cpe proteins described in the current study. Multiple sequence alignments using the predicted amino acid sequences of Cpe proteins listed in S4 Fig. and class A β-lactamases including Bla (from the *S. clavuligerus* cephamycin C biosynthetic gene cluster, CAA90895.1) and TEM-1 (from *E.coli*, AMM70781.1) were used to prepare the tree, and bootstrap analyses were performed using 100 replicates. All positions containing gaps were eliminated during the analysis and the number next to each node represents the percentage of trees in which the respective topologies were observed.
(PDF)

S6 Fig. HPLC analysis of 96 hour *S. clavuligerus* SA culture supernatants from the wt strain for comparison with the *Δcpe-INF* mutant expressing CpeSc or select single amino acid variants (Ser173Ala, Lys176Ala, Lys375Ala, Lys375Arg Ser234Ala) of the protein in trans using plasmid pHM11a. Clavulanic acid (CA) production was monitored at 311nm following imidazole derivatization.
(PDF)

S1 Table. Oligonucleotide primers used for cloning and sequencing in the current study.
(PDF)

S2 Table. Gene specific primer pairs used for RT-PCR analysis in the current study.
(PDF)

S3 Table. Primer pairs used for site directed mutagenesis of *cpe*^{Sc}.
(PDF)

S4 Table. LC-MS analysis of SA culture supernatants from wt *S. clavuligerus* and *Δcpe-INF* mutant strains for detecting clavulanic acid and pathway intermediates.
(PDF)

S5 Table. LC-MS analysis of soy medium culture supernatants from wt *S. clavuligerus* and *Δcpe-INF* mutant strains for detecting 5*S* clavam production.
(PDF)

Acknowledgments

The described work was funded by an operating grant from the Natural Science and Engineering Research Council of Canada (NSERC: 386417–2010 and 2018–05949) to KT. KSK, NFA and CJM were also the recipients of NSERC graduate/undergraduate student awards. In addition, we acknowledge Memorial University of Newfoundland for providing graduate student support to KSK, NFA and BMP.

Author Contributions

Conceptualization: Kapil Tahlan.

Data curation: Santosh K. Srivastava, Kelcey S. King.

Formal analysis: Kapil Tahlan.

Funding acquisition: Kapil Tahlan.

Investigation: Santosh K. Srivastava, Kelcey S. King, Nader F. AbuSara, Chelsea J. Malayny, Jaime A. Wilson, Kapil Tahlan.

Methodology: Santosh K. Srivastava, Kelcey S. King, Nader F. AbuSara, Kapil Tahlan.

Project administration: Kapil Tahlan.

Resources: Kapil Tahlan.

Software: Brandon M. Piercey.

Supervision: Kapil Tahlan.

Visualization: Brandon M. Piercey.

Writing – original draft: Santosh K. Srivastava, Kelcey S. King, Kapil Tahlan.

Writing – review & editing: Kapil Tahlan.

References

1. Tahlan K, Jensen SE. Origins of the β-lactam rings in natural products. J Antibiot (Tokyo). 2013; 66 (7):401–10. https://doi.org/10.1038/ja.2013.24 PMID: 23531986
2. Tomasz A. The mechanism of the irreversible antimicrobial effects of penicillins: how the β-lactam antibiotics kill and lyse bacteria. Annu Rev Microbiol. 1979; 33(1):113–37. https://doi.org/10.1146/annurev.mi.33.100179.000553 PMID: 40528
3. Lewis K. Platforms for antibiotic discovery. Nat Rev Drug Discov. 2013; 12(5):371–87. https://doi.org/10.1038/nrd3975 PMID: 23629505
4. Wright GD. The antibiotic resistome: the nexus of chemical and genetic diversity. Nat Rev Microbiol. 2007; 5(3):175–86. https://doi.org/10.1038/nrmicro1614 PMID: 17277795
5. Essack SY. The development of β-lactam antibiotics in response to the evolution of β-lactamases. Pharm Res. 2001; 18(10):1391–9. PMID: 11697463
6. Livermore DM. β-lactamase-mediated resistance and opportunities for its control. J Antimicrob Chemother. 1998; 41 PMID: 9688449

7. Heesemann J. Mechanisms of resistance to β-lactam antibiotics. Infection. 1993; 21 PMID: 8314292.
8. Reading C, Cole M. Clavulanic acid: a β-lactamase-inhibiting β-lactam from *Streptomyces clavuligerus*. Antimicrob Agents Chemother. 1977; 11(5):852–7. PMID: 879738
9. Maiti SN, Phillips OA, Micetich RG, Livermore DM. β-lactamase inhibitors: agents to overcome bacterial resistance. Curr Med Chem. 1998; 5(6):441–56. PMID: 9873109
10. Matsuura M, Nakazawa H, Hashimoto T, Mitsuhashi S. Combined antibacterial activity of amoxicillin with clavulanic acid against ampicillin-resistant strains. Antimicrob Agents Chemother. 1980; 17 (6):908–11. PMID: 6967713
11. Todd PA, Benfield P. Amoxicillin/clavulanic acid. An update of its antibacterial activity, pharmacokinetic properties and therapeutic use. Drugs. 1990; 39(2):264–307. https://doi.org/10.2165/00003495-199039020-00008 PMID: 2184003
12. Jensen SE, Paradkar AS. Biosynthesis and molecular genetics of clavulanic acid. Antonie Van Leeuwenhoek. 1999; 75(1–2):125–33. PMID: 10422585
13. Jensen SE. Biosynthesis of clavam metabolites. J Ind Microbiol Biotechnol. 2012; 39(10):1407–19. https://doi.org/10.1007/s10295-012-1191-0 PMID: 22948564
14. Ward JM, Hodgson JE. The biosynthetic genes for clavulanic acid and cephamycin production occur as a 'super-cluster'in three *Streptomyces*. FEMS Microbiol Lett. 1993; 110(2):239–42. https://doi.org/10.1111/j.1574-6968.1993.tb06326.x PMID: 8349096
15. Álvarez-Álvarez R, Martínez-Burgo Y, Pérez-Redondo R, Braña A, Martín JF, Liras P. Expression of the endogenous and heterologous clavulanic acid cluster in *Streptomyces flavogriseus*: why a silent cluster is sleeping. Applied Microbiology and Biotechnology. 2013; 97(21):9451–63. https://doi.org/10.1007/s00253-013-5148-7 PMID: 23974366
16. Tahlan K, Anders C, Wong A, Mosher RH, Beatty PH, Brumlik MJ, et al. 5*S* clavam biosynthetic genes are located in both the clavam and paralog gene clusters in *Streptomyces clavuligerus*. Chemistry & Biology. 2007; 14(2):131–42. https://doi.org/10.1016/j.chembiol.2006.11.012 PMID: 17317567
17. Hamed RB, Gomez-Castellanos JR, Henry L, Ducho C, McDonough MA, Schofield CJ. The enzymes of β-lactam biosynthesis. Nat Prod Rep. 2013; 30(1):21–107. https://doi.org/10.1039/c2np20065a PMID: 23135477
18. Egan LA, Busby RW, Iwata-Reuyl D, Townsend CA. Probable role of clavaminic acid as the terminal intermediate in the common pathway to clavulanic acid and the antipodal clavam metabolites. 1997; 119(10):2348–55. https://doi.org/10.1021/ja963107o
19. Tahlan K, Park HU, Wong A, Beatty PH, Jensen SE. Two sets of paralogous genes encode the enzymes involved in the early stages of clavulanic acid and clavam metabolite biosynthesis in *Streptomyces clavuligerus*. Antimicrob Agents Chemother. 2004; 48(3):930–9. https://doi.org/10.1128/AAC.48.3.930-939.2004 PMID: 14982786
20. MacKenzie AK, Kershaw NJ, Hernandez H, Robinson CV, Schofield CJ, Andersson I. Clavulanic acid dehydrogenase: structural and biochemical analysis of the final step in the biosynthesis of the β-lactamase inhibitor clavulanic acid. Biochemistry. 2007; 46(6):1523–33. https://doi.org/10.1021/bi061978x PMID: 17279617
21. Li R, Khaleeli N, Townsend CA. Expansion of the clavulanic acid gene cluster: identification and in vivo functional analysis of three new genes required for biosynthesis of clavulanic acid by *Streptomyces clavuligerus*. J Bacteriol. 2000; 182(14):4087–95. PMID: 10869089
22. Mellado E, Lorenzana LM, Rodriguez-Saiz M, Diez B, Liras P, Barredo JL. The clavulanic acid biosynthetic cluster of *Streptomyces clavuligerus*: genetic organization of the region upstream of the car gene. Microbiology. 2002; 148(Pt 5):1427–38. https://doi.org/10.1099/00221287-148-5-1427 PMID: 11988517
23. Jensen SE, Paradkar AS, Mosher RH, Anders C, Beatty PH, Brumlik MJ, et al. Five additional genes are involved in clavulanic acid biosynthesis in *Streptomyces clavuligerus*. Antimicrob Agents Chemother. 2004; 48(1):192–202. https://doi.org/10.1128/AAC.48.1.192-202.2004 PMID: 14693539
24. Álvarez-Álvarez R, Rodríguez-García A, Martínez-Burgo Y, Martín J, Liras P. Transcriptional studies on *Streptomyces clavuligerus oppA2*-deleted mutant: *N*-acetylglycyl-clavaminic acid is an intermediate of clavulanic acid biosynthesis. 2018: Appl Environ Microbiol. 01701–18. https://doi.org/10.1128/AEM.01701-18 PMID: 30194098
25. Valegård K, Iqbal A, Kershaw NJ, Ivison D, Genereux C, Dubus A, et al. Structural and mechanistic studies of the *orf12* gene product from the clavulanic acid biosynthesis pathway. Acta Crystallogr D Biol Crystallogr. 2013; 69(8):1567–79. https://doi.org/10.1107/S0907444913011013 PMID: 23897479
26. Cundliffe E. How antibiotic-producing organisms avoid suicide. Annu Rev Microbiol. 1989; 43(1):207–33. https://doi.org/10.1146/annurev.mi.43.100189.001231 PMID: 2679354

27. Hopwood DA. How do antibiotic-producing bacteria ensure their self-resistance before antibiotic biosynthesis incapacitates them? Mol Microbiol. 2007; 63(4):937–40. https://doi.org/10.1111/j.1365-2958.2006.05584.x PMID: 17238916

28. Ogawara H. Self-resistance in *Streptomyces*, with special reference to β-Lactam antibiotics. Molecules. 2016; 21(5):605. https://doi.org/10.3390/molecules21050605 PMID: 27171072

29. Coque JJ, Liras P, Martin JF. Genes for a β-lactamase, a penicillin-binding protein and a transmembrane protein are clustered with the cephamycin biosynthetic genes in *Nocardia lactamdurans*. EMBO J. 1993; 12(2):631–9. PMID: 8440253

30. Perez-Llarena F, Martin JF, Galleni M, Coque JJ, Fuente JL, Frere JM, et al. The *bla* gene of the cephamycin cluster of *Streptomyces clavuligerus* encodes a class A β-lactamase of low enzymatic activity. J Bacteriol. 1997; 179(19):6035–40. PMID: 9324249

31. Green MR, Sambrook J. Molecular cloning: A laboratory manual: Cold Spring Harbor Laboratory Press; 2012. www.cshlpress.com

32. Mosher RH, Paradkar AS, Anders C, Barton B, Jensen SE. Genes specific for the biosynthesis of clavam metabolites antipodal to clavulanic acid are clustered with the gene for clavaminate synthase 1 in *Streptomyces clavuligerus*. Antimicrob Agents Chemother. 1999; 43(5):1215–24. PMID: 10223939

33. Kieser T. Practical streptomyces genetics: John Innes Foundation; 2000. www.jic.ac.uk

34. Gust B, Challis GL, Fowler K, Kieser T, Chater KF. PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proc Natl Acad Sci U S A. 2003; 100(4):1541–6. https://doi.org/10.1073/pnas.0337542100 PMID: 12563033

35. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A. 2000; 97(12):6640–5. https://doi.org/10.1073/pnas.120163297 PMID: 10829079

36. Motamedi H, Shafiee A, Cai SJ. Integrative vectors for heterologous gene expression in *Streptomyces* spp. Gene. 1995; 160(1):25–31. PMID: 7628712

37. Bierman M, Logan R, O'brien K, Seno E, Rao RN, Schoner B. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. Gene. 1992; 116(1):43–9. PMID: 1628843

38. Zelyas NJ, Cai H, Kwong T, Jensen SE. Alanylclavam biosynthetic genes are clustered together with one group of clavulanic acid biosynthetic genes in *Streptomyces clavuligerus*. J Bacteriol. 2008; 190(24):7957–65. https://doi.org/10.1128/JB.00698-08 PMID: 18931110

39. Pradel N, Delmas J, Wu LF, Santini CL, Bonnet R. Sec- and Tat-dependent translocation of beta-lactamases across the *Escherichia coli* inner membrane. Antimicrob Agents Chemother. 2009; 53(1):242–8. https://doi.org/10.1128/AAC.00642-08 PMID: 18981261

40. Santamarta I, Rodriguez-Garcia A, Perez-Redondo R, Martin JF, Liras P. CcaR is an autoregulatory protein that binds to the *ccaR* and *cefD-cmcI* promoters of the cephamycin C-clavulanic acid cluster in *Streptomyces clavuligerus*. J Bacteriol. 2002; 184(11):3106–13. https://doi.org/10.1128/JB.184.11.3106-3113.2002 PMID: 12003953.

41. Thai W, Paradkar AS, Jensen SE. Construction and analysis of ss-lactamase-inhibitory protein (BLIP) non-producer mutants of *Streptomyces clavuligerus*. Microbiology. 2001; 147. https://doi.org/10.1099/00221287-147-2-325 PMID: 11158349

42. Tahlan K, Anders C, Jensen SE. The paralogous pairs of genes involved in clavulanic acid and clavam metabolite biosynthesis are differently regulated in *Streptomyces clavuligerus*. J Bacteriol. 2004; 186(18):6286–97. https://doi.org/10.1128/JB.186.18.6286-6297.2004 PMID: 15342599

43. Ghuysen JM. Serine β-lactamases and penicillin-binding proteins. Annu Rev Microbiol. 1991; 45(1):37–67. https://doi.org/10.1146/annurev.mi.45.100191.000345 PMID: 1741619

44. Laing E, Mersinias V, Smith CP, Hubbard SJ. Analysis of gene expression in operons of *Streptomyces coelicolor*. Genome Biol. 2006; 7(6). https://doi.org/10.1186/gb-2006-7-6-r46 PMID: 16749941

45. Charaniya S, Mehra S, Lian W, Jayapal KP, Karypis G, Hu WS. Transcriptome dynamics-based operon prediction and verification in *Streptomyces coelicolor*. Nucleic Acids Res. 2007; 35(21):7222–36. https://doi.org/10.1093/nar/gkm501 PMID: 17959654

46. Das A, Yanofsky C. Restoration of a translational stop-start overlap reinstates translational coupling in a mutant *trpB'-trpA* gene pair of the *Escherichia coli* tryptophan operon. Nucleic Acids Res. 1989; 17(22):9333–40. PMID: 2685759

47. Khosla C, Ebert-Khosla S, Hopwood DA. Targeted gene replacements in a S*treptomyces* polyketide synthase gene cluster: role for the acyl carrier protein. Mol Microbiol. 1992; 6(21):3237–49. PMID: 1453961

48. Pradhan P, Li W, Kaur P. Translational coupling controls expression and function of the DrrAB drug efflux pump. J Mol Biol. 2009; 385(3):831–42. https://doi.org/10.1016/j.jmb.2008.11.027 PMID: 19063901

49. Fukuda Y, Nakayama Y, Tomita M. On dynamics of overlapping genes in bacterial genomes. Gene. 2003; 323:181–7. PMID: 14659892

50. Raynal A, Karray F, Tuphile K, Darbon-Rongere E, Pernodet JL. Excisable cassettes: new tools for functional analysis of *Streptomyces* genomes. Appl Environ Microbiol. 2006; 72(7):4839–44. https://doi.org/10.1128/AEM.00167-06 PMID: 16820478

51. Løvdok L, Bentele K, Vladimirov N, Müller A, Pop FS, Lebiedz D, et al. Role of translational coupling in robustness of bacterial chemotaxis pathway. PLoS Biol. 2009; 7(8):e1000171. https://doi.org/10.1371/journal.pbio.1000171 PMID: 19688030

52. Goh S, Hohmeier A, Stone TC, Offord V, Sarabia F, Garcia-Ruiz C, et al. Silencing of essential genes within a highly coordinated operon in *Escherichia coli*. Appl Environ Microbiol. 2015; 81(16):5650–9. https://doi.org/10.1128/AEM.01444-15 PMID: 26070674

53. Cheng C, Nair AD, Jaworski DC, Ganta RR. Mutations in *Ehrlichia chaffeensis* causing polar effects in gene expression and differential host specificities. PLoS One. https://doi.org/10.1371/journal.pone.0132657 PMID: 26186429

54. Zhang G, Tian Y, Hu K, Feng C, Tan H. SCO3900, co-transcripted with three downstream genes, is involved in the differentiation of *Streptomyces coelicolor*. Curr Microbiol. 2010; 60(4):268–73. https://doi.org/10.1007/s00284-009-9536-2 PMID: 20012957

55. Sutcliffe IC, Harrington DJ. Lipoproteins of *Mycobacterium tuberculosis*: an abundant and functionally diverse class of cell envelope components. FEMS Microbiol Rev. 2004; 28(5):645–59. https://doi.org/10.1016/j.femsre.2004.06.002 PMID: 15539077

56. McDonough JA, Hacker KE, Flores AR, Pavelka MS, Braunstein M. The twin-arginine translocation pathway of *Mycobacterium smegmatis* is functional and required for the export of mycobacterial β-lactamases. J Bacteriol. 2005; 187(22):7667–79. https://doi.org/10.1128/JB.187.22.7667-7679.2005 PMID: 16267291

57. Bendtsen JD, Kiemer L, Fausboll A, Brunak S. Non-classical protein secretion in bacteria. BMC Microbiol. 2005; 5(1):58. https://doi.org/10.1186/1471-2180-5-58 PMID: 16212653

58. Lim FY, Keller NP. Spatial and temporal control of fungal natural product synthesis. Nat Prod Rep. 2014; 31(10):1277–86. https://doi.org/10.1039/c4np00083h PMID: 25142354

59. Tahlan K, Ahn SK, Sing A, Bodnaruk TD, Willems AR, Davidson AR, et al. Initiation of actinorhodin export in *Streptomyces coelicolor*. Mol Microbiol. 2007; 63(4):951–61. PMID: 17338074

60. Tahlan K, Yu Z, Xu Y, Davidson AR, Nodwell JR. Ligand recognition by ActR, a TetR-like regulator of actinorhodin export. J Mol Biol. 2008; 383(4):753–61. https://doi.org/10.1016/j.jmb.2008.08.081 PMID: 18804114

61. Liu C, Hermann T, Miller PA. Feedback inhibition of the synthesis of an antibiotic: aurodox (X-5108). J Antibiot (Tokyo). 1977; 30(3):244–51. PMID: 405357

62. Clarke SR, Dyke KG. The signal transducer (BlaRI) and the repressor (BlaI) of the *Staphylococcus aureus* β-lactamase operon are inducible. Microbiology. 2001; 147(Pt 4):803–10. https://doi.org/10.1099/00221287-147-4-803 PMID: 11283276

63. Lee D, Redfern O, Orengo C. Predicting protein function from sequence and structure. Nat Rev Mol Cell Biol. 2007; 8(12):995–1005. https://doi.org/10.1038/nrm2281 PMID: 18037900

64. Matagne A, Frere JM. Contribution of mutant analysis to the understanding of enzyme catalysis: the case of class A β-lactamases. Biochim Biophys Acta. 1995; 1246(2):109–27. PMID: 7819278

65. Maveyraud L, Pratt RF, Samama JP. Crystal structure of an acylation transition-state analog of the TEM-1 β-lactamase. Mechanistic implications for class A β-lactamases. Biochemistry. 1998; 37(8):2622–8. https://doi.org/10.1021/bi972501b PMID: 9485412

66. Vandavasi VG, Weiss KL, Cooper JB, Erskine PT, Tomanicek SJ, Ostermann A, et al. Exploring the mechanism of β-lactam ring protonation in the class A β-lactamase acylation mechanism using neutron and X-ray crystallography. J Med Chem. 2015; 59(1):474–9. https://doi.org/10.1021/acs.jmedchem.5b01215 PMID: 26630115

67. Jones M, Page MI. An esterase with β-lactamase activity. J. Chem Soc Chem. Commun. 1991;(5):316–7. https://doi.org/10.1039/C39910000316

68. Ryu BH, Ngo TD, Yoo W, Lee S, Kim BY, Lee E, et al. Biochemical and structural analysis of a novel esterase from *Caulobacter crescentus* related to Penicillin-Binding Protein (PBP). Sci Rep. 2016; 6:37978. https://doi.org/10.1038/srep37978 PMID: 27905486

69. Dubus A, Wilkin JM, Raquet X, Normark S, Frere JM. Catalytic mechanism of active-site serine beta-lactamases: role of the conserved hydroxy group of the Lys-Thr(Ser)-Gly triad. Biochem J. 1994; 301(Pt 2)(2):485–94. PMID: 8042993

70. Imtiaz U, Manavathu EK, Lerner SA, Mobashery S. Critical hydrogen bonding by serine-235 for cephalosporinase activity of TEM-1 β-lactamase. Antimicrob Agents Chemother. 1993; 37(11):2438–42. PMID: 8285630

71. Lamotte-Brasseur J, Knox J, Kelly JA, Charlier P, Fonze E, Dideberg O, et al. The structures and catalytic mechanisms of active-site serine β-lactamases. Biotechnol Genet Eng Rev. 1994; 12(1):189–230 PMID: 7727028

72. Ellerby LM, Escobar WA, Fink AL, Mitchinson C, Wells JA. The role of lysine-234 in β-lactamase catalysis probed by site-directed mutagenesis. Biochemistry. 1990; 29(24):5797–806. PMID: 1974463

73. Lenfant F, Labia R, Masson JM. Replacement of lysine-234 affects transition state stabilization in the active site of β-lactamase TEM1. J Biol Chem. 1991; 266(26):17187–94. PMID: 1910040

74. Banerjee S, Pieper U, Kapadia G, Pannell LK, Herzberg O. Role of the omega-loop in the activity, substrate specificity, and structure of class A β-lactamase. Biochemistry. 1998; 37(10):3286–96. https://doi.org/10.1021/bi972127f PMID: 9521648

75. Medeiros AA. β-Lactamases: quality and resistance. Clin Microbiol Infect. 1997; 3. PMID: 11869237