

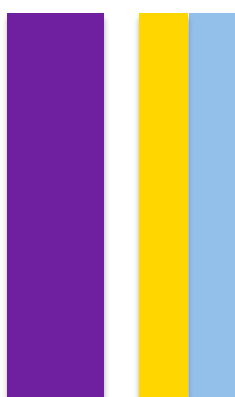
PROGRAMA DOUTORAL EM BIOTECNOLOGIA MOLECULAR E CELULAR APLICADA ÀS CIÊNCIAS DA
SAÚDE
BIOTECNOLOGIA MICROBIANA

Microbial factories based on *Corynebacterium glutamicum* for sustainable production of natural products

Filipe Castro Marques

D

2022



FILIPPE CASTRO MARQUES

**MICROBIAL FACTORIES BASED ON CORYNEBACTERIUM
GLUTAMICUM FOR SUSTAINABLE PRODUCTION OF NATURAL
PRODUCTS**

Tese de Candidatura ao grau de Doutor em
Biotecnologia Molecular e Celular Aplicada às
Ciências da Saúde;
Programa Doutoral da Universidade do Porto
(Instituto de Ciências Biomédicas Abel Salazar
e Faculdade de Farmácia)

Orientadora – Doutora Marta Vaz Mendes

Categoria – Professora Auxiliar Convidada

Afiliação – Instituto de Ciências Biomédicas
Abel Salazar

Co-orientador – Doutor Andriy Luzhetskyy

Categoria – Professor Catedrático

Afiliação – Universität des Saarlandes,
Alemanha

ACKNOWLEDGEMENTS

Acknowledgements are due to Prof. Dr. Paula Tamagnini, and Prof. Dr. Andriy Luzhetskyy, who accepted me as a member of their research groups at i3S – Instituto de Investigação e Inovação em Saúde, Portugal, and at Saarland University, Germany, respectively. My sincere gratitude goes to my doctoral supervisors, Prof. Dr. Marta Vaz Mendes, and Prof. Dr. Andriy Luzhetskyy, who push me forward and guided me through in research and beyond. A special word for Dr. Pedro Leão, who accepted to be my part of my PhD advisory committee, and to Prof. Dr. Maria Strecht Almeida, who encouraged me in exploring other perspectives of my work. Many thanks to the fellows of the BSM group at i3S, the AMEG at Saarland University, and the CNP group at CIIMAR for all the support and friendship. To all involved in this work – hosts, supervisors, co-authors and technicians, advisors, colleagues at the lab, partners outside the lab, friends, – thank you, and please read the next page.

This work was financially supported by National Funds through FCT – Fundação para a Ciência e a Tecnologia, I.P. under the project ERA-IB-2/0001/2015; fellowships SFRH/BD/131939/2016 and SFRH/BPD/95683/2013; and the FCT contract DL57/2016/CP1355/CT0023. The work was also supported by FEDER - Fundo Europeu de Desenvolvimento Regional through the COMPETE 2020 – Operacional Programme for Competitiveness and Internationalisation (POCI), Portugal 2020, and by Portuguese funds through FCT - Fundação para a Ciência e a Tecnologia/Ministério da Ciência, Tecnologia e Ensino Superior, and “Institute for Research and Innovation in Health Sciences” (POCI-01-0145-FEDER-007274), by the project NORTE-01-0145-FEDER-000012, Structured Programme on Bioengineering Therapies for Infectious Diseases and Tissue Regeneration, supported by Norte Portugal Regional Operational Programme (NORTE, 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF). This work was also supported by the DFG grant LU1524/4–1.

Host institutions



Main financial supporters



To all involved in this work

Who accepted my stay at their lab,

Just to discover a restless lad.

Who took my project hand in hand,

In Portugal and in Germany,

Guiding and supervising tirelessly.

Who helped me on the bench,

Controlling the machines,

Providing vitamins.

Who gave me advice,

Sweet licorice,

Showing new ways

In the bitter days.

Who were mates in research,

Sharing thoughts,

Concerns and laughs.

Who partnered outside:

Sing along! Go for a ride!

Hack that bug!

Green things up!

What a marvelous invention!

Give the students redemption!

Finally, who, dearest, patiently,

Fired up, inextinguishably,

My need for meaning,

Hushed my fears,

Shared joys and tears.

Thank you.

Aos meus pais

Que a despeito de finanças,

Erigiram biblioteca.

Vladimir: Let us not waste our time in idle discourse! (Pause. Vehemently.) **Let us do something** while we have the chance! It is not every day that we are needed. Not indeed that we personally are needed. Others would meet the case equally well, if not better. To all mankind they were addressed, those cries for help still ringing in our ears! But at this place, at this moment of time, **all mankind is us**, whether we like it or not. **Let us make the most of it**, before it is too late!

Samuel Beckett

O micróbio se não existisse era preciso inventá-lo.

Abel Salazar

LIST OF PUBLICATIONS

Ao abrigo do disposto no número 2 do Artigo 8.º do Decreto-Lei n.º 388/70

This dissertation is based on the publication **A)**, the manuscript in preparation **B)**, and some unpublished results. The chapter IV of this dissertation incorporates materials reported in publication **a)**. The results reported in chapter II of this dissertation supported the development of projects within the continuing training unit “Biolaboratory – Multidisciplinary Experimentation Project” (BioLab), at University of Porto. The pedagogical approaches of BioLab were described in publications **b)** and **c)**.

A) Marques, F., Luzhetskyy, A., & Mendes, M. V. (2020). Engineering *Corynebacterium glutamicum* with a comprehensive genomic library and phage-based vectors. *Metabolic Engineering*, 62, 221-234.

B) Marques, F., *et al.* (manuscript in preparation). Balancing growth and heterologous production of indigoidine in *Corynebacterium glutamicum* using a redox-sensing genetic switch.

Complementary publications:

a) Horbal, L., **Marques, F.**, Nadmid, S., Mendes, M. V., & Luzhetskyy, A. (2018). Secondary metabolites overproduction through transcriptional gene cluster refactoring. *Metabolic Engineering*, 49, 299-315.

b) Marques, F., Ramos, P. D., Lopes, M. M., Santos, J. B., Tavares, F., Olsson, I. A. S., & Almeida, M. S. (2022). BioLab – um espaço de experimentação e cruzamento disciplinar. In S. C. Soares, F. Remião, A. V. Martins & S. Nunes (Coords.), *Livro de Atas do 7.º Congresso Nacional de Práticas Pedagógicas no Ensino Superior* (pp. 295-301). UA Editora – Universidade de Aveiro.

c) Ramos, P. D., **Marques, F.**, Lopes, M. M., Santos, J. B., Tavares, F., Olsson, I. A. S., & Almeida, M. S. (2022). BioLab – uma proposta para o desenvolvimento de projetos em ambiente multidisciplinar. In João Veloso (Dir.), *Cadernos de Inovação Pedagógica, n.º 2 Unidades Curriculares INOVPED da Universidade do Porto. Transversalidade e criatividade* (pp. 125-139), U.Porto Press.

ABSTRACT

Natural products have been an important source of chemical scaffolds for the development of pharmaceuticals, especially antibiotics. There is a pressing need for novel drugs to tackle multidrug-resistant infections and, simultaneously, an urge for transitioning from a depletive fossil economy to an efficient bioeconomy. Microorganisms are a prolific source of bioactive natural products and a key tool for the bioproduction of chemicals from renewable feedstocks. The development of microbial factories with streamlined genomes and well-characterized molecular tools is a prerequisite to sustainably exploring new sources of natural products. The actinobacterium *Corynebacterium glutamicum* is acknowledged as a safe and genetically amenable microorganism, which is used to produce industrially amino acids, commodity chemicals, and, increasingly, bioactive products.

In this thesis, we present original research regarding the development of molecular tools and techniques to generate *C. glutamicum* strains suitable for the expression and sustainable production of active ingredients and their precursors.

Firstly, we developed a set of methodologies and molecular tools to metabolically engineer *C. glutamicum*. For that, we validated the usefulness of ϕ C31 and ϕ BT1 phage-based vectors for the delivery and expression of foreign biosynthetic genes in *C. glutamicum*. We validated a markerless genome editing method based on a Cre recombinase tool and a new comprehensive genomic library. In doing so, we generated the *C. glutamicum* BCA strain with a streamlined genome, devoid of secondary metabolites and capable of hosting the two different phage-based vectors. As proof-of-principle, we employed the novel *C. glutamicum* BCA strain as a heterologous host for the expression of a type III polyketide synthase and, for the first time, a non-ribosomal peptide synthetase. We observed a remarkable productivity of indigoidine, thereby confirming the metabolic capacity of the *C. glutamicum* BCA strain for the phosphopantetheinylation-dependent production of amino acid-derived natural products.

Secondly, having develop new tools for editing the genome of *C. glutamicum*, we envisioned a molecular device that would allow to balance the growth and production of small molecules. For that purpose, we selected the universal but not ubiquitous transcriptional regulator Rex, which is sensitive to the ratio of reduced and oxidized nicotinamide adenine dinucleotide coenzyme (NADH/NAD⁺). We generated a markerless *C. glutamicum* REX strain hosting the well-characterized *rex* gene of *Thermus aquaticus*, which was expressed at mRNA and protein levels. We determined that the REX strain displays a longer lag phase in an otherwise normal growth curve in defined media and a slight increase in susceptibility

to the oxidative stressor menadione. We implemented an indigoidine biosynthetic gene under the control of a Rex operator in *C. glutamicum* REX, resulting in lower target transcription and lower indigoidine production than in the parental BCA strain in defined medium. However, we showed that Rex regulation can be explored to increase the indigoidine content of producing cells when maltose is used as the carbon source in a complex medium with nitrogen limited conditions.

Thirdly, taking the developed strains and expression systems, we implemented biosynthetic pathways of complex natural products, so we can demonstrate the production of active ingredients and precursors thereof in *C. glutamicum*. We implemented a biosynthetic gene cluster for the production of the antibiotic bottromycin A2 into the streamlined *C. glutamicum* CR099attB_{C31} strain using a ϕ C31-based construct previously optimized for expression in *Streptomyces* spp. We obtained mass spectrometry data supporting the production of traces of bottromycin A2 in *C. glutamicum*. We implemented a synthetic operon for the production of an β -lactam intermediate of the antibiotic (5R)-carbapen-2-em-3-carboxylic acid into the *C. glutamicum* BCA strain using a ϕ BT1-based construct suitable for Rex-regulated expression. We ascertained the production of the target β -lactam by liquid chromatography and high-resolution mass spectrometry, attending to retention time, comparing to a commercial carbapenem, and the exact mass determination.

Finally, we examined our results in the context of the recent advances in genome editing, control of gene expression, and production of bioactive natural products in *C. glutamicum*. We discuss the relevance of microbial factories as pedagogical tools considering our participation in a course-based undergraduate research experience. We propose further scientific inquiries and technological demonstrations to advance *C. glutamicum* as a bioeconomy engine that runs on renewable resources to provide new chemical entities.

RESUMO

Os produtos naturais têm sido uma fonte relevante de estruturas químicas para o desenvolvimento de fármacos, especialmente antibióticos. Existe uma necessidade premente de novos medicamentos para debelar as infecções multirresistentes, enquanto se afigura urgente fazer a transição de uma economia baseada em combustíveis fósseis para uma bioeconomia eficiente. Os microorganismos são uma fonte abundante de produtos naturais bioactivos e são também considerados uma ferramenta essencial para a bioprodução de compostos químicos a partir de matérias-primas renováveis. O desenvolvimento de fábricas microbianas com genomas simplificados e ferramentas moleculares bem caracterizadas é um pré-requisito para o seu uso sustentado como novas fontes de produtos naturais. A actinobactéria *Corynebacterium glutamicum* é tida como um microorganismo de baixo risco biológico e passível de melhoramento genético, o qual tem vem sendo utilizado para a produção industrial de aminoácidos, compostos químicos de base e, cada vez mais, produtos bioactivos.

Nesta tese apresentamos investigação original sobre o desenvolvimento de ferramentas moleculares e técnicas para criar estirpes de *C. glutamicum* adequadas à expressão e à produção sustentável de princípios activos e seus precursores.

Em primeiro lugar, desenvolvemos um conjunto de metodologias e ferramentas moleculares para controlar o metabolismo de *C. glutamicum*. Para isso, validámos a utilidade de vectores baseados nos bacteriófagos ϕ C31 e ϕ BT1 na entrega e expressão de gene biossintéticos estranhos a *C. glutamicum*. Validámos também um método de edição genética para obter estirpes sem marcadores de selecção utilizando uma ferramenta baseada na recombinase Cre e numa nova biblioteca genómica de elevada abrangência. Durante esse processo, gerámos a estirpe *C. glutamicum* BCA com um genoma simplificado, livre de metabolitos secundários e com a capacidade de albergar os dois tipos de vectores baseados em bacteriófagos acima referidos. Como prova de conceito, tomámos a nova estirpe *C. glutamicum* BCA com organismo para expressão heteróloga de uma sintase de compostos policetónicos do tipo III e, de forma inédita, de uma sintetase de péptidos não-ribossomais. Observámos uma produtividade de indigoidina assinalável, confirmando desta forma a capacidade metabólica da estirpe *C. glutamicum* BCA para a produção dependente de fosfopanteteína de produtos naturais derivados de aminoácidos.

Em segundo lugar, tendo desenvolvido já novas ferramentas para edição do genoma de *C. glutamicum*, idealizámos um dispositivo molecular que permitisse controlar o crescimento e produção de moléculas de baixa massa molecular. Para isso, seleccionámos o regulador

de transcrição Rex, ubíquo, mas não universal, que detecta o rácio da coenzima nicotinamida adenina dinucleótido na forma oxidada e na forma reduzida (NADH/NAD⁺). Gerámos a estirpe *C. glutamicum* REX, sem marcador de selecção, albergando o gene *rex* bem caracterizado de *Thermus aquaticus*, o qual foi transcrito e traduzido *in vivo*. Determinámos que a estirpe REX apresenta uma fase de latência mais longa numa curva de crescimento ademais normal em meio de cultivo definido e apresenta um pequeno aumento de susceptibilidade ao agente oxidante menadiona. Implementámos o gene responsável pela biossíntese de indigoidina sob o controlo de um operador Rex em *C. glutamicum* REX, o que resultou numa menor transcrição do gene-alvo e uma menor produção de indigoidina do que na estirpe parental *C. glutamicum* BCA em meio de cultivo definido. Observámos que a regulação por Rex pode ser utilizada para aumentar o conteúdo em indigoidina das células produtoras quando a maltose é utilizada como fonte de carbono num meio de cultivo complexo com disponibilidade de azoto limitada.

Em terceiro lugar, valendo-nos das estirpes e sistemas de expressão desenvolvidos, implementámos vias biossintéticas de produtos naturais complexos, para assim demonstrar a produção de princípios activos e seus precursores em *C. glutamicum*. Implementámos o conjunto de genes biossintéticos necessários à produção de botromicina A2 na estirpe *C. glutamicum* CR099attB_{C31}, utilizando uma construção baseada num vector do tipo ϕ C31, previamente optimizada para expressão em *Streptomyces* spp. Deste modo, obtivemos dados de espectrometria de massa que sugerem a produção de quantidades vestigiais de botromicina A2 em *C. glutamicum*. Implementámos um operão sintético para a produção de um intermediário β -lactâmico do antibiótico ácido (5*R*)-carbapen-2-em-3-carboxílico na estirpe *C. glutamicum* BCA utilizando uma construção baseada num vector do tipo ϕ BT1 e regulável por Rex. Determinámos a produção do intermediário-alvo através de cromatografia líquida e espectrometria de massa de alta resolução, tendo em conta o tempo de retenção, em comparação com um carbapenemo comercial, e a determinação da massa exacta.

Para rematar, reexaminámos os nossos resultados no contexto dos progressos recentes nas áreas da edição do genoma, do controlo da expressão génica e da produção de produtos naturais bioactivos em *C. glutamicum*. Discutimos a relevância das fábricas microbianas como ferramentas pedagógicas tendo em consideração a nossa participação em experiências de investigação baseadas em cursos. Propusemos várias linhas de investigações científicas e demonstrações tecnológicas no sentido do desenvolvimento de *C. glutamicum* como força motriz da bioeconomia, que providencie novos compostos químicos a partir de matérias-primas renováveis.

TABLE OF CONTENTS

Acknowledgements	v
List of publications	xi
Abstract	xiii
Resumo	xv
Table of contents	xvii
List of abbreviations	xxi
CHAPTER I: General introduction	1
I.1 Natural products	3
I.2 Sustainable production	4
I.3 Microbial factories	5
I.4 <i>Corynebacterium glutamicum</i>	6
I.5 Aim of this work	8
CHAPTER II: Engineering <i>Corynebacterium glutamicum</i> with a comprehensive genomic library and phage-based vectors	9
Abstract	11
Introduction	11
Materials & Methods	12
Bacterial strains, plasmids, media and growth conditions	12
Recombinant DNA manipulations	13
Markerless integration of phage-based vectors at ϕ C31 bacterial attachment site	14
Generation of the cosmid-based genomic library	15
Markerless genome edition using modified cosmids	15
Integration of phage-based vectors at ϕ BT1 bacterial attachment site	15
Production and quantification of flaviolin and indigoidine	15
Nucleotide sequence accession number	16
Results	16

Characterization of a markerless integrative system based on actinophage ϕ C31 in <i>C. glutamicum</i> CR099attB _{C31}	16
Generation of an ordered cosmid library of <i>C. glutamicum</i> CR099attB _{C31}	17
Generation of a carotenoid-free <i>C. glutamicum</i> strain with an additional ϕ BT1 integration site	17
Production of the polyketide flaviolin in <i>C. glutamicum</i> using ϕ C31-based integration	19
Production of the non-ribosomal peptide indigoidine in <i>C. glutamicum</i> using ϕ BT1-based integration	20
Discussion	21
Declarations of interest	22
CRedit authorship contribution statement	23
Acknowledgements	23
Funding	23
References	23
SUPPLEMENTARY INFORMATION I	25
Material & Methods	27
Plasmid rescue analysis	27
Southern blot analysis	27
Episome isolation and detection	27
Analysis of unclonable genes	28
Design and application of a plasmid-based platform for editing of the <i>crt</i> locus	28
Comparative analysis of genomic sequences	28
Carotenoid extraction and UV-Vis quantification	29
HPLC-MS analysis of the supernatant of cultures	29
Figures	31
Data	36
Additional references	39
CHAPTER III: Balancing growth and heterologous production of indigoidine in <i>Corynebacterium glutamicum</i> using a redox-sensing genetic switch	41
Abstract	43
Introduction	44
Materials & Methods	46
Bacterial strains, plasmids, media and growth conditions	46
Bacterial manipulations	47

DNA manipulations	48
Construction of plasmids	49
RNA manipulations	50
Gene expression analysis	50
Protein crude extracts and SDS-PAGE analysis	51
Redox sensitivity bioassay	51
Indigoidine quantification	51
Bioinformatics analysis	52
Results	52
Implementation of a Rex-based regulatory network in <i>C. glutamicum</i>	52
Influence of the Rex regulatory system on the heterologous production of indigoidine in <i>C. glutamicum</i>	55
Effect of Rex regulation on the production of indigoidine in different carbon sources and carbon-to-nitrogen ratios	57
Quantitative determination of the effect of Rex regulation on the production of indigoidine in complex media with maltose	59
Discussion	61
Acknowledgements	66
Author contribution	66
Funding	66
References	67
SUPPLEMENTARY INFORMATION II	77
Materials & Methods	77
Determination of intracellular NADH/NAD ⁺ ratio	77
Figures	78
Additional references	84
CHAPTER IV: Producing pharmaceuticals and precursors thereof in <i>Corynebacterium glutamicum</i>	85
IV.1 Introduction	87
IV.2 Materials and Methods	92
IV.2.1 Bacterial strains, plasmids, media, and growth conditions	92
IV.2.2 Recombinant DNA manipulation	93
IV.2.3 Generation of transformant bacteria	94
IV.2.4 Production analysis by HPLC-MS	95
IV.3 Results	97
IV.3.1 Botromycin: generation of strains and production assessment.	97

IV.3.2 Carbapenam: generation of strains and production assessment.	99
IV.4 Discussion	102
IV.5 Conclusion and future perspectives	106
CHAPTER V: General discussion and perspectives	109
V.1 Prospects of a microbial factory of natural products based on <i>C. glutamicum</i>	111
V.2 Tools and methods to generate microbial factories based on <i>C. glutamicum</i> for production of natural products	113
V.3 Regulatory systems to control gene expression in microbial factories based on <i>C. glutamicum</i> for production of natural products	115
V.4 First steps towards the production of antimicrobials and precursors thereof in <i>C. glutamicum</i>	116
V.5 Microbial factories and their products as educational resources	116
V.6 Future perspectives	119
V.6 Conclusions	120
GENERAL REFERENCES	121
ANNEXES	139
COMPLEMENTARY PUBLICATION I: Secondary metabolites overproduction through transcriptional gene cluster refactoring	
COMPLEMENTARY PUBLICATION II: BioLab – um espaço de experimentação e cruzamento disciplinar	
COMPLEMENTARY PUBLICATION III: BioLab – uma proposta para o desenvolvimento de projetos em ambiente multidisciplinar	

LIST OF ABBREVIATIONS

- %GC – percentage of guanosine and cytidine nucleotides (in nucleic acids)
- 5' UTR – 5' untranslated region (of mRNA)
- ABC – adenosine triphosphate-binding cassette
- Am – apramycin
- Am^R – apramycin resistance marker
- ANOVA – analysis of variance
- Ap – ampicillin
- Ap^R – ampicillin resistance marker
- ATCC – American Type Culture Collection
- ATP – adenosine triphosphate
- attB* – bacterial attachment site
- attB_{BT1}* – bacterial attachment site for the phage ϕ BT1
- attB_{C31}* – bacterial attachment site for the phage ϕ C31
- attB_{C31(GG)}* – bacterial attachment site for the phage ϕ C31 with GG as core dinucleotides
- attL_{BT1}* – left scar of *attP/attB_{BT1}* recombination
- attL_{C31}* – left scar of *attP/attB_{C31}* recombination
- attP* – phage attachment site
- attP_{BT1}* – phage attachment site of phage ϕ BT1
- attP_{C31}* – phage attachment site of phage ϕ C31
- attP_{C31(GG)}* – phage attachment site of phage ϕ C31 with GG as core dinucleotides
- attR_{BT1}* – right scar of *attP/attB_{BT1}* recombination
- attR_{C31}* – right scar of *attP/attB_{C31}* recombination
- BAC – bacterial artificial chromosome
- BCA – BT1-C31-Albino [strain] or bicinchoninic acid [method]
- BGC – biosynthetic gene cluster
- BLASTP – protein basic local alignment search tool

bp – base pair

BpsA – non-ribosomal peptide synthetase blue pigment (indigoidine) synthetase (of *Streptomyces albus* (or *albidoflavus*) J1074)

BT1-Int – BT1 Integrase

C/N – carbon-to-nitrogen ratio

C31-Int – C31 Integrase

CAI – codon adaptation index

cDNA – (RNA) complementary deoxyribonucleic acid

CDW – cell dry weight

CFU – colony forming units

CGP – prophage (of *C. glutamicum*)

Cm – chloramphenicol

Cm^R – chloramphenicol resistance marker

CoA – coenzyme A

colE1 – plasmid origin of replication for *E. coli*

CRISPR – clustered regularly interspaced short palindromic repeats

crt – main cluster of biosynthetic genes of carotenoids in *C. glutamicum* (cgp_0717-0723)

CSPD – 3-(1-Chloro-3'-methoxyspiro[adamantane-4,4'-dioxetane]-3'-yl)phenyl] dihydrogen phosphate

CURE – course-based research experience

DDBJ – DNA Data Bank of Japan

ddH₂O – double distilled water

DIG – digoxigenin

DIY – do-it-yourself

DMSO – dimethyl sulfoxide

DNA – deoxyribonucleic acid

DSB – double-strand break

dsDNA – double stranded deoxyribonucleic acid

DSM – German Collection of Microorganisms (*Deutsche Sammlung von Mikroorganismen*)

EIC – extracted ion chromatogram

ESI – electrospray ionization

FAD – flavin adenine dinucleotide, fully oxidized form

FADH₂ – flavin adenine dinucleotide, fully reduced form

FTMS – Fourier transform mass spectrometry

FWHM – full width at half maximum

GMO – genetically modified organism

goi – gene of interest

GR – gene ruler

GRAS – generally regarded as safe

GTP – guanosine triphosphate

HILIC – hydrophilic interaction chromatography

HPLC – high performance liquid chromatography

HPLC-MS – high performance liquid chromatography coupled to mass spectrometry

Hyg – Hygromycin

Hyg^R – hygromycin resistance marker

IMES – iterative marker excision system

IS – insertion sequence

ISCg – insertion sequence (of *C. glutamicum*)

ISP4 – International *Streptomyces* Project medium 4

kbp – thousand base pairs

kDa – thousand Dalton

Km – kanamycin

Km^R – kanamycin resistance marker

LA – lysogeny agar

LB – lysogeny broth

LC – liquid chromatography

loxP – target site of Cre recombinase

m/z – mass-to-charge ratio

Mbp – million base pairs

MCS – multicloning site

mRNA – messenger RNA

MS – mass spectrometry

MS+ – mass spectrometry in positive mode

MS2 – tandem mass spectrometry

NA – not available

NAD+ – nicotinamide adenine dinucleotide, oxidized form

NADH – nicotinamide adenine dinucleotide, reduced form

NADPH – nicotinamide adenine dinucleotide phosphate, reduced form

NC – not covered (region)

NCBI – National Center for Biotechnology Information

NMR – nuclear magnetic resonance (spectroscopy)

NRP – non-ribosomal peptide

NRPS – non-ribosomal peptide synthetase

OD600 – optical density at 600 nm

OD800 – optical density at 800 nm

oriT – origin of transfer (of RP4/RK2)

P/O – phosphate/oxygen ratio

P69 – synthetic promoter P69

pA15A_{oriV} - plasmid origin of replication for *E. coli*

PCR – polymerase chain reaction

P_{JVR} – synthetic promoter with 4306 MU of activity

PK – polyketide

PKS – polyketide synthase

PMSF – phenylmethylsulfonyl fluoride

PPTase – 4'-phosphopantetheinyl transferase

pSG5_{oriV} – plasmid origin of replication for *C. glutamicum*

pUC ori – plasmid origin of replication for *E. coli*

QPS – qualified presumption of safety

r.p.m. – revolutions per minute

RBS – ribosome binding sequence

Rex – rex regulator (of *Thermus aquaticus*)

RiPP - Ribosomally-synthesized and post-translationally modified peptide

RNA – ribonucleic acid

RNAseq – ribonucleic acid sequencing

ROP – Rex operator

RppA – type III polyketide synthase 1,3,6,8-tetrahydroxynaphthalene synthase (of *Saccharopolyspora erythraea*)

RppA_{Cgl} – type III polyketide synthase 1,3,6,8-tetrahydroxynaphthalene synthase (of *Saccharopolyspora erythraea*) recoded for *C. glutamicum*

RppA_{CglStrept} – type III polyketide synthase 1,3,6,8-tetrahydroxynaphthalene synthase (of *Saccharopolyspora erythraea*) recoded for *C. glutamicum* and *Streptomyces* sp.

RR - *attB*_{C31(CC)}/*attP*_{C31(GG)} recombination scar

rt – retention time

RT-PCR – reverse transcription-polymerase chain reaction

RT-qPCR – reverse transcription-quantitative polymerase chain reaction

SacB – sucrose sensitivity marker

SD – Shine-Dalgarno sequence

SDS-PAGE – sodium dodecyl sulfate-polyacrilamide gel electrophoresis

sgRNA – single guided RNA

TCA – tricarboxylic acid cycle

THN - 1,3,6,8-tetrahydroxynaphthalene

TIC – total ion chromatogram

tRNA – transfer ribonucleic acid

TSA – tryptic soy agar

TSB – tryptic soy broth

TSS – transcriptional start site

UPLC – ultra high-performance liquid chromatography

UPLC-MS - ultra-high performance liquid chromatography coupled to mass spectrometry

UV-Vis – Ultraviolet and visible light

v/v – volume-to-volume ratio

w/v – weight-to-volume ratio

CHAPTER I

General introduction

General introduction

The title of this thesis points to the sustainable production of natural products using microbial factories based on *Corynebacterium glutamicum*. We shall give a general introduction on the definitions of natural products, sustainable production, and microbial factory, as well as introduce the adopted model organism *Corynebacterium glutamicum* before we can elaborate on the arguments put forward by this thesis.

I.1 Natural products

The definition we shall adopt for natural products relates to molecules produced by living organisms that serve essentially to improve survival fitness (Williams, Stone et al. 1989). In this sense, the concept of natural products coincides with the so-called secondary or specialized metabolites, which do not contribute to the internal economy of the cell but act on specific receptors of competing organisms to attain teleonomy (Monod 1970, Williams, Stone et al. 1989). We are not adopting the concept of natural product in the sense of being perishable (although some are unstable in ordinary conditions), being non-toxic (in fact, many are toxic at low concentrations), resulting exclusively from non-genetically altered organism or being somewhat connected to sacredness (Stanziani 2008). For the sake of simplicity, we shall extend our definition of natural products to include biosynthetic intermediates of the bioactive molecules, though excluding pseudonatural products partially composed by “unprecedented arrangements which are not available by current biosynthesis pathways” (Gally, Pahl et al. 2021).

The capability to interact with biological targets has attracted attention to natural products, leading to the commercial production of active pharmacological ingredients, such as the antibiotics penicillin and streptomycin, and the antiparasitic avermectins, to name only some Nobel-awarded discoveries. During its Golden Age (1940s-1970s), the screening for natural products provided abundant sources of bioactive, low molecular weight, structurally diverse and chemically complex molecules, mostly from bacterial sources (Katz and Baltz 2016). Despite the decrease of the number of novel classes discovered, natural products retain a key role as scaffold for drug development to this day: from 1981 to September 2019, natural products and their derivatives still represented 32% of all new small-molecule approved drugs, including 78 new antibacterial drugs (Newman and Cragg 2020).

The discovery of natural products generally encompasses obtaining sufficient biomass of the putative producer organisms, establishing assays to underpin interesting bioactivities, de-replicating bioactive extracts using the defined assays, physicochemical separations to determine the chemical composition and structure of new products, and validating their therapeutical efficiency (Mochales 1994). Despite the high-quality scaffolds it provided for

drug development, the empirical screening of natural products has faced a decrease in structural novelty of post-Golden Age screening programs and suffered from the advent of synthetic libraries as hit sources (Singh and Barrett 2006). While the discovery of natural products ceased to be a cost-effective process in drug development at traditional economic models, there was still an unmet need for new antimicrobials to tackle drug-resistant infectious diseases (Cook and Wright 2022).

The reinvigoration of the discovery of natural products happened in the beginning of the XXI century, as sequencing and bioinformatics analysis of microbial genomes helped to identify genes associated with known metabolites and predict biosynthetic pathways of cryptic (unobserved) products (Baltz 2019). In particular, the actinomycetes and other microorganisms with large genomes, which are known prolific producers of secondary metabolites, were shown to have up to 90% of their predicted clusters of biosynthetic genes silenced (not expressed) in conventional fermentation conditions (Baltz 2017). The development of molecular biology tools and methods for genetic manipulation has contributed for the screening of the biosynthetic potential of these organisms and implement a bio-based cost-effective production of natural products (Pawar, Ho et al. 2017).

I.2 Sustainable production

In the realm of chemicals, the Green Chemistry field established a set of synthesis rules to prevent the generation of hazardous wastes during manufacture, so that products and processes are profitable, “good for human health and the environment” (Anastas and Eghbali 2010). For instance, bio-based solvents have been studied as less toxic, economic and environmentally viable alternatives to petroleum-based solvents for the extraction of natural products (Rapinel, Claux et al. 2020). Within the framework of Green Chemistry, the biosynthesis of natural product has been considered an inspiration for the development of sustainable chemistry approaches (Richter, Vieira et al. 2021).

The biosynthesis of natural products fits into the concept of bioeconomy, which aims to use renewable biological resources for the efficient manufacturing of chemicals in a sustainable and economic way (Aguilar, Twardowski et al. 2019). Lignocellulose and other renewable resources represent potential substrates for biotransformation to produce added value chemicals (Zhang and Too 2019). The uncovered biosynthetic potential of sequenced organisms alongside the development of genetic manipulation techniques allows to envision the “construction of artificial cascades for the preparation of a range of molecules, which would be challenging to access using traditional synthetic chemical approaches” (Kuska and O'Reilly 2020).

In a time of reinvigorated discovery of natural products and antimicrobial resistance emergence, scientific consortia plead for additional funding to leverage programs of hit identification for drug development (Miethke, Pieroni et al. 2021). Fermentation can serve both the discovery of novel natural products and their sustainable production in large scale from sustainable feedstocks (Hoff, Plassmeier et al. 2021). The biosynthetic potential of microorganisms can be accessed by fermentation of genetically modified native “producers” (Wang, Zhao et al. 2019) or surrogate organisms carrying relevant foreign genes (Huo, Hug et al. 2019). The use of surrogate organisms, also known as heterologous hosts or chassis strains, has taken an important part into the systematic effort to detect and isolate new natural products (Miethke, Pieroni et al. 2021). Heterologous hosts are included in the concept of microbial factories, which will be introduced following section.

I.3 Microbial factories

The microbial production of chemicals, including natural products, was reenacted by the improvement in yields resulting from deeper understanding and the ability to manipulate metabolic pathways (Suntar, Cetinkaya et al. 2021). Microbial (cell) factories can be defined as microorganism which have been metabolically engineered to produce specific chemicals from renewable nutrient sources (Cho, Kim et al. 2022). The production of chemicals may be based on existing or new-to-nature biosynthetic pathways biosynthetic pathways, which require heterologous expression.

Microbial factories for heterologous expression of biosynthetic pathways of natural products include organisms such as *Escherichia coli*, cyanobacteria, *Saccharomyces cerevisiae*, myxobacteria and actinobacteria (Wenzel and Muller 2009, Myronovskyi and Luzhetskyy 2019, Rahmat and Kang 2020, Taton, Ecker et al. 2020, Yang, Prabowo et al. 2021). *E. coli* is a well-known organism with a vast array of molecular tools for metabolic engineering, being the most widely used microorganism in bioeconomy (Yang, Prabowo et al. 2021). Cyanobacteria can use sun light and atmospheric carbon dioxide as energy and carbon sources, emerging recently as hosts for nonnative natural products (Taton, Ecker et al. 2020, Ferreira, Pacheco et al. 2021). The eukaryote *Saccharomyces cerevisiae* is a model yeast frequently used in biotechnology, presenting efficient homologous recombination machinery (Rahmat and Kang 2020). The myxobacteria includes bacteria with unique morphology and large genomes, encoding polyketide synthases, non-ribosomal peptide synthetases and their hybrids (Wenzel and Muller 2009). Actinobacteria, especially the genus *Streptomyces*, *Saccharopolyspora* and *Salinispora*, are proficient producers and harbor significant genetic potential for secondary metabolites (Nett, Ikeda et al. 2009). Several *Streptomyces* spp. have been developed as hosts for heterologous expression of biosynthetic gene clusters (Myronovskyi and Luzhetskyy 2019).

The development of microbial factories for natural products is based on the premise that strains closely related to native producers (or the sources of silent genetic information) are more successful in producing novel chemistry (Miethke, Pieroni et al. 2021). Heterologous hosts based on native producers have been developed by applying molecular biology methods to remove their own biosynthetic gene clusters and introducing further optimizations to expression of foreign instructions (Miethke, Pieroni et al. 2021). Approaches based on systems biology and synthetic biology tools constitute the most advanced methodologies to develop such microbial cell factories (Chen, Banerjee et al. 2020). The use of synthetic biology strategies for the production of secondary metabolism are dependent on the development of microbial factories with not only the metabolic robustness of industrial strains but also streamlined with stable genomes that assure predictable behaviors of newly installed genetic information (Beites and Mendes 2015). This thesis will propose the development of an actinobacterial microbial factory based on the industrial amino acid producer *Corynebacterium glutamicum* and associated molecular tools for production of secondary metabolites.

1.4 *Corynebacterium glutamicum*

The name of the microorganism *Corynebacterium glutamicum* derives etymologically from the Greek name *coryne*, a club, and the latinized adjective *glutamicum*, pertaining to glutamic acid, hence describing a club-shaped bacterium that remarkably produces glutamic acid (Parte, Whitman et al. 2012). The organism was discovered in the 1950's in Japan during a screening of soil samples for overproducers of the said amino acid (Kinoshita, Nakayama et al. 2014). To date, no opposition has been deducted from claims that products derived from *C. glutamicum* are generally recognized as safe (GRAS) and have qualified presumption of safety (QPS) (FDA 2019, Hazards, Koutsoumanis et al. 2019). Besides the contributions to bioeconomy regarding the production of lysine and glutamic acid (6 million tons per year), lately, *C. glutamicum* has been explored as microbial factory of chemicals for cosmetics and healthcare applications, exceeding eighty different products as of 2021 (Becker and Wittmann 2021).

Foundational to targeted genome editing methods as well as systems and synthetic biology enabling extended bioproduction in *C. glutamicum* was the complete sequencing of its 3.3-million base pairs (Mbp) genome (Kalinowski, Bathe et al. 2003). The comprehensive mapping of the transcriptome of *C. glutamicum* provided additional insight into the transcriptional regulatory networks in this organism (Pfeifer-Sancar, Mentz et al. 2013). The available genomic and transcriptomic data has enabled the construction of validated genome-scale metabolic models of *C. glutamicum* to assist “sustainable enhancement of the microbe's productivity” (Feierabend, Renz et al. 2021).

The economic and environmental sustainability of using *C. glutamicum* as microbial factory has been demonstrated by the production of useful chemicals using agriculture and aquaculture side streams as nitrogen sources and non-glucose feedstocks, such as lignocellulosic acetate, as carbon sources (Kiefer, Merkel et al. 2021, Wendisch, Nampoothiri et al. 2022). Being a facultative anaerobic bacterium, *C. glutamicum* can also be cultivated at different oxygen levels, including in anaerobic conditions, thereby modulating its carbon flux to originate organic acids and other value-added compounds (Tsuge and Yamaguchi 2021).

The production of natural products in *C. glutamicum* is enabled by the overexpression or extension of endogenous metabolic pathways (e.g., carotenoids) and the heterologous expression of foreign metabolic pathways (e.g., plant polyphenols) (Kogure and Inui 2018). The portfolio of active ingredients produced by *C. glutamicum* includes non-proteinogenic amino acids, compatible solutes (extremolytes) and sun blockers agents (reviewed in (Wolf, Becker et al. 2021)). For projects involving heterologous gene expression, some streamlined *C. glutamicum* strains (Baumgart, Unthan et al. 2018), enhanced molecular tools and sensors (Bakkes, Ramp et al. 2020, Della Corte, van Beek et al. 2020), and bioinformatics resources are available for development of strategies based on synthetic biology to enhance production of natural products. In this thesis, we aim to develop these resources, enabling the discovery and production of novel classes of natural products in sustainable *C. glutamicum* based microbial factories.

I.5 Aim of this work

This thesis stems from research developed within the framework of the doctoral work plan associated with the fellowship SFRH/BD/131939/2016, funded by FCT – Fundação para a Ciência e a Tecnologia, I.P., entitled “Microbial factories based on *Corynebacterium glutamicum* for sustainable production of natural products”. It aimed to develop molecular tools and techniques to genetically manipulate *Corynebacterium glutamicum* to establish these bacteria as platforms for economically and environmentally sustainable production of active ingredients and their precursors.

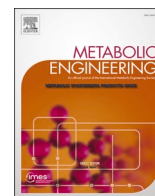
In that regard, we have defined three main objectives for this thesis:

- (a) to develop methodologies and molecular tools that enable metabolic engineering *Corynebacterium glutamicum* – Chapter II (Article I).
- (b) to balance growth and production of small molecules using non-native regulatory systems in *Corynebacterium glutamicum* – Chapter III (Manuscript in preparation).
- (c) to produce active ingredients and precursors thereof in *Corynebacterium glutamicum* – Chapter IV (Unpublished results).

CHAPTER II

Engineering *Corynebacterium glutamicum* with a comprehensive genomic library and phage-based vectors

Marques, F., Luzhetskyy, A., & Mendes, M. V. (2020). Engineering *Corynebacterium glutamicum* with a comprehensive genomic library and phage-based vectors. *Metabolic Engineering*, 62, 221-234.



Engineering *Corynebacterium glutamicum* with a comprehensive genomic library and phage-based vectors

Filipe Marques^{a,b,c,d}, Andriy Luzhetskyy^{d,e,**}, Marta V. Mendes^{a,b,c,*}

^a Bioengineering and Synthetic Microbiology Group, i3S- Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4200-135 Porto, Portugal

^b IBMC – Instituto de Biologia Molecular e Celular, Universidade do Porto, 4200-135 Porto, Portugal

^c ICBAS – Instituto de Ciências Biomédicas de Abel Salazar, Universidade do Porto, 4050-313 Porto, Portugal

^d Pharmazeutische Biotechnologie, Universität des Saarlandes, Campus C2.3, 66123 Saarbrücken, Germany

^e Helmholtz-Institut für Pharmazeutische Forschung Saarland, Campus E8.1, 66123 Saarbrücken, Germany

ARTICLE INFO

Keywords:

Actinobacteria
Genome editing
Natural products
Heterologous expression

ABSTRACT

The Gram-positive bacterium *Corynebacterium glutamicum* sustains the industrial production of chiral molecules such as L-amino acids. Through heterologous gene expression, *C. glutamicum* is becoming a sustainable source of small organic molecules and added-value chemicals. The current methods to implement heterologous genes in *C. glutamicum* rely on replicative vectors requiring lasting selection or chromosomal integration using homologous recombination. Here, we present a set of dedicated and transversal tools for genome editing and gene delivery into *C. glutamicum*. We generated a cosmid-based library suitable for efficient double allelic exchange, covering more than 94% of the chromosome with an average 5.1x coverage. We employed the library and an iterative marker excision system to generate the carotenoid-free *C. glutamicum* BT1-C31-Albino (BCA) host, featuring the attachment sites for actinophages ϕ C31 and ϕ BT1 for one-step chromosomal integration. As a proof-of-principle, we employed a ϕ C31-based integration and a Cre system for the markerless expression of the type III polyketide synthase RppA, and a ϕ BT1-based integration system for the expression of the phosphopantetheinylation-dependent non-ribosomal peptide synthetase BpsA in the *C. glutamicum* BCA host. The developed genomic library and microbial host, and the characterized molecular tools will contribute to the study of the physiology and the rise of *C. glutamicum* as a leading host for drug discovery.

1. Introduction

Corynebacterium glutamicum is a generally regarded as safe (GRAS) microorganism that provides a considerable part of the worldwide supply of bulk chemicals namely, amino acids (Kinoshita et al., 1957). Given its favorable physiological properties [reviewed in (Vertes et al., 2013)], *C. glutamicum* has soon become a multifaceted biorefinery for industrial biotechnology, e.g. on the production of organic acids, and pharmaceutical biotechnology, producing recombinant proteins and secondary metabolites (Becker et al., 2018a). The molecular basis of the industrial success of *C. glutamicum* has been unveiled by the analysis of the metabolic flux (Marx et al., 1996) and by the projects of genome and transcriptome sequencing (Bathe et al., 1996; Ikeda and Nakagawa, 2003; Kalinowski et al., 2003; Pfeifer-Sancar et al., 2013; Tauch et al.,

2002). Taken together, this knowledge paved the way for the development of genome-scale metabolic maps of *C. glutamicum* to guide systems-based engineering approaches (Kjeldsen and Nielsen, 2009; Shinfuku et al., 2009; Zhang et al., 2017).

The concept of a chassis for industrial biotechnology is centered on genome-minimized strains, whose physiology can be controlled, including the carbon flux, and for which there are well-characterized molecular tools (Vickers et al., 2010). The sequenced strain ATCC 13032 served as the starting point for the development of a chassis based on *C. glutamicum*. First, the removal of prophages (strain MB001) or insertion sequence (IS) elements (strains WJ004/WJ008) led to improvements in the production of recombinant proteins (Baumgart et al., 2013; Choi et al., 2015). Next, 26 clusters of the prophage- and IS element-free strain CR099 were found to be individually irrelevant for

* Corresponding author. Helmholtz-Institut für Pharmazeutische Forschung Saarland, Campus E8.1, 66123 Saarbrücken, Germany.

** Corresponding author. Bioengineering and Synthetic Microbiology Group, i3S- Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4200, Porto, MVM, Portugal.

E-mail addresses: a.luzhetskyy@mx.uni-saarland.de (A. Luzhetskyy), mvm@ibmc.up.pt (M.V. Mendes).

<https://doi.org/10.1016/j.ymben.2020.08.007>

Received 16 March 2020; Received in revised form 17 July 2020; Accepted 10 August 2020

Available online 20 August 2020

1096-7176/© 2020 International Metabolic Engineering Society. Published by Elsevier Inc. All rights reserved.

growth rate and growth yield in minimal media with glucose as carbon source (Unthan et al., 2015). Combinatorial knockout of those clusters from the strain CR099 led to the strain C1*, which presents a genome reduction of 13.2%, uncompromised growth rate, and stable genetic traits (Baumgart et al., 2018). All the genetic engineering steps during chassis development were performed by standard double homologous recombination using non-replicative vectors (Schafer et al., 1994) in an essentially error-free manner. The authors did not disclosed the efficiency of the process, but they stated that it took six years from evaluating the essentiality of genes until strain C1* was published (Baumgart et al., 2018).

The available methods for genome engineering in *C. glutamicum* take advantage of the native homologous recombination machinery, with or without help from heterologous recombineering and/or endonuclease systems. Traditionally, genome editing in *C. glutamicum* has been performed using non-replicative vectors as editing templates, and double crossovers have been isolated owing to selection and counter-selection markers (Kim et al., 2011; Ma et al., 2015; Schafer et al., 1994). These techniques rely exclusively on native homologous recombination machinery and are considered precise and virtually error-free. However, in the counter-selection step, cells can revert to the original genotype by excising the suicide vector due to lack of selection for mutants. Recently, strategies using endonucleases have gained momentum in *C. glutamicum* [reviewed in (Becker et al., 2018b)]. On the one hand, homologous recombination templates were coupled to catalytically active endonucleases Cas9 and Cpf1 programmed to introduce double-strand breaks (DSBs) at the target region in the genome of *C. glutamicum* (Cameron Coates et al., 2019; Cho et al., 2017; Jiang et al., 2017; Liu et al., 2017; Peng et al., 2017, 2019; Wang et al., 2018a; Zhang et al., 2019). These techniques impose DSBs that are lethal if left unrepaired; thus, most surviving clones carry the desired mutation. However, the efficiency of introduction of DSBs is dependent on the design of the sgRNA that guides the endonuclease. Moreover, the expression of endonucleases is deleterious for strains and imposes a great risk of unwanted genome rearrangements due to off-the-target DSBs and unexpected homologous recombination events. On the other hand, catalytically inactive endonucleases engineered with a deaminase domain were programmed to cause base mismatches in the genome of *C. glutamicum*, triggering native repair machinery to act in a predictable manner (Wang et al., 2018b, 2019). Since DSBs are not involved, these techniques are less likely to cause unexpected genome rearrangements. However, the efficiency of base editing is still dependent on the design of the sgRNAs, and there are no universal selection methods for clones carrying mutations. The limitation of the current methods for genome editing in *C. glutamicum* has been addressed with success in other actinobacteria, notably *Streptomyces*. An iterative excision marker system was employed to convert genomic libraries into platforms for genome engineering in *S. albidoflavus* (previously named *S. albus* J1074) and *S. lividans* TK24 (Myronovskiy et al., 2014; Rebets et al., 2018). This method employs the same principle as that of non-replicative vectors but is more efficient due to the larger homology regions and the presence of excisable selection markers in the mutant strains.

The biotechnological features and the availability of metabolic engineering methods potentiates *C. glutamicum* as a chassis for the heterologous production of secondary metabolites. Several classes of natural products have been produced in *C. glutamicum*, including polyphenols and terpenoids [recently reviewed in (Kogure and Inui, 2018)], as well as polyketides and non-ribosomal peptides. The reported polyketides produced in *C. glutamicum* are flaviolin (3.9 mg L⁻¹) encoded by the 1.1-kb *rppA* gene from *Streptomyces griseus* (Yang et al., 2018), and 6-methylsalicylic acid (41 mg L⁻¹) encoded by the 6.4-kb *chlB1* from *Streptomyces antibioticus* (Kallscheuer et al., 2019). The modest titers of most of the published attempts to produce secondary metabolites in *C. glutamicum* highlight the need for sophisticated metabolic analysis and engineering methods. On a positive note, native 4'-phosphopantetheinyl transferases (PPTase) of *C. glutamicum* were able to activate type

I polyketide synthases and non-ribosomal peptide synthetases (Kallscheuer et al., 2019). Furthermore, the reported production of 5.4 g L⁻¹ of violacein in an engineered *C. glutamicum* strain (Sun et al., 2016) clearly shows the potential for amino acid-based secondary metabolites in this organism.

The use of *C. glutamicum* as heterologous host to produce complex secondary metabolites requires appropriate vectors and methods to deliver large biosynthetic gene clusters in a straightforward manner. Current methods to integrate heterologous DNA at specific sites into *C. glutamicum* include non-replicative vectors with homologous regions to the chromosome and vectors derived from the integration systems of bacteriophages. Integration based on homologous recombination has been performed using non-replicative vectors with chromosomal homology regions spanning from 300 to 500 bp such as downstream of the *ppc* gene (Vasicová et al., 1998), at the *rrnD* loci (Amador et al., 2000), at repetitive insertion sequences (Correia et al., 1996), between *cg1121* and *cg1122* (Baumgart et al., 2013) or at the *crtEb* locus (Hoffmann et al., 2018). Being technically simple, this technique often presents low integration frequency and requires laborious screening of markerless strains. Alternatively, integration based on site-specific integrases of phages was performed, taking advantage of the native ϕ AAU2 (Marrec et al., 1994), ϕ 16 (Moreau et al., 1999a), ϕ 304 L (Moreau et al., 1999b), and β (Oram et al., 2007) attachment sites (*attB*) in *C. glutamicum*; ATCC 13032 and the genome minimized C1* strains retain all the native attachment sites described. However, all methods derived from corynephages rely on the permanent maintenance of recombinant *C. glutamicum* strains on selective antibiotics, restricting their use in large industrial fermentations. The recombination systems of bacteriophages without attachment sites in *C. glutamicum* have also been used for integration of exogenous DNA in engineered strains. The Cre recombinase of P1 phage was employed to catalyze the integration of *lox66*--bearing plasmids into the chromosome of the *C. glutamicum* CRRE strain harboring the recombination site *lox71*, with an efficiency of 3.1×10^4 colony forming units (CFU) per microgram (Suzuki et al., 2007). The integrase of the lactococci phage TP901-1, expressed in *cis*, was shown to promote integration of the *attP*_{TP901-1}-bearing plasmids onto the matching site of the engineered *C. glutamicum* JS34 strain with an efficiency of 1.3×10^3 CFU μ g⁻¹ (Shen et al., 2017). The authors combined this TP901-1-based system with Cre/*loxP* elements to develop a platform for repetitive and markerless chromosomal integration (Shen et al., 2017).

Herein, we present a general method and tools for genetic engineering of *C. glutamicum* grounded on a genome-wide cosmid library suitable for double allelic exchange and a pre-designed marker excision system. Starting from a pre-optimized chassis, we demonstrated the usefulness of the library by seamlessly replacing the carotenoid biosynthetic gene cluster of *C. glutamicum* by an additional genetic gateway derived from actinobacterial integration systems. The resulting strain was the host for heterologous expression of the polyketide synthase RppA and of the PPTase-dependent non-ribosomal peptide synthetase BpsA, demonstrating the feasibility of producing natural products via phage-based integrative systems in *C. glutamicum*.

2. Materials & methods

2.1. Bacterial strains, plasmids, media and growth conditions

All the strains and vectors used in this work are listed in Table 1. The *E. coli* strains were routinely cultivated in Lysogeny Broth (LB) with 5 g L⁻¹ NaCl (Lennox, 1955), unless stated otherwise. The *C. glutamicum* strains were routinely cultivated in LB or CGXII defined media (Keilhauer et al., 1993) with 4% (w/v) glucose, 0.2 mg L⁻¹ biotin and 30 mg L⁻¹ protocatechuic acid. When necessary, the following antibiotics were added to the media: ampicillin (Ap, 100 μ g mL⁻¹ to *E. coli*), apramycin (Am, 50 μ g mL⁻¹ to *E. coli*, 25 μ g mL⁻¹ to *C. glutamicum*), kanamycin (Km, 50 μ g mL⁻¹ to *E. coli*, 25 μ g mL⁻¹ to *C. glutamicum*), hygromycin

Table 1
Strains and vector used in this work used in this study.

Bacterial strains and vectors	Description	Source or reference
<i>E. coli</i>		
DH5 α	General cloning host	Hanahan (1983)
GB05-dir	Host for recombinering procedures	Fu et al. (2010)
EPI300-T1 ^R	Phage T1-resistant host for generation of cosmid library	Lucigen
<i>C. glutamicum</i>		
CR099attB _{C31}	<i>C. glutamicum</i> CR099 derivative for integration of ϕ C31-based vectors	Gift of J. Kalinowski (Present address: Center for Biotechnology, Bielefeld University, Bielefeld, Germany) [Based on CR099 strain (Baumgart et al., 2018; Unthan et al., 2015)]
BCA-Am	<i>C. glutamicum</i> CR099attB _{C31} derivative, in which the <i>crt</i> locus (<i>cgp.0717</i> - <i>cgp.0723</i>) has been replaced by <i>saac</i> (3')IV-attB _{BT1} cassette	This work
BCA	<i>C. glutamicum</i> BCA-Am derivative, in which the <i>saac</i> (3')IV marker has been excised.	This work
Plasmids		
pTES	Am ^R ; ϕ C31-based plasmid for stable and markerless integration via attB _{C31}	Herrmann et al. (2012)
cos15A-gus	Cm ^R , <i>gusA</i> ; vector for construction of cosmid genome library, suitable for allelic exchange in streptomycetes	Gift of M. Lopatniuk (Present address: Saarland University, Campus C2.3, 66123, Saarbrücken, Germany)
pK19mobsacB	Km ^R , <i>sacB</i> ; plasmid for allelic exchange in <i>C. glutamicum</i>	Schafer et al. (1994)
cos15A-sacB	cos15A-gus derivative, suitable for allelic exchange in <i>C. glutamicum</i>	This work
patt-saac	Ap ^R ; source of cassette for markerless deletions	Myronovskiy et al. (2014)
patt-saac-attBT1	patt-saac derivative, source of cassette for markerless deletions with attB _{BT1}	This work
pCG_4_F11	cos15A-sacB derivative containing the <i>crt</i> locus (<i>cgp.0717</i> – <i>0723</i>)	This work
pCG_4_F11_Δcrt::attB _{BT1} -saac(3')IV	pCG_4_F11 derivative for replacement of <i>crt</i> locus by <i>saac</i> (3')IV-attB _{BT1} cassette	This work
pKHInt31	Hyg ^R ; plasmid with temperature-sensitive replicon pSG5 for expression of C31-Int	Myronovskiy et al. (2014)
pEKEx2	Km ^R ; <i>C. glutamicum</i> /E. coli shuttle vector for inducible gene expression	Eikmanns et al. (1991)
pUC57-rppA _{Cgl}	Ap ^R ; source of SD_rppA _{Cgl}	This work
pUC57-rppA _{Cgl} Strept	Ap ^R ; source of SD_rppA _{Cgl} Strept	This work
pEKEx2-rppA _{Cgl}	pEKEx2 derivative for expression of flaviolin (for <i>C. glutamicum</i>)	This work
pEKEx2-rppA _{Cgl} Strept	pEKEx2 derivative for expression of flaviolin (for <i>C. glutamicum</i> and streptomycetes)	This work
pTES-rppA	pTES derivative for production of flaviolin (<i>C. glutamicum</i> /streptomycetes), via attB _{C31}	This work
pALCRE	Hyg ^R ; plasmid with temperature-sensitive	Herrmann et al. (2012)

Table 1 (continued)

Bacterial strains and vectors	Description	Source or reference
	replicon pSG5 for expression of Cre	
pRT802	Km ^R ; ϕ BT1-based integrative plasmid	Gregory et al. (2003)
pRT801	Am ^R ; ϕ BT1-based integrative plasmid	Gregory et al. (2003)
pMVM3100	pRT801 derivative with insulated multicloning site, constitutive P57 promoter (Siegl et al., 2013) and optimized 5' UTR N12 (Horbal et al., 2018)	M.V. Mendes (unpublished)
pFM69SD	pMVM3100 derivative with insulated multicloning site, constitutive P69 promoter (Zhang et al., 2018) and optimized 5' UTR of pEKEx2-rppA _{Cgl}	This work
pSMART-3B20	Cm ^R ; BAC of the genomic library of <i>S. albidoflavus</i> containing the <i>bpsA</i> gene	Gift of M. Myronovskiy (Present address: Saarland University, Campus C2.3, 66123, Saarbrücken, Germany)
pUC19	Ap ^R ; <i>lacZ</i> α , general cloning vector	Yanisch-Perron et al. (1985)
pUC19-opt-bpsA-H6	Ap ^R ; source of <i>bpsA</i> gene of <i>S. albidoflavus</i> with 5' codon optimization and encoding for a C-terminal His-tag	This work
pFM69SD-bpsA	pFM69SD derivative for expression of indigoidine	This work

Abbreviations: Am^R – apramycin resistance marker; Ap^R – ampicillin resistance marker; attB_{BT1} – bacterial attachment site for phage ϕ BT1; attB_{C31} – bacterial attachment site for phage ϕ C31; Cm^R – chloramphenicol resistance marker; Hyg^R – hygromycin resistance marker; Km^R – kanamycin resistance marker.

(Hyg, 80 μ g mL⁻¹ to *C. glutamicum*), chloramphenicol (Cm, 12.5 μ g mL⁻¹ to *E. coli*, 10 μ g mL⁻¹ to *C. glutamicum*). Transformation of *C. glutamicum* was performed by electroporation (Egeling and Bott, 2005), and selection of transformants was achieved in BHIS medium with half of the usual antibiotic concentration. Integration of plasmids into *C. glutamicum* genome was confirmed by PCR using the appropriate primer pairs. Unless stated otherwise, *E. coli* was grown at 37 °C and *C. glutamicum* was grown at 30 °C.

2.2. Recombinant DNA manipulations

The isolation of plasmid DNA was performed using NZYMiniprep kit following instruction of the manufacturer. The isolation of cosmid DNA was performed using alkaline lysis method (Sambrook and Russell, 2001). The isolation of genomic DNA of *C. glutamicum* was performed according to the salting-out method (Kieser et al., 2000), unless stated otherwise. Purification of enzymatic reactions and DNA fragments from agarose gels was performed using NZYGelpure kit following instruction of the manufacturer. PCR was performed using Q5® High-Fidelity DNA Polymerase (NEB), or GoTaq® G2 Flexi DNA Polymerase (Promega), following manufacturer's instructions. The oligonucleotides used in this study are listed in Table 2. Restriction enzymes, T4 ligase, alkaline phosphatase (FastAP), T4 polynucleotide kinase (T4 PNK) and GenEuler DNA Ladder Mix were all provided by ThermoFisher Scientific and employed following manufacturer's instructions. Genes were adapted to the codon usages of the specified organism and synthesized by GenScript (Piscataway, NJ, USA). Routine Sanger sequencing was conducted by STAB VIDA (Caparica, Portugal).

To produce a suitable genetic backbone for the cosmid library, the cos15A-sacB cosmid was constructed. For that, the *sacB* was amplified from pK19mobsacB (Schafer et al., 1994) using the primers sacB_F and

Table 2
Oligonucleotides and DNA fragments used in this study.

Primer	Sequence (5' to 3')	Use
pTESmcs_fwd	GCGAGGAAGCGGAAGAG	Sequence pTES-based plasmids and check integration of pTES derivatives at <i>attB_{C31}</i> locus
pTESmcs_rvs	CGATGTAGGAGGGCGTGTG	
attB-CR99_fwd	CCATCGCAGCCCGACTTG	
attB-CR99_rvs	AGCAGCCCCACCAACGTC	Sequence and check integration at <i>attB_{C31}</i> locus
sacB_F	CATTTGCGATCCAAATCGCGGGTTTGTAC	
sacB_R	GAAATTTGGATCCATCTCGTGATGGCAGGTTGG	Amplify <i>sacB</i> from pK19mobsacB
sacB-seq_F1	AGCGACACTGAATACGG	Check and sequence <i>sacB</i> in cos15A-sacB
sacB-seq_F2	GCATTAGCCGGAGATCC	
sacB-seq_F3	GATGAAACCGCTGATTGC	Sequence <i>PmeI</i> -cloned insert into cos15A-sacB
sacB-seq_R1	GACATTGGCCTATTGTAAGC	
cosmid-chk-For	GATCTCCATCGACTAAACGT	
cosmid-chk-Rev	GTAACTGCGGTCAAGATAT	
IMES_crt-aac_fwd	ACAACGATTTGGGAATTTTTCGAGATAACTGGCCGTGTGATACCTCGAATGTCACTATAGGCGCAATTGGC	
IMES_crt-aac_rvs	GCGTGGGTGGCTAGGCAAGTTACAGAACTGATCGGATAAAAGCAGAGTTACTGGAGCTGCTTCGCTAG	
crt_delchk2_fwd	GCGCAATGTCTTCTCCGAC	
crt_delchk2_rvs	GCCTACTTGATGAGGCCGAA	
pEKEx2MCS_F	ACTCCCGTCTGGATAATG	
pEKEx2MCS_R	ACTTCTGAGTTCGGCATG	
pRT_ORI_SEQ	GGATTCCCGTTGAGCACCGC	Sequence <i>crt</i> locus
P69-RBS_F	TGATTAAGCTTAAATATTATTTAAATTTCCGTTGACAAGACAATGCGCATGAGATATAATGGCA	
P69-RBS_R	TGTTATCATATGCTATCTCCTTACTTTCTGACAGGTACCGCACTATGCCATTATATCTCATGCGC	Sequence pEKEx2-based plasmids and produce <i>P_{tac}-SD rppA_{CglStrept}</i> fragment
bpsA-opt-fwd	TTGGTAGTCGACTGTCAGAAAAGTAAGGAGATAGCATATGTCCACCTCCACCCACCTCCAGCAACCCGCCAGCACGCCTCCCTG	
bpsA-H6-rvs	TAGAAAGGTACCTACTAGTTTAATGGTGATGGTGGGCGAAGAGTCCAAGGCGCGTA	Check integration of pRT801 and pRT802 derivatives at <i>attB_{BT1}</i> locus
bpsA1	AGAAGACGCCGATCAG	
bpsA2	CGACAACCAAGGTCAAAC	Create cassette with P69 promoter and 5' UTR of <i>rppA_{Cgl}</i>
bpsA3	AAGATCCACTTCTCTGGG	
bpsA4	CTCGTACTTCTCTGTGG	Amplify <i>bpsA</i> gene with 5' optimized codons and encoding a 3' His-tag
bpsA5	TACCATATGAACCTGCG	

sacB_R, digested with *Bam*HI and cloned into the *Bam*HI sites of cos15A-gus (gift of M. Lopatniuk). The orientation of *sacB* was confirmed by DNA restriction and the identity of the gene was confirmed by sequencing.

To generate a cassette suitable for iterative marker excision that would leave an *attB_{BT1}* locus as genetic scar, the **patt-saac-attBT1 plasmid** was constructed. For that, a synthetic 152-bp fragment harboring the *attB_{BT1}* locus flanked by terminators was cloned into the *Stu*I site of patt-saac plasmid.

To create a platform to delete the main cluster of carotenoid biosynthetic genes in *C. glutamicum* (*cgp_0717–0723*), the pCG_4.F11 cosmid from the genomic library of *C. glutamicum* CR099attB_{C31} was modified for appropriate allelic exchange. The cassette containing apramycin resistance and the *attB_{BT1}* sequence was amplified from patt-saac-attBT1 using the primers IMES_crt-aac_fwd and IMES_crt-aac_rvs. Red/ET-mediated recombination of pCG_4.F11 cosmid with the amplified fragment was performed in GB05-dir cells, following standard protocols (Datsenko and Wanner, 2000). The originated **pCG_4.F11_Δcrt::attB_{BT1}-saac(3')IV cosmid** was checked by PCR and confirmed by sequencing.

To express *rppA* gene in an inducible manner in *C. glutamicum*, the **pEKEx2-rppA_{Cgl}** and **pEKEx2-rppA_{CglStrept}** were constructed. The Shine-Dalgarno (SD) sequence and the *rppA* gene with adapted to the codon usage of *C. glutamicum* (*rppA_{Cgl}*) or to the mixed codon usages of *C. glutamicum* and *Streptomyces* (*rppA_{CglStrept}*) were synthesized and cloned by GenScript to originate the **pUC57-rppA_{Cgl}** and **pUC57-rppA_{CglStrept}** plasmids, respectively. The SD sequence and *rppA* variants were digested *Pst*I/*Bam*HI from the source plasmids and cloned into the respective sites of pEKEx2. The identity of the generated plasmids was confirmed by sequencing.

To express *rppA_{CglStrept}* from an ϕ C31-integrative plasmid, the **pTES-rppA** plasmid was constructed. The *P_{tac}* promoter, SD sequence and

rppA_{CglStrept} gene were amplified using the PNK-phosphorylated primers pEKEx2MCS_F and pEKEx2MCS_R. The amplified fragment was blunt cloned into the *Sca*I-digested pTES plasmid to yield pTES-rppA and confirmed by sequencing.

To create an ϕ BT1-integrative plasmid with insulated multicloning site, the **pFM69SD** plasmid was constructed. A cassette containing the P69 promoter (Zhang et al., 2018) and the 5' UTR of *rppA_{Cgl}* was created by PCR using the primers P69-RBS_F and P69-RBS_R. The amplified fragment was digested with *Hind*III and *Nde*I and cloned into the respective sites of pMVM3100 to yield pFM69SD. The identity of pFM69SD was confirmed by sequencing.

The *bpsA* gene (NCBI accession number CP004370.1, positions 6357383 to 6361213) was amplified from pSMART-3B20 using the primers bpsA-opt-fwd and bpsA-H6-rvs, digested with *Sal*I and *Kpn*I, and cloned into the respective sites of pUC19 to yield **pUC19-opt-bps-H6**. The optimization of the 5' sequence of *bpsA* gene (included in the primer bpsA-opt-fwd) corresponded to the most prevalent codons in *C. glutamicum* ATCC 13032, as described in the Codon Usage Database (Nakamura et al., 2000). The identity of pUC19-opt-bps-H6 was confirmed by sequencing.

To express *bpsA* from an ϕ BT1-integrative plasmid, the **pFM69SD-bpsA** plasmid was constructed. The 5'-modified *bpsA* gene (also encoding a C-terminal His-tag) was digested from pUC19-opt-bpsA-H6 plasmid with *Nde*I and *Spe*I and cloned into the respective sites of pFM69SD to yield pFM69SD-bpsA.

2.3. Markerless integration of phage-based vectors at ϕ C31 bacterial attachment site

To integrate the ϕ C31-based pTES (Herrmann et al., 2012) and derivative vectors at the ϕ C31 bacterial attachment site (*attB_{C31}*), *C. glutamicum* cells were transformed by electroporation and selected on

BHIS agar supplemented with $12.5 \mu\text{g mL}^{-1}$ apramycin for 48 h. The integration was confirmed PCR using the primers pTESmcs fwd and attB-CR99_fwd. To excise the backbone of pTES-based plasmids after integration, the relevant *C. glutamicum* strains were transformed with the pALCRE plasmid (Herrmann et al., 2012) and selected on BHIS agar supplemented with $40 \mu\text{g mL}^{-1}$ hygromycin for 48 h. Several transformants were individually grown in 5 mL LB for 16 h, then plated and grown in LB agar for 16 h. Several clones were picked and grown on Am25Hyg80-, Hyg80-, and LB agar for 16 h. The excision of pTES backbone and the loss of pALCRE plasmid were indicated by sensitivity to apramycin and hygromycin, respectively. The identity of the markerless strains was confirmed by PCR and sequencing using the primer pair attB-CR99_fwd/rvs.

2.4. Generation of the cosmid-based genomic library

The cosmid library of *C. glutamicum* CR099attB_{C31} was designed as a platform to perform two rounds of positive selection of allelic exchange with homologous regions of the chromosome, following the protocol of the MA171E CopyControl™ Fosmid & HTP Fosmid Kit (Lucigen), with minor modifications. The concentration and quality of genomic DNA isolated from *C. glutamicum* CR099attB_{C31} was analyzed by 0.4% (w/v) agarose gel electrophoresis, using GeneRuler High Range DNA Ladder (ThermoFisher Scientific) as reference. Ten micrograms of total genomic DNA were repaired to blunt, 5' phosphorylated sheared DNA fragments, using Lucigen's proprietary End-Repair Enzyme Mix (containing T4 DNA Polymerase and T4 Polynucleotide Kinase) and extending the reaction time to 6 h. The end-repaired DNA fragments were ligated to 0.7 μg of *PmeI*-digested and dephosphorylated cos15A-sacB vector. The ligated DNA was packaged into the phage particles by adding the MaxPlax Packaging Extract to the ligation reaction. Cells for transduction were precultured from a single colony in 10 mL LB supplemented with 10 mM MgSO₄, and 0.2% (w/v) maltose for 16 h, and then sub-cultured by diluting 1:100 of pre-culture in fresh medium, up to an OD₆₀₀ of 0.8. Transduction was performed by mixing 1 volume of packaged phage suspension with 10 vol of cells and incubating at 37 °C for 1 h. The titer of phage particles was determined by transducing the serial dilutions of packaged phages in Phage Dilution Buffer into *E. coli* EPI300-T1^R cells. To prepare the cosmid library, 250 μL of EPI300-T1^R cells were mixed with 25 μL of undiluted packaged phage suspension, and the transduced cells were plated in LB agar supplemented with $12.5 \mu\text{g}$ of chloramphenicol. After incubation for 20 h, a total of 576 clones of normal morphology were imaged and picked using a QPix2 Automatic Colony Picker and Imager (Genetix) to six 96-well plates containing 250 μL of Cm12.5-LB medium. After incubation at 37 °C for 24 h, at 800 rpm, the plates were replicated, and the cell suspensions were diluted in a glycerol solution, to a final concentration of 25% (w/v) glycerol. The sequencing of the genomic library was performed at Genewiz (Takeley, UK), by Sanger sequencing, using the primer cosmid-chk-For. To order the generated cosmid library, the Sanger reads were trimmed and mapped to the reference genome of *C. glutamicum* MB001 (Baumgart et al., 2013), using Geneious Prime® 2019.0.3 (Biomatters Ltd). The trimming and mapping parameters were the following: trimming region with more than a 5% chance of an error per base (5' and 3' ends), Geneious mapper with custom sensitivity and fine tuning up to 5 iterations, minimum mapping quality of 20 (imply multiple best matches are mapped to none), gaps not allowed. Only the mapped reads with perfect pairing to the 10 nucleotides of cos15A-sacB cloning region upstream the insert were unequivocally considered directional reads and used for calculation of library statistics. To predict the coverage and library statistics, 35-kbp sections were directionally extracted at the starting coordinates of the assigned reads and mapped to the reference genome, using Geneious Prime® 2019.0.3. Manual curing was performed to assess coverage of the attB_{C31} locus and other features.

2.5. Markerless genome edition using modified cosmids

To create a strain unable to produce carotenoids and featuring an attB_{BT1} integration site, deletion of the *crt* locus of *C. glutamicum* CR099attB_{C31} (*cgp_0717–0723*) was performed using a modified cosmid from the generated library and the IMES procedure (Myronovskiy et al., 2014). For that, CR099attB_{C31} cells were transformed with pCG_4_F11_Δ*crt::attB_{BT1}-saac(3')*IV cosmid and selected on BHIS agar supplemented with $12.5 \mu\text{g mL}^{-1}$ apramycin for 48 h. Apramycin-resistant clones were picked and grown on Am25-LB agar plates for 16 h. First crossover events were tested by colony PCR targeting the *sacB* gene of the cosmid, using primers sacB-seq_F3 and sacB-seq_R1, and by the ability to form colonies on Cm10-LB agar in 16 h. Several first crossovers were individually grown in 5 mL LB for 16 h, then plated and grown on LB agar with 10 g L^{-1} NaCl and 10 g L^{-1} sucrose for 16 h. Fast-growing colonies were picked on Am25- or Cm12.5-LB agar. Second crossover events were indicated by the chloramphenicol-sensitive, apramycin-resistant phenotype. Several second crossovers were picked and grown on Am25-LB agar for 16 h. A second crossover clone (*C. glutamicum* BCA-Am) was cultivated in 5 mL LB for 16 h, and its genomic DNA was isolated. The *crt* locus of the selected clone was amplified from the isolated gDNA using primers crt_delchk2_fwd and crt_delchk2_rvs and double allelic exchange was confirmed by sequencing. To excise the apramycin resistance marker according to the IMES protocol (Myronovskiy et al., 2014), *C. glutamicum* BCA-Am was transformed with pKHInt31 plasmid for transient expression of C31-Integrase. The process of strain isolation and plasmid curing proceed was described above for the pALCRE plasmid. The identity of the markerless strains was confirmed by PCR and sequencing using the primers crt_delchk2_fwd and crt_delchk2_rvs. A curated markerless clone dubbed *C. glutamicum* BCA was cultivated in 5 mL LB for 16 h, and its genomic DNA was isolated using the GeneJET Genomic DNA Purification Kit (ThermoFisher Scientific), following manufacturer's instructions for Gram-positive bacteria. The isolated DNA from the *C. glutamicum* BCA strain was sent to Novogene UK Company Ltd (Cambridge, UK) for PCR-free library construction (average size of 350 bp) and 150-bp paired-end Illumina (NovaSeq 6000) sequencing (Q30 of 92.58%). De novo assembly and reads mapping were carried out with CLC Genomics Workbench 20 (Qiagen) using as reference the genome of *C. glutamicum* MB001 strain (NCBI accession number CP005959.1).

2.6. Integration of phage-based vectors at ϕ BT1 bacterial attachment site

To integrate the ϕ BT1-based pRT802 (Gregory et al., 2003) and pFM69SD vectors at the ϕ BT1 bacterial attachment site (attB_{BT1}), *C. glutamicum* BCA cells were transformed by electroporation and selected on BHIS agar supplemented with $12.5 \mu\text{g mL}^{-1}$ kanamycin or apramycin, respectively, for 48 h. The integration was confirmed by PCR using the primers pRT_ORI_SEQ and crt_delchk2_rvs.

2.7. Production and quantification of flaviolin and indigoidine

The expression of flaviolin in *C. glutamicum* was based on the replicative pEKEx2-rppA and the ϕ C31 integrative pTES-rppA vectors. The expression of indigoidine in *C. glutamicum* was based on the ϕ BT1 integrative pFM69SD-bpsA. The pEKEx2 and pTES, and pFM69SD vectors served as negative controls for production of flaviolin and indigoidine, respectively. A single colony was inoculated and cultivated for 6–8 h in 5 mL LB medium. For pre-culture, 1 mL of this suspension was centrifuged at 5000g for 5 min, at room temperature, resuspended in 10 mL of CGXII medium and cultivated in a 100-mL flask for 12–16 h, at 200 rpm. For main culture, 25 mL of fresh CGXII medium were seeded with pre-culture to an OD₆₀₀ of 1 (or the equivalent OD₈₀₀ of 0.72 for indigoidine) and cultivated in a 250-mL flask for up to 96 h, at 200 rpm. In case of inducible systems (pEKEx2-based vectors), at 2 h of

cultivation, medium was supplemented with IPTG, to a final concentration of 500 μM . For cell dry weight (CDW) measurements, 1 mL of culture broth was collected, centrifuged at 16100g for 10 min, at room temperature and the pellet was dried at 55 °C for 48 h. For UV–Vis analysis, at defined timepoints, samples of culture broth were transferred to 2-mL tubes and centrifuged at 16100g for 10 min, at room temperature. Flaviolin was quantified directly from the supernatant of centrifuged samples. Indigoidine was extracted from the sediment of 100 μL of culture broth with 10 vol of dimethylsulfoxide (DMSO) (Chatterjee and Brown, 1981) and vigorous vortexing for 2 min, followed by centrifugation at 16100g for 10 min, at room temperature. The absorbance of flaviolin and indigoidine extracts were acquired in 1-cm path Optical Glass cuvettes (Hellma), in a UV Mini 1240 UV/Vis Spectrophotometer (Shimadzu) at 520 nm or 615 nm, respectively. The absorbance relative to the analyzed pigment (ΔA) was calculated by subtracting the correspondent absorbance of the negative control. The volumetric production of flaviolin was calculated using the published extinction coefficient (Krauser et al., 2012): flaviolin (mg L^{-1}) = $\Delta A_{520\text{nm}} / (1305 \text{ L mol}^{-1} \text{ cm}^{-1} \times 1 \text{ cm}) \times 206.153 \text{ g mol}^{-1} \times 10^3 \text{ mg g}^{-1}$. The volumetric production of indigoidine was calculated using the published extinction coefficient (Yumusak et al., 2019): indigoidine (g L^{-1}) = $10 \times \Delta A_{615\text{nm}} / (1570 \text{ L mol}^{-1} \text{ cm}^{-1} \times 1 \text{ cm}) \times 248.19 \text{ g mol}^{-1}$. The

data were parsed using Microsoft Office Excel and analyzed using GraphPad Prism 5 for Windows.

2.8. Nucleotide sequence accession number

All the referenced gene tags correspond to the nomenclature of *C. glutamicum* MB001 (Baumgart et al., 2013), deposited with the NCBI accession number CP005959.1. The correspondent cg numbers of the mentioned genes can be obtained by replacing the *cgp_* prefix by the *cg* prefix. The sequence of the genome of *C. glutamicum* BCA strain is available at NCBI (accession number CP059382).

3. Results

3.1. Characterization of a markerless integrative system based on actinophage ϕC31 in *C. glutamicum* CR099attB_{C31}

The *Corynebacterium glutamicum* CR099attB_{C31} host is derived from strain CR099 (Baumgart et al., 2018; Unthan et al., 2015) and harbors an attachment site for the ϕC31 phage (*attB*_{C31}) between loci *cgp_0223* and *cgp_0228* (Supp. data 1). The CR099attB_{C31} strain was derived from CR099 strain by double allelic exchange with a pK19mobsacB plasmid

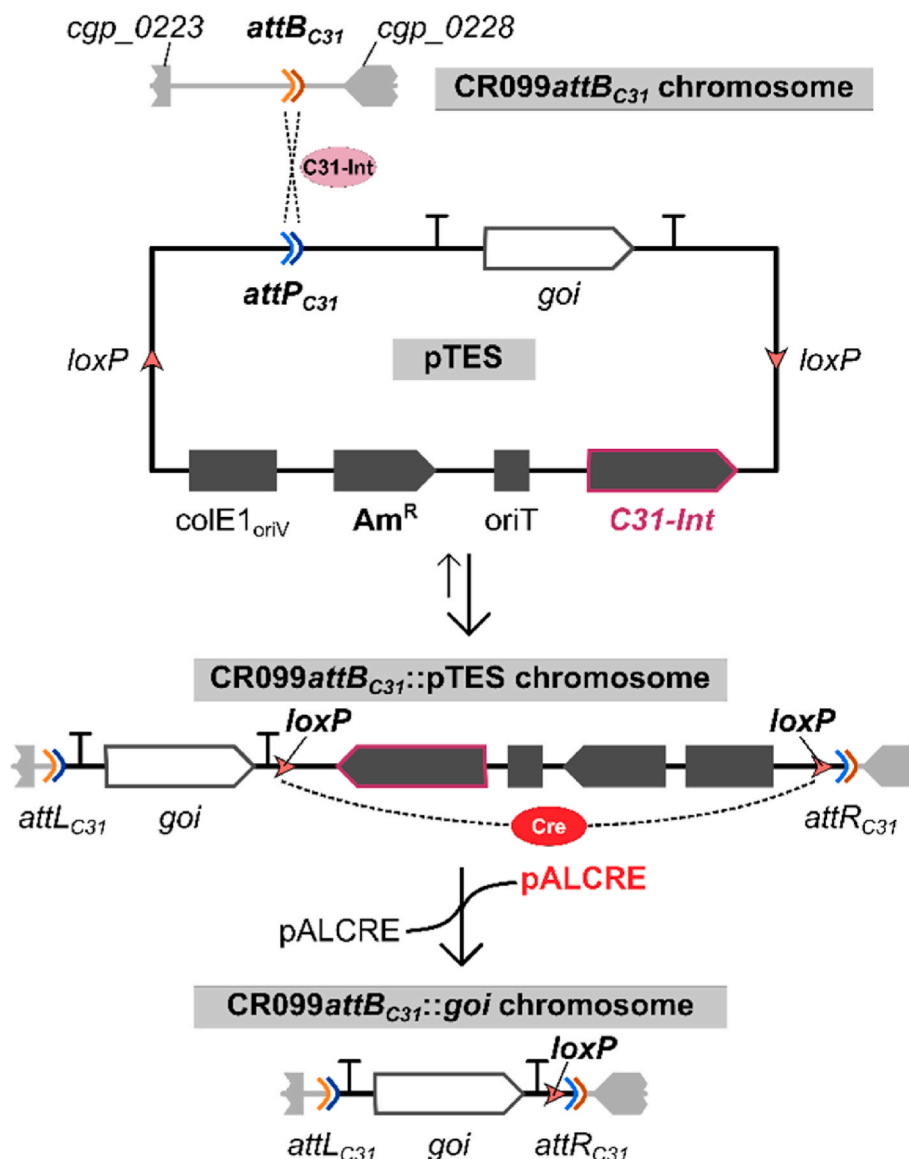


Fig. 1. Markerless integration of the ϕC31 -based pTES vector into the genome of CR099attB_{C31}. The site-directed integration of the *attP*_{C31}-harboring pTES vector (Herrmann et al., 2012) into *attB*_{C31} encoded in the genome of CR099attB_{C31} is catalyzed by C31-Integrase (C31-Int). Integration creates two short scars with low recombination potential, *attL*_{C31} and *attR*_{C31}. The gene of interest (*goi*) is carried in the multicloning site flanked by transcriptional terminators. The backbone elements of the integrated vector (*colE1*_{oriV}, *Am*^R, *oriT*, *C31-Int*) are flanked by *loxP* sites and can be excised by expressing Cre recombinase from the curable pALCRE plasmid (Herrmann et al., 2012).

carrying the *attB*_{C31} site of *Streptomyces lividans* TK24 (NCBI accession number CP009124.1, coordinates 4286958–4286995) (J. Kalinowski, personal communication). To test the functionality of the *attB*_{C31} site, the strain CR099*attB*_{C31} was transformed with the ϕ C31-based pTES vector (Herrmann et al., 2012). The pTES vector encodes the phage attachment site for ϕ C31 (*attP*_{C31}) and the serine recombinase C31-Int that catalyzes the site-directed, directional recombination into *attB*_{C31} sites (Fig. 1). Integration of pTES into the strain CR099*attB*_{C31} was verified by the appearance of apramycin-resistant colonies at up to 600 CFU μ g⁻¹ of plasmid. To confirm the site-specific integration of pTES, the genomic DNA of CR099*attB*_{C31} transformed with pTES was used for Southern blot analysis and plasmid rescue experiments (Supp. methods; Supp. Fig. 1). The results showed that pTES integration occurred at the expected *attB*_{C31} site. However, the results also suggested that the integration of pTES can be reversed *in vivo*, resulting in episomes. The frequency of pTES self-excision was accessed indirectly by transforming *E. coli* with plasmid DNA isolated from *C. glutamicum* CR099*attB*_{C31}::pTES (Supp. methods). The results showed that the number of pTES episomes per cell was less than 0.005% that of a pBL1-based replicative vector in *C. glutamicum*. Several functions of the integrated pTES vector, such as the apramycin resistance marker and C31-Integrase, are unnecessary for lasting expression of the delivered gene of interest (*goi*). To demonstrate the markerless integration of pTES, the *Cre/loxP* system was used to evict the backbone of the pTES vector from the *C. glutamicum* CR099*attB*_{C31}::pTES chromosome (Fig. 1), specifically using the pALCRE plasmid (Herrmann et al., 2012). Apramycin-sensitive transformants were obtained with high efficiency (16 sensitive clones out of 16 tested colonies). Three out of four apramycin-sensitive clones were shown to have undergone recombination of the *loxP* sites, resulting in the excision of the 5.5 kb-long backbone of the pTES vector and maintenance of the multi-cloning site. In summary, the *attB*_{C31} locus of the strain CR099*attB*_{C31} was shown to be functional and allowed site-specific and markerless integration of ϕ C31-based plasmids.

3.2. Generation of an ordered cosmid library of *C. glutamicum* CR099*attB*_{C31}

To facilitate the genetic manipulation of *C. glutamicum* CR099*attB*_{C31}, an ordered genomic library was generated. The library was built on the cos15A-sacB cosmid (Supp. data 2) that allows λ -mediated transfection into *Escherichia coli* and is suitable for double allelic exchange owing to selection with chloramphenicol and counter-selection with the levansucrase gene (*sacB*). High-molecular weight genomic DNA of CR099*attB*_{C31} was isolated by salting-out method, yielding detectable fragments over 24 kbp only. The isolated genomic DNA of *C. glutamicum* CR099*attB*_{C31} was directly end-repaired to blunt, 5'-phosphorylated fragments using an enzyme mix containing T4 DNA polymerase and T4 DNA kinase. The repaired DNA fragments were cloned into the cos15A-sacB cosmid and packaged into λ particles. After transduction, 576 *E. coli* clones were selected and placed into 96-well plates (identified as pCG_1–6_A1–H12). From the sequencing of the 5'-end of the inserts, 446 high-quality reads were unambiguously mapped to the *C. glutamicum* MB001 genome (Baumgart et al., 2013) (Fig. 2, Supp. data 3), the CR099 ancestor strain whose sequence was publicly available. Of the residual 130 clones that could not be mapped to MB001 genome, in 107 instances failed to one or several of the mapping parameters (namely, 93 instances could not be mapped due to poor base calling quality, 3 instances fell into repetitive regions, 8 instances mapped to the cosmid backbone, and 3 instances mapped to *E. coli* genome or miscellaneous vector regions) and in 23 instances, due to poor base calling quality of the 5' end, reads could not be assigned a direction.

Overall, the generated library achieved an average 5.1x coverage of 94.2% of the chromosome of *C. glutamicum* MB001. Notably, the regions not covered by our library were also unrepresented in the cosmid library created for the whole-genome sequencing project of *C. glutamicum* ATCC

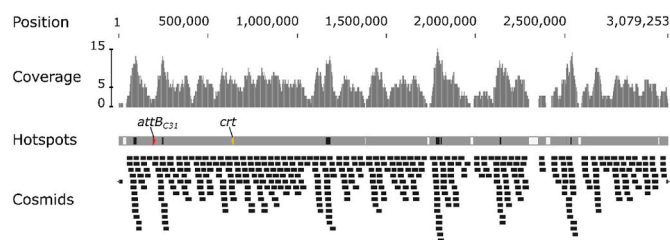


Fig. 2. Representation of the coverage of the generated genomic library of *C. glutamicum* CR099*attB*_{C31}. Position refers to the original numbering of strain MB001. The coverage graph indicates the number of cosmids covering a given region, and the average coverage is indicated in the vertical axis. Hotspots are regions with significant deviation from the average coverage: white strips indicate no-coverage areas, and black strips indicate areas with coverage at least two standard deviations higher than the average. Individual cosmids covering a given region are depicted on the bottom as black horizontal bars, covering 35 kbp-long stretches of the chromosome. Created using Geneious software. *attB*_{C31} – bacterial attachment site for phage ϕ C31; *crt* – main cluster of biosynthetic genes of carotenoids in *C. glutamicum*.

13032 (Bathe et al., 1996; Tauch et al., 2002) (Supp. Fig. 2). Thorough analysis of the unrepresented genes (Supp. data 4) against the PanDaTox gene collection (Kimelman et al., 2012) revealed that seven out of the eight regions without coverage in our library contain homologs reported as unclonable in *E. coli* (Supp. methods, Supp. data 5). Access to the bacterial clones of the library is available upon request.

3.3. Generation of a carotenoid-free *C. glutamicum* strain with an additional ϕ BT1 integration site

To demonstrate the potential of the cosmid library in genomic editing procedures, we generated a carotenoid-free *C. glutamicum* strain with an additional *attB*_{BT1} gateway replacing the main cluster of carotenoid biosynthetic genes (*crt* locus, *cgp*.0717–0723) (Fig. 3).

The sequence of the introduced *attB*_{BT1} site corresponds to that of *Streptomyces lividans* TK24 (NCBI accession number CP009124.1, coordinates 3212045–3212117). Initially, the *crt* locus found in the pCG_4_F11 cosmid was replaced by a cassette generated by PCR harboring the *attB*_{BT1} sequence and the necessary elements for iterative marker excision (Myronovskyi et al., 2014), in recombineering-proficient *E. coli* strain. In detail, the elements for iterative marker excision were the codon-optimized apramycin resistance marker *saac*(3')IV flanked by two engineered attachment sites *attB*_{C31} (CC) and *attP*_{C31}(GG) with central GG core dinucleotides, which are orthologous to the original *attB*_{C31} (Myronovskyi et al., 2014). For comparison purposes, the pK19mobsacB-*crt* plasmid was created, featuring the same cassette flanked by 432-bp left and 422-bp right homology arms targeting the *crt* locus (Supp. methods). The CR099*attB*_{C31} strain was transformed with the pCG_4_F11_Δ*crt*::*attB*_{BT1}-*saac*(3')IV cosmid or pK19mobsacB-*crt* plasmid, and the first crossover events were selected by apramycin resistance. The average frequency of first crossover was 9 CFU μ g⁻¹ (7×10^{-11} CFU per molecule) of pK19mobsacB-*crt* and 51 CFU μ g⁻¹ (2.0×10^{-9} CFU per molecule) of pCG_4_F11_Δ*crt*::*attB*_{BT1}-*saac*(3')IV cosmid. To obtain second crossover mutants, two first-crossover clones were grown without antibiotic selection and selected in the presence of sucrose. The second crossover events were screened by apramycin resistance, and kanamycin or chloramphenicol sensitivity phenotypes, for plasmid or cosmid-based manipulations, respectively. The percentage of second recombination events varied between 0 and 22% for manipulations using the plasmid and 44–94% using the cosmid. Interestingly, the percentage of events of reversion to the parental phenotype varied between 22 and 100% for manipulations using the plasmid and 6–56% using the cosmid. The identity of one of the clones that underwent double recombination using the cosmid as recombination template was confirmed by sequencing of

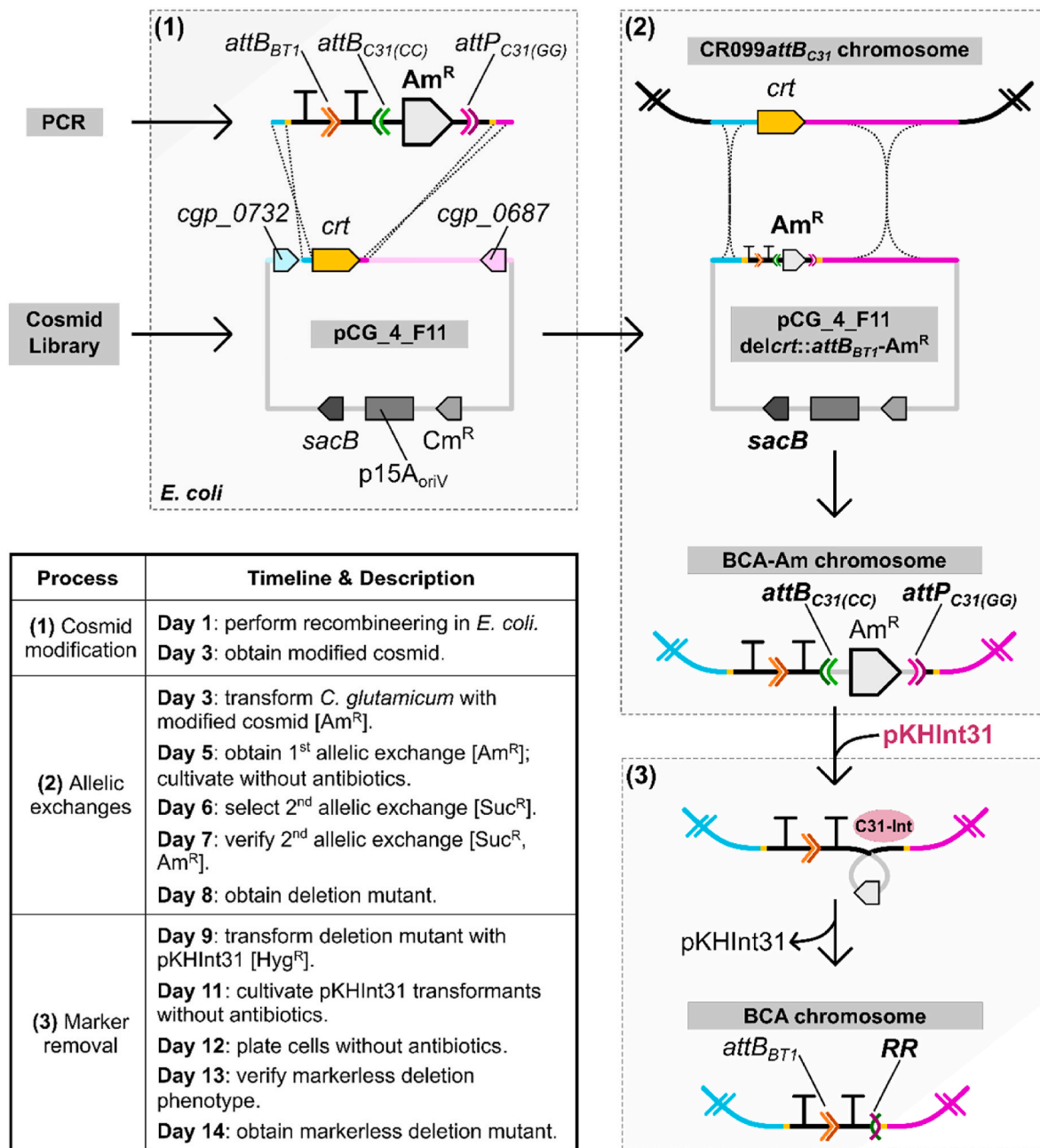


Fig. 3. Markerless replacement of the *crt* locus by the ϕ BT1 attachment site using the developed genomic library of *C. glutamicum* CR099attB_{C31}. The table represents the steps, procedure and time required for markerless genome editing using modified cosmids from the generated library. The panels (1), (2) and (3) show the detailed workflow of the markerless genome editing, focusing on (1) the cosmid modification, (2) the two consecutive allelic exchanges and (3) the marker excision. Only the selection markers relevant for each step are shown for clarity. The pKHInt31 plasmid used for marker excision was developed elsewhere (Myronovskiy et al., 2014). The referenced *cgp_0732* and *cgp_0687* are shown to illustrate the span of the pCG_4_F11 and pCG_4_F11_Δ*crt::attB_{BT1}-saac(3')IV* cosmids. Am^R – apramycin resistance marker; *attB_{BT1}* – attachment site for the ϕ BT1 phage; *attB_{C31(CC)}* – bacterial attachment site for the ϕ C31 phage with central GG core dinucleotides; *attP_{C31(GG)}* – phage attachment site for the ϕ C31 phage with central GG core dinucleotides; *C31-Int* – integrase of the ϕ C31 phage; Cm^R – chloramphenicol resistance marker; *crt* – main cluster of biosynthetic genes of carotenoids in *C. glutamicum*; Hyg^R – hygromycin resistance marker; p15A_{oriV} – plasmid origin of replication for *E. coli*; pSG5_{oriV} – plasmid origin of replication for *C. glutamicum*; RR – *attB_{C31(CC)}/attP_{C31(GG)}* recombination scar; *sacB* – sucrose sensitivity marker.

the replaced locus and the strain was dubbed *C. glutamicum* BCA-Am. To generate a markerless mutant strain, the apramycin resistance marker was excised from the genome of *C. glutamicum* BCA-Am by recombining the engineered *attB_{C31(CC)}* and *attP_{C31(GG)}* sites flanking the *saac(3')IV* gene. For that, the ϕ C31 integrase was transiently expressed from the pKHInt31 plasmid and marker-free mutants were screened by apramycin sensitivity. The identity of one apramycin-sensitive clone, named *C. glutamicum* BT1-C31-Albino (BCA), was confirmed by PCR and Sanger sequencing of the *attB_{BT1}* locus (Supp. data 6). The engineering process

from cosmid modification until the marker-free strain was performed in 14 days, with higher efficiency than conventional plasmid-based methods.

The genome of the *C. glutamicum* BCA strain was re-sequenced and the reads (a total of 7 668 172 reads) were mapped to the parental strain MB001 (NCBI accession number CP005959). In parallel, a *de novo* assembly of the reads was performed originating 28 contigs that were aligned with the MB001 genome using the Genome Finishing Module of the CLC Genomics Workbench package. The conflicts were manually

curated comparing the mapping of the reads to the MB001 genome with the contigs from the *de novo* assembly. 98.05% of the reads mapped to the MB001 genome and the average coverage was 374x. The BCA strain has a genome with a length of 3 061 143 bp and compared to MB001 strain it presents 74 nucleotide variants, the replacement of the *crt* cluster (*cgp_0717* to *cgp_0723*) for the *attB_{BT1}*, the replacement of the *cgp_0226* transposase encoding gene for the *attB_{C31}*, the deletion of 8 additional transposase encoding genes (*cgp_4030*, *cgp_0692*, *cgp_1213*, *cgp_2354*, *cgp_2426*, *cgp_2600*, *cgp_2725* and *cgp_2854*), 2 insertions of transposases at the intergenic region between *cgp_1046*–*cgp_1048* and at the *cgp_3195*. In summary, the genotype of the BCA strain is MB001 $\Delta crt::attB_{BT1}$ $\Delta cgp_{0226}::attB_{C31}$ $\Delta[cgp_{4030}; cgp_{0692}; cgp_{1213}; cgp_{2354}; cgp_{2426}; cgp_{2600}; cgp_{2725}]$ *cgp_{1046–1048}::tnp2b* *cgp_{3195}::tnp13b*. The genome of the BCA strain was deposited in the NCBI database with the accession number CP059382.

To assess the carotenoid production of the *C. glutamicum* BCA, the organic extract of the cells was analyzed by spectrophotometry (Supp. methods; Supp. Fig. 3). After 24 h of cultivation, the biomass of strain BCA showed less pigmentation than strain CR099*attB_{C31}*. The analysis of the methanolic extracts of the biomass showed that the visible spectrum corresponding to strain BCA lacked the characteristic absorption maxima present in the parental strain CR099*attB_{C31}*.

To test the functionality of the newly introduced *attB_{BT1}* locus, the strain BCA was transformed with the well-established ϕ BT1-based pRT802 vector (Gregory et al., 2003). The pRT802 vector encodes the

phage attachment site for ϕ BT1 (*attP_{BT1}*) and the serine recombinase BT1-Int that catalyzes the site-directed, directional recombination into the *attB_{BT1}* site. Integration of pRT802 into strain BCA was verified by the appearance of kanamycin-resistant colonies up to 10^5 CFU μ g⁻¹. No colonies were obtained upon transformation of the parental strain CR099*attB_{C31}* with pRT802, which corroborates the absence of an endogenous *attB_{BT1}* locus. To confirm the site-specific integration of pRT802, the genomic DNA of strain BCA transformed with pRT802 was isolated and digested for rescue plasmid experiments (Supp. methods; Supp. Fig. 4). The results showed that pRT802 integration occurred at the expected *attB_{BT1}* site. There was no evidence of episomal pRT802 in the assayed clones. The results show that the *attB_{BT1}* locus is functional and represents a unique locus for site-specific integration of ϕ BT1-based vectors into the genome of *C. glutamicum* BCA.

3.4. Production of the polyketide flaviolin in *C. glutamicum* using ϕ C31-based integration

To demonstrate that heterologous gene expression could be achieved using the ϕ C31 attachment site of the *C. glutamicum* strains, a sensitive and practical reporter system based on a polyketide synthase was designed. The candidate gene was *rppA* of *Saccharopolyspora erythraea*, encoding the type III polyketide synthase 1,3,6,8-tetrahydroxynaphthalene (THN) synthase responsible for the production of the diffusible red pigment flaviolin (Cortes et al., 2002). Despite our best efforts, the

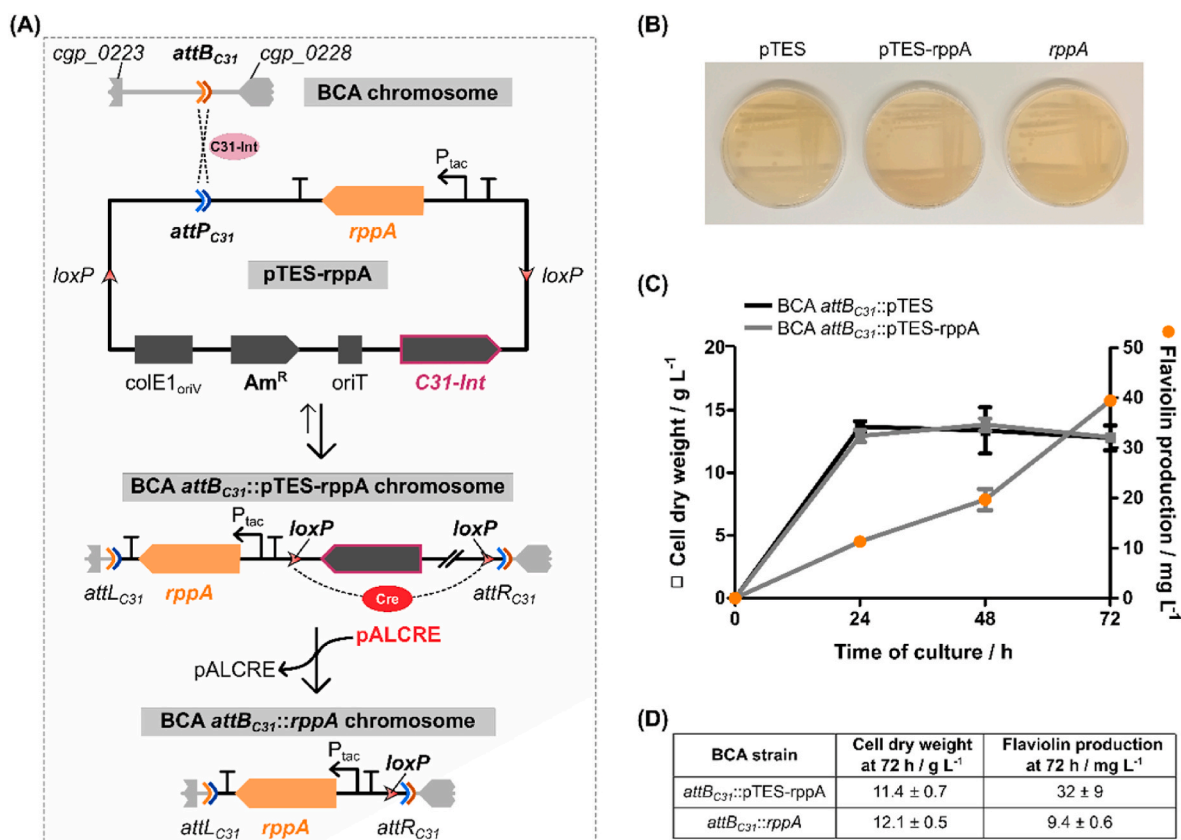


Fig. 4. Production of flaviolin in *C. glutamicum* using a ϕ C31-based integration system and *Cre*-mediated marker excision. The panel (A) shows a schematic representation of the pTES-rppA vector employed for ϕ C31-based integration of flaviolin biosynthetic gene *rppA*_{CglStrept}, its integration at the bacterial attachment site for phage ϕ C31 in the chromosome of the BCA strain, and the *Cre*-mediated excision of the *loxP*-flanked vector backbone. The picture (B) shows the phenotype of the control strain *BCA attB_{C31}::pTES* and strains expressing *rppA* reporter, *BCA attB_{C31}::pTES-rppA* and *BCA attB_{C31}::rppA*, after 4 days of culture, on LB agar, at 30 °C. The graph (C) represents the cell dry weight (left axis) and the volumetric production of flaviolin accumulated in the medium (right axis) of the control strain *BCA::pTES* and the strain harboring the pTES-rppA vector, *BCA attB_{C31}::pTES-rppA*. Table (D) shows the cell dry weight and the volumetric production of flaviolin accumulated in the medium, at 72 h of culture, without addition of antibiotics to the culture media, of the strain harboring the pTES-rppA vector strain, before (*BCA attB_{C31}::pTES-rppA*) and after (*BCA attB_{C31}::rppA*) the excision of the plasmid's backbone. The data points and error bars in graph (C), and the values and uncertainty in table (D) represent the average and the standard deviation of two experimental replicates, respectively.

expression of the native *rppA* gene from *Saccharopolyspora erythraea* was unable to sustain the production of flaviolin in *C. glutamicum*, even though transcription was detected (results not shown). Anticipating problems on the translational level, the *rppA* gene was adapted to the codon usage of *C. glutamicum* (*rppA_{Cgl}* Supp. data 7) and cloned into the replicative pEKEx2 vector for inducible expression. At 24 h of culture, the broth of the strain CR099attB_{C31} harboring pEKEx2-*rppA_{Cgl}* was orangish, suggesting the production of flaviolin while accumulating the same biomass as that of the control strain. The same behavior was observed in CR099attB_{C31} harboring the variant of the *rppA* gene adapted to the mixed codon usage of *C. glutamicum* and *Streptomyces* (*rppA_{CglStrept}* Supp. data 8) via pEKEx2-*rppA_{CglStrept}*. Noteworthy, both *rppA* variants have codon adaptation indexes (CAI) (Sharp and Li, 1987) over 0.9 and encode no rare codon in *C. glutamicum*. The *rppA_{CglStrept}* gene presents higher %GC content (70.8 vs 61.9), more biased codon usage (27 vs 35 different codons) and sustains 15% lower production of flaviolin (in replicative vector) than the *rppA_{Cgl}* gene. To confirm that the orange coloration was due to the production of flaviolin, the broth of the strain expressing *rppA_{Cgl}* was analyzed by tandem mass spectrometry (MS) coupled to liquid chromatography (Supp. methods, Supp. Fig. 5). In positive MS mode, the analysis of the orangish culture revealed a unique peak, eluted at retention time 2.5 min, corresponding to an MS analyte at 207.1 mass-to-charge ratio (m/z), and an MS² fragment at m/z 179, which were confidently assigned to the expected flaviolin molecule.

To study the viability of the ϕ C31 integration system to express heterologous genes in *C. glutamicum*, the *rppA_{CglStrept}* reporter was integrated into the genome of the BCA strain under the control of the Ptac promoter, using the pTES-*rppA* vector, originating the strain BCA attB_{C31}::pTES-*rppA* (Fig. 4A and B). The integration of pTES-*rppA* was confirmed by PCR and the frequency of pTES-*rppA* episomes per cell (Supp. methods) was estimated as less than 0.005% of a pBL1-based replicative vector in *C. glutamicum*. The supernatant of the culture of the strain BCA attB_{C31}::pTES-*rppA* showed increasing red pigmentation along the growth curve, representing a production of 39.5 mg L⁻¹ of flaviolin at 72 h of culture, without disturbing the biomass accumulation (Fig. 4C). To assess the effect of excising the backbone of pTES vector on production levels, the Cre recombinase was employed to generate the markerless strain BCA attB_{C31}::*rppA* (Fig. 4A and B). This strain maintained the ability to produce flaviolin, up to 9.6 mg L⁻¹ at 72 h of culture, although it was reduced to approximately one-quarter that of the strain BCA attB_{C31}::pTES-*rppA* (Fig. 4D). The variation of the flaviolin titer may be due the transcriptional readthrough of the promoter of the C31-Integrase, which would cause an overexpression of the *rppA_{CglStrept}* gene before the excision of the pTES-*rppA* backbone. This hypothesis is

supported by the observation that, when the *rppA_{CglStrept}* reporter, under the control of the Ptac promoter, was integrated using the ϕ BT1-based pRT802-*rppA* vector, the levels of production were similar to that of the strain BCA attB_{C31}::*rppA* (data not shown).

3.5. Production of the non-ribosomal peptide indigoidine in *C. glutamicum* using ϕ BT1-based integration

Indigoidine is an organic molecule that presents a unique combination of two conjugated pyridone rings functionalized with dicarboximide and enamine moieties, and is one of the few blue pigments occurring in Nature (Newsome et al., 2014). The indigoidine biosynthesis consist of the spontaneous condensation of two cyclized and oxidized derivatives of L-glutamine, formed by blue pigment synthetases of the IndC family of non-ribosomal peptide synthetases (NRPS) that require activation by phosphopantetheinylation (Reverchon et al., 2002).

To study the viability of the ϕ BT1 integration system to express heterologous genes in *C. glutamicum*, the cryptic IndC-like NRPS of *Streptomyces albidoflavus* (former *S. albus* J1074) encoded by *bpsA* gene (Olano et al., 2014) was expressed in the engineered strain BCA using a ϕ BT1-based integrative vector. For that purpose, the pFM69SD plasmid (Supp. data 9) with an insulated multicloning site, the P69 synthetic promoter for constitutive expression (Zhang et al., 2018) and an Shine-Dalgarno (SD) sequence was built upon the ϕ BT1 integration system of the pRT801/pRT802 vectors (Gregory et al., 2003). A preliminary attempt to produce indigoidine in *C. glutamicum* using the native *bpsA* gene (including native SD) from *S. albidoflavus* was unsuccessful, even though transcription had been detected (data not shown). To overcome eventual issues related to the beginning of translation, the thirty 5' nucleotides of *bpsA* were modified to correspond to the most prevalent codons in *C. glutamicum* ATCC 13032 (Nakamura et al., 2000). The modified *bpsA* gene (Supp. data 10) was cloned downstream the functional SD of the pFM69SD vector to yield the pFM69SD-*bpsA* construct and integrated at the attB_{BT1} locus of the BCA strain (Fig. 5A). The transformants of the strain BCA with pFM69SD-*bpsA*, dubbed BCA attB_{BT1}::pFM69SD-*bpsA*, developed a dark blue pigmentation from 24 h of culture in LB agar plates, that did not diffuse visibly to the agar even after 1 week (Fig. 5B). Upon cultivation in liquid CGXII defined media, with glucose as carbon source and ammonium sulfate and urea as nitrogen sources, the culture broth presented a blue pigmentation which mostly pelleted along with the biomass upon centrifugation. The blue pigment could not be extracted from the pellet with water nor with methanol, and N,N-dimethylformamide extracted it only to a small extent. In contrast, the pigment was readily extracted from the pellet with dimethyl sulfoxide (DMSO) and the extract

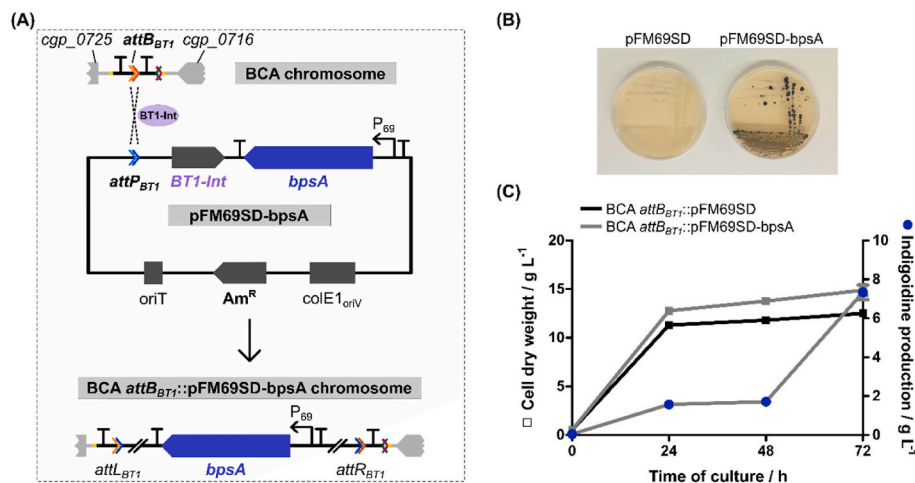


Fig. 5. Production of indigoidine in *C. glutamicum* using a ϕ BT1-based integration system. The panel (A) shows a schematic representation of the pFM69SD-*bpsA* vector employed for ϕ BT1-based integration of the indigoidine biosynthetic gene *bpsA*, and its integration at the bacterial attachment site for phage ϕ BT1 in the chromosome of the BCA strain. The picture (B) shows the phenotype of the control strain BCA attB_{BT1}::pFM69SD and the strain expressing *bpsA* gene, BCA attB_{BT1}::pFM69SD-*bpsA*, after 4 days of culture, on LB agar, at 30 °C. The graph (C) represents the cell dry weight (left axis) and the volumetric production of indigoidine extracted with dimethyl sulfoxide from the pellet of culture broth after centrifugation (right axis) of the control strain BCA attB_{BT1}::pFM69SD and the strain harboring the pFM69SD-*bpsA* vector, BCA attB_{BT1}::pFM69SD-*bpsA*. The data points and error bars in graph (C) represent the average and the standard deviation of two biological replicates of the strain BCA attB_{BT1}::pFM69SD-*bpsA*, respectively.

presented a clear UV–Vis peak at 614 nm. All the recorded physico-chemical properties were in accordance with the properties of indigoidine (Chatterjee and Brown, 1981; Kuhn et al., 1965) and that reasoned us to consider that the blue pigment was indigoidine. The monitoring of indigoidine production (Fig. 5C) showed that the greatest accumulation happened between 48 h and 72 h of cultivation. Quantification by spectrophotometry of the DMSO extracts, using the published extinction coefficient (Yumusak et al., 2019), indicated a volumetric production of indigoidine of $7.3 \pm 0.4 \text{ g L}^{-1}$, at 72 h of cultivation. The specific production of indigoidine at 72 h of culture was $0.492 \pm 0.006 \text{ g g}^{-1}$, which corresponded to about 50% of the cell dry weight. Noteworthy, the cell dry weight of the strain *BCA attB_{BT1}::pFM69SD-bpsA* was 119% that of the control strain *BCA attB_{BT1}::pFM69SD*. However, since the pigment precipitates along with the biomass, the cell dry weight includes the mass of indigoidine produced. Therefore, the biomass of cells of the strain *BCA attB_{BT1}::pFM69SD-bpsA* excluding indigoidine was about 70% that of the control strain *BCA attB_{BT1}::pFM69SD*. Considering the chemical formula of indigoidine, $\text{C}_{10}\text{H}_8\text{N}_4\text{O}_4$, the production of the blue pigment at 72 h of culture had theoretically absorbed 22% of the carbon atoms available from glucose and 25% of the nitrogen atoms available from ammonium sulfate and urea. Assuming glucose as the limiting reactant, the titer of indigoidine is in good agreement with the reduced biomass of cells of the strain *BCA attB_{BT1}::pFM69SD-bpsA*, excluding the mass of indigoidine.

4. Discussion

The success of synthetic biology-based approaches relies on the development of predictable and efficient chassis as well as stable and robust expression vector tools. (Beites and Mendes, 2015). The industrial workhorse *Corynebacterium glutamicum* has been engineered using top-down approaches towards a general chassis for industrial biotechnology. Here, we demonstrate the utility of a genomic library for chassis engineering by developing a *Corynebacterium*-based chassis for phage-based integration of heterologous genes. The development and characterization of compatible phage-based vectors resulted in the proof-of-principle production of polyketides and non-ribosomal peptides in the generated chassis.

The development of a chassis for synthetic biology application requires genetic editing platforms that encompass comprehensive ranges, traceable methods and targetable tools. The use of programmable endonucleases such as Cas9 and Cpf1 offers great flexibility in genome editing but substantially increases the risk of unwanted off-target effects [reviewed in (Kim et al., 2019)]. In our quest to enhance *C. glutamicum* capabilities as expression chassis, we generated a low-copy number cosmid library of the strain CR099attB_{C31} suitable for genome editing by double allelic exchange. The *C. glutamicum* strain CR099attB_{C31} is a derivative of the strain CR099 (Baumgart et al., 2018; Unthan et al., 2015) and harbors an attachment site for ϕ C31 phage-based DNA integration (attB_{C31}). Apart from the genome reductions, the regions represented in the generated genomic library were coincident with the much larger library of *C. glutamicum* ATCC 13032 prepared in multi-copy vectors (Tauch et al., 2002). Notably, the *relA* gene that regulates the stringent response was absent from both libraries. This result supports the idea that certain regions of the genome of *C. glutamicum* may be unclonable in *E. coli* even in low copy vectors, due to metabolic disruption. Our library includes all necessary features such as two bacterial selection markers for allelic exchange, a *loxP* site for Cre-mediated excision of genomic regions, and a eukaryotic selection marker for manipulation in yeast. In comparison with conventional plasmid-based allelic exchange platforms, such as pK19mobsacB (Schafer et al., 1994), the use of a cosmid library presents the following advantages: (i) a streamlined procedure to generate allelic exchange platforms, since it relies on pre-designed cassettes and recombineering, and avoids custom primer design and cloning; (ii) higher efficiency of chromosomal integration at first and second crossover steps, with lower

occurrence of reversion to parental genotype; (iii) an excisable selection marker, that simplifies phenotypic screening of second crossover events, and allows the selection of knockout mutations that may decrease the fitness of the strain; (iv) and a platform to perform point mutations in essential genes, as exemplified in (Lopatniuk et al., 2019). In summary, the generated library may be advantageous over conventional genome editing methods for the study of the physiology and for metabolic engineering in *C. glutamicum*.

The genomic library developed in this work was employed to further expand the capabilities of *C. glutamicum* as heterologous host, namely by introducing a site for ϕ BT1 phage-based DNA integration. We used the developed library to replace the 7.6-kb *crt* operon encoding the main carotenoid biosynthetic pathway in *C. glutamicum* by a markerless 0.3-kb region harboring the attachment site for ϕ BT1-based vectors (attB_{BT1}). The generated strain *C. glutamicum* BCA has an albino phenotype, due to the lack of carotenoids, and features the functional attB_{C31} and attB_{BT1} sites, for one-step orthogonal chromosomal integration of phage-based vectors. The high frequency of integration and the swiftness of introducing exogenous DNA using ϕ C31 and ϕ BT1 phage-based vectors can greatly impact on the user-friendliness of *C. glutamicum* BCA as routine heterologous host for biosynthetic pathways, especially for large gene clusters related to specialized metabolites. Furthermore, the lack of endogenous specialized metabolites of *C. glutamicum* BCA is also appealing for the sensitive detection and downstream purification of trace amounts of novel natural products. In conclusion, the development of the strain *C. glutamicum* BCA expands the capacities of the well-established strain ATCC 13032 with novel genetic gateways and turns it into a cutting-edge host for drug discovery.

The standardization of tools and techniques in the synthetic biology field intends transversal and predictable outcomes from genetic circuits from one organism to another. In this work, we aimed to extend the range of applications of tools already developed for other actinobacteria into *C. glutamicum*, specifically the prolific set of tools that has been developed for *Streptomyces* spp. We demonstrated that *C. glutamicum* strains harboring the appropriate attachment sites can integrate ϕ BT1 and ϕ C31-based vectors, originally developed for *Streptomyces* (Gregory et al., 2003; Herrmann et al., 2012), although with different efficiencies. While the ϕ C31-based system integrated at efficiencies one order of magnitude below those of the TP901-1 system (Shen et al., 2017), the ϕ BT1-based system became the most effective system of “non-native” phage-based integration for *C. glutamicum*, surpassing the best Cre-based integrase system (Suzuki et al., 2007) by one order of magnitude. In addition, systems based on Cre and C31-Integrase recombination systems, also originally designed for streptomycetes, have proven to be functional in *C. glutamicum*. The Cre expression vector pALCRE (Herrmann et al., 2012) was used for the excision of the plasmid backbone after integration, allowing markerless gene expression. On the other hand, the IMES technology (Myronovskiy et al., 2014) was used to generate the markerless strain *C. glutamicum* BCA. Importantly, the excision system can be used iteratively, so cumulative engineering stages can be performed using this method. This strategy can be used to knock-in features or genes into strains or even to deliver promoters, terminators and other short sequences. Overall, a set of molecular tools originally developed for *Streptomyces* were successfully transferred to *C. glutamicum*, bringing closer the prolific sources of novel biosynthetic gene clusters and the debutant heterologous host for specialized metabolites.

Integrative systems are desirable for stable inheritance of the introduced genetic traits. Rare events of instability in the modified strains can be avoided by only expressing the integrase function during introduction of the new traits (Petersen et al., 2013). The chromosomal integration of the ϕ BT1-based vectors (Gregory et al., 2003) caused the disruption of the promoter region controlling the expression of the vector BT1-Integrase. In accordance, a single rescue pattern was observed for the strains bearing the ϕ BT1-based vectors (Supp. Fig. 4). However, in the integration of the ϕ C31-based pTES vector, the

promoter region of the C31-Integrase encoding gene is not disrupted after integration. Even though it is described that the C31-Integrase requires an additional phage factor to recombine the *attL_{C31}* and *attR_{C31}* recombination scars (Khaleel et al., 2011), the patterns found by Southern blot and rescue plasmid experiments match exactly an eventual pTES episome excised by *attL_{C31}/attR_{C31}* recombination. However, the frequency of isolation of episomal pTES using a commercial miniprep kit was 0.005% that of a pBL1-based plasmid. Furthermore, owing to the *loxP* sites in pTES, the integrase function can be readily excised using Cre, rendering the genetic traits stable. In summary, the integrative vectors based on ϕ BT1 and ϕ C31 systems tested in this work can assure stable and inheritable integration of exogenous DNA without lasting selection, which is highly desirable for the development of industrial strains.

Flaviolin production was used as a proof-of-principle for the heterologous expression of specialized metabolites in *C. glutamicum* BCA host, namely polyketides, using the ϕ C31-based integration system. The production of flaviolin in *C. glutamicum* had been previously achieved (Yang et al., 2018), albeit only in replicative vectors. We succeeded on the production of flaviolin in *C. glutamicum* BCA host using the ϕ C31-based pTES-rppA platform carrying a codon-optimized version of the heterologous *rppA* gene of *S. erythraea*. The expression of *rppA* using the pTES-rppA vector sustained a flaviolin production of 40 mg L⁻¹, at 72 h of culture, significantly improving the published titer and making room for a more sensitive malonyl-CoA sensor (Yang et al., 2018). The expression of flaviolin using the pTES-rppA also served as a test for the Cre/*loxP* system, originally developed for *Streptomyces* (Herrmann et al., 2012) for the development of markerless heterologous expression platforms. Unexpectedly, upon excision of the backbone of pTES-rppA vector, the titer of flaviolin dropped considerably relatively to the previous situation. Interestingly, the titer of flaviolin obtained by the ϕ BT1-based integration of pRT802-rppA, carrying the same promoter and *rppA* variant, was similar to that after the excision of pTES-rppA backbone. From these results, we reasoned that the observed effect was likely caused by overexpression of *rppA* in the strain BCA *attB_{C31}*::pTES-rppA, due to the readthrough of the promoter of C31-Integrase. In spite of the presence of the *tfd* transcriptional terminator between *rppA* and the integrase gene, transcription can only be partially blocked by the intrinsic terminator, up to 85% in this particular case (Bilyk et al., 2017). To prevent the influence of *C31-Int* promoter in the pTES vector, it is, therefore, advisable to clone the genes of interest in the opposite direction. Of note, the highly efficient Cre/*loxP* system in combination with ϕ C31-based pTES vector (Herrmann et al., 2012) provide an universal platform for markerless heterologous expression in the *C. glutamicum* BCA host and an alternative to conventional non-replicative vectors, especially for delivery of large gene clusters or genes that decrease cell fitness. In summary, the pTES-rppA platform for production of flaviolin clearly demonstrated the effectiveness of the ϕ C31 phage-based system for heterologous expression in *C. glutamicum* BCA host and showcased the possibility to obtain markerless producers using site-specific recombinases.

Indigoidine production was used as a proof-of-principle for the heterologous expression of specialized metabolites in *C. glutamicum* BCA host, namely non-ribosomal peptides, using the ϕ BT1-based integration system. Indigoidine occurs naturally in many organisms, some of which phytopathogens (Reverchon et al., 2002), and several attempts to produce it in heterologous host have been reported. In *Streptomyces coelicolor* CH999, an indigoidine titer of 0.5935 g L⁻¹ was reported (Yu et al., 2013). After several rounds of optimization, the production of indigoidine in *E. coli* strains reached 8.8 g L⁻¹ (Xu et al., 2015; Yu et al., 2013). The fungi that served as hosts for indigoidine production were *Saccharomyces cerevisiae* and *Rhodospiridium toruloides*, and titers in these hosts ranged from 0.210 g L⁻¹ up to 86 g L⁻¹ in glucose-based fed-batch cultivations (Wehrs et al., 2018, 2019a, 2019b). Taking advantage of well-established capacity of *C. glutamicum* to produce compounds based on amino acids, we sought to redirect L-glutamine to the production of

the blue dye indigoidine. We succeeded on the production of indigoidine in *C. glutamicum* BCA host using the developed ϕ BT1-based pFM69SD-*bpsA* platform carrying the a slightly modified version of the heterologous *bpsA* gene of *Streptomyces albidoflavus*. Notably, the titer of indigoidine produced by *C. glutamicum* BCA strain harboring a single copy *bpsA* gene in standard cultivation conditions is comparable to that of the above described heterologous hosts. Importantly, the production of indigoidine in *C. glutamicum* confirmed that native PPTases are able to activate non-ribosomal peptide synthases *in vivo* (Kallscheuer et al., 2019).

The phage-based integrative vectors and the proof-of-principle biosynthetic enzymes employed in this work relied heavily on GC-rich heterologous genes, which imposed significant challenges to the translation machinery of *C. glutamicum*. Our attempt to express the *rppA* and *bpsA* genes with native codon usage and 5' UTRs resulted in no production of flaviolin nor indigoidine, respectively, although transcription was detected in both cases (data not shown). Even though *rppA* gene of *S. erythraea* presented a global CAI of 0.8, it contained fifteen rare codons and started with the GTG codon, a known factor for diminished translation in *C. glutamicum* (Becker et al., 2010). Thus, we developed *rppA*-based reporters for usage in *C. glutamicum* by employing variants of the *rppA* gene from *S. erythraea* optimized according to the coding preferences of *C. glutamicum* and/or *Streptomyces*. The integration of the codon-optimized *rppA_{CglStrept}* in *C. glutamicum* BCA led to the production of flaviolin at significantly better yields than those previously reported for the expression of *rppA* from *S. griseus* modified with an N-terminal hexa-histidine tag in *C. glutamicum* ATCC 13032 (Yang et al., 2018), even though the copy number was about 30-fold higher in the latter. Our results using two differently codon-optimized *rppA* genes (*rppA_{Cgl}* and *rppA_{CglStrept}*) also suggests that a codon usage biased to *Streptomyces* and with higher GC content may reduce translation in *C. glutamicum*. However, the major barrier to heterologous expression seems to reside in improper 5' UTRs and unfit starting codons. For instance, a simple upgrade of SD and of the first ten codons (old CAI: 0.76; new CAI: 0.96) was enough to convert the inert *bpsA* gene (overall CAI: 0.76) into a gene able to sustain high titer of indigoidine production in *C. glutamicum*. Likewise, the lower frequency of integration of the ϕ C31 phage-based vectors when comparing to the ϕ BT1 phage-based vectors could also be explained by an impairment on the beginning of translation of the respective integrases. In fact, despite the similar overall CAI of both genes (0.7), the adaptation index of the first ten codon of the *C31-Int* (0.5) is lower than in *BT1-Int* (0.7), and the former includes two rare codons. These examples showcase the challenges that can arise when expressing codon-biased, GC-rich genes from *Streptomyces* in the *C. glutamicum* host and pinpoint the 5'-UTR and the initial codons as interesting targets for metabolic engineering.

Overall, we developed a comprehensive platform for efficient genome editing of *C. glutamicum* based on a cosmid library. Using this platform, we generated the *C. glutamicum* BCA host, featuring the attachment sites for ϕ C31 and ϕ BT1 phage-based vectors. We confirmed the efficacy of the iterative marker excision system, the ϕ C31 and ϕ BT1 phage-based vectors and other tools originally developed for other actinobacteria. We improved the output of a described malonyl-CoA sensor derived from the polyketide synthase *RppA* and first demonstrated that the phosphopantetheinylation-dependent biosynthesis of non-ribosomal peptides is feasible in *C. glutamicum*. The microbial chassis and molecular toolbox characterized in this work are prepared to boost drug discovery efforts with forefront capabilities for heterologous expression and effective platforms for further metabolic engineering.

Declarations of interest

None.

CRedit authorship contribution statement

Filipe Marques: Investigation, Data curation, Methodology, Visualization, Writing - original draft. **Andriy Luzhetskyy:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing - review & editing. **Marta V. Mendes:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Supervision, Writing - review & editing.

Acknowledgements

The authors are grateful to Nikolas Eckert and Liliya Horbal for the help on producing the cosmid library, and to Anja Paluszczak for the help on HPLC-MS analysis. The authors would like to acknowledge Jörn Kalinowski and Christian Rückert for providing *C. glutamicum* strain CR099attB_{C31}, Giacomo Giacomelli for providing the pK19mobsacB and the pEKEx2 vectors, Maria Lopatniuk for providing cos15A-gus vector, and Birgit Rosenkränzer and Maksym Myronovskiy for providing pTES, pALCRE, pSMART-3B20 and the IMES plasmids.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jmben.2020.08.007>.

Funding

This work was supported by National Funds through FCT – Fundação para a Ciência e a Tecnologia, I.P. under the project ERA-IB-2/0001/2015; fellowships SFRH/BD/131939/2016 (FM) and SFRH/BPD/95683/2013 (MVM) and the FCT contract DL57/2016/CP1355/CT0023. The work was also supported by FEDER - Fundo Europeu de Desenvolvimento Regional through the COMPETE 2020 - Operacional Programme for Competitiveness and Internationalisation (POCI), Portugal 2020, and by Portuguese funds through FCT - Fundação para a Ciência e a Tecnologia/Ministério da Ciência, Tecnologia e Ensino Superior, and “Institute for Research and Innovation in Health Sciences” (POCI-01-0145-FEDER-007274), by the project NORTE-01-0145-FEDER-000012, Structured Programme on Bioengineering Therapies for Infectious Diseases and Tissue Regeneration, supported by Norte Portugal Regional Operational Programme (NORTE, 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF). This work was also supported by the DFG grant LU1524/4-1. The funders of the study had no role in study design, data collection, data analysis, data interpretation, writing of the report or decision to submit the article for publication.

References

- Amador, E., Martin, J.F., Castro, J.M., 2000. A *Brevibacterium lactofermentum* 16S rRNA gene used as target site for homologous recombination. *FEMS Microbiol. Lett.* 185, 199–204.
- Bathe, B., Kalinowski, J., Puhler, A., 1996. A physical and genetic map of the *Corynebacterium glutamicum* ATCC 13032 chromosome. *Mol. Gen. Genet.* 252, 255–265.
- Baumgart, M., Unthan, S., Kloss, R., Radek, A., Polen, T., Tenhaef, N., Müller, M.F., Kuberl, A., Siebert, D., Bruhl, N., Marin, K., Hans, S., Kramer, R., Bott, M., Kalinowski, J., Wiechert, W., Seibold, G., Frunzke, J., Rückert, C., Wendisch, V.F., Noack, S., 2018. *Corynebacterium glutamicum* chassis C1*: building and testing a novel platform host for synthetic biology and industrial biotechnology. *ACS Synth. Biol.* 7, 132–144.
- Baumgart, M., Unthan, S., Rückert, C., Sivalingam, J., Grunberger, A., Kalinowski, J., Bott, M., Noack, S., Frunzke, J., 2013. Construction of a prophage-free variant of *Corynebacterium glutamicum* ATCC 13032 for use as a platform strain for basic research and industrial biotechnology. *Appl. Environ. Microbiol.* 79, 6006–6015.
- Becker, J., Buschke, N., Bückner, R., Wittmann, C., 2010. Systems level engineering of *Corynebacterium glutamicum* - reprogramming translational efficiency for superior production. *Eng. Life Sci.* 10, 430–438.
- Becker, J., Rohles, C.M., Wittmann, C., 2018a. Metabolically engineered *Corynebacterium glutamicum* for bio-based production of chemicals, fuels, materials, and healthcare products. *Metab. Eng.* 50, 122–141.
- Becker, J., Rohles, C.M., Wittmann, C., 2018b. Metabolically engineered *Corynebacterium glutamicum* for bio-based production of chemicals, fuels, materials, and healthcare products. *Metab. Eng.* 50, 122–141.
- Beites, T., Mendes, M.V., 2015. Chassis optimization as a cornerstone for the application of synthetic biology based strategies in microbial secondary metabolism. *Front. Microbiol.* 6, 906.
- Bilyk, B., Horbal, L., Luzhetskyy, A., 2017. Chromosomal position effect influences the heterologous expression of genes and biosynthetic gene clusters in *Streptomyces albus* J1074. *Microb. Cell Factories* 16, 5.
- Cameron Coates, R., Blaskowski, S., Szyjka, S., van Rossum, H.M., Vallandingham, J., Patel, K., Serber, Z., Dean, J., 2019. Systematic investigation of CRISPR-Cas9 configurations for flexible and efficient genome editing in *Corynebacterium glutamicum* NRRL-B11474. *J. Ind. Microbiol. Biotechnol.* 46, 187–201.
- Chatterjee, A.K., Brown, M.A., 1981. Chromosomal location of a gene (idg) that specifies production of the blue pigment indigoidine in *Erwinia chrysanthemi*. *Curr. Microbiol.* 6, 269–273.
- Cho, J.S., Choi, K.R., Prabowo, C.P.S., Shin, J.H., Yang, D., Jang, J., Lee, S.Y., 2017. CRISPR/Cas9-coupled recombineering for metabolic engineering of *Corynebacterium glutamicum*. *Metab. Eng.* 42, 157–167.
- Choi, J.W., Yim, S.S., Kim, M.J., Jeong, K.J., 2015. Enhanced production of recombinant proteins with *Corynebacterium glutamicum* by deletion of insertion sequences (IS elements). *Microb. Cell Factories* 14, 207.
- Correia, A., Martin, J.F., Castro, J.M., 1996. Targeted integration of foreign genes into repetitive sequences of the *Brevibacterium lactofermentum* chromosome. *FEMS Microbiol. Lett.* 142, 259–264.
- Cortes, J., Velasco, J., Foster, G., Blackaby, A.P., Rudd, B.A., Wilkinson, B., 2002. Identification and cloning of a type III polyketide synthase required for diffusible pigment biosynthesis in *Saccharopolyspora erythraea*. *Mol. Microbiol.* 44, 1213–1224.
- Datsenko, K.A., Wanner, B.L., 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6640–6645.
- Eggeling, L., Bott, M., 2005. *Handbook of Corynebacterium Glutamicum*. Taylor & Francis, Boca Raton.
- Eikmanns, B.J., Kleinertz, E., Liebl, W., Sahm, H., 1991. A family of *Corynebacterium glutamicum*/*Escherichia coli* shuttle vectors for cloning, controlled gene expression, and promoter probing. *Gene* 102, 93–98.
- Fu, J., Teucher, M., Anastasiadis, K., Skarnes, W., Stewart, A.F., 2010. A recombineering pipeline to make conditional targeting constructs. In: *Methods in Enzymology*, Vol 477: Guide to Techniques in Mouse Development, Part B: Mouse Molecular Genetics, second ed., vol. 477, pp. 125–144.
- Gregory, M.A., Till, R., Smith, M.C., 2003. Integration site for *Streptomyces* phage phiBT1 and development of site-specific integrating vectors. *J. Bacteriol.* 185, 5320–5323.
- Hanahan, D., 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166, 557–580.
- Herrmann, S., Siegl, T., Luzhetskyy, M., Petzke, L., Jilg, C., Welle, E., Erb, A., Leadlay, P.F., Bechthold, A., Luzhetskyy, A., 2012. Site-specific recombination strategies for engineering actinomycete genomes. *Appl. Environ. Microbiol.* 78, 1804–1812.
- Hoffmann, S.L., Jungmann, L., Schiefelbein, S., Peyriga, L., Cahoreau, E., Portais, J.-C., Becker, J., Wittmann, C., 2018. Lysine production from the sugar alcohol mannitol: design of the cell factory *Corynebacterium glutamicum* SEA-3 through integrated analysis and engineering of metabolic pathway fluxes. *Metab. Eng.* 47, 475–487.
- Horbal, L., Siegl, T., Luzhetskyy, A., 2018. A set of synthetic versatile genetic control elements for the efficient expression of genes in *Actinobacteria*. *Sci. Rep.* 8, 491.
- Ikedo, M., Nakagawa, S., 2003. The *Corynebacterium glutamicum* genome: features and impacts on biotechnological processes. *Appl. Microbiol. Biotechnol.* 62, 99–109.
- Jiang, Y., Qian, F., Yang, J., Liu, Y., Dong, F., Xu, C., Sun, B., Chen, B., Xu, X., Li, Y., Wang, R., Yang, S., 2017. CRISPR-Cpf1 assisted genome editing of *Corynebacterium glutamicum*. *Nat. Commun.* 8, 15179.
- Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., Dusch, N., Eggeling, L., Eikmanns, B.J., Gaigalat, L., Goesmann, A., Hartmann, M., Huthmacher, K., Kramer, R., Linke, B., McHardy, A.C., Meyer, F., Mockel, B., Pfefferle, W., Puhler, A., Rey, D.A., Rückert, C., Rupp, O., Sahm, H., Wendisch, V.F., Wiegrabe, I., Tauch, A., 2003. The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. *J. Biotechnol.* 104, 5–25.
- Kallscheuer, N., Kage, H., Milke, L., Nett, M., Marienhagen, J., 2019. Microbial synthesis of the type I polyketide 6-methylsalicylate with *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 103, 9619–9631.
- Keilhauer, C., Eggeling, L., Sahm, H., 1993. Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB-ilvN-ilvC* operon. *J. Bacteriol.* 175, 5595–5603.
- Khaleel, T., Younger, E., McEwan, A.R., Varghese, A.S., Smith, M.C., 2011. A phage protein that binds ϕ C31 integrase to switch its directionality. *Mol. Microbiol.* 80, 1450–1463.
- Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., Hopwood, D.A., John Innes, F., 2000. *Practical Streptomyces Genetics*. John Innes Foundation, Norwich.
- Kim, D., Luk, K., Wolfe, S.A., Kim, J.S., 2019. Evaluating and enhancing target specificity of gene-editing nucleases and deaminases. *Annu. Rev. Biochem.* 88, 191–220.
- Kim, I.K., Jeong, W.K., Lim, S.H., Hwang, I.K., Kim, Y.H., 2011. The small ribosomal protein S12P gene *rpsL* as an efficient positive selection marker in allelic exchange mutation systems for *Corynebacterium glutamicum*. *J. Microbiol. Methods* 84, 128–130.

- Kimelman, A., Levy, A., Sberro, H., Kidron, S., Leavitt, A., Amitai, G., Yoder-Himes, D.R., Wurtzel, O., Zhu, Y., Rubin, E.M., Sorek, R., 2012. A vast collection of microbial genes that are toxic to bacteria. *Genome Res.* 22, 802–809.
- Kinoshita, S., Udaka, S., Shimono, M., 1957. Studies on the amino acid fermentation. *J. Gen. Appl. Microbiol.* 3, 193–205.
- Kjeldsen, K.R., Nielsen, J., 2009. In silico genome-scale reconstruction and validation of the *Corynebacterium glutamicum* metabolic network. *Biotechnol. Bioeng.* 102, 583–597.
- Kogure, T., Inui, M., 2018. Recent advances in metabolic engineering of *Corynebacterium glutamicum* for bioproduction of value-added aromatic chemicals and natural products. *Appl. Microbiol. Biotechnol.* 102, 8685–8705.
- Krauser, S., Kiefer, P., Heinzle, E., 2012. Multienzyme whole-cell *InSitu* biocatalysis for the production of flavolin in permeabilized cells of *Escherichia coli*. *ChemCatChem* 4, 786–788.
- Kuhn, R., Starr, M.P., Kuhn, D.A., Bauer, H., Knackmuss, H.J., 1965. Indigoidine and other bacterial pigments related to 3,3'-bipyridyl. *Arch. Mikrobiol.* 51, 71–84.
- Lennox, E.S., 1955. Transduction of linked genetic characters of the host by bacteriophage P1. *Virology* 1, 190–206.
- Liu, J., Wang, Y., Lu, Y., Zheng, P., Sun, J., Ma, Y., 2017. Development of a CRISPR/Cas9 genome editing toolbox for *Corynebacterium glutamicum*. *Microb. Cell Factories* 16, 205.
- Lopatniuk, M., Myronovskiy, M., Nottebrock, A., Busche, T., Kalinowski, J., Ostash, B., Fedorenko, V., Luzhetskyy, A., 2019. Effect of “ribosome engineering” on the transcription level and production of *S. albus* indigenous secondary metabolites. *Appl. Microbiol. Biotechnol.* 103, 7097–7110.
- Ma, W., Wang, X., Mao, Y., Wang, Z., Chen, T., Zhao, X., 2015. Development of a markerless gene replacement system in *Corynebacterium glutamicum* using up as a counter-selection marker. *Biotechnol. Lett.* 37, 609–617.
- Marrec, C.L., Michotey, V., Blanco, C., Trautwetter, A., 1994. aU2, a temperate bacteriophage specific for ‘*Arthrobacter aureus*’, whose integrative functions work in other corynebacteria. *Microbiology* 140, 3071–3077.
- Marx, A., de Graaf, A.A., Wiechert, W., Eggeling, L., Sahm, H., 1996. Determination of the fluxes in the central metabolism of *Corynebacterium glutamicum* by nuclear magnetic resonance spectroscopy combined with metabolite balancing. *Biotechnol. Bioeng.* 49, 111–129.
- Moreau, S., Blanco, C., Trautwetter, A., 1999a. Site-specific integration of coryneophage phi16: construction of an integration vector. *Microbiology* 145 (Pt 3), 539–548.
- Moreau, S., Le Marrec, C., Blanco, C., Trautwetter, A., 1999b. Analysis of the integration functions of phi304L: an integrase module among coryneophages. *Virology* 255, 150–159.
- Myronovskiy, M., Rosenkranz, B., Luzhetskyy, A., 2014. Iterative marker excision system. *Appl. Microbiol. Biotechnol.* 98, 4557–4570.
- Nakamura, Y., Gojobori, T., Ikemura, T., 2000. Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* 28, 292.
- Newsome, A.G., Culver, C.A., van Breemen, R.B., 2014. Nature's palette: the search for natural blue colorants. *J. Agric. Food Chem.* 62, 6498–6511.
- Olano, C., Garcia, I., Gonzalez, A., Rodriguez, M., Rozas, D., Rubio, J., Sanchez-Hidalgo, M., Brana, A.F., Mendez, C., Salas, J.A., 2014. Activation and identification of five clusters for secondary metabolites in *Streptomyces albus* J1074. *Microbial biotechnol.* 7, 242–256.
- Oram, M., Woolston, J.E., Jacobson, A.D., Holmes, R.K., Oram, D.M., 2007. Bacteriophage-based vectors for site-specific insertion of DNA in the chromosome of *Corynebacteria*. *Gene* 391, 53–62.
- Peng, F., Liu, X., Wang, X., Chen, J., Liu, M., Yang, Y., Bai, Z., 2019. Triple deletion of *clpC*, *porB*, and *mepA* enhances production of small ubiquitin-like modifier-N-terminal pro-brain natriuretic peptide in *Corynebacterium glutamicum*. *J. Ind. Microbiol. Biotechnol.* 46, 67–79.
- Peng, F., Wang, X., Sun, Y., Dong, G., Yang, Y., Liu, X., Bai, Z., 2017. Efficient gene editing in *Corynebacterium glutamicum* using the CRISPR/Cas9 system. *Microb. Cell Factories* 16, 201.
- Petersen, K.V., Martinussen, J., Jensen, P.R., Solem, C., 2013. Repetitive, marker-free, site-specific integration as a novel tool for multiple chromosomal integration of DNA. *Appl. Environ. Microbiol.* 79, 3563–3569.
- Pfeifer-Sancar, K., Mentz, A., Rückert, C., Kalinowski, J., 2013. Comprehensive analysis of the *Corynebacterium glutamicum* transcriptome using an improved RNAseq technique. *BMC Genom.* 14, 888.
- Rebets, Y., Tsois, K.C., Guethmundsdottir, E.E., Koepff, J., Wawiernia, B., Busche, T., Bleidt, A., Horbal, L., Myronovskiy, M., Ahmed, Y., Wiechert, W., Rückert, C., Hamed, M.B., Bilyk, B., Anne, J., Friethjonnsson, O., Kalinowski, J., Oldiges, M., Economou, A., Luzhetskyy, A., 2018. Characterization of sigma factor genes in *Streptomyces lividans* TK24 using a genomic library-based approach for multiple gene deletions. *Front. Microbiol.* 9, 3033.
- Reverchon, S., Rouanet, C., Expert, D., Nasser, W., 2002. Characterization of indigoidine biosynthetic genes in *Erwinia chrysanthemi* and role of this blue pigment in pathogenicity. *J. Bacteriol.* 184, 654–665.
- Sambrook, J., Russell, D.W., 2001. *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Schafer, A., Tauch, A., Jager, W., Kalinowski, J., Thierbach, G., Puhler, A., 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* 145, 69–73.
- Sharp, P.M., Li, W.-H., 1987. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* 15, 1281–1295.
- Shen, J., Chen, J., Jensen, P.R., Solem, C., 2017. A novel genetic tool for metabolic optimization of *Corynebacterium glutamicum*: efficient and repetitive chromosomal integration of synthetic promoter-driven expression libraries. *Appl. Microbiol. Biotechnol.* 101, 4737–4746.
- Shinfuku, Y., Sorpitiporn, N., Sono, M., Furusawa, C., Hirasawa, T., Shimizu, H., 2009. Development and experimental verification of a genome-scale metabolic model for *Corynebacterium glutamicum*. *Microb. Cell Factories* 8, 43.
- Siegl, T., Tokovenko, B., Myronovskiy, M., Luzhetskyy, A., 2013. Design, construction and characterisation of a synthetic promoter library for fine-tuned gene expression in actinomycetes. *Metab. Eng.* 19, 98–106.
- Sun, H., Zhao, D., Xiong, B., Zhang, C., Bi, C., 2016. Engineering *Corynebacterium glutamicum* for violacein hyper production. *Microb. Cell Factories* 15, 148.
- Suzuki, N., Inui, M., Yukawa, H., 2007. Site-directed integration system using a combination of mutant *lox* sites for *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* 77, 871–878.
- Tauch, A., Homann, I., Mormann, S., Rüberg, S., Billault, A., Bathe, B., Brand, S., Brockmann-Gretz, O., Rückert, C., Schischka, N., Wrenger, C., Hoheisel, J., Möckel, B., Huthmacher, K., Pfefferle, W., Pühler, A., Kalinowski, J., 2002. Strategy to sequence the genome of *Corynebacterium glutamicum* ATCC 13032: use of a cosmid and a bacterial artificial chromosome library. *J. Biotechnol.* 95, 25–38.
- Unthan, S., Baumgart, M., Radek, A., Herbst, M., Siebert, D., Bruhl, N., Bartsch, A., Bott, M., Wiechert, W., Marin, K., Hans, S., Kramer, R., Seibold, G., Frunzke, J., Kalinowski, J., Rückert, C., Wendisch, V.F., Noack, S., 2015. Chassis organism from *Corynebacterium glutamicum*—a top-down approach to identify and delete irrelevant gene clusters. *Biotechnol. J.* 10, 290–301.
- Vasicová, P., Abrahámová, Z., Nesvera, J., Pátek, M., Sahm, H., Eikmanns, B., 1998. Integrative and autonomously replicating vectors for analysis of promoters in *Corynebacterium glutamicum*. *Biotechnol. Tech.* 12, 743–746.
- Vertes, A.A., Inui, M., Yukawa, H., 2013. The Biotechnological Potential of *Corynebacterium glutamicum*, from Umami to Chemurgy. *Corynebacterium glutamicum*. Springer, pp. 1–49.
- Vickers, C.E., Blank, L.M., Krömer, J.O., 2010. Grand challenge commentary: chassis cells for industrial biochemical production. *Nat. Chem. Biol.* 6, 875–877.
- Wang, B., Hu, Q., Zhang, Y., Shi, R., Chai, X., Liu, Z., Shang, X., Zhang, Y., Wen, T., 2018a. A RecET-assisted CRISPR-Cas9 genome editing in *Corynebacterium glutamicum*. *Microb. Cell Factories* 17, 63.
- Wang, Y., Liu, Y., Li, J., Yang, Y., Ni, X., Cheng, H., Huang, T., Guo, Y., Ma, H., Zheng, P., Wang, M., Sun, J., Ma, Y., 2019. Expanding targeting scope, editing window, and base transition capability of base editing in *Corynebacterium glutamicum*. *Biotechnol. Bioeng.* 116, 3016–3029.
- Wang, Y., Liu, Y., Liu, J., Guo, Y., Fan, L., Ni, X., Zheng, X., Wang, M., Zheng, P., Sun, J., Ma, Y., 2018b. MACBETH: multiplex automated *Corynebacterium glutamicum* base editing method. *Metab. Eng.* 47, 200–210.
- Wehrs, M., Gladden, J.M., Liu, Y., Platz, L., Pahl, J.-P., Moon, J., Papa, G., Sundstrom, E., Geiselman, G.M., Tanjore, D., Keasling, J.D., Pray, T.R., Simmons, B. A., Mukhopadhyay, A., 2019a. Sustainable bioproduction of the blue pigment indigoidine: expanding the range of heterologous products in *R. toruloides* to include non-ribosomal peptides. *Green Chem.* 21, 3394–3406.
- Wehrs, M., Pahl, J.P., Moon, J., Li, Y., Tanjore, D., Keasling, J.D., Pray, T., Mukhopadhyay, A., 2018. Production efficiency of the bacterial non-ribosomal peptide indigoidine relies on the respiratory metabolic state in *S. cerevisiae*. *Microb. Cell Factories* 17, 193.
- Wehrs, M., Pahl, J.P., Moon, J., Li, Y., Tanjore, D., Keasling, J.D., Pray, T., Mukhopadhyay, A., 2019b. Correction to: production efficiency of the bacterial non-ribosomal peptide indigoidine relies on the respiratory metabolic state in *S. cerevisiae*. *Microb. Cell Factories* 18, 218.
- Xu, F., Gage, D., Zhan, J., 2015. Efficient production of indigoidine in *Escherichia coli*. *J. Ind. Microbiol. Biotechnol.* 42, 1149–1155.
- Yang, D., Kim, W.J., Yoo, S.M., Choi, J.H., Ha, S.H., Lee, M.H., Lee, S.Y., 2018. Repurposing type III polyketide synthase as a malonyl-CoA biosensor for metabolic engineering in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* 115, 9835–9844.
- Yanisch-Perron, C., Vieira, J., Messing, J., 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33, 103–119.
- Yu, D., Xu, F., Valiente, J., Wang, S., Zhan, J., 2013. An indigoidine biosynthetic gene cluster from *Streptomyces chromofuscus* ATCC 49982 contains an unusual *IndB* homologue. *J. Ind. Microbiol. Biotechnol.* 40, 159–168.
- Yumusak, C., Prochazkova, A.J., Apaydin, D.H., Seelajaroen, H., Sariciftci, N.S., Weiter, M., Krajcovic, J., Qin, Y., Zhang, W., Zhang, J., Kovalenko, A., 2019. Indigoidine – biosynthesized organic semiconductor. *Dyes Pigments* 171, 107768.
- Zhang, J., Yang, F., Yang, Y., Jiang, Y., Huo, Y.X., 2019. Optimizing a CRISPR-Cpf1-based genome engineering system for *Corynebacterium glutamicum*. *Microb. Cell Factories* 18, 60.
- Zhang, S., Liu, D., Mao, Z., Mao, Y., Ma, H., Chen, T., Zhao, X., Wang, Z., 2018. Model-based reconstruction of synthetic promoter library in *Corynebacterium glutamicum*. *Biotechnol. Lett.* 40, 819–827.
- Zhang, Y., Cai, J., Shang, X., Wang, B., Liu, S., Chai, X., Tan, T., Zhang, Y., Wen, T., 2017. A new genome-scale metabolic model of *Corynebacterium glutamicum* and its application. *Biotechnol. Biofuels* 10, 169.

SUPPLEMENTARY INFORMATION I

SUPPLEMENTARY INFORMATION I

Material & Methods

Plasmid rescue analysis

To identify the flanking regions of the bacterial attachment sites upon transformation with integrative vectors, isolation of rescue plasmids was performed as described before (Bilyk, Horbal et al. 2017). Briefly, chromosomal DNA of strains transformed with pTES or pRT802 was isolated and digested with the endonucleases *Bam*HI or *Eco*RV, respectively. The resulting DNA fragments were purified by precipitation, self-ligated using T4 DNA ligase and transformed into DH5 α cells. The transformants containing rescue plasmids were selected by the respective resistance marker. The obtained plasmids were mapped by restriction and sequenced using the primer pair pTESmcs_fwd/rvs or the M13R universal primer, for pTES- and pRT802-derived vectors, respectively.

Southern blot analysis

One microgram of the genomic DNA of *C. glutamicum* CR099attB_{C31} transformed with pTES was digested with the endonucleases *Bam*HI, *Sca*I, *Sph*I or *Bgl*II. Standard protocols were used for separation and transfer of the DNA fragments onto an Hybond-N membrane (Amersham). The DNA probe was obtained by PCR using pTES vector and the primer pair aac_south410_fwd (5'-ACACCGGACCTTGGAGTTG-3') and aac_south410_rvs (5'-AAGGCATGACCGACTGGAC-3'). The DNA probe was labelled with the DIG High Prime DNA Labelling and Detection Starter Kit II (Roche) and hybridization was performed at 65 °C. Immunological detection was carried out by chemiluminescence as result of alkaline phosphatase activity in the presence of CSPD reagent (Roche), and recorded in a ChemiDoc XRS+ Imager (Bio-Rad) using an exposition period of 7 seconds.

Episome isolation and detection

Single colonies of *C. glutamicum* harboring pEKEx2 (two clones), pTES (five clones) and pTES-rppA (five clones) were cultivated for 16 h, in 5 mL LB, 200 rpm, at 30 °C, with the appropriate antibiotics. The equivalent to 1 mL of each culture at OD600 of 8 was centrifuged at 16100g for 5 min, at room temperature. The cells were resuspended in 250 μ L of A1 buffer (NZYMiniprep kit) with 15 mg/mL lysozyme and incubated at 37°C, with shaking, for 3.5 h. The remaining protocol for isolation of plasmid DNA proceeded according to manufacturer's instructions and elution was done in 50 μ L of distilled water. One hundred microliters of competent DH5 α cells prepared by chemical methods were transformed with 5 μ L of each of the isolated DNAs by heat-shock, added to 900 μ L of LB medium and allowed to recover at 37 °C, with shaking, for 1 h. Fifty microliters of each cell suspension

were plated on LB agar with appropriate antibiotics and incubated at 37 °C, for 16 h. The relative frequency of episomes was calculated as the ratio of average number of colony-forming units (CFU) using DNA isolated from strains harboring integrative vectors to the average CFUs from DNA isolated from strains harboring pEKEx2 plasmid.

Analysis of unclonable genes

For discovery of genes unclonable in *E. coli*, the annotated open reading frames of MB001 strain not represented in the generated genomic library were submitted for analysis against the PanDaTox gene collection (Kimelman, Levy et al. 2012), considering E values less than 10^{-8} . Query sequences with at least one unclonable homolog (either in plasmids or fosmids) were considered positive candidates for toxic effects in *E. coli*. The analysis of the metabolic pathways and functions of the selected genes was performed by submitting the list of the annotated genes of interest to BLASTKoala (Kanehisa, Sato et al. 2016), using the taxonomy group of prokaryotes and the KEGG database species_prokaryotes.

Design and application of a plasmid-based platform for editing of the *crt* locus

To design a platform to edit the *crt* locus based on pK19mobsacB plasmid (Schafer, Tauch et al. 1994), for comparative purposes, the template for homologous recombination was amplified from the pCG_4_F11_Δ*crt*::*attB_{BT1}*-*saac*(3')/IV cosmid using the primers *crt*_pK19_fwd (5'-CTCGCATATCAGGGGCAC-3') and *crt*_pK19_rvs (5'-CATGTGTGGGAGGCTTCG-3'). The amplified region contained 432-bp left and 422-bp right homology arms flanking the *crt* locus and the same cassette than the template cosmid for marker excision procedure. The **pK19mobsacB-crt** plasmid for homologous recombination was generated by cloning the amplified DNA fragment at the *Sma*I site of pK19mobsacB plasmid. The transformation of *C. glutamicum* with the pK19mobsacB-crt plasmid, the selection of transformants, and the obtention and phenotypic screening of second crossover events were performed in parallel and using the same methods of pCG_4_F11_Δ*crt*::*attB_{BT1}*-*saac*(3')/IV cosmid.

Comparative analysis of genomic sequences

The alignment of the whole genomes of *C. glutamicum* MB001 (accession CP005959.1) and the generated *C. glutamicum* BCA strains was performed using the progressive Mauve algorithm (Darling, Mau et al. 2004) plugin for the Geneious Prime software (Biomatters). The following parameters were used: full alignment, automatic seed weight, computed Locally Collinear Blocks (LCB), automatic LCB score.

Carotenoid extraction and UV-Vis quantification

For production of carotenoids, *C. glutamicum* strains were cultivated as described for flavin production, up to 24 h. The assessment of carotenoid production was performed by UV-Vis analysis, according to the protocol described by Heider *et al.* (Heider, Peters-Wendisch *et al.* 2012). Briefly, 20 mL of culture broth were transferred to a 50-mL tube and centrifuged at 4000g, for 15 min, at room temperature. The medium was discarded, and the biomass was washed with 1 volume of ddH₂O, then centrifuged at 4000g, for 15 min, at room temperature. The supernatant was discarded, and the washed biomass was resuspended in 20 mL of a 7:3 (V/V) mixture of methanol and acetone, vortexed for 1 min, and incubated at 80 °C, for 80 min. The extraction was vortexed for 1 min, every 20 mins during the extraction. The phases were separated by centrifuging at 4500g, for 15 min, at room temperature. Two 1.8-mL samples from the organic phase were collected into 2-mL tubes. The absorbance spectra (400-800 nm) was acquired in 1-cm path Quartz Glass High Performance cuvettes (Hellma), in an UV2401-PC recording spectrophotometer (Shimadzu), using 7:3 (v/v) mixture of methanol and acetone as reference. The spectrophotometer was controlled from a computer using the UV-Probe software (Shimadzu), and using the following parameters: wavelength 400 – 800 nm, readings every 2 nm, slow mode. The data were parsed and graphs of absorbance as function of the wavelength were created using Microsoft Office Excel and GraphPad Prism 5 for Windows.

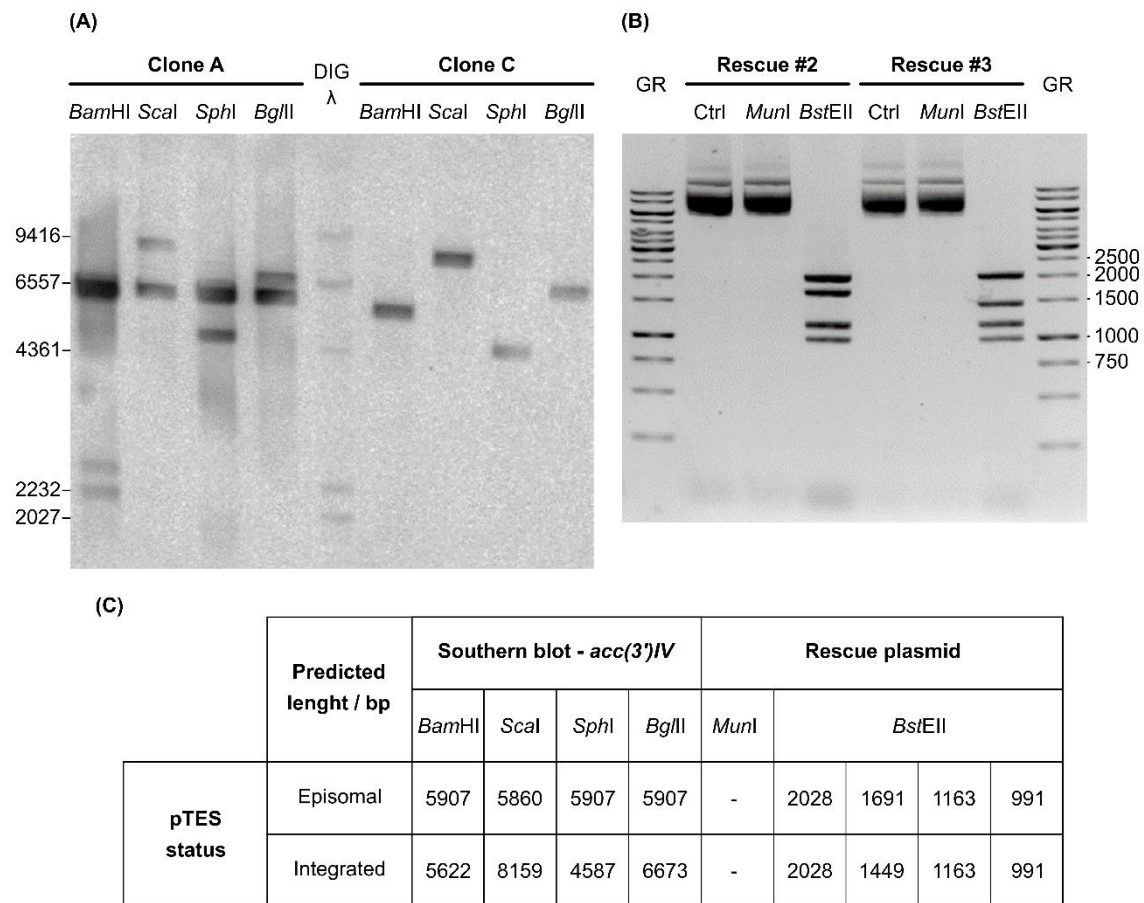
HPLC-MS analysis of the supernatant of cultures

For HPLC-MS analysis, 200 µL of the samples' supernatant were diluted 1:1 in 0.2% formic acid in methanol, vortexed, centrifuged at 16000g, for 10 min, at 4 °C; 50 µL of the mixtures' supernatant were transferred to polypropylene HPLC vials and kept at 4 °C until injection. The samples (4 µL) were injected by inline split loop well plate sampler, at a dispense rate of 20 µL/s, after needle flushing with 200 µL 50% (v/v) isopropanol in ddH₂O. The samples were analyzed using a UPLC ACQUITY column (Waters), part number 176001304, for Reverse Phase separation, with a column matched BEH C18 pre-column VanGuard (Waters). The column had 30 mm in length, 2.1 mm internal diameter, 1.7 µm diameter Ethylene Bridged Hybrid C18 spherical particles, and 130Å pore size. The chromatographic system was the Dionex UltiMate 3000 Rapid Separation LC system, with binary pump, autosampler, thermostated column compartment and diode array detector modules (Thermo Scientific). The mobile phase A was ddH₂O with 0.1% formic acid and mobile phase B was acetonitrile with 0.1% formic acid, both freshly prepared with HPLC-grade components. The total runtime was 7.7 minutes, with the following proportions of mobile phases: min 0 – 0.5, 5% B; min 0.5 – 5.9, gradient 5% to 95% B; min 5.9 – 6.9, 95% B; min 6.9 – 7.2, gradient 95% to 5% B; min 7.2 – 7.7, 5% B. The flow rate was 0.6 mL/min, and

the column compartment was kept at 45 °C. The column was preconditioned during 5 min, 95% B, flow rate 0 – 0.6 mL/min followed by 5 min, 5% B, flow rate 0.6 mL/min, at 45°C. Before the test samples, one injection of 10 µL methanol and one injection of 1 µL “Sulfamix” (mixture of sulfonamide compounds [Agilent P/N 59987-20033], used as LC/MS checkout standard) were performed. After the last test sample, one injection of 10 µL methanol was performed. Column rinsing, and storage was performed with 80% (v/v) acetonitrile in ddH₂O solution, at 0.6 mL/min, for 20 minutes. The detection was performed by the DAD module of the chromatographic system, on 200 – 600 nm wavelength range, from runtime 0 to 6.9 min, at 20 Hz acquisition rate. The UV-Vis channels were populated with the 340 nm and 520 nm wavelengths, for live tracking of the eluate. The column output was also detected by mass spectrometry (Quadrupole Ion Trap). The MS analysis of HPLC eluate was performed using an amaZon speed Iontrap mass spectrometer (Bruker). The ions were produced by an electrospray ionization (ESI) Apollo II source (Bruker), fed by the UPLC column output. Capillary at 3500V, nebulizer activated. The post-source components were a dual ion funnel transfer line, an octapole ion guide, and a (Paul) quadrupole ion trap analyzer. The tandem mass spectrometry fragmentation was performed by collision induced dissociation (CID) in the quadrupole ion trap, using He 5.0 as damping and collision gas. The instrument was managed with the trapControl V7.2 and HyStar V3.2 software (Bruker). The acquisition parameters were the following: positive mode, save line spectra only, scan m/z 50 – 250, enhanced resolution (8100 Da/s @ 0.3 Da FWHM) scan mode, automatic ion charge control (set at 1000000, max. accumulation time of 100 ms), auto tune for target mass 207 m/z, auto MS2 for the 2 most abundant ions (absolute threshold 2500, relative threshold 5%), CID cut-off 27% of precursors' mass, CID Smartfrag at 60% - 180% amplitude for 40 ms. The HPLC-MS data were analyzed using DataAnalysis V4.0 software (Bruker). The MS data were processed into extracted ion chromatogram (EIC), using built-in features of the software; the relevant UV-Vis and MS² data were gathered by manually selection of chromatographic peaks.

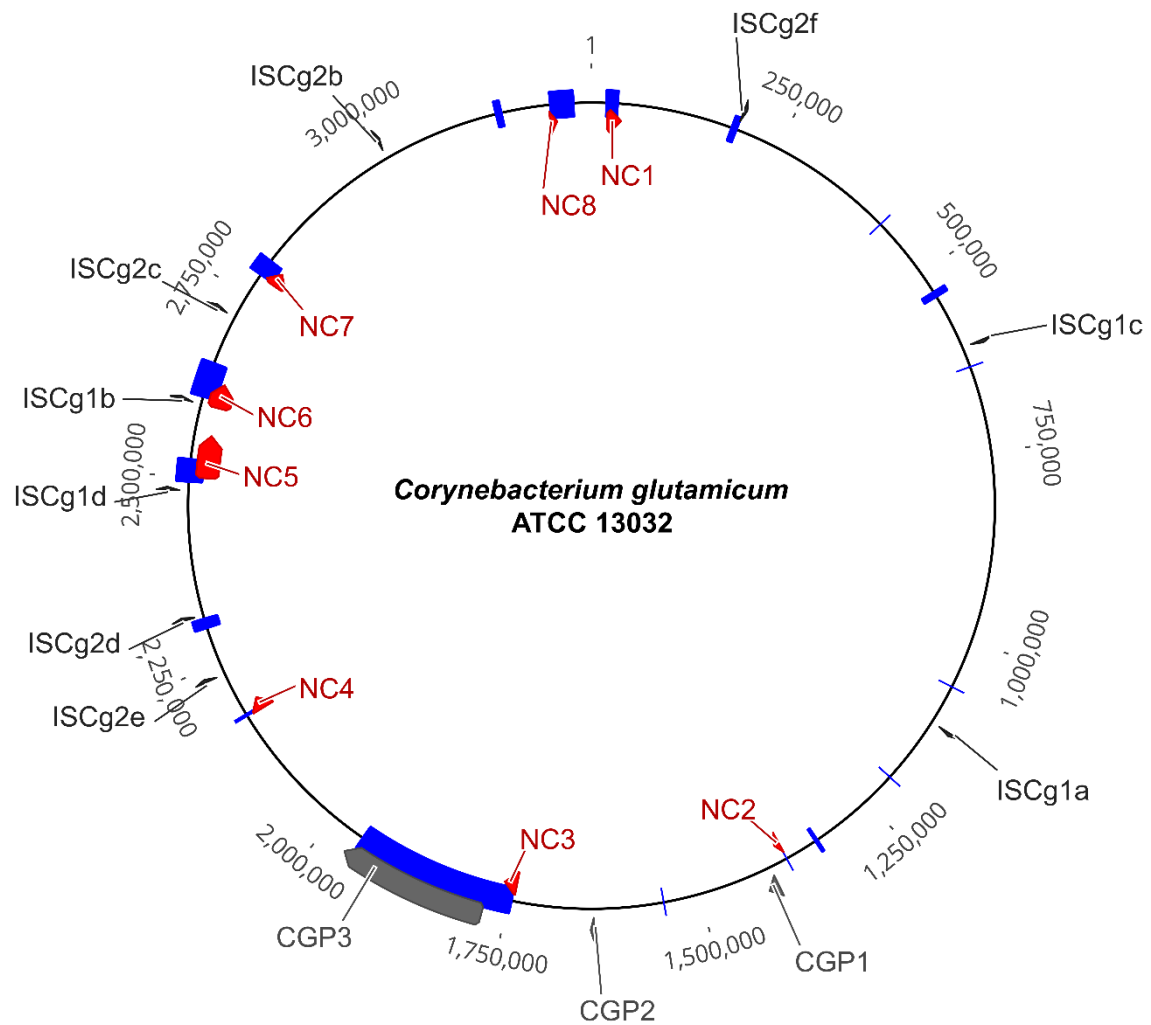
Figures

Supp. fig. 1



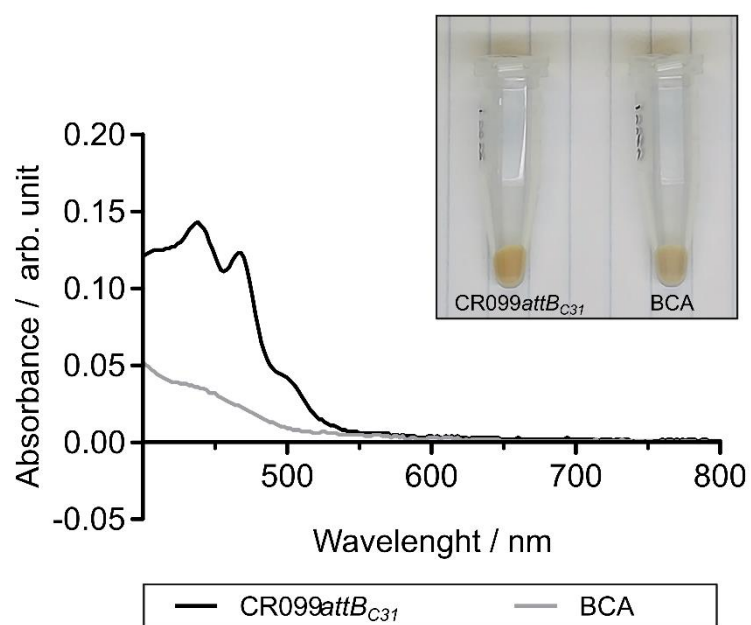
Supplementary figure 1. Analysis of the integration of ϕ C31-based plasmid pTES into the genome of *C. glutamicum* CR099attB_{C31} strain. In panel (A), Southern blot targeting the *aac(3')IV* gene, on clone A and C of CR099attB_{C31} transformed with pTES vector. Clone C presents the expected patterns for pTES being integrated in the attB_{C31} locus; clone A present extra bands that match the digestion of pTES plasmid with the indicated enzymes. In panel (B), rescue of pTES-derived plasmids from CR099attB_{C31} transformed with pTES vector. Rescue plasmid #3 presented the expected digestion pattern and sequence corresponding to the flanking region of the attB_{C31} locus; rescue plasmid #2 corresponded to the pattern and sequence of pTES plasmid digested by *Bam*HI endonuclease. Patterns like those of rescue plasmid #3 and #2 occurred at a ratio of about 3:50, respectively. In panel (C), predicted lengths for the described southern blot and rescue plasmid experiments. DIG λ – Digoxigenin-labelled λ DNA-*Hind*III digest (DNA Molecular Weight Marker II, Roche); GR – GeneRuler 1 kb DNA ladder (Thermo Scientific).

Supp. fig. 2



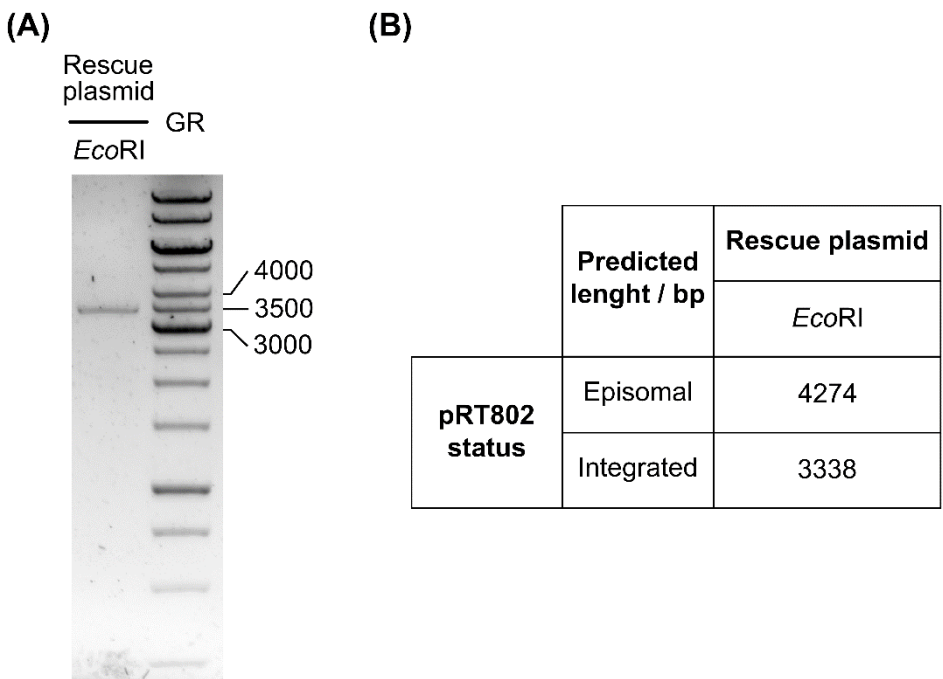
Supplementary figure 2. Regions of the genome of *C. glutamicum* ATCC 13032 not covered by cosmid libraries. Insertion sequences (ISCg) and prophages (CGP) removed during genome streamlining from ATCC 13032 to CR099attB_{C31} represented by gray arrows. Regions not represented in the cosmid library of CR099attB_{C31} represented by red arrows (NC1-8). Regions not represented in the cosmid library of ATCC 13032 (Bathe, Kalinowski et al. 1996, Tauch, Homann et al. 2002) represented in blue.

Supp. fig. 3



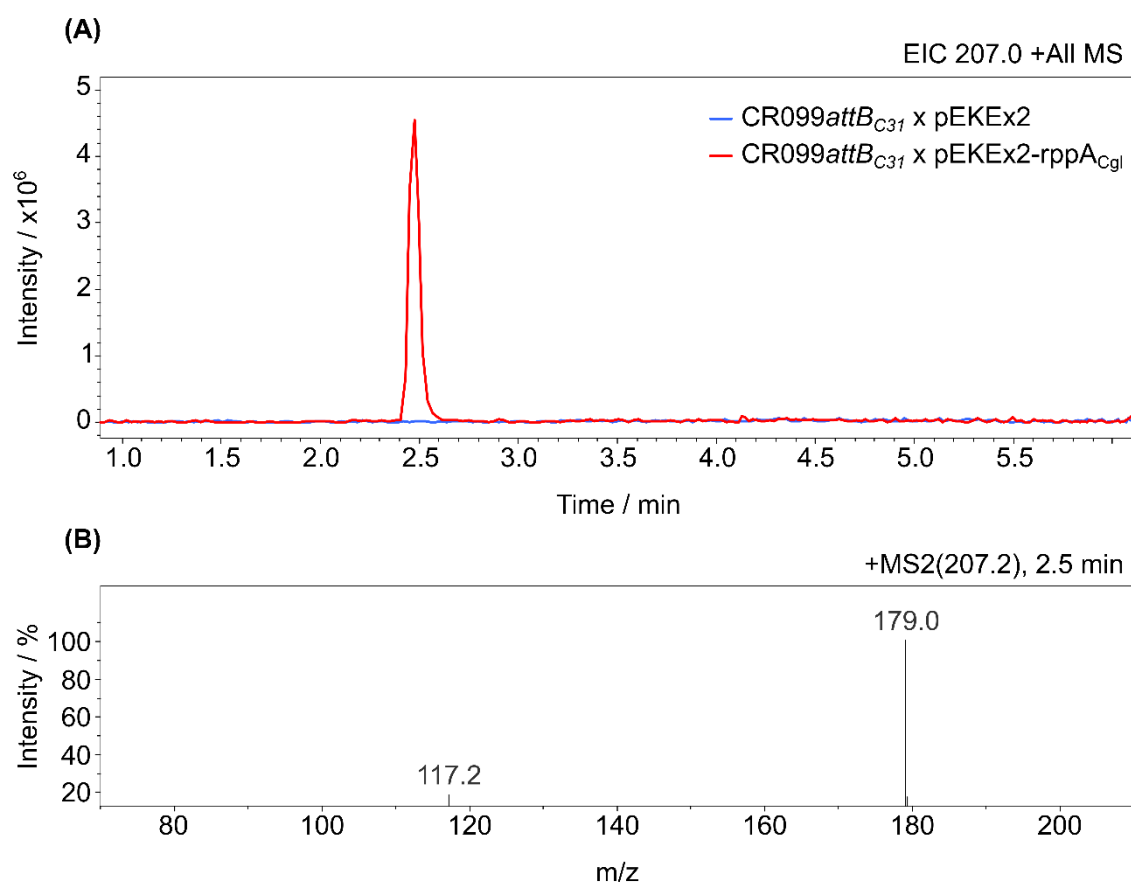
Supplementary figure 3. Metabolic profile of *C. glutamicum* strains. Visible spectrum of the methanolic extracts of the biomass of the CR099attB_{C31} and BCA strains. Inset: macroscopic aspect of the biomass of CR099attB_{C31} and the BCA strains.

Supp. fig. 4



Supplementary figure 4. Analysis of the integration of ϕ BT1-based plasmid pRT802 into the genome of *C. glutamicum* BCA strain. In panel (A), rescue of pRT802-derived plasmid from the BCA-101 strain. The rescue plasmid presented the expected digestion pattern and sequence corresponding to the flanking region of the *attB_{BT1}* locus. All the four rescue plasmids tested presented the pattern shown. In panel (B), predicted lengths for the described rescue plasmid experiment. GR – *GeneRuler 1 kb DNA ladder* (Thermo Scientific).

Supp. fig. 5



Supplementary figure 5. Mass spectrometry (MS) analysis of flaviolin in supernatant of culture broths. In panel **(A)**, extracted ion chromatograms at m/z 207 of the liquid chromatography eluates of the supernatants of the indicated strains. The selected mass-to-charge ratio (m/z) corresponds to the molecular ion of flaviolin $[C_{10}H_6O_5+H]^+$. In panel **(B)**, tandem MS spectrum of the parent 207 m/z ions eluted at 2.5 min from the supernatant of the culture of *C. glutamicum* CR099attB_{C31} x pEKEx2-rppA_{Cgl} strain.

Data

Supp. data 1

Supplementary data 1. Sequence and features of the attachment site for the ϕ C31 phage (*attB*_{C31}) in *C. glutamicum* CR099*attB*_{C31}. (See attached file.) The selected region is located between the genomic coordinates 192030 and 195000 in the original numbering of MB001 strain, NCBI accession CP005959.1.

Supp. data 2

Supplementary data 2. Sequence of the cosmid cos15A-sacB, used as backbone for genomic library of *C. glutamicum* CR099*attB*_{C31} strain. See attached file.

Supp. data 3

Supplementary data 3. Contents of the genomic library of *C. glutamicum* CR099*attB*_{C31}. (See attached file) Includes the position of each bacterial clone in the physical library and the correspondent genomic coordinates of the regions covered by each cosmid. The genomic coordinates correspond to the genome of *C. glutamicum* MB001, NCBI accession CP005959.1.

Supp. data 4

Supplementary data 4. List of genes not represented in the *C. glutamicum* CR099*attB*_{C31} cosmid library. (See attached file.) The gene tags correspond to the genome of *C. glutamicum* MB001, NCBI accession CP005959.1.

Supp. data 5

Supplementary data 5. Candidate genes of *C. glutamicum* related to unclonability into p15A-based cosmids in *E. coli* host. The listed genes of *C. glutamicum* MB001 have homologs ($E < 10^{-8}$) that are known to be unclonable in *E. coli* (Kimelman, Levy et al. 2012). Functional category according to BlastKOALA (Kanehisa, Sato et al. 2016). NC# – number of the region not represented in the cosmid library of CR099attB_{C31} (refer to supp. fig. 2).

NC #	Locus tag	Type	Protein ID	Name	Product	Main functional category (KEGG)
1	cgp_t5003	tRNA	-	tRNA-Ala	tRNA-Ala	-
1	cgp_t5004	tRNA	-	tRNA-Ile	tRNA-Ile	-
1	cgp_t5005	tRNA	-	tRNA-Ala	tRNA-Ala	-
2	cgp_1482	CDS	AGT05296.1	putative Zn-dependent hydrolase	putative Zn-dependent hydrolase	Carbohydrate metabolism
2	cgp_1487	CDS	AGT05301.1	leuC	3-Isopropylmalate dehydratase, large subunit	Carbohydrate metabolism
2	cgp_1488	CDS	AGT05302.1	leuD	3-Isopropylmalate dehydratase, small subunit	Carbohydrate metabolism
3	cgp_1852	CDS	AGT05602.1	L-Serine dehydratase	L-Serine dehydratase	Amino acid metabolism
3	cgp_1861	CDS	AGT05609.1	GTP diphosphokinase	GTP diphosphokinase	Nucleotide metabolism
4	cgp_2291	CDS	AGT05824.1	pyk	pyruvate kinase	Carbohydrate metabolism
5	cgp_2620	CDS	AGT06115.1	clpX	ATP-dependent clp protease ATPase subunit	Cellular processes
5	cgp_2645	CDS	AGT06139.1	clpP1	endopeptidase Clp, proteolytic subunit	Cellular processes
5	cgp_t5045	tRNA	-	tRNA-Pro	tRNA-Pro	
5	cgp_t5046	tRNA	-	tRNA-Gly	tRNA-Gly	
6	cgp_2753	CDS	AGT06221.1	rph	tRNA nucleotidyltransferase	Genetic information processing
6	cgp_t5051	tRNA	-	tRNA-Leu	tRNA-Leu	
7	cgp_2924	CDS	AGT06358.1	cysS	cysteine-tRNA ligase	Genetic information processing
7	cgp_2925	CDS	AGT06359.1	ptsS	phosphotransferase system (PTS), sucrose-specific enzyme IIBCA component	Carbohydrate metabolism
7	cgp_2937	CDS	AGT06369.1	ABC-type putative dipeptide/oligopeptide transporter, substrate-binding lipoprotein	ABC-type putative dipeptide/oligopeptide transporter, substrate-binding lipoprotein	Cellular processes

Supp. data 6

Supplementary data 6. Sequence and features of the attachment site for the ϕ BT1 phage (attB_{BT1}) in *C. glutamicum* BCA. (file attached). The selected region is located between the genomic coordinates 634378 and 642628 in the original numbering of MB001 strain, NCBI accession CP005959.1.

Supp. data 7

Supplementary data 7. Sequence of the synthesized *rppA*_{Cgl} gene. See attached file

Supp. data 8

Supplementary data 8. Sequence of the synthesized *rppA*_{CglStrept} gene. See attached rppA_CglStrept.gb file.

Supp. data 9

Supplementary data 9. Sequence of the plasmid pFM69SD, used as ϕ BT1-based integration into the genome of *C. glutamicum* BCA strain. See attached file.

Supp. data 10

Supplementary data 10. Sequence of the modified *bpsA* gene of *S. albidoflavus* for expression of indigoidine in *C. glutamicum*. See attached file.

Additional references

- Bathe, B., J. Kalinowski and A. Puhler (1996). "A physical and genetic map of the *Corynebacterium glutamicum* ATCC 13032 chromosome." Mol Gen Genet **252**(3): 255-265.
- Bilyk, B., L. Horbal and A. Luzhetskyy (2017). "Chromosomal position effect influences the heterologous expression of genes and biosynthetic gene clusters in *Streptomyces albus* J1074." Microb Cell Fact **16**(1): 5.
- Darling, A. C., B. Mau, F. R. Blattner and N. T. Perna (2004). "Mauve: multiple alignment of conserved genomic sequence with rearrangements." Genome research **14**(7): 1394-1403.
- Heider, S. A., P. Peters-Wendisch and V. F. Wendisch (2012). "Carotenoid biosynthesis and overproduction in *Corynebacterium glutamicum*." BMC Microbiol **12**: 198.
- Kanehisa, M., Y. Sato and K. Morishima (2016). "BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences." J Mol Biol **428**(4): 726-731.
- Kimelman, A., A. Levy, H. Sberro, S. Kidron, A. Leavitt, G. Amitai, D. R. Yoder-Himes, O. Wurtzel, Y. Zhu, E. M. Rubin and R. Sorek (2012). "A vast collection of microbial genes that are toxic to bacteria." Genome Res **22**(4): 802-809.
- Schafer, A., A. Tauch, W. Jager, J. Kalinowski, G. Thierbach and A. Puhler (1994). "Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*." Gene **145**(1): 69-73.
- Tauch, A., I. Homann, S. Mormann, S. Rüberg, A. Billault, B. Bathe, S. Brand, O. Brockmann-Gretza, C. Rückert, N. Schischka, C. Wrenger, J. Hoheisel, B. Möckel, K. Huthmacher, W. Pfefferle, A. Pühler and J. Kalinowski (2002). "Strategy to sequence the genome of *Corynebacterium glutamicum* ATCC 13032: use of a cosmid and a bacterial artificial chromosome library." Journal of Biotechnology **95**(1): 25-38.

CHAPTER III

Balancing growth and heterologous production of indigoidine in *Corynebacterium glutamicum* using a redox-sensing genetic switch.

Marques, F., *et al.* (manuscript in preparation). Balancing growth and heterologous production of indigoidine in *Corynebacterium glutamicum* using a redox-sensing genetic switch.

Balancing growth and heterologous production of indigoidine in *Corynebacterium glutamicum* using a redox-sensing genetic switch

Filipe Marques ^{1,2,3,4} *et al.*

¹ i3S- Instituto de Investigação e Inovação em Saúde, Universidade do Porto, 4200-135 Porto, Portugal

² IBMC – Instituto de Biologia Molecular e Celular, Universidade do Porto, 4200-135 Porto, Portugal.

³ ICBAS – Instituto de Ciências Biomédicas de Abel Salazar, Universidade do Porto, 4050-313 Porto, Portugal.

⁴ Pharmazeutische Biotechnologie, Universität des Saarlandes, Campus C2.3, 66123 Saarbrücken, Germany.

Abstract

Biotechnological production of fine chemicals aims at the highest conversion rate of substrates into valuable molecules while minimizing by-products and cell biomass. In the industrial workhorse *Corynebacterium glutamicum*, the transition from aerobic to microaerobic (and anaerobic) conditions is associated with the overflow metabolism and has been used for zero-growth processes, such as the production of organic acids and amino acids. In this work we developed a redox-sensing genetic switch based on the Rex repressor from *Thermus aquaticus*, to integrate the transition to overflow metabolism in *C. glutamicum* and the expression of fine chemicals encoded by heterologous genes.

Using phage-based methodology previously developed, the Rex-based genetic switch was implemented in the *C. glutamicum* BCA strain generating the REX strain that constitutively expressed the Rex protein. The transcriptional repressors of the Rex family are well-described effectors of the adaptation to overflow metabolism in response to the intracellular NADH/NAD⁺ ratio. The implemented Rex system effectively delayed the onset of the expression of the non-ribosomal peptide synthetase reporter *bpsA* under the control of its cognate operator, following the pattern of intracellular NADH/NAD⁺ poise. In defined media with glucose, the presence of Rex was associated with the transcriptional repression of the reporter gene at early cultivation stages. In complex media with maltose as the (main) carbon source, the Rex regulation was associated with effective metabolic shifting from cell biomass to indigoidine production, especially at high carbon-to-nitrogen ratios. The redox-sensing expression system based on the Rex regulator will accelerate the optimization of yields of natural products in *C. glutamicum* and other fast-growing hosts.

Keywords: *Corynebacterium glutamicum*; heterologous expression; Rex; indigoidine; genetic switch.

Introduction

Corynebacterium glutamicum is a Gram-positive soil-dwelling bacteria best known for being the platform microorganism for the industrial production of amino acids, organic acids and other products (Hermann 2003, Wendisch, Bott et al. 2006, Smith, Cho et al. 2010, Litsanov, Kabus et al. 2012). Most of the production processes of commodity chemicals by *C. glutamicum* is performed under aerobic conditions and relies on the balance between the final product yield and the redirection of substrates for the generation of cell biomass (Wendisch, Bott et al. 2006). This balance on the carbon flux distribution is intimately linked to the energy status of the cell. During the exponential growth phase, in glucose-sufficient conditions, the core metabolism is devoted to cell growth and the metabolic intermediates are used as building blocks for generating biomass and the energy needed. At late-exponential growth phase, oxygen becomes limiting and there is a decrease in NADH reoxidation and ATP generation via the respiratory chain. To counteract these effects, the cell favours the production of lactate and, in a less extent, succinate, as a way to re-oxidize NADH, and the production of acetate from acetyl-CoA as a way to generate ATP by substrate level phosphorylation (Koch-Koerfges, Kabus et al. 2012, Koch-Koerfges, Pfelzer et al. 2013). Cell growth decreases as the biomass precursors are being redirected to re-establishing the intracellular energetic homeostasis. This metabolic redirection under oxygen-limiting conditions is determined at the transcriptional level through the up-regulation of genes involved in glycolysis, lactate biosynthesis, C3 carboxylation and the reductive direction of the TCA cycle (Yamamoto, Sakai et al. 2011). The trade-off between biomass formation, energy metabolism and product yield is key for the development of efficient industrial strains.

Optimization of the biotechnological production processes based on *C. glutamicum* led to the generation of extensive knowledge on its physiology, biochemistry and genetics (Inui and Toyoda 2020). *C. glutamicum* is a facultative anaerobic organism that has a respiratory type of energy metabolism with oxygen or nitrate as terminal electron acceptors. Electron transfer from menaquinol to oxygen is catalysed either by a cytochrome *bc1-aa3* complex or by the cytochrome *bd* oxidase (Kusumoto, Sakiyama et al. 2000, Niebisch and Bott 2003). The terminal oxidases have different energetic efficiencies being the *bc1-aa3* branch more efficient on proton translocating (Kabashima, Kishikawa et al. 2009) and predominant in oxygen-sufficient conditions (Bott and Niebisch 2003, Koch-Koerfges, Pfelzer et al. 2013). Generation of ATP through oxidative phosphorylation is not essential for *C. glutamicum* although there is a decrease in the growth rate and reduced biomass in the absence of the F1-F0-ATP synthase (Koch-Koerfges, Kabus et al. 2012). Alternatively, ATP can be generated by substrate-level phosphorylation in glycolysis, in the TCA cycle and in the

conversion of acetyl-CoA to acetate. In fact, up to 55% of total ATP in *C. glutamicum* can be generated by substrate-level phosphorylation (Koch-Koerfges, Kabus et al. 2012, Zelle, Pfelzer et al. 2021).

When the electron chain complex reduces its activity in oxygen-limited conditions (e.g. at the onset of the stationary phase) the NADH levels increase due to a decrease of the activity of the non-proton-pumping NADH dehydrogenase Ndh (Maeda, Koch-Koerfges et al. 2020). However, catabolic NADH reoxidation is crucial to sustain carbon flux in central metabolism and *C. glutamicum* possess two alternative quinone dependent NADH-oxidizing enzymes able to complement Ndh activity: the L-lactate dehydrogenase LdhA and the malate dehydrogenase Mdh (Maeda, Koch-Koerfges et al. 2020). At the transcriptional level, *ldhA* expression is repressed by the L-lactate utilization genes regulator LldR (Toyoda, Teramoto et al. 2009) and by SugR, a global repressor of genes of the phosphoenolpyruvate-dependent phosphotransferase system (PTS) and glycolysis (Engels, Lindner et al. 2008). In this way, *C. glutamicum* is able to control the intracellular NADH/NAD⁺ ratio, maintain the cytosol redox balance and sustain the metabolic flux.

In microorganisms such as *Streptomyces* sp. or *Bacillus subtilis*, the intracellular NADH/NAD⁺ balance is monitored by the redox-sensing repressor Rex (Brekasis and Paget 2003). In fact, in *Bacillus subtilis*, Rex regulates the expression of the LdhA homolog encoding gene (Larsson, Rogstam et al. 2005). The determination of the crystal structures of *Thermus aquaticus* Rex (T-Rex) bound to NADH, NAD⁺ and/or DNA elucidated the molecular mechanism of Rex (Sickmier, Brekasis et al. 2005, McLaughlin, Strain-Damerell et al. 2010). Both NAD⁺ and NADH compete for binding to the C-terminal domain of Rex proteins but only NADH causes Rex to lose its affinity for DNA (Brekasis and Paget 2003). In this way, Rex proteins are able to regulate the expression of target genes in response to changes in the intracellular NADH/NAD⁺ redox balance. Rex represses gene expression by binding to a conserved inverted repeat sequence, known as Rex operator (ROP) site, located at the promoter regions of the target genes (Ravcheev, Li et al. 2012). Characterization of the Rex regulon revealed that Rex plays a central role not only in regulating energy metabolism (Brekasis and Paget 2003, Gyan, Shiohira et al. 2006, Pagels, Fuchs et al. 2010) but also in central carbon metabolism (Ravcheev, Li et al. 2012) and secondary metabolism (Liu, Cheng et al. 2017) among other metabolic pathways. Although Rex homolog proteins are present in a wide variety of organisms, no encoding gene is found in the *C. glutamicum* genome.

Heterologous expression has proven to be a valuable strategy not only for the characterization of novel microbial metabolites but also to improve their production titre (Yamanaka, Reynolds et al. 2014, Boecker, Gratz et al. 2018). However, the production of

heterologous products is often associated with the reallocation of the host cell resources (carbon and energy) that exerts a metabolic burden on the host and negatively affects the host productivity (Wu, Yan et al. 2016). A strategy that can mitigate the metabolic burden induced by the heterologous expression involves the decoupling of the metabolite production phase from cell growth using dynamic metabolic regulation where the metabolic flux is redirected in response to internal or external signals (Burg, Cooper et al. 2016). This strategy of dynamic metabolic regulation has been successfully implemented in many microorganisms including *C. glutamicum* (Lo, Chng et al. 2016, Durante-Rodriguez, de Lorenzo et al. 2018, Kobayashi, Kawaguchi et al. 2020, Kasari, Kasari et al. 2022). For instance, the implementation of an oxygen-responsive metabolic switch in *C. glutamicum* allowed to redirect the metabolic flux for the production of 1,5-diaminopentane and succinate (Kobayashi, Kawaguchi et al. 2020).

In this work, we aimed at decoupling the expression of heterologous genes from *C. glutamicum* growth using a Rex-based genetic switch able to sense the metabolic status of the cell and regulate gene expression accordingly. The onset of the production of microbial secondary metabolites occurs at the late-exponential phase where metabolites from the central carbon metabolism can be diverted to secondary metabolism without affecting cell fitness. In addition, several studies have pointed out the correlation between the metabolic switch from primary to secondary metabolism and the intracellular redox state (Beites, Oliveira et al. 2015, Oliveira, Bush et al. 2020). By implementing a Rex-based switch in *C. glutamicum* we were able to delay the heterologous expression of *bpsA* in a way that the host metabolic resources were first directed for cell growth, ensuring cell fitness, and later for the production of indigoidine.

Materials & Methods

Bacterial strains, plasmids, media and growth conditions

All the strains and plasmids used in this work are listed in Table 1. *Escherichia coli* was routinely cultivated in Lysogeny Broth (LB) with 5 g L⁻¹ NaCl (Lennox 1955) at 37°C. *Corynebacterium glutamicum* was routinely cultivated in LB, or CGXII defined media (Keilhauer, Eggeling et al. 1993) with 4% (w/v) glucose, 0.2 mg L⁻¹ biotin, and 30 mg L⁻¹ protocatechuic acid at 30°C. When necessary, media was supplemented with ampicillin (Ap; 100 µg mL⁻¹ to *E. coli*), apramycin (Am; 50 µg mL⁻¹ to *E. coli*, 25 µg mL⁻¹ to *C. glutamicum*), hygromycin (Hyg; 120 µg mL⁻¹ to *E. coli*, 80 µg mL⁻¹ to *C. glutamicum*), maltose [1-3% (w/v)], glucose [1-3% (w/v)], and urea [0.5% (w/v)]. As a pre-cultivation step, *C. glutamicum* colonies from fresh plates were inoculated in 5 mL LB for 6-8 h at 200 rpm. For cultivation

in complex media, pre-cultures were diluted 1:100 in 25 mL of fresh LB in 250-mL unbaffled flasks and incubated for 72 h at 200 rpm. For cultivation in a defined medium, cells from 1 mL of pre-culture were firstly transferred to 10 mL of CGXII medium in 100-mL flasks, incubated for 12-16 h at 200 rpm; then, cells were transferred to 25 mL of fresh CGXII medium in 250-mL unbaffled flasks to an initial OD800 of 0.72 and incubated for 72 h at 200 rpm. Bacterial growth was determined by cell dry weight, as described previously (Marques, Luzhetskyy et al. 2020).

Table 1. Strains and plasmids used in this work

Bacterial strains and vectors	Description ^a	Source or reference
Bacteria		
<i>E. coli</i> DH5α	General cloning host	(Hanahan 1983)
<i>C. glutamicum</i> BCA	Prophage-free and IS-reduced <i>C. glutamicum</i> strain, featuring <i>attB_{BT1}</i> and <i>attB_{C31}</i> , with albino phenotype	(Marques, Luzhetskyy et al. 2020); NCBI acc. CP059382.1
<i>C. glutamicum</i> REX	<i>C. glutamicum</i> BCA derived strain, <i>attB_{BT1}::rex</i>	This work
Plasmids		
pTES	Am ^R ; ϕC31-based integrative plasmid, suitable for marker excision	(Herrmann, Siegl et al. 2012)
pTES-ΔermEp	pTES derivative, lacking the <i>ermE</i> promoter	This work
pALCRE	Hyg ^R ; temperature-sensitive plasmid for expression of Cre	(Herrmann, Siegl et al. 2012)
pUC57-T-Rex-Prop-CreiLOV	Ap ^R ; source of <i>t-rex</i> gene (NCBI acc. WP_011173693)	This work (GenScript)
pTES-rsx	pTES-ΔermEp derivative for constitutive expression of T-Rex regulator	This work
pFM69SD	Am ^R ; ϕBT1-based integrative plasmid for high-level gene expression	(Marques, Luzhetskyy et al. 2020)
pFM69SD-bpsA	pFM69SD derivative for expression of indigoidine	(Marques, Luzhetskyy et al. 2020)
pFM69ROP	pFM69SD derivative featuring a Rex operator (ROP) at the multicloning site	This work
pFM69ROP-bpsA	pFM69ROP derivative for Rex-controlled expression of indigoidine	This work

^aAbbreviations: Am^R – apramycin resistance marker; Ap^R – ampicillin resistance marker; *attB_{BT1}* – bacterial attachment site for phage ϕBT1; *attB_{C31}* – bacterial attachment site for phage ϕC31; Hyg^R – hygromycin resistance marker.

Bacterial manipulations

Transformation of *E. coli* was performed using chemically competent cells (Renzette 2011). Transformation of *C. glutamicum* was performed by electroporation (Eggeling and Bott

2005). Integration of phage-based plasmids into the chromosome of *C. glutamicum* was confirmed by PCR using the appropriate primer pairs (Table 2). Excision of the backbone of ϕ C31-based integrative plasmids was performed using the *cre* function of pALCRE plasmid as described previously (Herrmann, Siegl et al. 2012, Marques, Luzhetskyy et al. 2020).

DNA manipulations

Plasmid DNA was isolated with the NZYMiniprep kit (NZYTech, Lisbon, Portugal). Genomic DNA of *C. glutamicum* was isolated following the salting-out method (Kieser, Bibb et al. 2000). DNA was purified from enzymatic reactions and agarose gels with the NZYGelpure kit. PCR was performed using Q5® High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA) or GoTaq® G2 Flexi DNA Polymerase (Promega, Madison, MA, USA). The oligonucleotides used in this study are listed in Table 2. Restriction enzymes, T4 ligase, alkaline phosphatase (FastAP), T4 polynucleotide kinase (T4 PNK), and GeneRuler DNA Ladder Mix were provided by Thermo Fisher Scientific (Waltham, MA, USA). Customized dsDNA fragments were synthesized by GenScript (Piscataway, NJ, USA). Sanger sequencing was performed by STAB VIDA (Caparica, Portugal).

Table 2. Oligonucleotides and DNA fragments used in this work.

Oligonucleotide	Sequence (5' to 3')	Use
pTESmcs_fwd	GCGAGGAAGCGGAAGAG	Sequence pTES-based plasmids
pTESmcs_rvs	CGATGTAGGAGGGCGTG	
attB-CR99_fwd	CCATCGCAGCCCGACTTG	Sequence and check markerless integration at <i>attB_{C31}</i> locus
attB-CR99_rvs	AGCAGCCCACCAAACGTC	
crt_delchk2_rvs	GCCTACTTGATGAGGCCGAA	Check integration of pFM69SD derivatives at <i>attB_{BT1}</i> locus
pRT_ORI_SEQ	GGATTCCCGTTGAGCACCGC	
ROP_F	CCGCTGTGAACGCGTTCACAGCGA GGCCTTGTCAGAAAAGTAAGGAG ATAGCA	Generate a Rex operator and SD sequence (ROP in bold)
ROP_R	TATGCTATCTCCTTACTTTTCTGACA AGGCCTCGCTGTGAACGCGTTCA CAGCGGGTAC	
bpsA_S	CACCGACAACCAGGTCAAAC	RT-qPCR <i>bpsA</i>
bpsA_AS	CCTGGAGCTTGGAGGACTTG	
pccB_S	ATGCACCTGTGACCACTGAC	RT-qPCR <i>pccB</i>
pccB_AS	AGGTCAAAGGCAGGAAGTCG	
gyrB_S	GTTGAAGAGACCACCGATGC	RT-qPCR <i>gyrB</i>
gyrB_AS	CTTCGCGGGTGTAAATGGTG	
rex_S	AGGTTGCAAGGATCTGTCC	RT-PCR
rex_AS	TGGGTCCACATCGAAGAAGC	

NdeI_carB_fwd	CCCAAACATATGGTGTTCGAAGAG AACTCC	Generate fragment encoding <i>carB</i> gene and <i>carBA</i> operon (restriction site underlined)
carB_rvs	ATTCTATCTCCTTACTTGGTACTGG ACCGGTTTAGTACTTCTTGC	
carA_fwd	CCAGTACCAAGTAAGGAGATAGAA TATGTCCAACCTCCTTCTGCGTG	Generate fragment encoding <i>carA</i> gene and <i>carBA</i> operon (SD sequence in bold, restriction site underlined)
SpeI_carA_rvs	AAACCTACTAGTTTAGTCCTTCTTG ATGAGATCC	

Construction of plasmids

Integration at the *attB_{C31}* was achieved using a derivative of pTES vector lacking the *ermE* promoter. pTES- Δ ermEp plasmid was obtained by digestion of pTES with *ScaI* and self-ligation.

The codon-optimized sequence of the well-characterized Rex protein of *Thermus aquaticus* (NCBI acc. WP_011173693) (McLaughlin, Strain-Damerell et al. 2010) was synthesized according to preferred codon usage of *C. glutamicum* (CAI 0.93), with a designed ribosome binding site (RBS) (Salis, Mirsky et al. 2009), an optimized Shine-Dalgarno (SD) sequence (5'-TCAGGAGGT-3', 4-nt spacer to ATG start codon) (Seo, Yang et al. 2013), and the medium-high strength promoter JVR4.3 from a library of synthetic promoters (Rytter, Helmark et al. 2014). To constitutively express *rex* in *C. glutamicum*, the *AvrII* and *EcoRI* 757 bp *rex* containing fragment from pUC57-T-Rex-Prop-CreiLOV plasmid (Table 1) was cloned into the *XbaI* and *EcoRI* sites of pTES- Δ ermEp generating pTES-*rex* (Fig. S1A). The identity of pTES-*rex* was confirmed by sequencing.

Integration at the *attB_{BT1}* site of *C. glutamicum* BCA was performed using derivatives of the pFM69SD expression vector (Marques, Luzhetskyy et al. 2020). Rex-controlled expression of indigoidine was achieved by introducing a DNA fragment harboring the Rex operator site [ROP, 5'-CGCTGTGAACGCGTTCACAGCG-3', as described in the crystallographic elucidation of the interaction between Rex and the ROP (McLaughlin, Strain-Damerell et al. 2010)] and an SD sequence, downstream the strong P69 promoter (Zhang, Liu et al. 2018) and upstream to the non-ribosomal peptide synthase gene *bpsA* of *S. albidoflavus*. The ROP-SD fragment was generated by “touchdown” annealing of equimolar amounts of the oligonucleotides ROP_F and ROP_R (Table 2). The *KpnI*- and *NdeI*-compatible overhangs of the generated fragment were 5'-phosphorylated and cloned into the respective sites of pFM69SD to yield pFM69ROP (Fig. S1B) or into pFM69SD-*bpsA* (Marques, Luzhetskyy et al. 2020) to yield pFM69ROP-*bpsA* (Fig. S1C). The identity of both generated plasmids was confirmed by sequencing.

RNA manipulations

Total RNA was purified by sample stabilization, adsorption to silica-based membranes, and digestion of residual genomic DNA. Samples of cultures (1.5 mL at 12 h, 1 mL at 24 h and 48 h) were collected by centrifugation (5000g, 4°C, 5 min). Samples were stabilized using RNA protect Bacteria Reagent (QIAGEN, Venlo, Netherlands), following the manufacturer's instructions, and stored at -80°C. RNA extraction was performed using rNeasy® Mini Kit (QIAGEN), following the manufacturer's instructions. Briefly, cells were resuspended in 350 µL of RLT buffer with 2-mercaptoethanol (1% v/v) and disrupted using a Branson Sonifier S-250A (3 cycles of 10 seconds, duty cycle 50%, power setting 3) on ice. Cell lysates were cleared by centrifugation (16100g, 4°C, 3 min) and employed for RNA extraction. The optional on-column DNA digestion was performed following the manufacturer's instructions. RNA was eluted using 40 µL of RNase-free water (Ambion, Life technologies) and stored at -80°C. RNA concentration was assessed using a NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific). Isolated RNA (up to 8 µg) was treated with DNA-free™ Kit (Ambion, Life technologies) to remove DNA contaminations, following the manufacturer's instructions. The absence of genomic DNA was confirmed by PCR using *gyr_S/AS* primers (Table 2).

Gene expression analysis

Gene expression was determined by two-step reverse transcription and quantitative PCR (RT-qPCR) or RT-PCR. Purified RNA (750 ng) was reverse transcribed to cDNA using iScript™ Reverse Transcription Supermix (Bio-Rad, Hercules, CA, USA) with the supplied random primers in a final volume of 20 µL, following the manufacturer's instructions.

For qPCR analysis the reactions were constituted by 0.25 µM of specific primers (Table 2), 5 µL of iTaq Universal SYBR Green Supermix (Bio-Rad), and 1 µL of template cDNA (dilution 1/5), in a final volume of 10 µL, in 384-well plates. qPCR was carried out in a CFX384 Touch Real-Time PCR Detection System (Bio-Rad) with the following conditions: 95 °C for 3 min, 40 cycles of 95 °C for 30 s, and 60 °C for 30 s. The resulting amplicons were subjected to melting curve analysis (65 to 95°C, with 0.5°C increments every 10 s) to exclude the formation of non-specific products. The efficiency and the quality of each primer pair were determined using standard serial dilutions of the cDNA (1 – 1/125 dilution). No-template controls were included. The software *Bio-Rad CFX Maestro* 2.3 was used to determine the quantification cycle (C_q), the stability of reference genes (geNorm M value) (Vandesompele, De Preter et al. 2002), and the relative normalized expression of target genes using the method described by Pfaffl (2001). For each analysis *pccB* and *gyrB* were used for normalization.

For RT-PCR analysis, the synthesized cDNA was used as template for PCR reactions constituted by 0.25 μ M of gene specific primers (Table 2), 1.5 mM $MgCl_2$, 0.2 mM dNTP, 1.25 U of GoTaq® G2 Flexi DNA Polymerase (Promega, Madison, MA, USA), and 2 μ L of template cDNA, in a final volume of 20 μ L. PCR was carried out with the following conditions: 95 °C for 10 min, 30 cycles of 95 °C for 30 s, 55 °C for 30 s and 72°C for 20 s, and a final extension at 72°C for 1 min. The resulting amplicons were analyzed by gel electrophoresis. Genomic DNA and non-template reactions were included and positive and negative controls, respectively.

Protein crude extracts and SDS-PAGE analysis

Heterologous protein expression was evaluated by polyacrylamide gel electrophoresis in denaturing conditions (SDS-PAGE). Cells were harvested by centrifugation (5000g, 4°C, 5 min), washed with 1 volume of lysis buffer (10 mM Tris-HCl, pH 8.0, 25 mM $MgCl_2$, 200 mM NaCl). Cells were resuspended in 400 μ L of lysis buffer with 1 mM PMSF protease inhibitor and disintegrated using a Branson Sonifier S-250A (7 min, duty cycle 30%, power setting 3) on ice. Cell lysates were cleared by centrifugation (16100g, 4°C, 15 min) and the supernatants were used further. The concentration of soluble protein was quantified using BCA Protein Assay Kit (Pierce™, Thermo Fisher Scientific, Rockford, IL, USA), using bovine serum albumin standards, following manufacturer's instructions for microplate procedure. Twenty micrograms of soluble proteins were separated by electrophoresis in SDS-polyacrylamide gel (polyacrylamide concentration (v/v): stacking gel 4%, running gel 12%) using standard protocol (Laemmli 1970), at 25 mA. Gels were stained using PageBlue™ Protein Staining Solution (Thermo Fisher Scientific, Rockford, IL, USA) following manufacturer's instructions and imaged on a GS-800 Calibrated Densitometer (Bio-Rad, Hercules, CA, USA) using the Image Lab software (Bio-Rad, Hercules, CA, USA).

Redox sensitivity bioassay

Bacterial strains were inoculated in 5 mL LB, in 50-mL tubes, and incubated at 30°C, overnight at 200 rpm. In the morning, 100 μ L of each inoculum (adjusted to OD600 3.5 with fresh media) was mixed with molten lysogeny top agar (10 mL, 0.9% agar) and spread on pre-solidified lysogeny agar (10 mL, 2% agar). Filter paper disks (6-mm diameter; M3) impregnated with menadione sodium bisulfate (1.1 mg, dissolved in water; Sigma-Aldrich) were placed on the top agar. Control disks were prepared with water only. The plates were inspected after incubation at 30°C, for 24 h.

Indigoidine quantification

The quantification of indigoidine in the cultures was carried out as described previously (Marques, Luzhetskyy et al. 2020). Briefly, sediments of samples (50-100 μ L) were

extracted with 1 mL of dimethyl sulfoxide by vigorous vortex mixing and cleared by centrifugation (16100g, 10 min). Absorbance at 615 nm was determined in a UV Mini 1240 UV/Vis Spectrophotometer (Shimadzu). Absorbance at 615 nm subtracted the background from control samples (*bpsA*⁻) was converted to mass concentration using the molecular weight of indigoidine and the extinction coefficient, 248.19 g mol⁻¹ and 1570 L mol⁻¹ cm⁻¹, respectively (Yumusak, Prochazkova et al. 2019).

Bioinformatics analysis

DNA annotation was performed using Geneious Prime 2022 software (Dotmatics, Boston, MA, USA). Search and alignment of nucleotide and protein sequences in databases (e.g., non-redundant nucleotide collection (Mizrachi 2007), UniProtKB/SwissProt sequences (The UniProt Consortium 2017)) were performed using BLAST webtool (Altschul, Gish et al. 1990). Operator sequences were retrieved from the RegPrecise database (Novichkov, Kazakov et al. 2013). Prediction of DNA-binding regions was performed using the Virtual Footprint webtool at PRODORIC (Dudek and Jahn 2022). Selected sequences were annotated with promoters and terminators found in *C. glutamicum* ATCC 13032 (Pfeifer-Sancar, Mentz et al. 2013). The statistical analysis was performed using GraphPad Prism 5 (GraphPad software, Inc.). One-way ANOVA test with Bonferroni's multiple comparison test were performed to compare a variable in three or more groups. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Results

Implementation of a Rex-based regulatory network in *C. glutamicum*

Rex proteins are redox-sensing transcriptional regulators widely spread among Actinobacteria (Ravcheev, Li et al. 2012). However, no Rex ortholog is encoded in the *Corynebacterium glutamicum* MB001 genome. To identify Rex-like proteins in *C. glutamicum* or phylogenetically close organisms, an alignment-based search on the non-redundant protein databases was performed using the BLASTP tool and the sequence of Rex protein from *Thermus aquaticus* (accession number WP_011173693) as query. The results showed that the lowest taxonomic rank related to the *Corynebacterium* genus that retrieved significant results (E-value < 0.05) for the presence of Rex-like proteins was the Mycobacteriales (Janke 1924) order (Gupta 2019) (Fig. S2A). The search yielded 643 significant results (E < 0.05), all described as redox-sensing transcriptional repressor Rex, highly represented in genus *Nocardia* (245 hits), *Rhodococcus* (241 hits), *Gordonia* (80 hits) and *Williamsia* (23 hits).

Although *C. glutamicum* lacks a Rex-like regulator, its genome may still retain Rex operator (ROP) sites that could modulate the transcription of genes in the presence of Rex protein. To find putative ROP sites in *C. glutamicum* MB001, whole-genome motif search was performed using Virtual Footprint feature at PRODORIC (Dudek and Jahn 2022) and the matrix for the ROP site motif reconstructed from 88 manually curated ROP sequences of Actinomycetales available at RegPrecise (Novichkov, Kazakov et al. 2013) (Fig. S2B). The search retrieved one intergenic occurrence between the *whcE* gene (*cgp_0878*) and the hypothetical protein-coding sequence *cgp_0879* (Fig. S2C). The putative ROP-like motif presented 66.7% identity to strict consensus (conserved in >50% sequences of the matrix), including the motif NNGTGNNNNNNNNCACNN present in >90% of the sequences. This result suggested that the expression of these genes could be modulated by the presence of Rex protein. Worth noting that *C. glutamicum*'s WhcE presents homology to *S. avermitilis* WblE (72.9 % identity), which is regulated by Rex (Liu, Cheng et al. 2017). The intergenic region between *whcE* and *cgp_0879* showed two divergent promoters (*PwhcE* and *Pcgp_0879*) originating a *whcE* transcript with a ribosome binding site and a leaderless *cgp_0879* transcript, respectively (Pfeifer-Sancar, Mentz et al. 2013) (Fig. S2C). The ROP-like motif is located between the -51 and the -68 bp region upstream the transcriptional start site of *whcE* transcript, partially overlapping the annotated *whcE* promoter region. Given the distance of the ROP-like sequence to the *whcE* transcriptional start site, it is unlikely Rex expression would significantly affect *whcE* expression.

In silico analysis suggested that a NADH-dependent Rex-based regulatory network in *C. glutamicum* would have a limited impact on cells physiology. To implement a gene expression system regulated by Rex in *C. glutamicum*, we used the previously developed BCA strain that harbors the *attB* sites for ϕ C31 and ϕ BT1 (Marques, Luzhetskyy et al. 2020). The Rex protein encoding gene from *Thermus aquaticus* was integrated into the *C. glutamicum* chromosome at the *attB*_{C31} site whereas Rex-controlled expression was carried out at the *attB*_{BT1} site.

The transcriptional unit containing the JVR promoter (Rytter, Helmark et al. 2014), the optimized RBS and the codon-optimized coding sequence of the Rex protein (NCBI acc. WP_011173693) was integrated at the ϕ C31 *attB* site of *C. glutamicum* BCA genome using a markerless phage-based methodology as previously described (Marques, Luzhetskyy et al. 2020) (Fig. S1A). The generated markerless strain was confirmed by sequencing and it was thereafter named *C. glutamicum* REX strain (Fig. S1A).

For the Rex-controlled gene expression at the *attB*_{BT1} locus, the Rex operator (ROP) site was cloned in-between the strong P69 promoter (Zhang, Liu et al. 2018) and the RBS of the ϕ BT1-based integrative vector pFM69SD (Marques, Luzhetskyy et al. 2020). The

resulting vector was named pFM69ROP (Fig. S1B) and features a transcriptionally insulated cloning region with the ROP site located at 11-32 bp downstream to the putative transcriptional start site and 53-22 bp upstream to the start codon ATG.

To unveil ectopic effects due to the presence of the heterologous Rex system in *C. glutamicum*, the strains BCA and REX were transformed with the pFM69ROP plasmid, integration was confirmed by PCR (data not shown) and the originated strains, BCA [pFM69ROP] and REX [pFM69ROP] (Fig. 1A), characterized. The results showed that the growth pattern of REX [pFM69ROP] was not significantly affected when compared to BCA [pFM69ROP] (Fig. 1B). The constitutive expression of the Rex encoding gene in the REX [pFM69ROP] strain was confirmed by RT-PCR along the growth curve (Fig. 1C). In addition, the analysis of total protein cell-free extracts of REX [pFM69ROP] strain by SDS-PAGE showed the presence of a band with a molecular mass similar to the predicted molecular mass of Rex protein (23.2 kDa) which was not observed in the protein profile of the BCA [pFM69ROP] strain (Fig. 1D).

In silico analysis for the presence of ROP-like sequences in the genome of *C. glutamicum* revealed the occurrence of a putative ROP site within the promoter region of the *whcE* gene, a transcriptional factor involved in redox sensing (Kim, Park et al. 2005). The influence of the presence of Rex protein in the expression of *whcE*, was assessed indirectly by a disk-diffusion assay to assert the susceptibility of both strains to the oxidative stressor menadione sodium bissulfate (Fig. 1E). The assay showed that REX strain was slightly more susceptible to oxidative stress caused by exposition to menadione sodium bissulfate than the BCA strain. This result suggests that Rex might recognize the identified ROP-like site in *C. glutamicum* located upstream *whcE*.

Overall, the results showed that implementation of a Rex-based system in *C. glutamicum* did not affect significantly the physiology of the cell and suggested it can have an orthogonal behaviour on the regulation of the expression of heterologous genes. Moreover, the [NADH]/[NAD⁺] variations (Fig S3) along the growth curve are compatible to the induction of the Rex system.

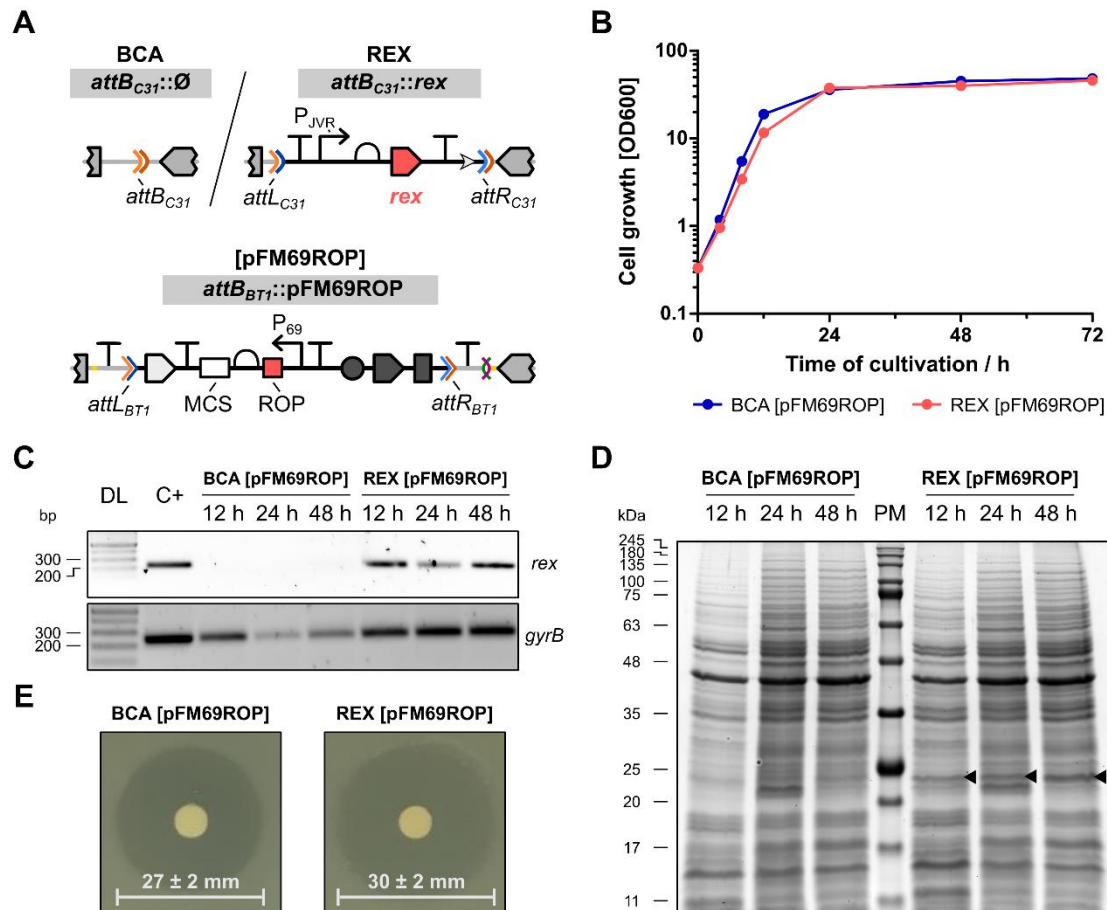


Fig. 1 - Implementation of a Rex regulatory system in *C. glutamicum*. (A) Schematic representation of the *attB*_{C31} locus in strains BCA and REX, and the *attB*_{BT1} locus in strains harboring pFM69ROP. For details on the generation of the strains, refer to Figure S1. (B) Growth curves of BCA [pFM69ROP] and REX [pFM69ROP] strains grown in CGXII medium with 4% (w/v) glucose, at 30°C. The data represents the average of two experimental replicates; standard deviation plotted as error bars, but not visible. (C) RT-PCR analysis of the target gene *rex* and the reference gene *gyrA* visualized by DNA gel electrophoresis. DL - DNA ladder GeneRuler DNA Ladder Mix; C+ - positive control where genomic DNA of REX [pFM69ROP] strain was used as template. (D) SDS-PAGE analysis of cell-free protein extracts of BCA [pFM69ROP] and REX [pFM69ROP] strains. Twenty micrograms of total soluble proteins were loaded per lane. PM - protein marker NZYColour Protein Marker II. (E) Disk diffusion test to assess the susceptibility to menadione sodium bisulfate (1.1 mg per disk) after incubation at 30°C for 24 h. Diameter and error represent the average and standard deviation of two experimental replicates.

Influence of the Rex regulatory system on the heterologous production of indigoidine in *C. glutamicum*

To determine the effectiveness of the heterologous Rex-based system to regulate gene expression *in vivo*, a reporter system based on the indigoidine biosynthetic gene BpsA was implemented in *C. glutamicum* REX. The reporter activity of BpsA is based on the capacity of ATP-dependent conversion of L-glutamine to an intermediate which spontaneously

dimerizes to yield the blue DMSO-soluble pigment indigoidine (Fig. S4). Indigoidine remains inside *C. glutamicum* cells and contributes to the total biomass of the culture (Ghiffary, Prabowo et al. 2021).

The plasmid pFM69ROP-bpsA (Fig. S1C) encoding the P69 promoter, the ROP site and the *bpsA* gene was integrated into the *attB_{BT1}* locus of the *C. glutamicum* BCA and REX strains (Fig. 2A). Site-specific integration of pFM69ROP-bpsA was confirmed by PCR (data not shown). Both generated strains, BCA [pFM69ROP-bpsA] and REX [pFM69ROP-bpsA], were cultivated in defined CGXII media with 4% (w/v) glucose as carbon source, and characterized regarding total biomass accumulation (cell and indigoidine production) and *bpsA* transcription (Fig. 2B, 2C).

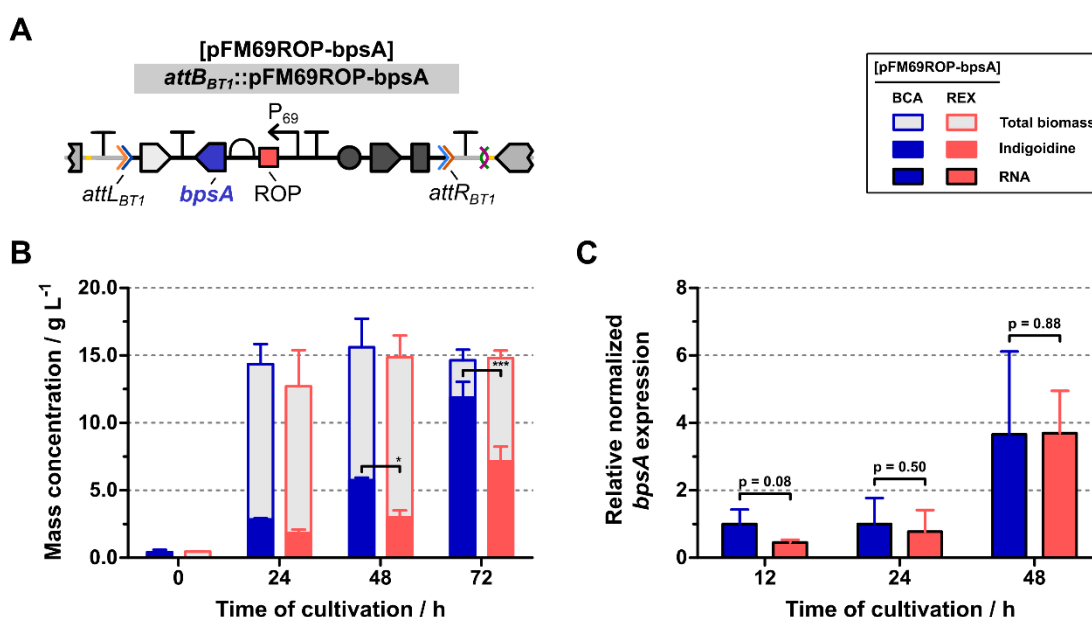


Fig. 2 - Role of the Rex regulatory system on the production of indigoidine in *C. glutamicum*.

Strains BCA [pFM69ROP-bpsA] (blue bars) and REX [pFM69ROP-bpsA] (red bars) were cultivated in CGXII medium with 4% (w/v) glucose at 30°C for 72 h. **(A)** Schematic representation of the genomic arrangement at the *attB_{BT1}* locus in BCA [pFM69ROP-bpsA] and REX [pFM69ROP-bpsA] strains. For details on the generation of the strains, refer to Figure S1. **(B)** Total biomass accumulation (grey bars) and indigoidine content (blue/red bars) of BCA [pFM69ROP-bpsA] and REX [pFM69ROP-bpsA] strains along the growth curve. **(C)** Relative fold expression of *bpsA*, determined by RT-qPCR, normalized to the reference genes *gyrB* and *pccB* (M value 0.396) and to strain BCA [pFM69ROP-bpsA] at 12 h as the control situation. Bars and error bars represent the average and the standard deviation of two experimental/biological replicates, respectively. Statistically significant differences: * $p < 0.05$, *** $p < 0.001$

The results show that both strains presented similar total biomass accumulation along the growth curve (Fig. 2B). To infer the transcriptional activity of the Rex-controlled promoter in *C. glutamicum* along the growth curve, the production of indigoidine catalyzed by the

regulated reporter gene *bpsA* was monitored (Fig. 2B). The results showed that, in these cultivation conditions, the REX [pFM69ROP-*bpsA*] strain achieved 39%, 51% and 41% lower yields of indigoidine at 24 h, 48 h and 72 h, respectively, in comparison to the BCA [pFM69ROP-*bpsA*] strain.

To further investigate the impact of Rex on the heterologous expression of *bpsA*, samples were collected along the growth curve to assess expression of *bpsA* by RT-qPCR (Fig. 2C). The results showed that the relative abundance of *bpsA* transcripts increased along the time of cultivation and most importantly, the REX [pFM69ROP-*bpsA*] strain displayed a 0.5- and a 0.8-fold down-regulation of *bpsA* expression at 12 h and 24 h, respectively. At 48 h, both strains presented essentially the same relative abundance of *bpsA* transcript.

The results suggest that, in these cultivation conditions, Rex down-regulates the transcription of genes downstream ROP site at early cultivation stages of *C. glutamicum*, leading to lower production of the reporter molecule indigoidine. This observation is in accordance with the [NADH]/[NAD⁺] ratios determined for *C. glutamicum* (Fig. S3) with lower ratios at early growth stages. Moreover, integration of pFM69ROP-*bpsA* in *Streptomyces tsukubaensis* showed that the synthetic ROP site is functional in other actinobacteria that encode Rex regulators natively (Mendes et al., unpublished results / personal communication).

Effect of Rex regulation on the production of indigoidine in different carbon sources and carbon-to-nitrogen ratios

The regulation of gene expression by Rex regulators is influenced by the redox status of cells. Different carbons sources and different carbon-to-nitrogen ratios play a role in the redox state of cells, and, importantly, on the availability of building blocks for primary and secondary metabolism (Wehrs, Prahel et al. 2018).

To study the influence of the carbon sources on the Rex-regulated reporter system, the production of indigoidine was qualitatively evaluated in complex solid media, supplemented with glucose or maltose (Fig. 3). For that, generated *C. glutamicum* strains (BCA [pFM69ROP], REX [pFM69ROP], BCA [pFM69ROP-*bpsA*] and REX [pFM69ROP-*bpsA*]) together with the previously generated strain REX [pFM69SD-*bpsA*] (Marques, Luzhetskyy et al. 2020) that harbours plasmid pFM69SD-*bpsA* encoding the *bpsA* gene without the ROP site (i.e. the *bpsA* expression is not regulated by REX), were streaked as patches on lysogeny agar (LA) supplemented with 1 or 3% (w/v) of selected carbon sources and incubated for 72 h at 30°C. After 72 hours of incubation, all *bpsA* positive strains become dark blue in LA (Fig. 3A), LA with 1% (w/v) glucose (Fig. 3B), and LA with 1% (w/v) maltose (Fig. 3D). In LA with 3% (w/v) glucose (Fig. 3C), all *bpsA* positive strains presented a very

light blue hue suggesting a impaired indigoidine production. In LA with 3% (w/v) maltose (Fig. 3E), the *bpsA* positive strains with an incomplete Rex system (either ROP only or REX only) displayed also an impaired indigoidine production. However, the strain with complete Rex system (REX [pFM69ROP-*bpsA*]) presented a dark blue coloration, indicative of normal indigoidine production. Interestingly, after additional time (13 days) at room temperature, the REX [pFM69ROP-*bpsA*] strain was the only of the *bpsA* positive strains not originating a dark blue patch in LA with 3% (w/v) glucose (Fig. S5). Overall, the presence of the complete Rex-based regulatory system seems to be important to enable production of indigoidine in complex media with high maltose concentrations.

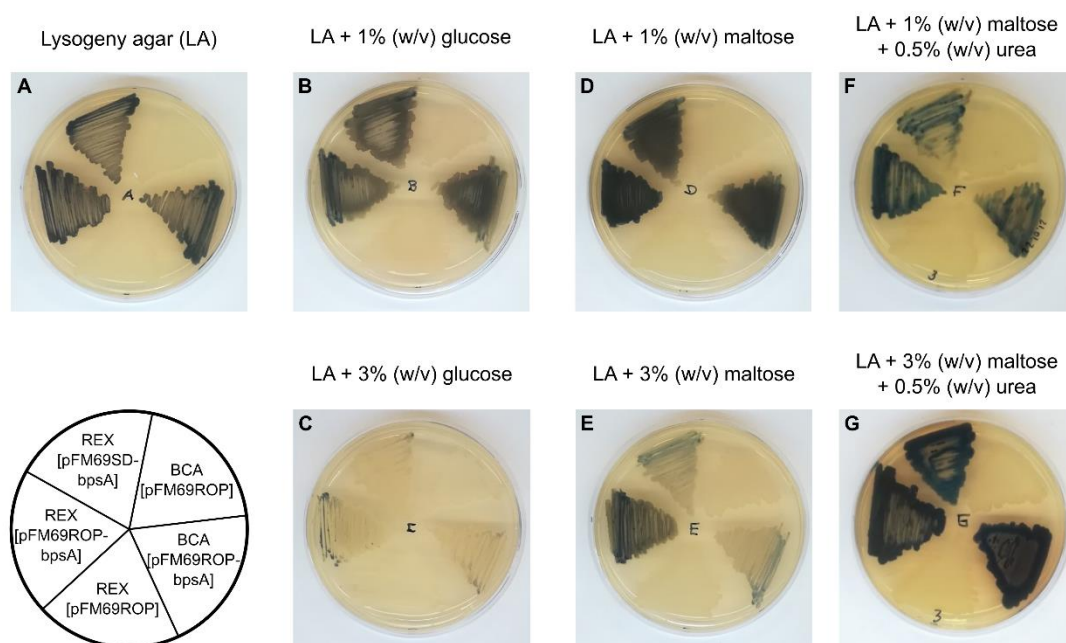


Fig. 3 - Effect of carbon and nitrogen sources on the Rex-dependent production of indigoidine by *C. glutamicum*. Strains of *C. glutamicum* were inoculated on lysogeny agar (LA) supplemented with selected carbon and nitrogen sources and incubated at 30°C for 72 h. Blue pigmentation indicates the production of indigoidine.

Previous works in *R. toruloides* showed an optimal 8:1 carbon-to-nitrogen (C/N) ratio for the production of indigoidine; at C/N 4:1 and 40:1, the cell growth observed was similar to C/N 8:1, but indigoidine production was reduced by 50% (Wehrs, Gladden et al. 2019). Considering our experimental setup and based on the quantification by (Sezonov, Joseleau-Petit et al. 2007), the C/N ratio of non-supplemented LB medium is 3:1, of LB supplemented with 1% (w/v) glucose or maltose is about 5:1 and of LB supplemented with 3% (w/v) maltose is 12:1. These values suggested that the impaired production of indigoidine in LB

supplemented with 3% (w/v) glucose or maltose was due to a non-optimal C/N ratio for the production of indigoidine. To corroborate this association between the C/N ratio and the production of indigoidine the strains were streaked as patches on lysogeny agar with 1 or 3% (w/v) maltose supplemented with 0.5% (w/v) urea, and incubated for 72 hours, at 30°C (Fig. 3F-G). The addition of urea reduced the C/N ratios, in particular, the C/N ratio of LB supplemented with 3% (w/v) maltose and 0.5% (w/v) urea that shifted from 12:1 to about 5:1. The results show the addition of urea to media with high maltose concentration recovered the production of indigoidine in the *bpsA* positive strains with incomplete Rex system. Notably, urea supplementation in media with low maltose concentrations lead to a phenotype closer to high maltose concentrations without urea supplementation. The results suggested that the presence of a full Rex system (Rex protein and ROP site) is associated with the maintenance of indigoidine production in wider ranges of carbon-to-nitrogen ratios. Overall, the presence of the complete Rex-based regulatory system seems to be important to enable production of indigoidine in complex media with high maltose concentrations, especially at earlier cultivation stages.

Quantitative determination of the effect of Rex regulation on the production of indigoidine in complex media with maltose

To quantify the effects of the Rex system in high-maltose fermentations observed in agar plates, the strains were cultivated in shake flasks with monitorization of biomass accumulation and indigoidine production (Fig. 4). Preliminary studies in LB medium without supplementation of additional carbon and/or nitrogen sources had shown that the BCA [pFM69ROP-*bpsA*] and REX [pFM69ROP-*bpsA*] presented already some differences in biomass accumulation and the latter achieved 1.7-fold higher indigoidine production at 72 h (Fig. S6).

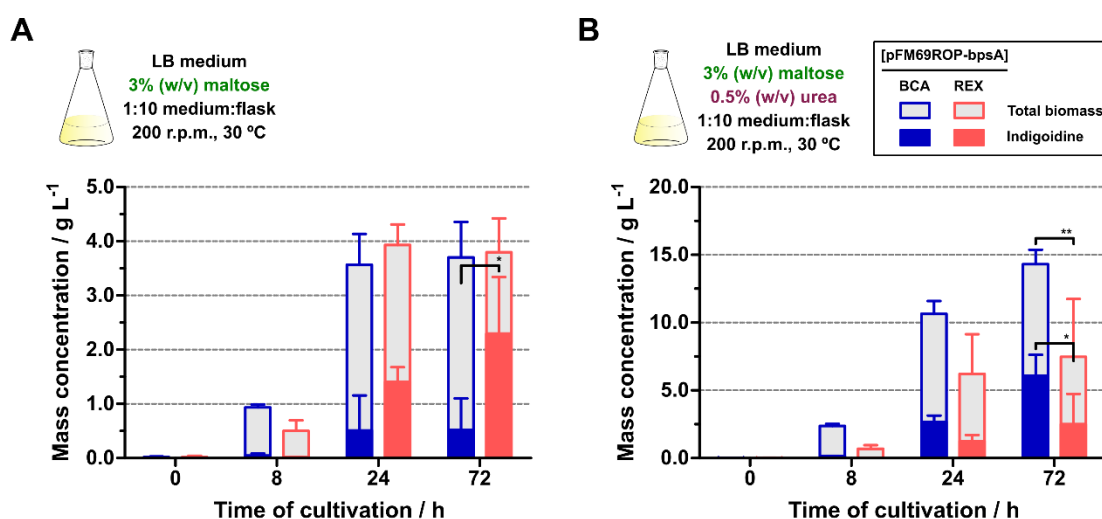


Fig. 4 - Effect of Rex regulation on the production of indigoidine in *C. glutamicum* in complex media with different carbon-to-nitrogen ratios. Strains BCA [pFM69ROP-bpsA] (blue bars) and REX [pFM69ROP-bpsA] (red bars) were cultivated in lysogeny broth supplemented with 3% (w/v) maltose **(A)** and 0.5% (w/v) urea **(B)**, at 30°C for 72 h. Stacked bar graphs represent the total biomass (grey bars) and the indigoidine content (blue/red bars) of the strains along the growth curve. Total biomass (grey bars) and indigoidine (blue/red bars) were determined by cell dry weight and by spectrophotometry of DMSO extracts, respectively. Bars and error bars represent the average and the standard deviation of two experimental replicates, respectively. Statistically significant differences: * $p < 0.05$; ** $p < 0.01$.

The results show that the BCA and REX strains harbouring pFM69ROP-bpsA grow alike in LB + 3% (w/v) maltose (Fig. 4A) and furthermore, do not growth differently from *bpsA* negative strains (data not shown). In comparison to LB without additional carbon source supplementation (Fig. S6), strains BCA and REX harbouring pFM69ROP-bpsA accumulated 2.1- and 2.8-fold more biomass in the presence of maltose, respectively. Regarding indigoidine production, in LB with 3% (w/v) maltose, the results showed that, at 72 h, the REX [pFM69ROP-bpsA] achieves a 4.6-fold higher indigoidine production than the BCA counterpart (Fig. 4A). This result is in-line with the qualitative evaluation of indigoidine production in lysogeny agar with 3% (w/v) maltose (Fig. 3). Overall, maltose addition seemed to enhance the overall biomass accumulation of both strains in complex media, of which indigoidine accounted for a greater percentage in case of REX strain (61%) than in BCA strain (14%) at 72h.

When cultivated in LB supplemented with 3% (w/v) maltose and 0.5% (w/v) urea, the strain BCA [pFM69ROP-bpsA] achieved 1.9-fold higher biomass than the REX [pFM69ROP-bpsA] strain at 72 h (Fig. 4B). Notably, the supplementation with urea led to an increase of total biomass for both strains at 72 h. Regarding indigoidine production, in LB with 3% (w/v) maltose and 0.5% (w/v) urea, the results showed that, at 72 h, the REX [pFM69ROP-bpsA] achieves 0.4-fold the indigoidine production than the BCA counterpart (Fig. 4B). Noteworthy, the strain BCA [pFM69ROP-bpsA] had accumulated substantially more indigoidine (12-fold) in LB supplemented with urea compared with no added urea. These results are in-line with the qualitative observations in solid medium (Fig. 3). In case of REX [pFM69ROP-bpsA], the addition of urea did not led to a significant increase in indigoidine production. In other words, about 52% of the gain of biomass of BCA [pFM69ROP-bpsA] associated to urea supplementation corresponded to indigoidine while in the case of REX [pFM69ROP-bpsA] strain only 6% of the additional biomass associated to urea supplementation corresponded to indigoidine, the remaining 94% was associated to cell growth.

Overall, the results suggested that REX strain produce more indigoidine than the BCA counterpart in media with higher carbon-to-nitrogen ratio, and, conversely, that additional nitrogen sources did not increase indigoidine production in REX strain (as it does in BCA strain). Furthermore, the presence of Rex system seems to improve the allocation of limited nitrogen sources for both cell growth and indigoidine biosynthesis, as additional nitrogen does not increase production but only non-indigoidine biomass.

Discussion

In this study we have developed a Rex-based switch that is able to integrate *C. glutamicum* redox status for the optimization of the heterologous production of indigoidine in non-optimal culture conditions. Although Rex regulators are widely spread among bacteria (Ravcheev, Li et al. 2012), no Rex ortholog encoding gene was identified in members of the *Corynebacteriaceae* family. Unlike other facultative anaerobes, such as *Staphylococcus aureus*, where Rex is a central regulator of anaerobic metabolism (Pagels, Fuchs et al. 2010), or *Bacillus subtilis*, where Rex plays a role in coordinating anaerobic and high-affinity aerobic mechanisms of regenerating NAD⁺ (Larsson, Rogstam et al. 2005), *C. glutamicum* seems to have not adopted Rex as regulator for adaptation to low oxygen availability nor for other purposes. The absence of a native Rex protein would favour the implementation of a Rex-based redox-sensitive regulatory system in *C. glutamicum* with minimal interference with the host's genetic background. However, *in silico* analysis of the *C. glutamicum* genome identified a ROP-like site in the promoter region of the *whcE* gene that codes for a WhiB-like regulator. WhiB-like regulators are conserved targets of Rex in other actinobacteria (Ravcheev, Li et al. 2012) and the occurrence of a ROP-like sequence in the promoter region of *whcE* in *C. glutamicum* suggested that *whcE* transcription could be regulated in the presence of a Rex protein. In *C. glutamicum*, the expression of *whcE* is regulated by the sigma factor SigH (Kim, Park et al. 2005) and *WhcE* deficiency was linked to prolonged lag phase in minimal media, early death in complex media, and increased susceptibility to thiol-specific oxidants and redox-cycling compounds (Kim, Park et al. 2005). Upon implementation of the genetic device to express Rex and a mock expression vector encoding a ROP sequence in *C. glutamicum* BCA (Fig. 1A), the resulting strain presented a slightly longer lag phase in CGXII media with 4% glucose with normal behaviour in late stationary phase in complex media (Fig. 1B), and slight increase in susceptibility to menadione sodium bisulfite (Fig. 1E). The observed phenotypes of REX [pFM69ROP] strain correlate with the downregulation of *whcE* expression at early growth stages but not at late cultivation stages. This behaviour agrees with the Rex regulation of *whcE* expression dependent on the intracellular NADH/NAD⁺ ratio (Fig. S3) i.e., the lower intracellular

NADH/NAD⁺ ratio at early timepoints would favour DNA-binding activity of Rex and conversely, the increase of the NADH/NAD⁺ ratio at 48h would reduce Rex affinity to DNA, derepressing gene expression. These results suggested an interference of the non-native Rex protein with the host's gene expression; however, the results also indicate that the implemented Rex-based switch is able to sense the *C. glutamicum* intracellular NADH/NAD⁺ redox balance and regulate gene expression accordingly (Brekasis and Paget 2003). Overall, the interaction of the implemented Rex system with the host genetic background caused little disturbance to cells' physiology showing that the redox responsive Rex-based switch can be used for controlling foreign genetic circuits in *C. glutamicum*.

Experimental evidence that the Rex-based switch is able to regulate the expression of heterologous genes in *C. glutamicum* in response to the intracellular NADH/NAD⁺ ratio was also provided by the expression of the indigoidine synthetase encoding gene *bpsA* in *C. glutamicum* BCA and REX strains (Fig. 2). The *BpsA* encoding gene was placed downstream a constitutive promoter and a Rex operator (ROP) sequence, integrated into the *attB_{BT1}* locus and the generated strains grown on CGXII defined medium with 4 % (w/v) glucose for 72 h. Characterization of both strains showed a down-regulation of *bpsA* expression at early time points in REX [pFM69ROP-*bpsA*] when compared to BCA [pFM69ROP-*bpsA*] strain where the Rex protein is absent. Taking in consideration the intracellular NADH/NAD⁺ profile of *C. glutamicum* (Fig. S3), this result is in agreement with a redox-responsive regulation of *bpsA* expression in the REX strain. At early time points where the NADH/NAD⁺ ratio is lower, Rex would repress *bpsA* transcription while as the NADH/NAD⁺ ratio increases along the growth curve, the DNA-binding activity of Rex would be inhibited allowing *bpsA* transcription. The results on *bpsA* expression are reflected in the indigoidine production as the REX [pFM69ROP-*bpsA*] strain displays a lower indigoidine production rate at early time points when compared to BCA [pFM69ROP-*bpsA*]. Overall, the downregulation of *bpsA* expression in REX [pFM69ROP-*bpsA*] strain resulted in a lower accumulation of indigoidine at 72h when compared to BCA [pFM69ROP-*bpsA*] (7.16 g L⁻¹ vs 11.90 g L⁻¹ respectively – Fig. 2B). Curiously, as indigoidine remains inside *C. glutamicum* cells and hence it contributes to the total biomass of the culture (Ghiffary, Prabowo et al. 2021), the similar values of total biomass (cells plus indigoidine) in both strains at 72 h (Fig. 2B) suggested that the REX [pFM69ROP-*bpsA*] strain culture accumulates more cells (i.e. non-indigoidine biomass) at 72 h. We suggest that the constitutive expression of *bpsA* in BCA [pFM69ROP-*bpsA*] strain, and consequently of indigoidine production, redirects glutamine (an essential building block of proteins and nucleic acids) towards the biosynthesis of indigoidine and drained the glutamine pool needed for the biosynthesis of cell biomass precursors. The outcome is that 81 % (w/w) of

the total biomass of BCA [pFM69ROP-bpsA] strain at 72 h corresponds to indigoidine *versus* 49 % (w/w) in strain REX [pFM69ROP-bpsA]. By itself, the delay on the onset of indigoidine production in the presence of Rex could be considered unnecessary. However, by repressing *bpsA* transcription at early cultivation stages, Rex allows the use of glutamine pool for the synthesis of proteins and nucleotides, much needed during the exponential growth phase of bacteria. When cells enter stationary phase and Rex DNA-binding activity is inhibited by the increase in the NADH/NAD⁺ ratio, BpsA is produced and indigoidine accumulates at fast pace while maintaining cell fitness. In summary, Rex seems to be suitable to repress (heterologous encoded) secondary metabolism during exponential growth phase, saving building blocks for cellular components, and to induce expression of secondary metabolites in response to the reductive intracellular state at early stationary phase.

The high indigoidine production levels of strain BCA [pFM69ROP-bpsA] led us to revisit our previous work where we also expressed, in the same culture conditions, the *bpsA* gene in the BCA strain but without including the ROP site in the promoter region (Marques, Luzhetskyy et al. 2020). The addition of the ROP site in-between the P69 promoter and the Shine-Dalgarno sequence was associated with the 1.8-, 3.3-, and 1.6-fold higher indigoidine titer at 24, 48 and 72 hours of cultivation, respectively, than the parental genetic circuit pFM69SD-bpsA (Marques, Luzhetskyy et al. 2020). One explanation for the additional indigoidine production is the extension of the half-life of the transcript. Upon transcription, the 5'-untranslated region (5'-UTR) of ROP-*bpsA* originates an hairpin, where the inverted repeats of ROP form the stem, followed by a 14-bp spacer and the Shine-Dalgarno sequence. Interestingly, the described 5'-UTR closely matches the configuration of the engineered 5'-UTR "2h", characterized in *E. coli*, that extends 3-fold the mRNA half-life and increases 5-fold the expression level of a downstream gene in comparison to an 5'-UTR without harpin and spacer upstream SD sequence (Viegas, Apura et al. 2018). In this sense, the inclusion of ROP as transcriptional regulatory element in the pFM69ROP-series plasmids affected translation by extending the half-life of the mRNA.

For the routine manipulation and propagation of *C. glutamicum* strains, we used Lysogeny Broth or Agar (LB or LA) medium. We noticed that regarding indigoidine production, the behaviour of BCA [pFM69ROP-bpsA] and REX [pFM69ROP-bpsA] strains in LB/LA medium did not agree with the results obtained for the defined medium CGXII. In LB/LA medium, REX [pFM69ROP-bpsA] strain reached higher titres (Fig. 3A and Fig. S6).

The ATP-dependent production of indigoidine by the non-ribosomal peptide synthetase BpsA requires glutamine as biosynthetic precursor and FAD as coenzyme (Fig. S4). In *Saccharomyces cerevisiae*, indigoidine heterologous production was correlated with the

respiratory metabolism and the activity of the TCA cycle. The use of fermentable carbon sources was associated with a delay of indigoidine production (Wehrs, Prahl et al. 2018). Additionally, expression of *bpsA* in *R. toruloides* showed that indigoidine production was dependent on the carbon-to-nitrogen (C/N) ratio. An optimal C/N for production of indigoidine was found to be 8:1 (Wehrs, Gladden et al. 2019). Based on the quantification by Sezonov, Joseleau-Petit et al. (2007), we calculated that non-supplemented LB medium has a C/N of about 3:1 which is non-optimal for indigoidine production. In complex medium, the results suggested that the REX [pFM69ROP-*bpsA*] strain was able to produce indigoidine more efficiently. Nonetheless, the *C. glutamicum* strains expressing *bpsA* constitutively can produce indigoidine at 72 hours of cultivation on LA with no or low supplementation of carbon sources, but not in media supplemented with high concentration of sugars (Fig. 3). The increase of the C/N ratio to non-optimal conditions in LA supplemented with high concentration of carbon sources suppressed the indigoidine production due to the limited availability of nitrogen for growth and indigoidine biosynthesis (Wehrs, Gladden et al. 2019). The reduction of the C/N ratio by co-supplementing urea reversed the suppression of indigoidine production. These observations are in-line with the previously reported importance of the C/N ratio for the production of indigoidine.

Curiously, the replacement of glucose for maltose as carbon source in LA medium, resulted in a different indigoidine production phenotype in the REX [pFM69ROP-*bpsA*] strain when compared to the strains constitutively expressing *bpsA* i.e., REX [pFM69SD-*bpsA*] and BCA [pFM69ROP-*bpsA*] strains. In complex media with high maltose concentration the *C. glutamicum* strain carrying *bpsA* under the control of Rex produced indigoidine at 72 hours while the indigoidine production was visibly downregulated on the strains where *bpsA* expression was not controlled by Rex (Fig. 3E and Fig. 4A). As described before, a high C/N ratio seems to be important to pinpoint the differences between low and high glucose concentration but offers a rather insufficient explanation for the difference between the carbon sources. We hypothesize that the assimilation of maltose would alter redox status of the cell at early timepoints, conditioning Rex-dependent expression of indigoidine biosynthetic machinery.

In *C. glutamicum*, maltose is known to be transported into *C. glutamicum* by the ATP-dependent ABC transporter system MusEFGK2I and metabolized to glucose in the process of formation of maltodextrins and, later, in the process of degradation of maltodextrins (Seibold, Wurst et al. 2009, Henrich, Kuhlmann et al. 2013). The rate at which *C. glutamicum* converts maltose to cell biomass was reported to be similar to glucose during early and mid-exponential growth phases (Krause Felix, Henrich et al. 2010). Despite being assimilated at the same rate maltose undergoes additional transformations that may impose

flux bottlenecks (Seibold, Wurst et al. 2009) or may divert it to short-term reserve or functional carbohydrates such as glycogen and trehalose (Wolf, Krämer et al. 2003, Seibold and Eikmanns 2013). Thus, maltose could be assimilated at a high P/O ratio, mostly based on electron transport chain-dependent energy generation and with low intracellular NADH/NAD⁺ ratios. In consequence, the Rex-dependent expression of indigoidine synthetase (BpsA) would be down-regulated during exponential (aerobic) growth phase. Counterintuitively, by suppressing expression of BpsA at exponential phase, Rex could favour indigoidine overproduction at later stages of cultivation, especially in high C/N ratio media. By delaying BpsA expression, especially in nitrogen limited media, Rex system saves building block fundamental for cell growth at exponential phase and for indigoidine biosynthesis at the stationary phase. This is especially important as NRPS in general are known to have low turnover rates (Brown, Robins et al. 2017), and, so, anticipating its expression would imply a huge effort in terms of protein synthesis (including building blocks and energy) but translated to moderate gains in product accumulation. This hypothesis explains why strains expressing *bpsA* constitutively (BCA [pFM69ROP-*bpsA*] and REX [pFM69SD-*bpsA*]) visibly produced less indigoidine than the strain where *bpsA* expression was controlled by Rex. The early expression of BpsA imposes an intense channelling of the limited nitrogen supplied by LA/LB with 3% maltose to heterologous protein expression which compromises the pool of precursors for the actual production of indigoidine. In fact, the increased indigoidine production in BCA [pFM69ROP-*bpsA*] strain upon urea supplementation (Fig. 4B) suggests a relief on the bottleneck for nitrogen sources supporting the idea that the presence of the Rex system can aid the cell in managing the nutrient flux in nitrogen-limited conditions. Curiously, strains BCA [pFM69ROP-*bpsA*] and REX [pFM69SD-*bpsA*] recovered indigoidine production on long-term cultivation in high carbon sources (Fig. S5). This could be explained by the fact BpsA enzyme continued to be expressed and precursors become available by the degradation of other proteins under primary carbon source limitation (Paczia, Nilgen et al. 2012). Conversely, the inability of REX [pFM69ROP-*bpsA*] to sustain indigoidine production on the long-term in LA with 3% glucose, suggests that carbon sources (glucose/maltose) do affect differently the NADH/NAD⁺ ratio, conditioning BpsA expression. In case of LA media with 3% glucose, BpsA expression controlled by Rex system would be induced at early stages (due to high NADH/NAD⁺) with limited success in producing indigoidine (due to lack of nitrogen, as explored above) and would be repressed at late stages (due to low NADH/NAD⁺), when remobilized nitrogen could be used for indigoidine synthesis.

For industrial purposes, the use of maltose and its polymeric versions (dextrins and starch, with additional amylases) may prove useful as carbon sources to obtain natural products

from sustainable feedstocks. Our results show that the redox-sensing Rex-based genetic switch could be helpful to delay the expression of heterologous products in such a way that their biosynthesis does not compete with cell growth and fitness, i.e., mimicking the onset of secondary metabolism of native producers at late exponential phase. In short, Rex regulation in *C. glutamicum* helps the cell to manage its own metabolic resources to face the metabolic burden exerted by heterologous expression.

Acknowledgements

The authors acknowledge the technical support of the i3S Scientific Platform “Cell Culture and Genotyping” in the RT-qPCR experiments.

Author contribution

Filipe Marques (FM), Andriy Luzhetskyy (AL) and Marta Vaz Mendes (MVM) conceived and designed the experimental work. FM and MVM conducted experiments, analysed data and wrote the manuscript. Paula Tamagnini (PT), AL and MVM were responsible for funding acquisition. AL and MVM were responsible for supervision. AL and PT reviewed and edited the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by National Funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the project ERA-IB-2/0001/2015; fellowships SFRH/BD/131939/2016 (FM) and SFRH/BPD/95683/2013 (MVM) and the FCT contract DL57/2016/CP1355/CT0023. The work was also supported by FEDER - Fundo Europeu de Desenvolvimento Regional funds through the COMPETE 2020 - Operacional Programme for Competitiveness and Internationalisation (POCI), Portugal 2020, and by Portuguese funds through FCT - Fundação para a Ciência e a Tecnologia/Ministério da Ciência, Tecnologia e Ensino Superior in the framework of the project POCI-01-0145-FEDER-031045. The funders of the study had no role in study design, data collection, data analysis, data interpretation, writing of the report or decision to submit the article for publication.

Declarations

Ethics approval

This article does not contain any studies with animals or with human participants.

Conflict of interest

The authors declare no competing interests.

References

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-410.
- Beites, T., P. Oliveira, B. Rioseras, S. D. S. Pires, R. Oliveira, P. Tamagnini, P. Moradas-Ferreira, A. Manteca and M. V. Mendes (2015). "*Streptomyces natalensis* programmed cell death and morphological differentiation are dependent on oxidative stress." Scientific Reports **5**: 12887.
- Boecker, S., S. Gratz, D. Kerwat, L. Adam, D. Schirmer, L. Richter, T. Schutze, D. Petras, R. D. Sussmuth and V. Meyer (2018). "Aspergillus niger is a superior expression host for the production of bioactive fungal cyclodepsipeptides." Fungal Biol Biotechnol **5**: 4.
- Bott, M. and A. Niebisch (2003). "The respiratory chain of *Corynebacterium glutamicum*." J Biotechnol **104**(1-3): 129-153.
- Brekasis, D. and M. S. Paget (2003). "A novel sensor of NADH/NAD⁺ redox poise in *Streptomyces coelicolor* A3(2)." EMBO J **22**(18): 4856-4865.
- Brown, A. S., K. J. Robins and D. F. Ackerley (2017). "A sensitive single-enzyme assay system using the non-ribosomal peptide synthetase BpsA for measurement of L-glutamine in biological samples." Scientific Reports **7**(1): 41745.
- Burg, J. M., C. B. Cooper, Z. Ye, B. R. Reed, E. A. Moreb and M. D. Lynch (2016). "Large-scale bioprocess competitiveness: the potential of dynamic metabolic control in two-stage fermentations." Current Opinion in Chemical Engineering **14**: 121-136.
- Dudek, C. A. and D. Jahn (2022). "PRODORIC: state-of-the-art database of prokaryotic gene regulation." Nucleic Acids Res **50**(D1): D295-D302.
- Durante-Rodriguez, G., V. de Lorenzo and P. I. Nikel (2018). "A Post-translational Metabolic Switch Enables Complete Decoupling of Bacterial Growth from Biopolymer Production in Engineered *Escherichia coli*." ACS Synth Biol **7**(11): 2686-2697.
- Eggeling, L. and M. Bott (2005). Handbook of corynebacterium glutamicum. Boca Raton, Taylor & Francis.
- Engels, V., S. N. Lindner and V. F. Wendisch (2008). "The global repressor SugR controls expression of genes of glycolysis and of the L-lactate dehydrogenase LdhA in *Corynebacterium glutamicum*." J Bacteriol **190**(24): 8033-8044.
- Ghiffary, M. R., C. P. S. Prabowo, K. Sharma, Y. Yan, S. Y. Lee and H. U. Kim (2021). "High-Level Production of the Natural Blue Pigment Indigoidine from Metabolically

Engineered *Corynebacterium glutamicum* for Sustainable Fabric Dyes." ACS Sustainable Chemistry & Engineering **9**(19): 6613-6622.

Gupta, R. S. (2019). "Commentary: Genome-Based Taxonomic Classification of the Phylum Actinobacteria." Front Microbiol **10**: 206.

Gyan, S., Y. Shiohira, I. Sato, M. Takeuchi and T. Sato (2006). "Regulatory loop between redox sensing of the NADH/NAD(+) ratio by Rex (YdiH) and oxidation of NADH by NADH dehydrogenase Ndh in *Bacillus subtilis*." J Bacteriol **188**(20): 7062-7071.

Hanahan, D. (1983). "Studies on transformation of *Escherichia coli* with plasmids." J Mol Biol **166**(4): 557-580.

Henrich, A., N. Kuhlmann, W. Eck Alexander, R. Krämer and M. Seibold Gerd (2013). "Maltose Uptake by the Novel ABC Transport System MusEFGK2I Causes Increased Expression of ptsG in *Corynebacterium glutamicum*." Journal of Bacteriology **195**(11): 2573-2584.

Hermann, T. (2003). "Industrial production of amino acids by coryneform bacteria." J Biotechnol **104**(1-3): 155-172.

Herrmann, S., T. Siegl, M. Luzhetska, L. Petzke, C. Jilg, E. Welle, A. Erb, P. F. Leadlay, A. Bechthold and A. Luzhetskyy (2012). "Site-specific recombination strategies for engineering actinomycete genomes." Appl. Environ. Microbiol. **78**(6): 1804-1812.

Inui, M. and K. Toyoda (2020). *Corynebacterium glutamicum* - Biology and Biotechnology, Springer Nature Switzerland AG.

Kabashima, Y., J. Kishikawa, T. Kurokawa and J. Sakamoto (2009). "Correlation between proton translocation and growth: genetic analysis of the respiratory chain of *Corynebacterium glutamicum*." J Biochem **146**(6): 845-855.

Kasari, M., V. Kasari, M. Karmas and A. Joers (2022). "Decoupling Growth and Production by Removing the Origin of Replication from a Bacterial Chromosome." ACS Synth Biol **11**(8): 2610-2622.

Keilhauer, C., L. Eggeling and H. Sahm (1993). "Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the ilvB-ilvN-ilvC operon." J Bacteriol **175**(17): 5595-5603.

Kieser, T., M. J. Bibb, M. J. Buttner, K. F. Chater, D. A. Hopwood and F. John Innes (2000). Practical Streptomyces genetics. Norwich, John Innes Foundation.

Kim, T.-H., J.-S. Park, H.-J. Kim, Y. Kim, P. Kim and H.-S. Lee (2005). "The whcE gene of *Corynebacterium glutamicum* is important for survival following heat and oxidative stress." Biochemical and Biophysical Research Communications **337**(3): 757-764.

Kobayashi, S., H. Kawaguchi, T. Shirai, K. Ninomiya, K. Takahashi, A. Kondo and Y. Tsuge (2020). "Automatic Redirection of Carbon Flux between Glycolysis and Pentose Phosphate Pathway Using an Oxygen-Responsive Metabolic Switch in *Corynebacterium glutamicum*." ACS Synth Biol **9**(4): 814-826.

Koch-Koerfges, A., A. Kabus, I. Ochrombel, K. Marin and M. Bott (2012). "Physiology and global gene expression of a *Corynebacterium glutamicum* DeltaF(1)F(O)-ATP synthase mutant devoid of oxidative phosphorylation." Biochim Biophys Acta **1817**(2): 370-380.

Koch-Koerfges, A., N. Pflzer, L. Platzen, M. Oldiges and M. Bott (2013). "Conversion of *Corynebacterium glutamicum* from an aerobic respiring to an aerobic fermenting bacterium by inactivation of the respiratory chain." Biochim Biophys Acta **1827**(6): 699-708.

Krause Felix, S., A. Henrich, B. Blombach, R. Krämer, J. Eikmanns Bernhard and M. Seibold Gerd (2010). "Increased Glucose Utilization in *Corynebacterium glutamicum* by Use of Maltose, and Its Application for the Improvement of L-Valine Productivity." Applied and Environmental Microbiology **76**(1): 370-374.

Kusumoto, K., M. Sakiyama, J. Sakamoto, S. Noguchi and N. Sone (2000). "Menaquinol oxidase activity and primary structure of cytochrome bd from the amino-acid fermenting bacterium *Corynebacterium glutamicum*." Arch Microbiol **173**(5-6): 390-397.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." nature **227**(5259): 680-685.

Larsson, J. T., A. Rogstam and C. von Wachenfeldt (2005). "Coordinated patterns of cytochrome bd and lactate dehydrogenase expression in *Bacillus subtilis*." Microbiology (Reading) **151**(Pt 10): 3323-3335.

Lennox, E. S. (1955). "Transduction of linked genetic characters of the host by bacteriophage P1." Virology **1**(2): 190-206.

Litsanov, B., A. Kabus, M. Bocker and M. Bott (2012). "Efficient aerobic succinate production from glucose in minimal medium with *Corynebacterium glutamicum*." Microb Biotechnol **5**(1): 116-128.

Liu, X., Y. Cheng, M. Lyu, Y. Wen, Y. Song, Z. Chen and J. Li (2017). "Redox-sensing regulator Rex regulates aerobic metabolism, morphological differentiation, and avermectin production in *Streptomyces avermitilis*." Sci Rep **7**: 44567.

- Lo, T. M., S. H. Chng, W. S. Teo, H. S. Cho and M. W. Chang (2016). "A Two-Layer Gene Circuit for Decoupling Cell Growth from Metabolite Production." Cell Syst **3**(2): 133-143.
- Maeda, T., A. Koch-Koerfges and M. Bott (2020). "Relevance of NADH Dehydrogenase and Alternative Two-Enzyme Systems for Growth of *Corynebacterium glutamicum* With Glucose, Lactate, and Acetate." Front Bioeng Biotechnol **8**: 621213.
- Marques, F., A. Luzhetskyy and M. V. Mendes (2020). "Engineering *Corynebacterium glutamicum* with a comprehensive genomic library and phage-based vectors." Metab Eng **62**: 221-234.
- McLaughlin, K. J., C. M. Strain-Damerell, K. Xie, D. Brekasis, A. S. Soares, M. S. Paget and C. L. Kielkopf (2010). "Structural basis for NADH/NAD⁺ redox sensing by a Rex family repressor." Mol Cell **38**(4): 563-575.
- Mizrachi, I. (2007). "Genbank: the nucleotide sequence database." The NCBI handbook [Internet], updated **22**.
- Niebisch, A. and M. Bott (2003). "Purification of a cytochrome bc-aa3 supercomplex with quinol oxidase activity from *Corynebacterium glutamicum*. Identification of a fourth subunit of cytochrome aa3 oxidase and mutational analysis of diheme cytochrome c1." J Biol Chem **278**(6): 4339-4346.
- Novichkov, P. S., A. E. Kazakov, D. A. Ravcheev, S. A. Leyn, G. Y. Kovaleva, R. A. Sutormin, M. D. Kazanov, W. Riehl, A. P. Arkin, I. Dubchak and D. A. Rodionov (2013). "RegPrecise 3.0 - a resource for genome-scale exploration of transcriptional regulation in bacteria." BMC Genomics **14**: 745.
- Oliveira, R., M. J. Bush, S. Pires, G. Chandra, D. Casas-Pastor, G. Fritz and M. V. Mendes (2020). "The novel ECF56 SigG1-RsfG system modulates morphological differentiation and metal-ion homeostasis in *Streptomyces tsukubaensis*." Scientific Reports **10**(1): 21728.
- Paczia, N., A. Nilgen, T. Lehmann, J. Gätgens, W. Wiechert and S. Noack (2012). "Extensive exometabolome analysis reveals extended overflow metabolism in various microorganisms." Microbial Cell Factories **11**(1): 122.
- Pagels, M., S. Fuchs, J. Pané-Farré, C. Kohler, L. Menschner, M. Hecker, P. J. McNamarra, M. C. Bauer, C. von Wachenfeldt, M. Liebeke, M. Lalk, G. Sander, C. von Eiff, R. A. Proctor and S. Engelmann (2010). "Redox sensing by a Rex-family repressor is involved in the regulation of anaerobic gene expression in *Staphylococcus aureus*." Mol Microbiol **76**(5): 1142-1161.
- Pfaffl, M. W. (2001). "A new mathematical model for relative quantification in real-time RT-PCR." Nucleic Acids Res. **29**(9): e45.

Pfeifer-Sancar, K., A. Mentz, C. Ruckert and J. Kalinowski (2013). "Comprehensive analysis of the *Corynebacterium glutamicum* transcriptome using an improved RNAseq technique." BMC Genomics **14**: 888.

Ravcheev, D. A., X. Li, H. Latif, K. Zengler, S. A. Leyn, Y. D. Korostelev, A. E. Kazakov, P. S. Novichkov, A. L. Osterman and D. A. Rodionov (2012). "Transcriptional regulation of central carbon and energy metabolism in bacteria by redox-responsive repressor Rex." J Bacteriol **194**(5): 1145-1157.

Renzette, N. (2011). "Generation of Transformation Competent *E. coli*." Current Protocols in Microbiology **22**(1): A.3L.1-A.3L.5.

Rytter, J. V., S. Helmark, J. Chen, M. J. Lezyk, C. Solem and P. R. Jensen (2014). "Synthetic promoter libraries for *Corynebacterium glutamicum*." Appl Microbiol Biotechnol **98**(6): 2617-2623.

Salis, H. M., E. A. Mirsky and C. A. Voigt (2009). "Automated design of synthetic ribosome binding sites to control protein expression." Nat Biotechnol **27**(10): 946-950.

Seibold, Gerd M. and Bernhard J. Eikmanns (2013). "Inactivation of the phosphoglucomutase gene *pgm* in *Corynebacterium glutamicum* affects cell shape and glycogen metabolism." Bioscience Reports **33**(4): e00059.

Seibold, G. M., M. Wurst and B. J. Eikmanns (2009). "Roles of maltodextrin and glycogen phosphorylases in maltose utilization and glycogen metabolism in *Corynebacterium glutamicum*." Microbiology **155**(2): 347-358.

Seo, S. W., J.-S. Yang, I. Kim, J. Yang, B. E. Min, S. Kim and G. Y. Jung (2013). "Predictive design of mRNA translation initiation region to control prokaryotic translation efficiency." Metabolic Engineering **15**: 67-74.

Sezonov, G., D. Joseleau-Petit and R. D'Ari (2007). "Escherichia coli Physiology in Luria-Bertani Broth." Journal of Bacteriology **189**(23): 8746-8749.

Sickmier, E. A., D. Brekasis, S. Paranawithana, J. B. Bonanno, M. S. Paget, S. K. Burley and C. L. Kielkopf (2005). "X-ray structure of a Rex-family repressor/NADH complex insights into the mechanism of redox sensing." Structure **13**(1): 43-54.

Smith, K. M., K. M. Cho and J. C. Liao (2010). "Engineering *Corynebacterium glutamicum* for isobutanol production." Appl Microbiol Biotechnol **87**(3): 1045-1055.

The UniProt Consortium (2017). "UniProt: the universal protein knowledgebase." Nucleic Acids Research **45**(D1): D158-D169.

Toyoda, K., H. Teramoto, M. Inui and H. Yukawa (2009). "The *ldhA* gene, encoding fermentative L-lactate dehydrogenase of *Corynebacterium glutamicum*, is under the control of positive feedback regulation mediated by *LldR*." J Bacteriol **191**(13): 4251-4258.

Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe and F. Speleman (2002). "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes." Genome Biol **3**(7): RESEARCH0034.

Viegas, S. C., P. Apura, E. Martínez-García, V. de Lorenzo and C. M. Arraiano (2018). "Modulating Heterologous Gene Expression with Portable mRNA-Stabilizing 5'-UTR Sequences." ACS Synthetic Biology **7**(9): 2177-2188.

Wehrs, M., J. M. Gladden, Y. Liu, L. Platz, J.-P. Prah, J. Moon, G. Papa, E. Sundstrom, G. M. Geiselman, D. Tanjore, J. D. Keasling, T. R. Pray, B. A. Simmons and A. Mukhopadhyay (2019). "Sustainable bioproduction of the blue pigment indigoidine: Expanding the range of heterologous products in *R. toruloides* to include non-ribosomal peptides." Green Chemistry **21**(12): 3394-3406.

Wehrs, M., J. P. Prah, J. Moon, Y. Li, D. Tanjore, J. D. Keasling, T. Pray and A. Mukhopadhyay (2018). "Production efficiency of the bacterial non-ribosomal peptide indigoidine relies on the respiratory metabolic state in *S. cerevisiae*." Microb Cell Fact **17**(1): 193.

Wendisch, V. F., M. Bott and B. J. Eikmanns (2006). "Metabolic engineering of *Escherichia coli* and *Corynebacterium glutamicum* for biotechnological production of organic acids and amino acids." Curr Opin Microbiol **9**(3): 268-274.

Wolf, A., R. Krämer and S. Morbach (2003). "Three pathways for trehalose metabolism in *Corynebacterium glutamicum* ATCC13032 and their significance in response to osmotic stress." Molecular Microbiology **49**(4): 1119-1134.

Wu, G., Q. Yan, J. A. Jones, Y. J. Tang, S. S. Fong and M. A. G. Koffas (2016). "Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications." Trends Biotechnol **34**(8): 652-664.

Yamamoto, S., M. Sakai, M. Inui and H. Yukawa (2011). "Diversity of metabolic shift in response to oxygen deprivation in *Corynebacterium glutamicum* and its close relatives." Appl Microbiol Biotechnol **90**(3): 1051-1061.

Yamanaka, K., K. A. Reynolds, R. D. Kersten, K. S. Ryan, D. J. Gonzalez, V. Nizet, P. C. Dorrestein and B. S. Moore (2014). "Direct cloning and refactoring of a silent lipopeptide

biosynthetic gene cluster yields the antibiotic taromycin A." Proc Natl Acad Sci U S A **111**(5): 1957-1962.

Yumusak, C., A. J. Prochazkova, D. H. Apaydin, H. Seelajaroen, N. S. Sariciftci, M. Weiter, J. Krajcovic, Y. Qin, W. Zhang, J. Zhan and A. Kovalenko (2019). "Indigoidine – Biosynthesized organic semiconductor." Dyes and Pigments **171**: 107768.

Zelle, E., N. Pflzer, M. Oldiges, A. Koch-Koerfges, M. Bott, K. Noh and W. Wiechert (2021). "An energetic profile of *Corynebacterium glutamicum* underpinned by measured biomass yield on ATP." Metab Eng **65**: 66-78.

Zhang, S., D. Liu, Z. Mao, Y. Mao, H. Ma, T. Chen, X. Zhao and Z. Wang (2018). "Model-based reconstruction of synthetic promoter library in *Corynebacterium glutamicum*." Biotechnol Lett **40**(5): 819-827.

SUPPLEMENTARY INFORMATION II

SUPPLEMENTARY INFORMATION II

Materials & Methods

Determination of intracellular NADH/NAD⁺ ratio

Samples of bacterial cultures (600 µL) were added into tubes with 1.2 mL of 60% (v/v) methanol, inverted rapidly, and kept on ice. Cells were collected by centrifugation (16100g, 4°C, 10 min) and washed with cold phosphate buffer saline. Intracellular NAD(H) was extracted and quantified using the NAD/NADH Quantification Kit (Sigma-Aldrich, Burlington, MA, USA), following the manufacturer's instructions. Specifically, cells were resuspended in 400 µL NAD/NADH Extraction Buffer and disrupted using a Branson Sonifier S-250A (3 cycles of 20 seconds, duty cycle 50%, power setting 3) on ice. Cell lysates were cleared by centrifugation (16100g, 4°C, 15 min). The resulting supernatants were deproteinized using 3000 MW cutoff spin filters (Amicon Ultra-4). The total NAD(H) and NADH were quantified following the manufacturer's instructions in clear flat-bottom 96-well plates. Upon color development resulting from specific enzymatic reactions, the absorbance at 450 nm (A₄₅₀) was read in an iMark™ Microplate Absorbance Reader (Bio-Rad). The amounts of total NAD(H) nucleotides and NADH in the samples were interpolated from a calibration curve prepared with fresh NADH standards (0 - 100 pmol). The NADH/NAD⁺ ratio was calculated as follows: $\text{NADH/NAD}^+ = \text{NADH} / (\text{NAD(H)}_{\text{total}} - \text{NADH})$.

Figures

Figure S1

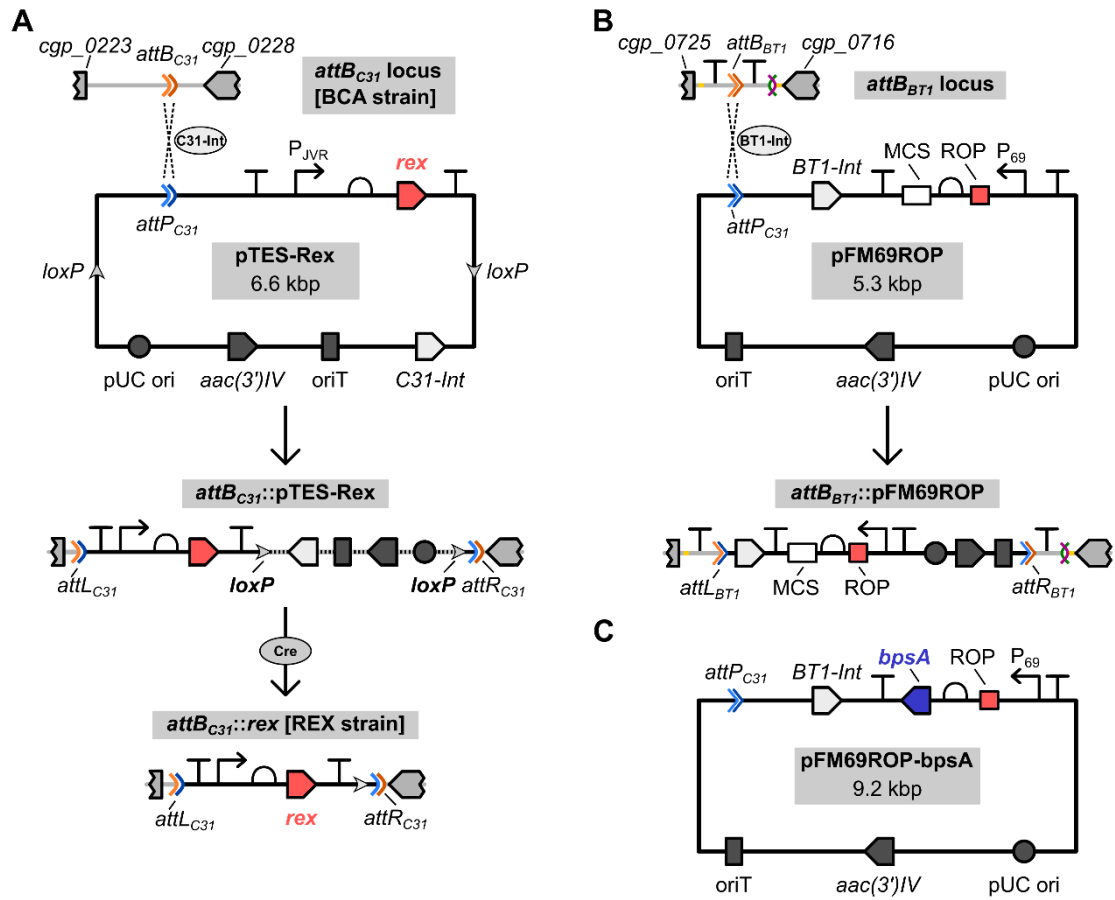


Fig. S1 - Maps and integration of phage-based plasmids in the genome of *C. glutamicum*. **(A)** Schematic representation of the pTES-Rex plasmid and its markerless integration into the genome of *C. glutamicum* BCA strain to originate the *C. glutamicum* REX strain. The process involved the sequential ϕ C31-based integration of pTES-Rex at the *attB_{C31}* locus, and the Cre-mediated excision of vector's backbone catalyzed by Cre (expressed in *cis* from pALCRE (Herrmann, Siegl et al. 2012)). **(B)** Schematic representation of the ϕ BT1-based integration of pFM69ROP and derivative plasmids at the *attB_{BT1}* of *C. glutamicum* BCA/REX strains. **(C)** Schematic representation of pFM69ROP-bpsA plasmid. *aac(3')/IV* – apramycin resistance marker; *attB/P_{BT1}* – bacterial/phage attachment site for the ϕ BT1 phage; *attL/R_{BT1}* – left/right scar of recombined attachment sites of the ϕ BT1 phage; *attB/P_{C31}* – bacterial/phage attachment site for the ϕ C31 phage; *attL/R_{C31}* – left/right scar of recombined attachment sites of the ϕ BT1 phage; *bpsA* – modified coding sequence for the *bpsA* of *Streptomyces albidoflavus* J1074 (Marques, Luzhetskyy et al. 2020); *BT1-Int* – integrase of the ϕ BT1 phage; *C31-Int* – integrase of the ϕ C31 phage; *loxP* – *loxP* site; MCS – multicloning site; *oriT* – transfer origin; *P₆₉* – synthetic promoter P69 (Zhang, Liu et al. 2018); *P_{JVR}* – synthetic promoter with 4306 MU of activity (Rytter, Helmark et al. 2014); pUC ori – plasmid origin of replication for *E. coli*; *rex* – coding sequence for *rex* gene of *Thermus aquaticus*.

Figure S2

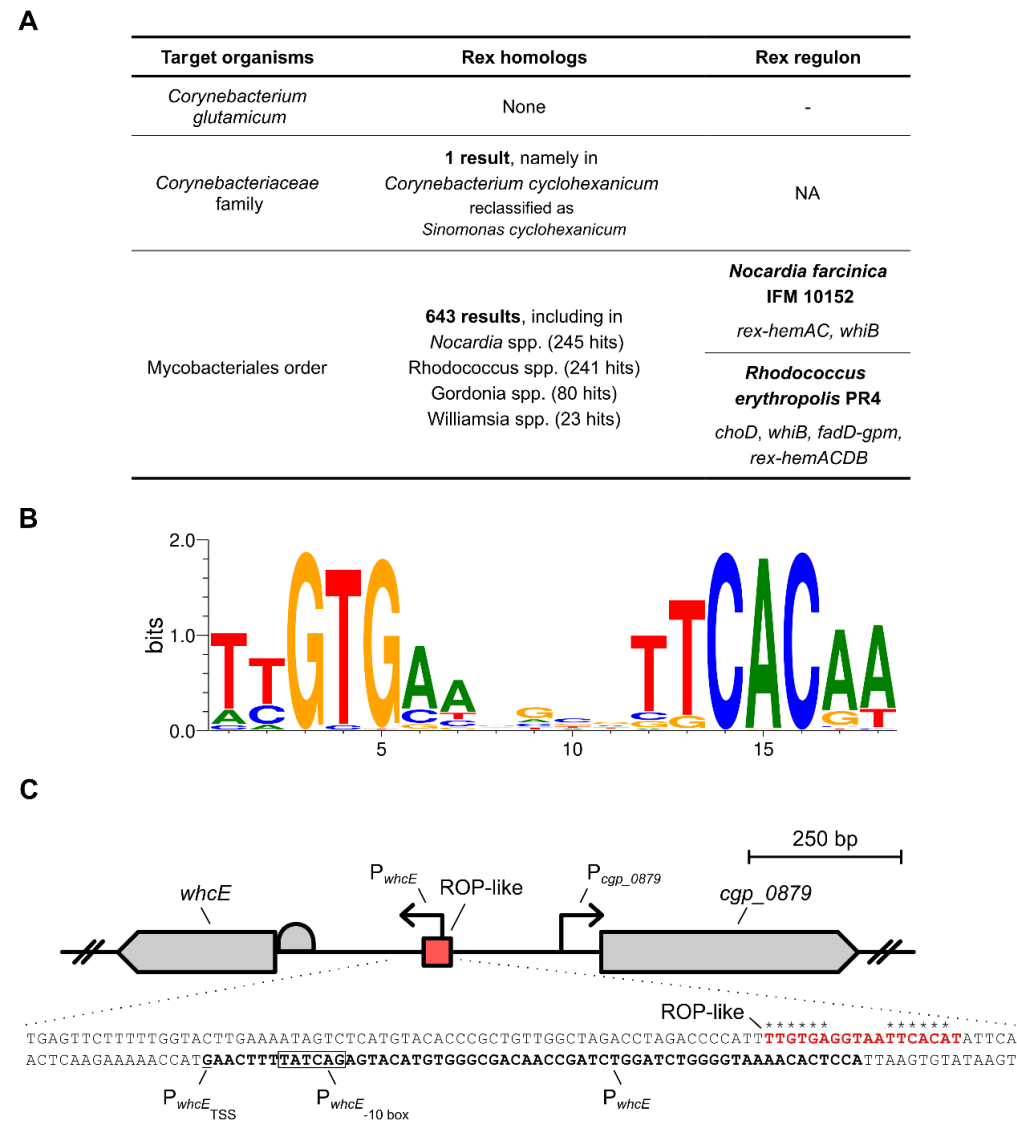


Fig. S2 - *In silico* studies on the suitability of a Rex-based regulatory system in *Corynebacterium glutamicum*. (A) Hits of proteins with homology to *Thermus aquaticus*' Rex protein in non-redundant protein databases and examples of Rex-regulated operons. Note that *Corynebacterium cyclohexanicum* has been reclassified as *Sinomonas cyclohexanicum* of the *Micrococcaceae* family (Yamamoto, Hasegawa et al. 2021) upon genome sequencing and 16 rRNA analysis (DDBJ acc. number AP024525). (B) Sequence logo of Rex operator (ROP) site reconstructed from 88 manually curated sequences from Actinomycetales (Novichkov, Kazakov et al. 2013). (C) Schematic representation of the genomic context and the detailed sequence of the ROP-like sequence found in the genome of *C. glutamicum*. The annotations of coding sequences refer to that of strain MB001 (NCBI acc. CP005959), ancestor of BCA strain. The promoters and ribosome binding site were annotated according to the RNAseq study of strain ATCC 13032 (Pfeifer-Sancar, Mentz et al. 2013). Occurrences of ROP-like sequences were analyzed using PRODORIC webtool (Dudek and Jahn 2022). NA – not available, TSS – transcriptional start site.

Figure S3

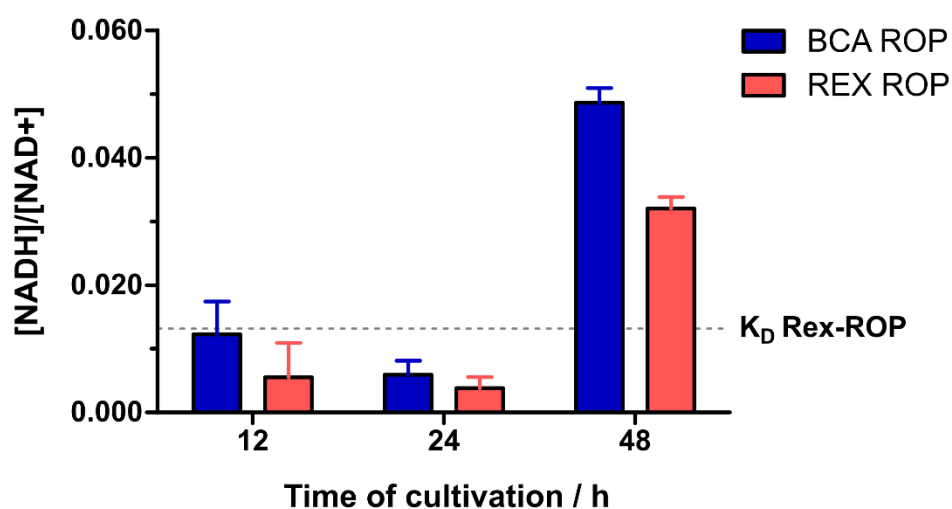


Fig. S3 - Intracellular ratio of the concentration of reduced and oxidized nicotinamide dinucleotide co-factors along growth curve of *C. glutamicum*. Strains BCA and REX hosting pFM69ROP (BCA ROP and REX ROP, respectively) were grown in defined media CGXII with 4% (w/v) glucose for 72 h. At defined times of cultivation, cell metabolism was quenched, NAD(H) was extracted and the [NADH]/[NAD⁺] ratio was calculated using a cycling enzymatic assay. The dotted line labeled “K_D Rex-ROP” correspond to the inferred half maximal dissociation ratio of Rex of *Thermus aquaticus*, from a prior SPR experiment (Sickmier, Brekasis et al. 2005). Bars and error bars represent the average and the standard deviation of two technical replicates, respectively.

Figure S4

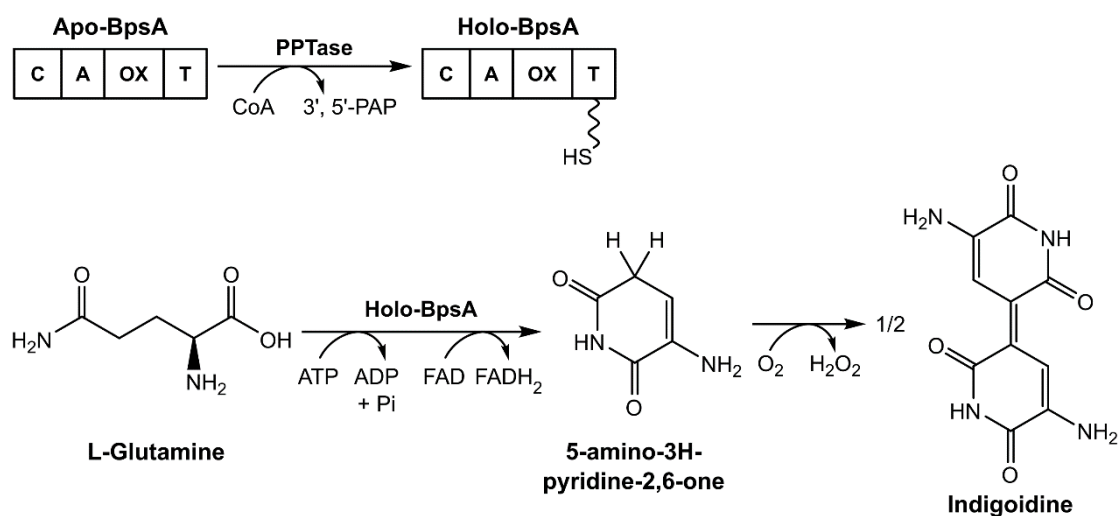


Fig. S4 - Reporter system based on the production of the non-ribosomal peptide indigoidine. The apo-BpsA reporter is firstly activated by phosphopantetheinylation to the active form holo-BpsA. The holo-BpsA catalyzes the conversion of L-glutamine to the intermediate 5-amino-3H-pyridine-2,6-dione, at expense of ATP and originating the reduced coenzyme FADH₂ (Walsh and Wenciewicz 2013). The intermediate compound oxidizes spontaneously and dimerizes to originate the blue pigment indigoidine.

Figure S5

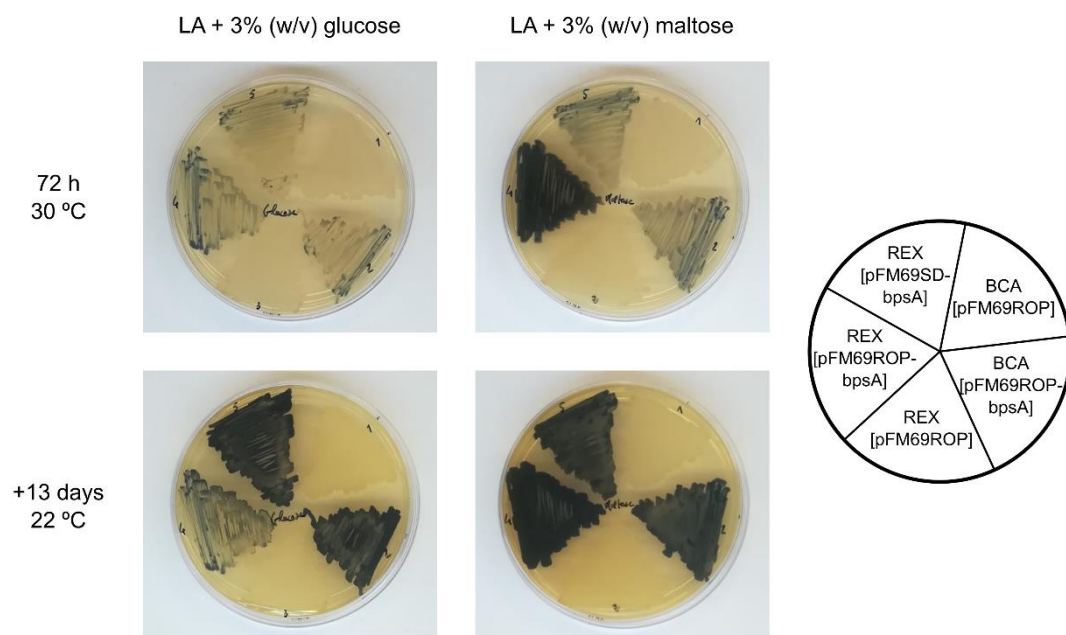


Fig. S5 - Long-term effect of carbon sources on the Rex-dependent production of indigoidine in *C. glutamicum*. Strains of *C. glutamicum* were inoculated on lysogeny agar (LA) supplemented with specified carbon source and incubated at 30°C for 72 h. After that, the strains were incubated at 22°C (controlled room temperature) for additional 13 days. Blue pigmentation indicates the production of indigoidine.

Figure S6

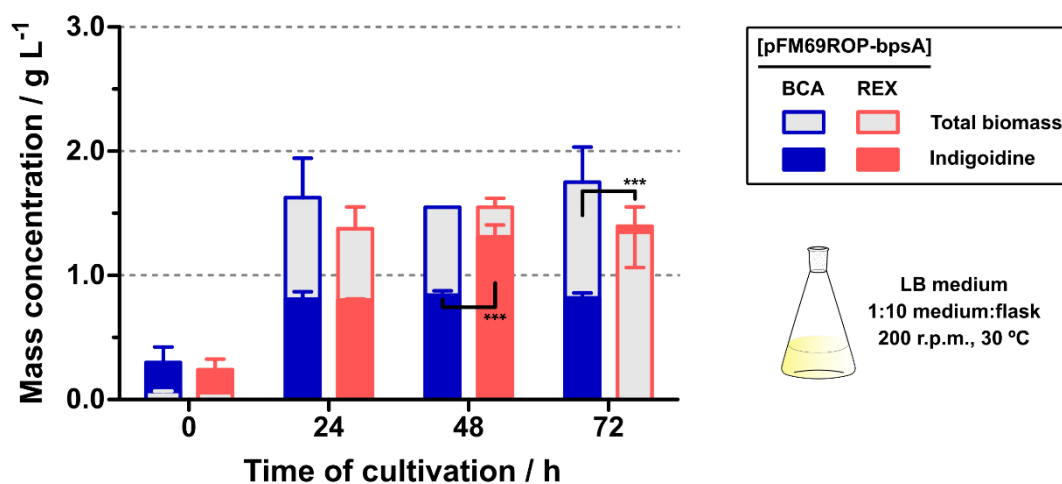


Fig. S6 - Cell growth and production of indigoidine by *C. glutamicum* in complex media. Strains BCA[pFM69ROP-bpsA] (blue bars) and REX[pFM69ROP-bpsA] (red bars) were cultivated in lysogeny broth (LB) for 72 h. Total biomass (grey bars) and indigoidine content (blue/red bars) were determined by cell dry weight and by spectrophotometry of DMSO extracts, respectively. Bars and error bars represent the average and the standard deviation of two experimental replicates, respectively. Statistically significant differences: *** $p < 0.001$.

Additional references

- Dudek, C.-A. and D. Jahn (2022). "PRODORIC: state-of-the-art database of prokaryotic gene regulation." Nucleic Acids Research **50**(D1): D295-D302.
- Herrmann, S., T. Siegl, M. Luzhetska, L. Petzke, C. Jilg, E. Welle, A. Erb, P. F. Leadlay, A. Bechthold and A. Luzhetskyy (2012). "Site-specific recombination strategies for engineering actinomycete genomes." Appl Environ Microbiol **78**(6): 1804-1812.
- Marques, F., A. Luzhetskyy and M. V. Mendes (2020). "Engineering *Corynebacterium glutamicum* with a comprehensive genomic library and phage-based vectors." Metab Eng **62**: 221-234.
- Novichkov, P. S., A. E. Kazakov, D. A. Ravcheev, S. A. Leyn, G. Y. Kovaleva, R. A. Sutormin, M. D. Kazanov, W. Riehl, A. P. Arkin, I. Dubchak and D. A. Rodionov (2013). "RegPrecise 3.0--a resource for genome-scale exploration of transcriptional regulation in bacteria." BMC Genomics **14**: 745.
- Pfeifer-Sancar, K., A. Mentz, C. Ruckert and J. Kalinowski (2013). "Comprehensive analysis of the *Corynebacterium glutamicum* transcriptome using an improved RNAseq technique." BMC Genomics **14**: 888.
- Rytter, J. V., S. Helmark, J. Chen, M. J. Lezyk, C. Solem and P. R. Jensen (2014). "Synthetic promoter libraries for *Corynebacterium glutamicum*." Appl Microbiol Biotechnol **98**(6): 2617-2623.
- Sickmier, E. A., D. Brekasis, S. Paranawithana, J. B. Bonanno, M. S. B. Paget, S. K. Burley and C. L. Kielkopf (2005). "X-ray structure of a Rex-family repressor/NADH complex insights into the mechanism of redox sensing." Structure **13**(1): 43-54.
- Walsh, C. T. and T. A. Wenciewicz (2013). "Flavoenzymes: versatile catalysts in biosynthetic pathways." Nat Prod Rep **30**(1): 175-200.
- Yamamoto, T., Y. Hasegawa, P. C. K. Lau and H. Iwaki (2021). "Identification and characterization of a *chc* gene cluster responsible for the aromatization pathway of cyclohexanecarboxylate degradation in *Sinomonas cyclohexanica* ATCC 51369." Journal of Bioscience and Bioengineering **132**(6): 621-629.
- Zhang, S., D. Liu, Z. Mao, Y. Mao, H. Ma, T. Chen, X. Zhao and Z. Wang (2018). "Model-based reconstruction of synthetic promoter library in *Corynebacterium glutamicum*." Biotechnol Lett **40**(5): 819-827.

CHAPTER IV

**Producing pharmaceuticals and precursors
thereof in *Corynebacterium glutamicum***

IV.1 Introduction

Antimicrobials are medicines used to prevent and treat infectious diseases caused by virus, bacteria, fungi, and protozoan parasites. Antimicrobial resistance occurs when infections no longer respond to antimicrobials: bacteria can transiently adapt to sub-inhibitory concentrations of antibiotics, be intrinsically resistant to certain antibiotics, or acquire resistance mechanisms due to mutations or horizontal gene transfer. The use of antimicrobials contributes to the emergence of pathogens with acquired resistance mechanisms (Costelloe, Metcalfe et al. 2010), so, among other initiatives, such as the parsimonious use of antimicrobials, the development of novel antibacterial therapies is required to tackle the emerging resistant pathogens. The World Health Organization prioritized the development of antibacterial drugs against *Mycobacterium tuberculosis*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacteriaceae*, which present multidrug resistant variants (World Health Organization 2017).

Natural products are a prolific source and inspiration for novel antibacterial drugs. In the period from 1981 to 2019, 48% of new antibacterial drugs introduced in the market were natural products or derivatives (Newman and Cragg 2020). Despite only three new antibacterial classes entered the market in the period from 1970 to 2005, medicinal chemistry can take known natural products as scaffolds for the development of novel drugs with extended spectrum of action, oral availability, among other optimizations (Butler and Buss 2006). For instance, tetracyclines and beta-lactams scaffolds have been extensively derivatized for better pharmacokinetics and to overcome resistance mechanisms (Grossman 2016, Lima, Silva et al. 2020). The biological production of scaffolds of natural products for semi-synthesis of antibacterial drugs is an opportunity to produce pharmaceuticals in an economically and environmentally sustainable way.

Bottromycin A2 is a macrocyclic peptide ribosomally synthesized and modified post-translationally (**Figure 1**). It was first isolated from *Streptomyces bottropensis* and described as an effective agent against Gram-positive bacteria and mycoplasma *in vitro* (Waisvisz, van der Hoeven et al. 2002). The mechanism of action of bottromycin A2 consists in the inhibition of bacterial protein synthesis by blocking access of amino acid-loaded tRNA to A the site of 50S unit of ribosome (Otaka and Kaji 1976). However, bottromycin A2 showed poor activity *in vivo* due to its chemical instability (Tanaka, Nishimura et al. 1966) that was overcome by derivatives which do not undergo hydrolysis at the terminal methyl ester moiety (Miller, Chalet et al. 1968, Kobayashi, Ichioka et al. 2010). The characterization of bottromycin biosynthesis and the obtention of bottromycin for semi-synthesis of derivatives was initially carried out using native producers, namely *Streptomyces* sp. KM-9459 (at a titer of 30.3 mg.L⁻¹) (Kobayashi, Ichioka et al. 2010), *Streptomyces scabies* (for

biosynthesis elucidation) (Crone, Leeper et al. 2012), and *Streptomyces* sp. DSM 17025 (also known as *Streptomyces* sp. BC16019, approximately producing 0.1 - 0.4 mg L⁻¹) (Lerchen, Schiffer et al. 2006). Bottromycin derives from the *botA*-encoded pro-peptide, which is successively modified, trimmed, and cyclized to originate the mature form of the antibiotic. Altogether, *botA* and the genes encoding the post-translation modification enzymes, regulators and transporters span over 17 kbp (**Figure 1**). Heterologous expression and production of bottromycin was achieved in *Streptomyces albus* (0.001 mg L⁻¹) and in *Streptomyces coelicolor* (up to 0.08 mg L⁻¹) carrying the 35-kb DG2-kan cosmid, which encodes the biosynthetic cluster of *Streptomyces* sp. BC16019 (Huo, Rachid et al. 2012). The modification of the DG2-kan cosmid with a bidirectional cassette of random promoters (around a hygromycin resistance gene) originated the DG2-kan-P41hyg cosmid, which sustained the heterologous production of bottromycin in *Streptomyces lividans* TK24 (up to 50-fold increase compared to the initial cosmid) (Horbal, Marques et al. 2018). Enhanced heterologous production of bottromycin was achieved by employing the bottromycin resistant *S. lividans* BotR-2 host carrying a DG2-kan derivative modified with defined promoters (up to 3-fold increase vs original TK24 host) (Horbal, Marques et al. 2018) or by adding talc microparticles to cultivation media (up to 60 mg L⁻¹ bottromycin A2, corresponding to a 1.4-fold increase vs no talc added) (Kuhl, Ruckert et al. 2021). Finally, *in vivo* derivatization and overproduction of bottromycins with alternative methylation patterns or amino acid sequences was achieved in *Streptomyces* sp. BC16019 using DG2-based cosmids featuring synthetic promoters and mutated peptide-encoding genes (Horbal, Marques et al. 2018). On the one hand, the BGC of bottromycin cloned in the DG2 cosmid was shown to be a promising platform to overproduce bottromycin A2 (fundamental for semi-synthesis) and test derivatization *in vivo*. On another hand, these endeavors highlighted the difficulties in manipulating native producers (so that the use of heterologous hosts was preferred), the susceptibility of heterologous hosts to heterologous products (which implied the selection of resistant strains and fine-tuning production and export genes), still relatively low production yields and infrequent *in vivo* derivatization (two observed substitutions in 12 attempts).

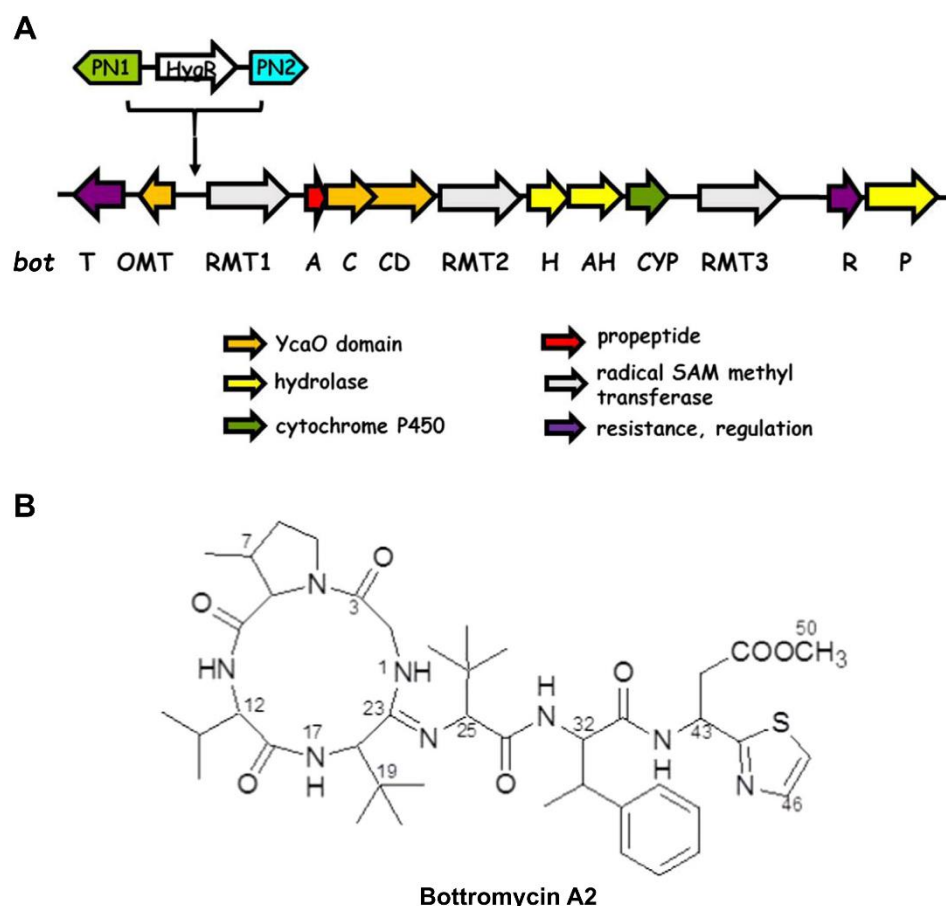


Figure 1. Genetics and chemistry of bottromycin antibiotics. Panel (A) represents the organization of the *bot* gene cluster, responsible for the biosynthesis of bottromycins in *Streptomyces* sp. BC16019 and in heterologous hosts carrying derivatives of the DG2 cosmid (Huo, Rachid et al. 2012), and the cassette of promoters (PN1, PN2) and the hygromycin resistance marker (HygR) included in the refactored DG2-kan-P41hyg cosmid (Horbal, Marques et al. 2018). Panel (B) represents the chemical structure of bottromycin A2. Adapted from (Horbal, Marques et al. 2018).

The (5R)-carbapen-2-em-3-carboxylic acid (also known as SQ 27860) is bicyclic β -lactam of bacterial origin (**Figure 2**). It was firstly isolated from species of *Serratia* and *Erwinia* (later, *Pectinobacterium*) and displayed *in vitro* activity against *E. coli* and *Staphylococcus aureus* (Parker, Rathnum et al. 1982). SQ 27860 presents features common to the carbapenem thienamycin, isolated from *Streptomyces cattleya* (Kahan, Kahan et al. 1979), including a broad spectrum of action (targeting several penicillin binding proteins involved in peptidoglycan formation), low susceptibility to the activity of inactivating beta-lactamases but also low chemical stability. Another common trait of SQ 27860 and thienamycin is the biosynthetic pathway: unlike the tripeptide origin of penicillin and cephalosporin β -lactams, the carbapenem nucleus derives from acetate (malonyl-CoA) and glutamate (glutamate

semialdehyde) (Bycroft, Maslen et al. 1988). The genetic basis of SQ 27860 was studied by heterologous expression of the candidate BGC in *E. coli*, revealing a distinct β -lactam biosynthetic pathway (McGowan, Sebahia et al. 1996). The BGC of SQ 27860 in *Pectinobacterium carotovora* is composed of 8 genes (*carA-H*), of which *carABC* are necessary for biosynthesis and *carFG* are involved in a carbapenem resistance mechanism (McGowan, Sebahia et al. 1997, Monson, Jones et al. 2018). The biosynthesis of SQ 27860 was determined to be initiated with the condensation of malonyl-CoA and glutamate semialdehyde catalyzed by carboxymethylproline synthase (CarB), followed by the ATP/Mg²⁺-dependent formation of the β -lactam ring catalyzed by carbapenam synthetase (CarA, originating (3S,5S)-carbapenam-3-carboxylic acid), and concluded with the Fe²⁺-dependent C5 epimerization and C2-C3 desaturation catalyzed by carbapenam synthase (CarC) (Li, Stapon et al. 2000). The proteins CarD and CarE were associated with the formation of pyrrolidine 5-carboxylate (the spontaneous cyclization product of glutamate semialdehyde) from L-proline and the regeneration of the Fe(II) co-factor of CarC, respectively (Shomar, Gontier et al. 2018). Enhanced productivity of the heterologous production of SQ 27860 (about 5.7 mg L⁻¹ h⁻¹) was achieved by co-expressing *carABCE* in an engineered *E. coli* strain with feedback-insensitive production of pyrrolidine 5-carboxylate and induced tolerance to beta-lactams, in media supplemented with 5 g L⁻¹ potassium glutamate (Shomar, Gontier et al. 2018). The bioproduction and purification of SQ 27860 is challenging due to its toxicity to the bacterial heterologous host, its chemical instability leading to hydrolysis of the beta-lactam function, and lack of standards for accurate quantification. Due to low titers and difficulties in purification, the carbapenem core became commercially produced by chemical synthetic routes (reviewed in (Southgate 1994)). Among those, carbapenems could be semi-synthesized starting from chiral biomolecules, such as L-aspartate (Salzmann, Ratcliffe et al. 1980), D-glucosamine (Miyashita, Chida et al. 1982), and 6-aminopenicillanic acid (Karady, Amato et al. 2002). The stable and innocuous intermediate (3S,5S)-carbapenam-3-carboxylic acid resulting from the activity of CarB and CarA is a candidate precursor for semi-synthesis of carbapenems.

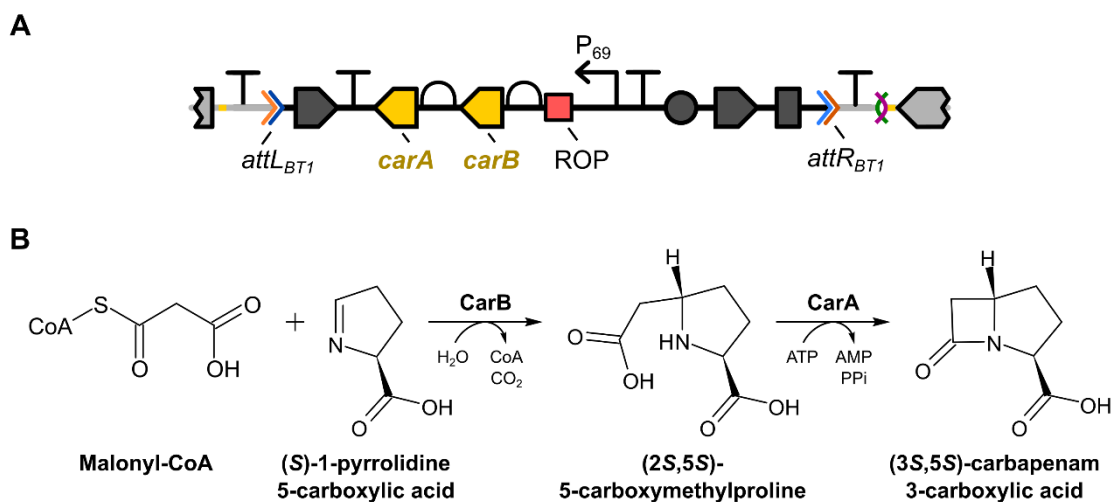


Figure 2. Genetics and biosynthesis of (3S,5S)-carbapenam-3-carboxylic acid, a precursor of carbapenem antibiotics. Panel (A) represents the organization of the synthesized *car* operon, inspired in the gene cluster responsible for the biosynthesis of carbapenem in *Pectinobacterium carotovora* (McGowan, Sebaihia et al. 1997). The operon is under the control of P69 promoter (Zhang, Liu et al. 2018) and the Rex operator (ROP) (McLaughlin, Strain-Damerell et al. 2010). The implementation of the operon using the ϕ BT1-based pFM69ROP-carBA plasmid originates the *attL*_{BT1} and *attR*_{BT1} scars. Panel (B) represents the biosynthetic steps and chemical structure of precursors of carbapenem antibiotics (adapted from (Shomar, Gontier et al. 2018)).

To enable the cost-effective and sustainable bioproduction of intermediates for discovery and production of antibiotics, suitable microbial factories for heterologous expression are required. For that purpose, we tested the potential of engineered *C. glutamicum* strains to produce bottromycin A2 and (3S,5S)-carbapenam-3-carboxylic acid. Regarding bottromycin, *C. glutamicum* CR099attB_{C31} was transformed with a ϕ C31-based cosmid encoding the refactored BGC of bottromycin A2 from *Streptomyces* sp. BC16019 used for heterologous expression in *S. lividans* TK24 (Horbal, Marques et al. 2018). Regarding (3S,5S)-carbapenam-3-carboxylic acid (henceforth, carbapenam), *C. glutamicum* BCA was transformed with a ϕ BT1-based plasmid encoding a recoded and engineered *carBA* operon of *Pectinobacterium carotovora* (McGowan, Sebaihia et al. 1997). Production of the compounds was assessed by mass spectrometry coupled to liquid chromatography of extracts of liquid cultures of *C. glutamicum* transformants.

IV.2 Materials and Methods

IV.2.1 Bacterial strains, plasmids, media, and growth conditions

All the strains and plasmids used in this work are listed in **Table 1**. *Escherichia coli* and *Corynebacterium glutamicum* strains were routinely cultivated in lysogeny broth/agar (Lennox 1955), at 37°C and 30°C, respectively. For bottromycin production, *C. glutamicum* colonies were inoculated in 10 mL tryptic soy broth (TSB) for 16 h, in 100-mL baffled flasks, at 200 rpm; grown cells were used to inoculate 50 mL of fresh TSB medium to a final OD₆₀₀ of 0.1, in 500-mL baffled flasks, and were cultivated for 6 days, at 200 rpm. For carbapenam production, *C. glutamicum* colonies were inoculated in 5 mL lysogeny broth for 16 h, in 50-mL tubes, at 200 rpm; cells from 1 mL of previous culture were inoculated in 10 mL of CGXII medium (Keilhauer, Eggeling et al. 1993) [containing with 4% (w/v) glucose, 0.2 mg L⁻¹ biotin, and 30 mg L⁻¹ protocatechuic acid], in 100-mL flasks, and were incubated for 16 h, at 200 rpm; finally, cells grown in CGXII were used to inoculate 25 mL of fresh CGXII medium to a final OD₆₀₀ of 0.5, in 250-mL flasks, and were cultivated for 3 days, at 200 rpm. When necessary, media was supplemented with ampicillin (Ap; 100 µg mL⁻¹ to *E. coli*), apramycin (Am; 50 µg mL⁻¹ to *E. coli*, 25 µg mL⁻¹ to *C. glutamicum*) or hygromycin (Hyg; 120 µg mL⁻¹ to *E. coli*, 80 µg mL⁻¹ to *C. glutamicum*). Bacterial growth was determined by cell dry weight, as described previously (Marques, Luzhetskyy et al. 2020).

Table 1. Strains and vectors used in this study.

Bacteria and vectors	Description ^a	Source or reference
Bacteria		
<i>E. coli</i> DH5α	General cloning host	(Hanahan 1983)
<i>C. glutamicum</i> CR099attB _{C31}	Prophage-free and IS-reduced <i>C. glutamicum</i> strain, featuring attB _{C31}	Gift of J. Kalinowski, based on <i>C. glutamicum</i> CR099 (Unthan, Baumgart et al. 2015, Baumgart, Unthan et al. 2018)
<i>C. glutamicum</i> BCA	Prophage-free and IS-reduced <i>C. glutamicum</i> strain, featuring attB _{BT1} and attB _{C31} , with albino phenotype	(Marques, Luzhetskyy et al. 2020)
Vectors		
DG2-kan-P41hyg	Am ^R , Km ^R , Hyg ^R ; ϕC31-based integrative cosmid encoding bottromycin biosynthetic gene cluster with P41 pair of promoters integrated in-between two operons	(Horbal, Marques et al. 2018)
pUC57-carC-carBA	Ap ^R ; source of <i>carB</i> and <i>carA</i> genes (translation: NCBI acc. AAD38230 and AAD38229, respectively)	This work (GenScript)
pFM69ROP	Am ^R ; ϕBT1-based integrative plasmid for high-level gene expression, repressible by Rex regulator	(Marques et al., manuscript in preparation)
pFM69ROP-carBA	pFM69ROP derivative for expression of carbapenam	This work

^a**Abbreviations:** Am^R – apramycin resistance marker; Ap^R – ampicillin resistance marker; attB_{BT1} – bacterial attachment site for phage ϕBT1; attB_{C31} – bacterial attachment site for phage ϕC31; Hyg^R – hygromycin kanamycin resistance marker; Km^R – kanamycin resistance marker.

IV.2.2 Recombinant DNA manipulation

Standard recombinant DNA methods were followed (Sambrook and Russell 2001). Genomic DNA of *C. glutamicum* strains was isolated by the salting-out method according to (Kieser, Bibb et al. 2000). Polymerase chain reaction (PCR) for routine amplification and cloning purposes was performed using the GoTaq® G2 Flexi DNA Polymerase (Promega) and the Q5® High-Fidelity DNA Polymerase (NEB), respectively, following manufacturers' instructions. The lists of plasmids and oligonucleotides used in this work are provided in **Table 1** and **Table 2**, respectively.

Table 2. Oligonucleotides used in this study. Ribosome binding site in bold, restriction sites underlined.

Oligonucleotide	Sequence (5' to 3')	Use
DG2_intchk_F	GGTCGGCGGCGTATGACTGG	Check integration of DG2-kan-P41hyg at <i>attB_{C31}</i> locus
DG2_intchk_R	TGGCAGCATTCCTCCGCGAC	
botA_intchk_F	CCGGTTCCGCGACCTCCAAG	Check presence of <i>botA</i> locus (from DG2-kan-P41hyg)
botA_intchk_R	CACACCCGCCGGCTACTCAC	
NdeI_carB_fwd	CCCAAACATATGGTGTTCGAAGAGA ACTCC	Generate fragment encoding <i>carB</i> gene and assemble the <i>carBA</i> operon
carB_rvs	ATTCTATCTCCTTACTTGGTACTGG ACCGGTTTAGTACTTCTTGC	
carA_fwd	CCAGTACCAAGTAAGGAGATAGAA TATGTCCAACCTCTTCTGCGTG	Generate fragment encoding <i>carA</i> gene and assemble the <i>carBA</i> operon
SpeI_carA_rvs	AAACCTACTAGTTTAGTCCTTCTTGA TGAGATCC	
crt_delchk2_rvs	GCCTACTTGATGAGGCCGAA	Check integration of pFM69ROP derivatives at <i>attB_{BT1}</i> locus
pRT_ORI_SEQ	GGATTCCCGTTGAGCACCGC	
pUCR	ACACAGGAAACAGCTATGAC	Check presence of P69-UTR- <i>carB</i> locus (from pFM690ROP- <i>carBA</i>)
carSPL-chk-R	TCCTTGACGGAGGTTTCC	

To create an ϕ BT1-integrative plasmid for Rex-controlled expression of carbapenam, the **pFM69ROP-carBA** plasmid was constructed. Firstly, fragments encoding the *carB* and *carA* genes were amplified from pUC57-carC-carBA by PCR using the primer pairs NdeI_carB_fwd/carB_rvs and carA_fwd/SpeI_carA_rvs, respectively. Secondly, equimolar amounts of the generated *carB* and *carA* fragments were assembled and amplified by PCR using the primer pairs NdeI_carB_fwd/ SpeI_carA_rvs to yield a fragment encoding the full *carBA* operon. Finally, the *carBA* operon was cloned into the *NdeI* and *SpeI* sites of pFM69ROP to yield pFM69ROP-carBA. The identity of pFM69ROP-carBA was confirmed by sequencing.

IV.2.3 Generation of transformant bacteria

Preparation of RbCl competent *E. coli* and transformation with recombinant DNA was performed using standard methods (Sambrook and Russell 2001). Preparation of electrocompetent *C. glutamicum* cells and transformation with recombinant DNA was

performed according to (Eggeling and Bott 2005). Electroporation was performed using two different apparatus: an Eporator® (Eppendorf), for bottromycin construct, and a Gene Pulser™ plus Pulse Controller (Bio-Rad), for carbapenam construct. Integration of ϕ C31- and ϕ BT1-integrative vectors into the respective sites of the genome of *C. glutamicum* was verified by PCR using the primer pairs crt_delchk2_rvs/pRT_ORI_SEQ and DG2_intchk_F/R, respectively. The presence of the *botA* locus was assessed by PCR using the primer pair botA_intchk_F/R.

IV.2.4 Production analysis by HPLC-MS

For bottromycin extraction, cultivation broth was immediately extracted with 1 volume of ethyl acetate. The organic phase was collected, and the solvent was evaporated to dryness at low pressure, at 30 °C. For bottromycin analysis, extracts were resuspended in 450 μ L liquid chromatography (LC)-grade methanol and cleared by centrifugation (20000g, 10 min, 4°C). Bottromycin detection was performed by ultra-high performance liquid chromatography coupled to mass spectrometry (UPLC-MS), adapting a published method (Horbal, Marques et al. 2018). Briefly, 1 μ L of the extracts was injected onto a UPLC ACQUITY column (Waters), for Reverse Phase separation, with a column matched BEH C18 pre-column VanGuard (Waters). The column had 100 mm in length, 2.1 mm internal diameter, 1.7 μ m diameter Ethylene Bridged Hybrid C18 spherical particles, and 130 Å pore size. The chromatographic system was a Dionex UltiMate 3000 Rapid Separation LC system (Thermo Fisher Scientific), with binary pump, autosampler, and thermostatted column compartment. The mobile phase A was ddH₂O with 0.1% formic acid and mobile phase B was acetonitrile with 0.1% formic acid, both freshly prepared with LC-grade components. The total runtime was 9 minutes, with the following proportions of mobile phases: min 0 – 9, gradient 5% to 95% B; 9 – 10.5 min, gradient 95% - 5% B. The flow rate was kept at 0.6 mL/min. The column compartment was kept at 45 °C. The column was preconditioned with 5% B, flow rate 0 – 0.6 mL/min until backpressure was stable (~15 min). Before the first test sample and after the last test sample, one injection of 10 μ L methanol was performed. The MS analysis of UPLC eluate was performed using an amaZon speed ion trap mass spectrometer (Bruker). The ions were produced by an electrospray ionization (ESI) Apollo II source (Bruker), fed by the LC column output. The post-source components were dual ion funnel transfer line, an octapole ion guide, and a (Paul) quadrupole ion trap analyzer. Acquisition was performed in positive mode with automatic gain control (target: 1 x 10⁶). MS analysis was performed at a resolution 0.5 u FWHM, at a scan speed of 32,000 u/s, and a scan range m/z 200 – 2000. The instrument was managed with Chromeleon Xpress software (Dionex), and the data was analyzed using DataAnalysis software (Bruker).

For carbapenem extraction, cultivated cells were collected by centrifugation (4500g, 10 min, 4°C), frozen at -80°C, freeze-dried, and stored away from light, at room temperature. Dried cells were percolated with 20 mL of LC-grade methanol for one hour, at room temperature, with agitation. Cell extracts were cleared by centrifugation (5000g, 10 min, 4°C) and solvent was evaporated to dryness at low pressure, at 30°C. Extraction yield was quantified by gravimetry. For carbapenem analysis, 10 mg mL⁻¹ solutions of the extracts were prepared in LC-grade methanol with 0.1% (v/v) formic acid and filtered through 0.2 µm regenerated cellulose filter (Minisart® RC4 Syringe Filter, Sartorius). Carbapenem detection was performed by mass spectrometry coupled to ultra-high performance liquid chromatography, adapting a published method (Shomar, Gontier et al. 2018). Briefly, 5 µL of the extracts were injected onto a Kinetex HILIC 100A column (Phenomenex) for hydrophilic interaction liquid chromatography, without pre-column. The column had 150 mm in length, 2.1 mm internal diameter, 2.6 µm diameter particles and 100 Å pore size. The chromatographic system was a Dionex UltiMate 3000 Rapid Separation LC system (Thermo Fisher Scientific), with binary pump, autosampler, and thermostatted column compartment. The mobile phase A was 25 mM ammonium formate and mobile phase B was 100% acetonitrile, both freshly prepared with LC-grade components. The total runtime was 11 minutes, with the following proportions of mobile phases: min 0 – 5, gradient 95% to 60% B, at a flow rate of 0.5 mL/min; min 5 – 7, gradient 60% to 50% B, at a flow rate of 0.5 mL/min to 0.6 mL/min; min 7 – 9, gradient 50% to 95% B, at a flow rate of 0.7 mL/min; min 9 – 11, 95% B, at a flow rate of 0.5 mL/min. The column compartment was kept at 30 °C. The column was preconditioned with 95% B, flow rate 0 – 0.5 mL/min until backpressure was stable (~15 min). Before the first test sample, before standards, and after the last sample, one injection of 5 µL methanol with 0.1% formic acid was performed. As approximate standard, one injection of 5 µL of aqueous solution of imipenem (10 mg/mL, European Pharmacopoeia Reference Standard) diluted 1:40 in acetonitrile with 0.1% formic acid was performed after the test samples. The MS analysis of UPLC eluate was performed using an Q Exactive Focus mass spectrometer (Thermo Fisher Scientific). The ions were produced by an Ion Max API source with H-ESI II probe, fed by the LC column output. The post-source components were RF lens, a bent flatapole ion guide, a quadrupole mass filter, a curved linear trap (C-Trap), a straight multipole collision cell, and an Orbitrap mass analyzer for Fourier Transform Mass Spectrometry (FTMS) analysis. Acquisition was performed in switching polarity mode with automatic gain control (target: 1 x 10⁶). Full MS analysis was performed at a resolution of 70000 (@ m/z 200) and scan range m/z 50 – 500. The instrument management and data analysis were done with XCalibur software suite (Thermo Fisher Scientific).

IV.3 Results

IV.3.1 Bottromycin: generation of strains and production assessment.

The heterologous production of bottromycin in *S. lividans* host was established by intergeneric conjugation of 36-kbp DG2-kan-P41hyg cosmid encoding a refactored bottromycin biosynthetic gene cluster of *S. sp.* BC16019 (Horbal, Marques et al. 2018). To determine whether transformation was a suitable method to obtain *C. glutamicum* strains carrying the bottromycin BGC, the strain CR099attB_{C31} (featuring a cognate site for ϕ C31-based integration) was electroporated with DG2-kan-P41hyg cosmid and selected in the presence of apramycin and hygromycin. To confirm the identity of the transformants, the genomic DNA of two clones (A & B) was isolated and the integration of DG2-kan-P41 was verified by PCR (Figure 3). The results showed that the tested clones were carrying the bottromycin BGC, which had integrated at the attB_{C31} locus, as expected.

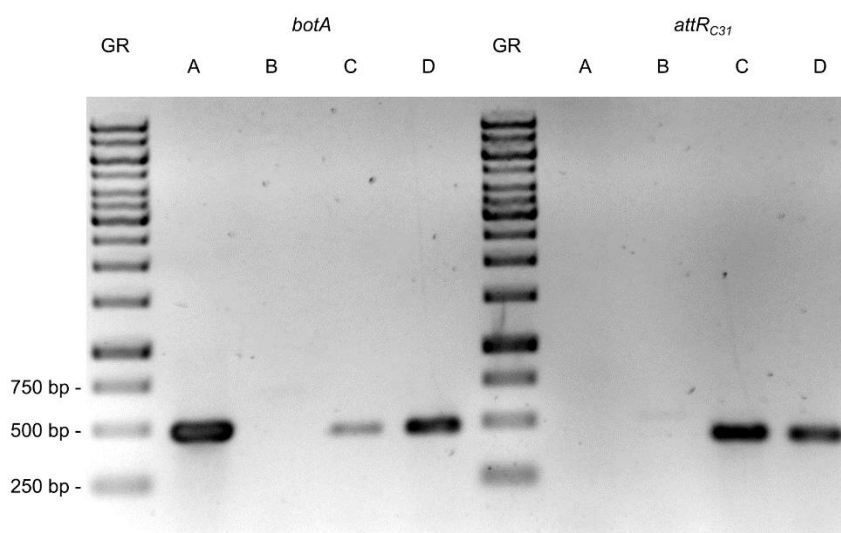


Figure 3. Integration of ϕ C31-based DG2-kan-P41hyg vector carrying the biosynthetic gene cluster of bottromycin into the genome of *C. glutamicum* CR099attB_{C31}. Integration of cosmid DG2-kan-P41hyg was confirmed by PCR amplification of the cosmid-borne *botA* locus (primer pair: botA_intchk_F/R) and the attR_{C31} scar originated upon integration (primer pair: DG2_intchk_F/R), using genomic DNA of two biological replicates (#1, #2) of strain CR099attB_{C31} [DG2-Km-P41hyg] (lanes C and D, respectively). DNA of DG2-Km-P41hyg cosmid (lanes A) and genomic DNA of CR099attB_{C31} (lanes B) were used as controls. GR: GeneRuler 1 kb DNA Ladder.

The higher yields of bottromycin obtained from *S. lividans* carrying the refactored DG2-kan-P41hyg cosmid comparing to parental DG2-kan cosmid are thought to be related to more balanced production and export of bottromycin provided by a fortunate combination of promoters with the appropriate strength (Horbal, Marques et al. 2018, Kuhl, Ruckert et al.

2021). To determine if the refactored BGC is also functional in *C. glutamicum*, the transformants carrying DG2-kan-P41 were cultivated in rich media (TSB) for seven days and culture broths were analyzed by liquid chromatography coupled to mass spectrometry (Figure 4).

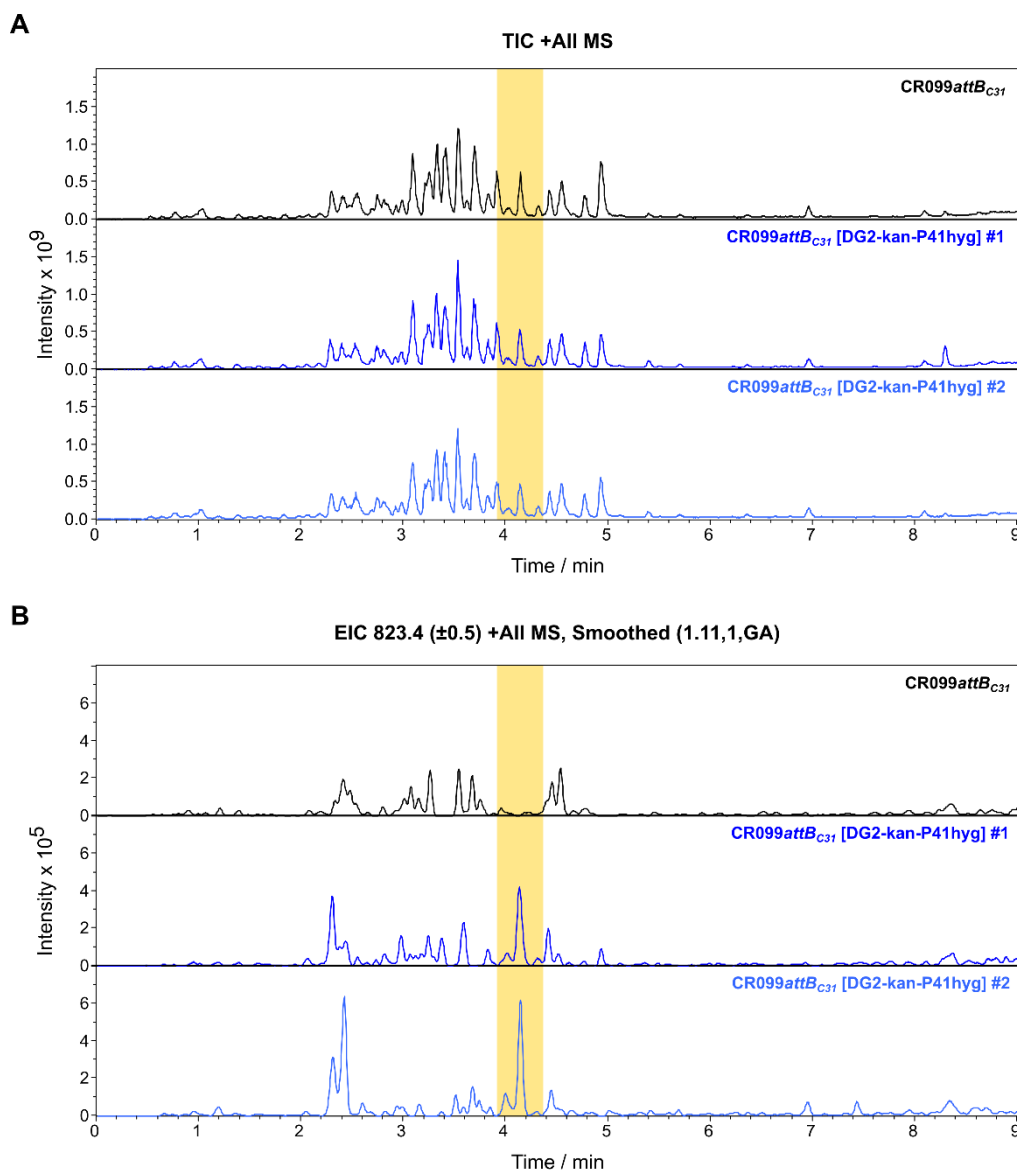


Figure 4 . Detection of bottromycin production in *C. glutamicum* CR099attB_{C31} derivatives harbouring the DG2-kan-P41hyg cosmid. Results depict two biological replicates (#1, #2) of strain CR099attB_{C31} [DG2-kan-P41hyg] and the parental strain CR099attB_{C31} as control. Panel (A) represents the total ion chromatogram (TIC). Panel (B) represents the extracted ion chromatogram (EIC) at m/z 823.4, which corresponds to bottromycin A2 molecular ion, in positive mode. The yellow stripe indicates the retention period of the novel peaks at m/z 823.4 in the strains carrying DG2-kan-P41hyg.

The total ion chromatograms of the analyzed strains had no major differences in their metabolic profile. The extract ion chromatograms at mass-to-charge (m/z) 823.4, corresponding to monoprotonated bottromycin A2 ion, showed a low-intensity MS⁺ signal at min 4.1 for the extracts of clones #1 and #2 of *C. glutamicum* CR099attB_{C31} carrying DG2-kan-P41hyg, which was not present in the control strain. Unfortunately, no further determination of exact mass nor fragmentation pattern were performed to confirm the identity of the detected signal. No clear signals of pre-bottromycins, non-cyclized precursor peptides or incomplete precursor peptides (as defined in (Kuhl, Ruckert et al. 2021)) were found in the tested strains.

IV.3.2 Carbapenam: generation of strains and production assessment.

The production of the carbapenam intermediate of SQ 27860 requires the sequential participation of the carboxymethylproline synthase (CarB) and the carbapenam synthase (CarA) from *P. carotovora* (Li, Stapon et al. 2000). Instead of cloning the carbapenam biosynthetic genes from the native producer, the *carB* and *carA* genes were synthesized according to the codon usage of *C. glutamicum*. The synthetic genes have a superior codon adaptation index (above 0.95), start with ATG (the most frequent start codon in *C. glutamicum*) and are free of common regulatory sequences. The synthetic genes were cloned as *carBA* operon into the ϕ BT1-based integrative vector pFM69ROP, downstream its strong synthetic promoter, Rex operator, and ribosome binding site (**see chapter III**). This procedure decoupled gene expression from the regulation of the original BGC, which was responsive to quorum sensing, high temperature and anaerobiosis (Barnard, Bowden et al. 2007). The resulting 7.6-kbp pFM69ROP-carBA plasmid is suitable for Rex-dependent expression in suitable strains (**see chapter III**). In the absence of Rex protein, it behaves as a constitutive expression vector. To implement the generated construct, the *C. glutamicum* strain BCA (featuring a site for ϕ BT1-based integration) was transformed with pFM69ROP-carBA and transformants were selected in the presence of apramycin. The identity of the obtained transformants was verified by PCR (**Figure 5**). The results showed the correct integration of pFM69ROP-carBA at the *attB_{BT1}* locus of the genome of *C. glutamicum* BCA.

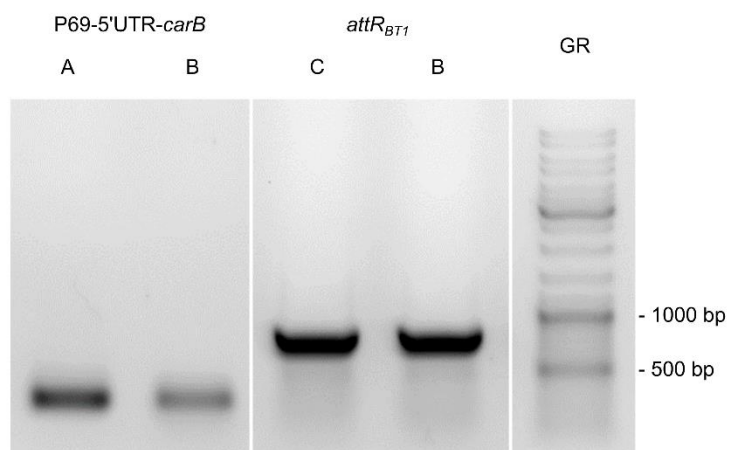


Figure 5. Integration of ϕ BT1-based pFM69ROP-carBA vector carrying the biosynthetic genes of carbapenam into the genome of *C. glutamicum* BCA. Integration of plasmid pFM69ROP-carBA was confirmed by PCR amplification of the plasmid-borne *carB* locus (including the P69 promoter and 5' UTR, primer pair: pUCR/carSPL-chk-R) and the *attR*_{C31} scar generated upon integration (primer pair: crt_delchk2_rvs/pRT_ORI_SEQ) using genomic DNA of strain BCA [pFM69ROP-carBA] #1 (lanes B) as template. DNA of pFM69ROP-carBA plasmid (lane A) and genomic DNA of BCA strain carrying pFM69ROP (lane C) were used as controls. GR: GeneRuler 1 kb DNA Ladder.

The high-yield heterologous production of SQ 27860 in *E. coli* required the optimization of precursors supply by adding glutamate to the medium and by genetically manipulating the metabolic pathways related to proline and malonyl-CoA use (Shomar, Gontier et al. 2018). To determine if carbapenam could be produced in *C. glutamicum* in the absence of these optimizations, the BCA strain carrying pFM69ROP-carBA was cultivated in the standard defined media CGXII with glucose as carbon source for 3 days. For the detection of carbapenam production, cells extracts were analyzed by liquid chromatography coupled to mass spectrometry (**Figure 6**). The decision to analyze cell extracts rather than supernatant of cultures (as in (Shomar, Gontier et al. 2018)) took in consideration that carbapenam do not cause cell lysis (in contrast to the bioactive carbapenam), no dedicated transporter was expressed, and no homologs of the carbapenam-exporting AcrB multidrug efflux pumps exist in *C. glutamicum* (Atzori, Malviya et al. 2019).

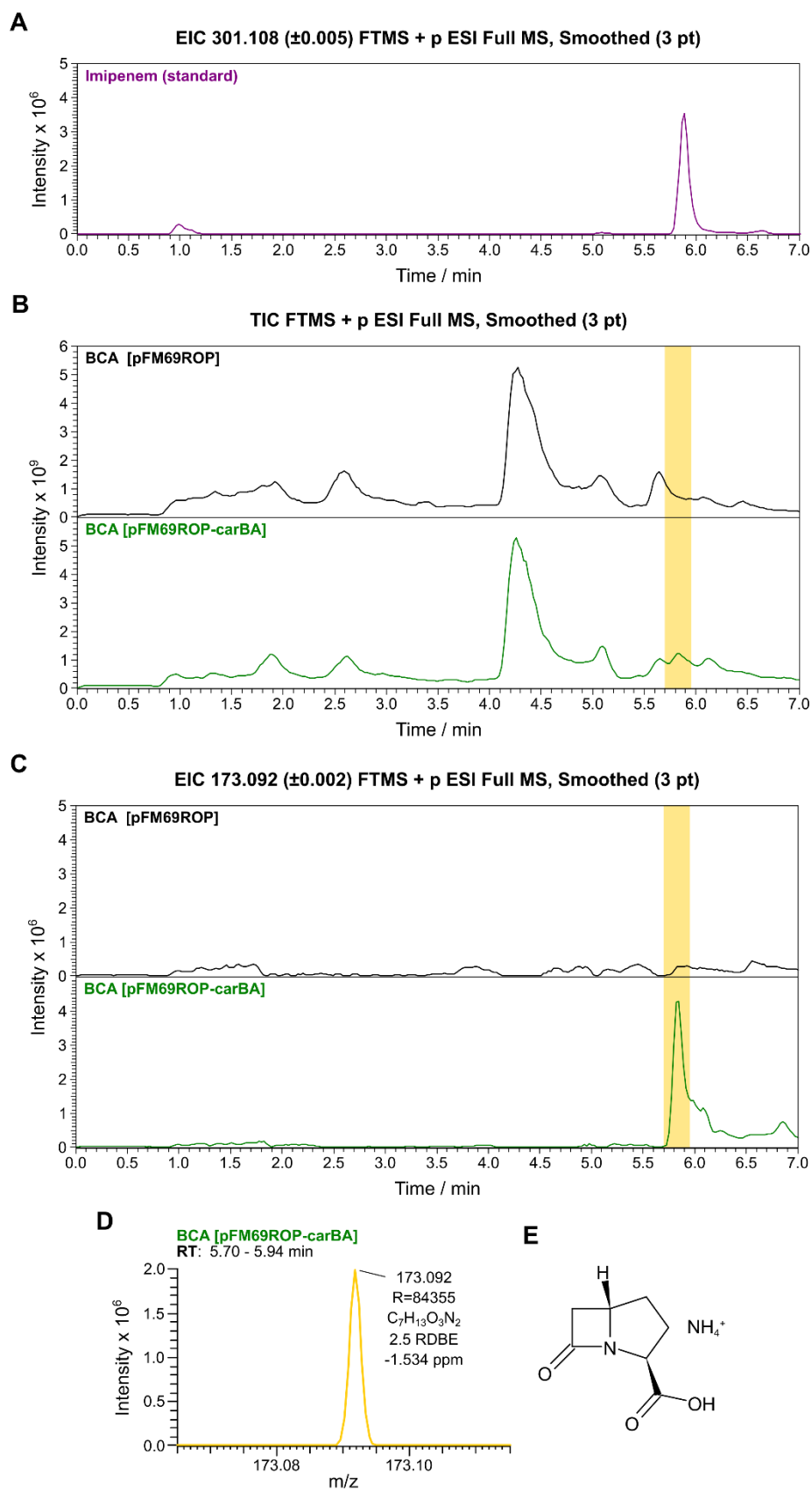


Figure 6. Detection of carbapenem production in *C. glutamicum* BCA strain harboring pFM69ROP-carBA plasmid. Results depict one biological replicate of strain BCA [pFM69ROP-carBA], and the

strain BCA [pFM69ROP] as control. Panel (A) represents the extracted ion chromatogram (EIC) at m/z 301.108 of the imipenem standard ($[M+H]^+ = 301.108$ Da). Panel (B) represents the total ion chromatogram (TIC), and panel (C) represents the EIC at m/z 173.092, which corresponds to adduct of carbapenam and an ammonium ion, in positive mode. Control strain corresponds to *C. glutamicum* BCA strain while clone C correspond to a single biological replicate of *C. glutamicum* BCA carrying pFM69ROP-carBA plasmid. The yellow stripe indicates the retention period of the novel peak at m/z 173.092 in the strain carrying pFM69ROP-carBA. Panel (D) represents the mass spectrum of the novel peak and associated details (m/z , resolution, molecular formula, ring double bond equivalents, and m/z accuracy). Panel (E) represents the molecular structure of the adduct of carbapenam with an ammonium ion.

Given that no commercial standard of carbapenam was available, the clinically used imipenem was used as reference; the monoprotonated imipenem ion (m/z 301.108) showed a retention time of 5.9 min. The total ion chromatograms of the analyzed strains had some differences in their metabolic profile, coincidentally, in BCA [pFM69ROP-carBA], a new shoulder at about min 5.9. The extract ion chromatogram at m/z 173.092, corresponding to the adduct of carbapenam with an ammonium ion, showed a clear FTMS⁺ signal at min 5.8 in the extract of *C. glutamicum* BCA carrying pFM69ROP-carBA, which was not present in the control strain. The FTMS⁺ spectrum showed that the identified signal was within 2 ppm of the theoretical monoisotopic mass of the said adduct. The MS² fragmentation pattern of the precursor ion at m/z 173.092 included signal matching monoprotonated carbapenam ion (m/z 156.065) and the decarboxylated carbapenam (5S)-7-oxo-1-azabicyclo[3.2.0]hept-2-ylum ion (m/z 110.060) (data not shown). The EIC at m/z 173.076, corresponding to the monoprotonated carbapenam precursor carboxymethylproline, also showed a low intensity FTMS⁺ signal at 6.0 min in the extract of BCA [pFM69ROP-carBA], but not in the control strain (data not shown).

IV.4 Discussion

In this chapter, the potential of engineered *C. glutamicum* strains to produce bottromycin A2 and (3S,5S)-carbapenam-3-carboxylic acid was addressed. The discussion on the progress achieved will encompass focus on four topics: (1) implementation of the BGC into the host, (2) the attempt to produce bottromycin, (3) the attempt to produce carbapenam, (4) the contribution for cost-effective and sustainable bioproduction of intermediates for discovery and production of antibiotics.

Two strains of *C. glutamicum* were employed as heterologous expression hosts: CR099attB_{C31} for bottromycin and BCA for the carbapenam. The BCA strain had been derived from CR099attB_{C31} by replacing the main carotenoid biosynthetic gene cluster with an attB_{BT1} locus for phage-based integrations (Marques, Luzhetskyy et al. 2020). BCA strain has a clearer metabolic profile in MS analysis than the ancestor and can accept integrative vectors based in the ϕ C31 and ϕ BT1 phages. Despite these advantages, for timing reasons the bottromycin BGC was expressed in CR099attB_{C31}: BCA strain did not exist yet at the time the experiment was performed. Regardless, both CR099attB_{C31} and BCA strains have a high transformation efficacy and heterologous protein production because of prior removal of their restriction system and a reduction of transposable elements in the parental strain ATCC 13032 (Baumgart, Unthan et al. 2013, Unthan, Baumgart et al. 2015). Accordingly, transformants of *C. glutamicum* carrying the selected BGC were obtained by electroporation. However, the frequency of transformation was 2-3 orders of magnitude lower than plasmids based in the same phage-based integration systems (Marques, Luzhetskyy et al. 2020). Even though electroporation is an efficient method for transformation of *C. glutamicum* with plasmids (Dunican and Shivnan 1989), physical constraints likely reduce the frequency of transformation with large vectors, such those carrying BGC. To tackle restriction systems and introduce large BGC into the cells, intergeneric conjugation has been widely adopted for *Streptomyces* and other organisms dedicated to industrial applications (reviewed in (Musiol-Kroll, Tocchetti et al. 2019)). Conjugation of vectors between *E. coli* and *C. glutamicum* is a well-established technique (Schafer, Kalinowski et al. 1990), but it is infrequently used due to high transformation efficiency with routine vectors. Even so, for heterologous expression of BGC, our results indicate that conjugation could be a valuable resource to transfer large vectors to *C. glutamicum*. Noteworthy, the DG2-kan-P41hyg cosmid and most vectors used in Actinomycetes already feature the required *oriT* locus for mobilization. In addition, the low frequency of transformation of *C. glutamicum* BCA with the small pFM69ROP-carBA plasmid could also be related to a metabolic burden caused by high-level expression of *carBA* operon under the influence of the strong constitutive P69 promoter (Zhang, Liu et al. 2018) and an unrepressed Rex operator (McLaughlin, Strain-Damerell et al. 2010) of the pFM69ROP vector. If so, the use of a host capable of temporarily downregulating *carBA* expression, such as the *C. glutamicum* REX strain (**see chapter III**), might improve the frequency of transformation.

Bottromycin production in *C. glutamicum* CR099attB_{C31} harboring DG2-kan-P41hyg could not be confirmed from the low-intensity and low-resolution MS data available. In fact, at first sight, the novel signal was completely overlooked, and, therefore, the studies on

bottromycin production were prematurely abandoned. Several reasons were likely to hamper the production of bottromycin in *C. glutamicum*. The regulatory elements used to refactor DG2-kan-P41hyg for the expression in *Streptomyces* (Horbal, Marques et al. 2018) could have hindered transcription and translation by the machinery of *C. glutamicum*. Strong promoters that led to toxic effects due to overproduction of bottromycin in a susceptible *S. lividans* host were nevertheless cloneable in an *E. coli* host (Horbal, Marques et al. 2018), highlighting the particularities of promoter recognized by *Streptomyces*' machinery. Follow-up studies using the DG2-kan-P41hyg determined that the actual promoters responsible for bottromycin overproduction were different from the expected, one of them (P_{AS}) being a previously cryptic promoter (Kuhl, Ruckert et al. 2021). The lack of characterization of the randomly generated promoters used in DG2-kan-P41hyg contribute to uncertain outcomes in *C. glutamicum*. In addition, the highly biased codon usage of the bottromycin BGC (sourced in *Streptomyces* sp. BC16019) could hinder its translation in *C. glutamicum*. The median %G+C, effective number of codon (base: 61) (Wright 1990) and adaptation index (base: 1, relatively to *C. glutamicum*) (Sharp and Li 1987) of the 12 genes of bottromycin BGC are 73%, 32 and 0.72, respectively; these values indicate a quite skewed and distant codon usage relatively to *C. glutamicum* own genes. Moreover, half of these biosynthetic genes have GTG as start codon, which is known to reduce expression relatively to ATG in *C. glutamicum* and *Streptomyces* spp. alike (Becker, Buschke et al. 2010, Myronovskiy, Welle et al. 2011). Altogether, these transcription and translation efficiencies could lead to insufficient or unbalanced expression of biosynthetic and transporter genes of bottromycin BGC. Lastly, even though bottromycin derives from a ribosomal peptide, insufficiency of coenzyme or co-factors associated to post-translation modifications (e.g., S-adenosylmethionine, Co(II) ion) could lead to production of intermediates but not its mature form (Crone, Leeper et al. 2012). Our observations indicate that the most common intermediates of bottromycin do not accumulate in the culture broth. May the detected MS signal be confirmed as bottromycin, minimal levels of coenzymes and co-factors should be available for post-translation modifications. Overall, the genetic and metabolic complexity of bottromycin BGC would probably require systematic troubleshooting of the appointed checkpoints, which was beyond the scope of this work. Furthermore, additional sensitivity to the detection of bottromycin by chromatographic techniques could be achieved by employing the carotenoid-free BCA strain.

The implementation of the biosynthetic operon *carBA* in *C. glutamicum* BCA strain led to a novel signal in MS chromatogram that matched the exact mass and fragmentation pattern of a carbapenam adduct. The evidence of production of carbapenam is not definitive though, as the lack of commercial standards hindered the identification by HPLC-MS.

Alternatively, purification and NMR analysis could provide evidence for production, but these efforts were not yet pursued. The probable production of carbapenam can be interpreted in the light of the implemented genetic circuits and the metabolic background of the *C. glutamicum* host. Comparatively to bottromycin BGC, the reduced length of the instructions required for carbapenam biosynthesis allowed for a bottom-up approach, where all components were carefully chosen for higher chances of expression. Regarding biosynthetic genes, the coding sequences *carA* and *carB* of *P. carotovora* were synthesized selecting the most frequent codons in *C. glutamicum* (ATCC 13032 entry at (Nakamura, Gojobori et al. 2000)) and organized as a *carBA* operon with a ribosome binding sequence per gene (inspired by (Sun, Zhao et al. 2016)). Regarding the expression vector, the ϕ BT1-based integrative plasmid selected for *carBA* expression has an insulated cloning site with a well-characterized promoter (Zhang, Liu et al. 2018) and a Rex operator (McLaughlin, Strain-Damerell et al. 2010), for strong and controllable gene expression in special strains (**see chapter III**). Regarding the expression host, the BCA strain was selected for its simpler metabolic profile and its gateway for ϕ BT1-based plasmids (Marques, Luzhetskyy et al. 2020). Finally, regarding the cultivation conditions, the CGXII medium with defined carbon and nitrogen sources was chosen as a convenient recipe to meet the specific requisites of carbapenam production (Keilhauer, Eggeling et al. 1993).

Upon the encouraging results regarding carbapenam, it seems appropriate to think about strategies to optimize titers and productivity for cost-effective bioproduction of this beta-lactam. Considering prior attempts of heterologous expression of carbapenems, namely in *E. coli* (Shomar, Gontier et al. 2018), the availability of precursors malonyl-CoA and pyrrolidine 5-carboxylate is likely a limiting factor to the production. In fact, the low titers of the polyketide flaviolin obtained in *C. glutamicum* (in comparison to other hosts) is likely to be due to limited malonyl-CoA supply (Yang, Kim et al. 2018, Marques, Luzhetskyy et al. 2020). A quick fix for enhanced malonyl-CoA could be the addition of cerulenin, an inhibitor of fatty acid elongation, to the media; however, the toxicity of the compound to *C. glutamicum* and its high cost renders this strategy economically unsustainable (Yang, Kim et al. 2018). Alternatively, an integrated metabolic engineering approach to overproduce malonyl-CoA in *C. glutamicum* could include de-regulating the expression of acetyl-CoA carboxylase subunits BCD1, de-regulating glucose transport via iolT, and reducing carbon flux through tricarboxylic acid cycle by downregulating citrate synthase and blocking pyruvate carboxylation, as proposed by Milke, Kallscheuer et al. (2019). The availability of the pyrroline 5-carboxylate precursor of carbapenam could be enhanced by downregulating pyrroline 5-carboxylate reductase (*ProC*); deletion of *proC* leads to the overproduction and secretion of pyrroline 5-carboxylate to the medium but provokes L-proline autotrophy (Ankri,

Serebrijski et al. 1996). Alternative, proline could be added to the media and then converted to the necessary precursors by the activity of pyrroline 5-carboxylate reductase (ProC, NADPH-forming reaction) or the putative membrane-bound proline dehydrogenase (PutA, reduced menaquinone-forming reaction, mentioned in (Bott and Niebisch 2003)). Metabolic engineered strains of *C. glutamicum* producing high titers of carbapenam could represent a cost-effective source of precursors for discovery and production of bioactive antibiotics.

The results presented in this chapter demonstrate that the developed *C. glutamicum* strains, molecular tools and engineering techniques can be put together to generate a *C. glutamicum*-based platform to produce natural products. Microbial production of bioactive natural products and their precursors fits into the ecopharmacognosy and green chemistry concepts (Anastas and Eghbali 2010, Cordell 2015), as it contributes for waste prevention, atom economy, energy efficiency, and the use of renewable feedstocks (rather than depleting resources). Natural product remain important assets for lead discovery, presenting good complementarity to ligands and balanced properties, and usually gaining potency during hit-to-lead optimization (Keseru and Makara 2009). The cues of the capacity to produce bottromycins and beta-lactams in *C. glutamicum* paves the way for the use of this well-established industrial organism as sustainable source of natural scaffolds for drug discovery and development.

IV.5 Conclusion and future perspectives

We tested the potential of engineered *C. glutamicum* strains to produce bottromycin A2 and (3S,5S)-carbapenam-3-carboxylic acid. The *C. glutamicum* hosts CR099attB_{C31} and BCA were successfully transformed with phage-based integrative vectors encoding the biosynthetic gene cluster of bottromycin (DG2-kan-P41hyg cosmid) and the biosynthetic operon of (3S,5S)-carbapenam-3-carboxylic acid (pFM69ROP-carBA plasmid), respectively. A novel, low-intensity mass spectrometry signal (m/z 823.4, rt 4.1 min) suggests that *C. glutamicum* CR099attB_{C31} [DG2-kan-P41hyg] may produce traces of bottromycin A2. A unique high-resolution mass spectrometry signal (m/z 173.092, rt 5.9 min) and its MS² pattern strongly suggest that *C. glutamicum* BCA [pFM69ROP-carBA] can produce (3S,5S)-carbapenam-3-carboxylic acid.

Improvements in the frequency of delivery of large BGC into *C. glutamicum* would require testing additional transformation and conjugation protocols. In the case of biosynthetic operons bearing strong promoters, it might be worthy considering regulatory systems, such as Rex and Rex operator, to cope with the transcriptional burden. Additional techniques would be required to confirm identity of the putative bottromycin signal, such as high-

resolution and tandem mass spectrometry. Improvements in bottromycin titers would likely involve refactoring the respective BGC with regulatory elements (promoters, RBS) previously characterized in *C. glutamicum*. The quantification of carbapenam would involve the upscaling of production, the purification of standards and preparation of a calibration curve. For that, improvements in carbapenam production might include metabolic rewiring for optimal supply of malonyl-CoA and pyrrolidine 5-carboxylate precursors.

CHAPTER V

General discussion and perspectives

General discussion and perspectives

The research presented in this work aimed at the development of methodologies and molecular tools to enable metabolic engineering *Corynebacterium glutamicum* for the production of active ingredients and their precursors. This work fits into the framework of multi-chassis engineering, which sustains that a diversified portfolio of host chassis would contribute to augment the probability of successful heterologous expression of biosynthetic gene clusters of natural products (Ke and Yoshikuni 2020). Firstly, we addressed metabolic engineering through the development of molecular tools and methods for genome editing, the validation of methods for delivery of heterologous biosynthetic genes, and the proof-of-principle production of natural products in *C. glutamicum* (chapter II). Secondly, we addressed sustainable production through the implementation of a genetic switch to balance cell growth and production of natural products in *C. glutamicum* (chapter III). Thirdly, we addressed production of active ingredients and their precursors through the implementation and expression of biosynthetic genes of pharmacologically relevant natural products in *C. glutamicum* (chapter IV). In this section, we discuss our results in the context of the recent advances in genome editing, control of gene expression, and production of bioactive natural products in *C. glutamicum*. Additionally, we give a perspective on the relevance of microbial factories as educational resources for the development of research experiences for undergraduates and the general public. In the end, we present some future perspectives for this research topic and reinforce the main conclusions of this thesis.

V.1 Prospects of a microbial factory of natural products based on *C. glutamicum*

Coming to this discussion, we may start by recalling why we believe that *Corynebacterium glutamicum* suits as a host organism to develop microbial factories for natural products. In our brief general introduction, we have attempted to showcase that *C. glutamicum* has been isolated specifically to produce amino acids and, as knowledge increased and ingenuity allowed, it become an excellent source of other commodity and fine chemicals. We also tried to highlight that the understanding of genetics and metabolism have enabled ongoing efforts to generate fully predictable *C. glutamicum* strains and respective molecular tools, for systems and synthetic biology approaches. Building on that groundwork, we considered that our research should focus on expanding the capacities of *C. glutamicum* as reliable microbial factory for the sustainable production of pharmaceuticals and precursors thereof. Even for an industrially used, genetically amenable, and well-described microorganism such as *C. glutamicum*, the development of full-blown microbial factories for valuable chemicals

is challenging: issues can arise from genetic instability of wild-type strains, significant metabolic burden or drainage associated to poor control of biosynthetic genes, or unexpected susceptibility to high concentrations of intermediates or products (Czajka, Wang et al. 2017). Concerning genetic instability, a relieving reassurance comes from studies where the genome of *C. glutamicum* has been stripped away of its mobile genetic elements in a stepwise manner (Baumgart, Unthan et al. 2013, Unthan, Baumgart et al. 2015, Baumgart, Unthan et al. 2018), leading to the development of strains completely free of prophages and transposase genes (Linder, Haak et al. 2021). Our work was based on the genome reduced CR099 strain, the stabler strain available at the beginning of our research, which is devoid of prophages and ISCg1-2 transposable elements (Unthan, Baumgart et al. 2015). Concerning the metabolic burden or drainage, it had been pointed out that the development of the next-generation industrial *C. glutamicum* strains would require “manipulating fluxes and the intricate regulatory networks (...), rather than genes and protein levels” (Vertes, Inui et al. 2012). Decisively, the manipulation of metabolic flows in *C. glutamicum* has been addressed by expanding its molecular toolbox, including precise genetic editing tools, tuneable transcription and translation regulatory elements, and compatible vectors; however, some of these tools were designed having in mind the development of microbial factories for C2-C6 platform chemicals (Baritugo, Kim et al. 2018). Our work focused heavily on developing a molecular toolbox that would enable the generation of optimized strains and the implementation of biosynthetic genes for complex chemical scaffolds, namely natural products with application as pharmaceuticals and their precursors. The polyketides (PK), the non-ribosomal peptides (NRP), and the ribosomally synthesized and post-translationally modified peptides (RiPP) were pharmaceutically relevant classes of natural products still underexplored in *C. glutamicum*. The only native polyketide synthase described in *C. glutamicum* is involved in the biosynthesis of corynomycolic acids (Gande, Gibson et al. 2004, Portevin, De Sousa-D'Auria et al. 2004), and no native biosynthetic gene clusters for NRP and RiPP have being described in *C. glutamicum* so far. To the date of beginning of our study (2017), heterologous production of non-native PK, NRP and RiPP in *C. glutamicum* had not been reported in the literature. The ability to produce high amounts of amino acids and the capacity to sustain the synthesis corynomycolic acids prompted us to explore the potential of *C. glutamicum* to produce non-native PK, NRP and RiPP. For that, we sought inspiration in tools and methods developed for heterologous expression of biosynthetic gene clusters in *Streptomyces* sp. (Huo, Hug et al. 2019). By employing tools transversal to a set of regularly used microbial factories, we sought to bring *C. glutamicum* closer to academics and companies that are actively engaged in drug discovery by means of heterologous gene expression. Concerning detrimental effects of the production of natural products, it has been encouraging to find

that *C. glutamicum* is quite resilient to harsh industrial conditions, displaying high shear tolerance and hardiness enabling cell recycling, and relatively high tolerance to fermentation inhibitors (cited from (Vertès, Inui et al. 2013)). Altogether, the reassuring scientific background and industrial attributes led us to postulate that additional methods and tools would enable the development of microbial factories based on *C. glutamicum* for the sustainable production of natural products.

Optimizing microbial factories through the reduction of genomic complexity is a cornerstone in success heterologous production of natural products (Beites and Mendes 2015). The roadmap for optimized microbial factories includes the selection of a suitable (starting) host and the refining the host for production using optimization of metabolic flux (Xu, Liu et al. 2020). The selection of the starting microbial platform is often a choice or compromise between a well-described model organism with developed engineering methods and a less amenable microorganism with a proven metabolic capacity to produce the class of natural products of interest. In that regard, *C. glutamicum* is a model microorganism with some developed tools and an efficient producer of amino acids and other added-value compounds. After host selection, the refinement of microbial factories enters the phase of metabolic flux optimization, which involves gene editing, the implementation of biosynthetic genes, and the introduction of regulatory systems (Xu, Liu et al. 2020). Our work aimed to develop the necessary tools and methods to contribute to the refinement of *C. glutamicum* as a microbial factory for natural products, especially PK, NRP, and RiPP. For that purpose, we divided our work into three objectives, namely: (i) to develop methods and tools to engineer *C. glutamicum* and test production of PK and NRP; (ii) to balance growth and production of natural products on renewable feedstocks; and (iii) to demonstrate the production of pharmaceutically relevant natural products. In the following sections, we will discuss our results concerning these objectives in light of recent developments.

V.2 Tools and methods to generate microbial factories based on *C. glutamicum* for production of natural products

Our priority in the development of a microbial factory based on *C. glutamicum* was the generation of strains and associated tools that comply with requirements of the production of natural products. Firstly, we needed to develop strategies to integrate the characteristic large biosynthetic gene clusters of natural products in *C. glutamicum*. For that, we got inspiration in the actinophage-based integrative vectors routinely used in *Streptomyces*. Our results support *C. glutamicum* could be transformed with ϕ C31-based vectors, provided that cognate integration site is encoded in the genome. Meantime, integrative vectors based on

the temperate corynephage $\phi 16$ were further developed for repeated introduction of foreign DNA into *C. glutamicum*, providing further methodologies for heterologous expression (Lobanova, Gorshkova et al. 2022). Secondly, we sought to develop a *C. glutamicum* strain with reduced metabolic footprint to ease detection of natural products produced in low abundance. Our target was the 7.6-kbp carotenoid biosynthetic gene cluster which is responsible for producing visible pigments of *C. glutamicum*. After some failed attempts to use existing CRISPR-Cas9 and adapted recombineering technologies (data not shown), we decided to generate a genomic library for homologous recombination and combine it with existing recombinase tools for markerless procedures. While designing our editing strategy, we took the opportunity to replace the carotenogenic genes by an attachment site for $\phi BT1$ -based vectors, for additional engineering flexibility. Our results indicate that the genome of *C. glutamicum* could be edited as intended using the generated library and excision system. Meantime, other research groups continued to mature genome editing techniques for *C. glutamicum*, including counter selection markers (Wu, Deng et al. 2018, Wang, Li et al. 2019), recombineering (Chang, Wang et al. 2021, Li, Swofford et al. 2021) and CRISPR-Cas technologies (reviewed in (Wang, Zhang et al. 2021)), enabling genome editing with increasing efficiency, precision or flexibility. Our results also indicated that the generated carotenoid-free *C. glutamicum* BCA strain showed a growth pattern similar to its parental strain, in defined media. Notably, a structural elucidation of cytochrome *bcc-aa₃* of *C. glutamicum* recently suggested a direct role of carotenoids in the protection from electron leaks and reactive oxygen species (Kao, Ortmann de Percin Northumberland et al. 2022); the effects of the albino phenotype in respiratory efficiency remain to be studied. Thirdly, we required proof-of-principle evidence that interesting classes of natural products could be produced in *C. glutamicum*. Our target classes of compounds were polyketides and non-ribosomal peptides, which encompass several pharmaceutically active ingredients and their precursors. For that, we chose the type III polyketide synthase RppA to produce red pigment flaviolin and the phosphopantetheinylation-dependent non-ribosomal peptide synthetase BpsA to produce the blue pigment indigoidine. The respective biosynthetic genes were purposely sourced from highly codon biased actinobacterial genomes to assess eventual genetic barriers. Worth mentioning, not before the genes were at least partially recoded and combined with well-characterized promoters and RBS could we obtain the expected products. Ultimately, our results indicate that the two selected classes of natural products could be produced in the *C. glutamicum* BCA strain using the developed phage-based vectors. Meantime, the production of the polyketide 6-methylsalicylic acid (Kallscheuer, Kage et al. 2019) and the overproduction of indigoidine (Ghiffary, Prabowo et al. 2021) were also demonstrated in genetically enhanced *C. glutamicum* strains, further confirming their metabolic capabilities to produce these classes of natural products. Overall, our results and

independent investigations have collectively laid groundwork for enhanced delivery of heterologous genes, more efficient and precise genome editing, and fruitful expression of polyketide and non-ribosomal peptides in *C. glutamicum*, hereby contributing to establish *C. glutamicum* as a microbial factory of natural products.

V.3 Regulatory systems to control gene expression in microbial factories based on *C. glutamicum* for production of natural products

Following the proof-of-principle production of natural products in *C. glutamicum* using the developed strain and tools, we focused on controlling the timing of expression of heterologous genes to optimize production. We sought for an orthologous transcriptional regulator that could serve to automatically adjust the expression of heterologous biosynthetic genes according to cell's energetic status. For that, we got inspiration in the universal Rex regulator, first characterized in *S. coelicolor*, which responds to intracellular NADH/NAD⁺ poise. Since *C. glutamicum* does not encode a native Rex protein nor an extended Rex regulon, we defined a strategy to implement a heterologous Rex system into the *C. glutamicum* BCA using the developed phage-based vectors. We showed that the resulting *C. glutamicum* REX strain effectively expresses the Rex regulator, albeit at the cost of a slightly longer lag phase and increased susceptibility to oxidative stressor menadione. Decisively, we also proved that the transcription of biosynthetic genes under control of Rex operator in the *C. glutamicum* REX strain follows a time-dependent pattern that coincides with the intracellular NADH/NAD⁺ poise. Thirdly, we determined that the usage of the Rex regulation system can increase the production of indigoidine up to 5-fold comparing to unregulated strains, in cultivations medium containing maltose as carbon source and limiting nitrogen supply. Meantime, other groups have developed genetic systems to synchronization of production to metabolic status of *C. glutamicum*. Straightforward switches that activate production of target compounds were based on systems responsive to phenolic compounds in the media (Siebert, Altenbuchner et al. 2021) or dissolved oxygen in the media (Kobayashi, Kawaguchi et al. 2020). Metabolic toggle switches that simultaneously activate production of target compounds and downregulate competing pathways have been develop based on quorum sensing systems (Liu, Shi et al. 2021), growth-phase responsive promoters combined to sequences targeting proteins for degradation (degrons) (Wei, Zhao et al. 2022), or carbon source-dependent regulators combined to xenogeneic silencing systems (Wiechert, Gatgens et al. 2020). In all cases, the timely control of production allowed for higher final content of interesting products on the provided, most valuable carbon source provided. Overall, our results and independent investigations pave the way for rational regulation of expression of heterologous

biosynthetic genes in *C. glutamicum*, hereby contributing to establish *C. glutamicum* as a sustainable microbial factory of natural products.

V.4 First steps towards the production of antimicrobials and precursors thereof in *C. glutamicum*

In possession of microbial factories based on *C. glutamicum* suitable for sustainable production of simple natural products, we attempted the production of complex molecules with pharmaceutical interest, namely antibiotics and precursors thereof. In that regard, we envisioned the implementation of biosynthetic genes for bottromycin A2 and carbapenam using phage-based vector in conjugation with streamlined *C. glutamicum* strains. In the case of bottromycin, we employed a refactored ϕ C31-based cosmid encoding the full BGC, which had been previously used for heterologous production in *Streptomyces* sp. Our results indicate that the biosynthetic genes of bottromycin could be easily installed in *C. glutamicum* CR099attB_{C31} strain by transformation. Moreover, we got hints of the production of bottromycin A2 in HPLC-MS chromatograms, though signals were of low intensity. In case of carbapenam, we recreated an existing biosynthetic operon with recoded genes and well-characterized regulatory elements into a ϕ BT1-based vector. Once again, our results indicate that carbapenam biosynthetic genes could easily be installed in *C. glutamicum* BCA strain. Importantly, we also observed hints of production of carbapenam in HPLC-MS chromatogram, namely using high resolution MS data and a standard compound for reference. Meantime, other groups have continued to report the production of diverse natural products in *C. glutamicum*, including flavonoids and other phenolic compounds (Milke, Mütz et al. 2020, Son, Jang et al. 2021, Wu, Liu et al. 2022), an vitamin B₂ analog with antibiotic activity (Mora-Lugo, Stegmüller et al. 2019), and the precursor of a lanthionine-containing antibiotic peptide (Weixler, Berghoff et al. 2022). Overall, our results and independent investigations indicate suitability of *C. glutamicum* to produce bioactive compounds, hereby contributing to establish *C. glutamicum* as a sustainable microbial factory for production of natural product, including antibiotics.

V.5 Microbial factories and their products as educational resources

Generalized science education is considered fundamental to enable active and responsible participation of all citizens in decision-making and innovation (Directorate-General for Research Innovation 2015). Back in the 1930s, Abel Salazar, patron of Instituto de Ciências Biomédicas Abel Salazar, had advocated that “creating a scientific mentality (...) is not the same thing as vulgarizing science”, and, thus, complementary to standard learning

environments, self-didacticism would be the way to shape free and conscious personalities (Salazar 1933). Recently, the rise of Do-It-Yourself biology (DIY Bio) movements and community labs contributed to bring the process of scientific discovery near to citizens, “serving as vehicles of public engagement and participation in scientific inquiry” (Scheifele and Burkett 2016). It has been argued that the collaboration between academics, students and “real world” communities can be fostered by a Connected Curriculum framework, which includes pedagogical approaches such as inquiry-based learning (Fung 2017). The very development of aptitude to scientific research of the sheer number of undergraduates requires novel formats, where “students have ownership of their individual projects, but the projects are done in parallel, enabling the faculty mentor to guide multiple students efficiently” (Carson 2007). Course-based Undergraduate Research Experiences (CURE) are laboratory learning environments in which “students address a research question or problem that is of interest to the broader community with an outcome that is unknown both to the students and to the instructor” (cited from (Auchincloss, Laursen et al. 2014)). Systematic review of literature on undergraduate research models provides evidence that CURE allow students to “engage intimately with scientific discovery and process” and “increase the persistence to pursue research experiences ahead in their career”, while fulfilling the prerequisite of having research experience for post-graduate studies (Ahmad and Al-Thani 2022). Furthermore, as CURE are designed to provide authentic research experience in collaborative context, they constitute an opportunity to convey core ethics and notions of responsible conduct of research (Olimpo, Diaz-Martinez et al. 2017). In a different perspective, U.S. teachers who collectively develop CURE frequently state tangible, faculty-centered motivations (e.g., pilot research projects; obtain grant money; form collaborations; recruit trained students) to establish these courses (Shortlidge, Banger et al. 2017). The development and implementation of a CURE is often perceived as a risky endeavour regarding costs, which can be eased by establishing collaborations across disciplines and using materials already available (Govindan, Pickett et al. 2020). In our view, microbial factories can be considered advanced educational resources for CURE.

Microbial factories are, by definition, useful resources for research and for industrial purposes. As genetically modified organisms (GMO), microbial factories also considered novel and non-obvious can be protected by patents for commercial purposes (Silver 2020). But GMO can certainly play other roles too. Memorably, back in 2016, we witnessed that the training in Molecular Biotechnology at Saarland University included the production of natural products in microbial factories based on *Streptomyces* sp. We noticed that the ease to manipulate, availability of molecular tools, fast growth, and robust heterologous expression of microbial factories brought evident pedagogical advantages in comparison to

unoptimized hosts. Later, in the course of our doctoral studies, we were invited to join a multidisciplinary team that developed and implemented the curricular unit of pedagogical innovation “Biolaboratory – Multidisciplinary Experimentation Project” (BioLab), hosted by Instituto de Ciências Biomédicas Abel Salazar, in collaboration with the Faculty of Sciences and the i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto (ICBAS 2020). The format of BioLab is inspired in that of CURE, where students and mentors co-designed and conducted research projects of interest to the broader community (Marques, Ramos et al. 2021, Ramos, Marques et al. 2022). During the first edition of BioLab, in 2021, one of the students’ projects was dedicated to investigating alternative bio-based dyes for clothing. In that project, the students took advantage of the blue pigment indigoidine produced by a microbial factory based on *C. glutamicum* developed in our research (Marques, Luzhetskyy et al. 2020). The BioLab project led to the isolation of the pigment from the bacteria, the development of protocols to dye cotton using indigoidine, and the identification of soil microbial isolates able to bleach tinted wastewaters (Marques, Mendes et al. 2022). The project also contributed to the development of a sense of critical thinking regarding the environmental and social impact of fashion, which was enthusiastically disseminated by the students during a public demonstration at European Researchers’ Night 2021 (Reis 2021).

Besides the BioLab course, other reported CURE involving microbiology and biotechnology support the use of microorganisms, their derivatives, or their products as valuable educational resources. Relevant scientific questions in microbiology have provided engaging topics for the development of CURE, such as investigations on the human microbiome, novel antibiotics, new genomes, or antibiotic resistance (cited in (Burmeister, Dickinson et al. 2021)). The optimization of the production of biofuels and fermented beverages using yeast, or the investigation on the viral immunity of bacteria involved in yogurt production are yet another examples of research topics where microbial factories were key resources for the development of CURE (Pedwell, Fraser et al. 2018, Petrie 2020). The biotechnology companies are aware of the value of CURE and, in some cases, have financially sustained these initiatives as a way to expose undergraduates to scientific thinking, while promoting soft skills such as communication skills and teamwork (Buxeda and Saliceti-Piazza 2010). Overall, biotechnologically valued microorganisms used as educational resources seem able to contribute significantly for the initial training of the next generation of scientists, for the involvement of whole communities into the scientific mindset and, consequently, to foster knowledge-based innovation. After all, it seems, not only patented microorganisms can generate profits for the society.

V.6 Future perspectives

Considering our results, we unpretentiously suggest some scientific inquiries and technological demonstrations that could contribute to advance *C. glutamicum* as a sustainable source of natural products. The inquiries deal with the unanswered or unexplored topics that arouse during this investigation, while the demonstrations relate to assessments of the applicability of *C. glutamicum* strains and methods to produce natural products in industrial settings.

- Continue to investigate the genetics of *C. glutamicum*, especially unknown coding sequences, unclonable loci, and regulators and their regulons, so we can inform the development of more robust and predictable microbial factories.
- Review and righteously compare the available molecular tools and methods for genome editing in *C. glutamicum*, so we can establish their efficiency and convenience for each intended purpose and beyond.
- Continue to investigate, expand, and assess the industrial value of catabolic capacities of *C. glutamicum*, so we can not only take advantage of diverse renewable feedstocks as nutrients but also prevent the degradation of added value products.
- Continue to investigate the metabolic triggers of *C. glutamicum* associated to certain growth phases and different dissolved oxygen levels, so we can improve the design of regulatory systems to synchronize production of natural products.
- Continue to investigate and model the metabolic capacity to provide for precursors, coenzymes, co-factors and post-translation modifications related to secondary metabolism in *C. glutamicum*, so we can infer metabolic bottlenecks to production of natural products.
- Systematically address the factors affecting the expression of heterologous biosynthetic gene clusters from various sources in *C. glutamicum*, so we can pinpoint the major barriers to transgenic production of natural products.
- Incorporate microbial factories based on *C. glutamicum* and its products as educational resources for the development of further course-based undergraduate research experiences and other training activities for undergrads and general public;
- Incorporate microbial factories based on *C. glutamicum* into the panel of strains routinely used for the screening promising genetic libraries, so we can gradually understand the classes of compounds and genetic sources that are better expressed in this organism.

V.6 Conclusions

In conclusion, we reiterate the three main arguments which, based in our research and other recent works, support our thesis that microbial factories based on *Corynebacterium glutamicum* can contribute for sustainable production of natural products:

- 1) The actinobacterium *C. glutamicum* is amenable to the metabolic engineering processes required to generate microbial factories, which have been facilitated by an increasing number of validated molecular tools and methods.
- 2) The metabolism of *C. glutamicum* is programmable for a sustainable production of added value compounds, as exemplified by high-productivity strains controlled with repurposed regulatory systems of gene expression.
- 3) Microbial factories based in *C. glutamicum* are suitable for the production of bioactive natural compounds and their precursors, enabled by the implementation of relevant biosynthetic genes into streamlined strains using high-capacity vectors.

General references

Aguilar, A., T. Twardowski and R. Wohlgemuth (2019). "Bioeconomy for Sustainable Development." Biotechnol J **14**(8): e1800638.

Ahmad, Z. and N. J. Al-Thani (2022). "Undergraduate Research Experience Models: A systematic review of the literature from 2011 to 2021." International Journal of Educational Research **114**: 101996.

Anastas, P. and N. Eghbali (2010). "Green chemistry: principles and practice." Chem Soc Rev **39**(1): 301-312.

Ankri, S., I. Serebrijski, O. Reyes and G. Leblon (1996). "Mutations in the *Corynebacterium glutamicum* proline biosynthetic pathway: a natural bypass of the proA step." J Bacteriol **178**(15): 4412-4419.

Atzori, A., V. N. Malviya, G. Malloci, J. Dreier, K. M. Pos, A. V. Vargiu and P. Ruggerone (2019). "Identification and characterization of carbapenem binding sites within the RND-transporter AcrB." Biochim Biophys Acta Biomembr **1861**(1): 62-74.

Auchincloss, L. C., S. L. Laursen, J. L. Branchaw, K. Eagan, M. Graham, D. I. Hanauer, G. Lawrie, C. M. McLinn, N. Pelaez, S. Rowland, M. Towns, N. M. Trautmann, P. Varma-Nelson, T. J. Weston and E. L. Dolan (2014). "Assessment of course-based undergraduate research experiences: a meeting report." CBE Life Sci Educ **13**(1): 29-40.

Bakkes, P. J., P. Ramp, A. Bida, D. Dohmen-Olma, M. Bott and R. Freudl (2020). "Improved pEKEx2-derived expression vectors for tightly controlled production of recombinant proteins in *Corynebacterium glutamicum*." Plasmid **112**: 102540.

Baltz, R. H. (2017). "Gifted microbes for genome mining and natural product discovery." J Ind Microbiol Biotechnol **44**(4-5): 573-588.

Baltz, R. H. (2019). "Natural product drug discovery in the genomic era: realities, conjectures, misconceptions, and opportunities." J Ind Microbiol Biotechnol **46**(3-4): 281-299.

Baritugo, K. A., H. T. Kim, Y. David, J. I. Choi, S. H. Hong, K. J. Jeong, J. H. Choi, J. C. Joo and S. J. Park (2018). "Metabolic engineering of *Corynebacterium glutamicum* for fermentative production of chemicals in biorefinery." Appl Microbiol Biotechnol **102**(9): 3915-3937.

Barnard, A. M., S. D. Bowden, T. Burr, S. J. Coulthurst, R. E. Monson and G. P. Salmond (2007). "Quorum sensing, virulence and secondary metabolite production in plant soft-rotting bacteria." Philos Trans R Soc Lond B Biol Sci **362**(1483): 1165-1183.

Baumgart, M., S. Unthan, R. Kloss, A. Radek, T. Polen, N. Tenhaef, M. F. Muller, A. Kuberl, D. Siebert, N. Bruhl, K. Marin, S. Hans, R. Kramer, M. Bott, J. Kalinowski, W. Wiechert, G. Seibold, J. Frunzke, C. Ruckert, V. F. Wendisch and S. Noack (2018). "Corynebacterium glutamicum Chassis C1*: Building and Testing a Novel Platform Host for Synthetic Biology and Industrial Biotechnology." ACS Synth Biol **7**(1): 132-144.

Baumgart, M., S. Unthan, C. Ruckert, J. Sivalingam, A. Grunberger, J. Kalinowski, M. Bott, S. Noack and J. Frunzke (2013). "Construction of a prophage-free variant of Corynebacterium glutamicum ATCC 13032 for use as a platform strain for basic research and industrial biotechnology." Appl Environ Microbiol **79**(19): 6006-6015.

Becker, J., N. Buschke, R. Bückner and C. Wittmann (2010). "Systems level engineering of Corynebacterium glutamicum - Reprogramming translational efficiency for superior production." Engineering in Life Sciences **10**(5): 430-438.

Becker, J. and C. Wittmann (2021). Metabolic Engineering of Corynebacterium glutamicum. Metabolic Engineering: 403-468.

Beites, T. and M. V. Mendes (2015). "Chassis optimization as a cornerstone for the application of synthetic biology based strategies in microbial secondary metabolism." Front Microbiol **6**: 906.

Bott, M. and A. Niebisch (2003). "The respiratory chain of Corynebacterium glutamicum." J Biotechnol **104**(1-3): 129-153.

Burmeister, A. R., K. Dickinson and M. J. Graham (2021). "Bridging Trade-Offs between Traditional and Course-Based Undergraduate Research Experiences by Building Student Communication Skills, Identity, and Interest." J Microbiol Biol Educ **22**(2): e00156-00121.

Butler, M. S. and A. D. Buss (2006). "Natural products--the future scaffolds for novel antibiotics?" Biochem Pharmacol **71**(7): 919-929.

Buxeda, R. and L. Saliceti-Piazza (2010). BIO-MINDS: AN INDUSTRY-ACADEMIA PROGRAM TO PROMOTE UNDERGRADUATE RESEARCH IN BIOTECHNOLOGY AND RELATED AREAS. ICERI2010 Proceedings, IATED.

Bycroft, B. W., C. Maslen, S. J. Box, A. Brown and J. W. Tyler (1988). "The biosynthetic implications of acetate and glutamate incorporation into (3R,5R)-carbapenam-3-carboxylic acid and (5R)-carbapen-2-em-3-carboxylic acid by *Serratia* sp." J Antibiot (Tokyo) **41**(9): 1231-1242.

Carson, S. (2007). "A new paradigm for mentored undergraduate research in molecular microbiology." CBE Life Sci Educ **6**(4): 343-349.

Chang, Y., Q. Wang, T. Su and Q. Qi (2021). "Identification of phage recombinase function unit in genus *Corynebacterium*." Appl Microbiol Biotechnol **105**(12): 5067-5075.

Chen, Y., D. Banerjee, A. Mukhopadhyay and C. J. Petzold (2020). "Systems and synthetic biology tools for advanced bioproduction hosts." Curr Opin Biotechnol **64**: 101-109.

Cho, J. S., G. B. Kim, H. Eun, C. W. Moon and S. Y. Lee (2022). "Designing Microbial Cell Factories for the Production of Chemicals." JACS Au **2**(8): 1781-1799.

Cook, M. A. and G. D. Wright (2022). "The past, present, and future of antibiotics." Sci Transl Med **14**(657): eabo7793.

Cordell, G. A. (2015). "Ecopharmacognosy and the responsibilities of natural product research to sustainability." Phytochemistry Letters **11**: 332-346.

Costelloe, C., C. Metcalfe, A. Lovering, D. Mant and A. D. Hay (2010). "Effect of antibiotic prescribing in primary care on antimicrobial resistance in individual patients: systematic review and meta-analysis." BMJ **340**: c2096.

Crone, W. J. K., F. J. Leeper and A. W. Truman (2012). "Identification and characterisation of the gene cluster for the anti-MRSA antibiotic bottromycin: expanding the biosynthetic diversity of ribosomal peptides." Chemical Science **3**(12): 3516-3521.

Czajka, J., Q. Wang, Y. Wang and Y. J. Tang (2017). "Synthetic biology for manufacturing chemicals: constraints drive the use of non-conventional microbial platforms." Appl Microbiol Biotechnol **101**(20): 7427-7434.

Della Corte, D., H. L. van Beek, F. Syberg, M. Schallmey, F. Tobola, K. U. Cormann, C. Schlicker, P. T. Baumann, K. Krumbach, S. Sokolowsky, C. J. Morris, A. Grunberger, E. Hofmann, G. F. Schroder and J. Marienhagen (2020). "Engineering and application of a biosensor with focused ligand specificity." Nat Commun **11**(1): 4851.

Directorate-General for Research Innovation (2015). Science education for responsible citizenship: report to the European Commission of the expert group on science education. LU, Publications Office of the European Union.

Dunican, L. K. and E. Shivnan (1989). "High Frequency Transformation of Whole Cells of Amino Acid Producing Coryneform Bacteria Using High Voltage Electroporation." Nature Biotechnology **7**(10): 1067-1070.

Eggeling, L. and M. Bott (2005). Handbook of corynebacterium glutamicum. Boca Raton, Taylor & Francis.

FDA, U. S. (2019, 2019/04/26). "GRN No. 792 Cultured corn starch fermented by Corynebacterium glutamicum." GRAS Notices, from https://www.cfsanappsexternal.fda.gov/scripts/fdcc/index.cfm?set=GRASNotices&id=792&sort=GRN_No&order=DESC&startrow=1&type=basic&search=corynebacterium.

Feierabend, M., A. Renz, E. Zelle, K. Noh, W. Wiechert and A. Dräger (2021). "High-Quality Genome-Scale Reconstruction of Corynebacterium glutamicum ATCC 13032." Front Microbiol **12**: 750206.

Ferreira, E. A., C. C. Pacheco, J. S. Rodrigues, F. Pinto, P. Lamosa, D. Fuente, J. Urchueguia and P. Tamagnini (2021). "Heterologous Production of Glycine Betaine Using Synechocystis sp. PCC 6803-Based Chassis Lacking Native Compatible Solutes." Front Bioeng Biotechnol **9**: 821075.

Fung, D. (2017). Connected Curriculum for Higher Education, UCL Press.

Gally, J. M., A. Pahl, P. Czodrowski and H. Waldmann (2021). "Pseudonatural Products Occur Frequently in Biologically Relevant Compounds." J Chem Inf Model **61**(11): 5458-5468.

Gande, R., K. J. Gibson, A. K. Brown, K. Krumbach, L. G. Dover, H. Sahm, S. Shioyama, T. Oikawa, G. S. Besra and L. Eggeling (2004). "Acyl-CoA carboxylases (accD2 and accD3), together with a unique polyketide synthase (Cg-pks), are key to mycolic acid biosynthesis in Corynebacteriaceae such as Corynebacterium glutamicum and Mycobacterium tuberculosis." J Biol Chem **279**(43): 44847-44857.

Ghiffary, M. R., C. P. S. Prabowo, K. Sharma, Y. Yan, S. Y. Lee and H. U. Kim (2021). "High-Level Production of the Natural Blue Pigment Indigoidine from Metabolically

Engineered *Corynebacterium glutamicum* for Sustainable Fabric Dyes." ACS Sustainable Chemistry & Engineering **9**(19): 6613-6622.

Govindan, B., S. Pickett and B. Riggs (2020). "Fear of the CURE: A Beginner's Guide to Overcoming Barriers in Creating a Course-Based Undergraduate Research Experience." J Microbiol Biol Educ **21**(2): 50.

Grossman, T. H. (2016). "Tetracycline Antibiotics and Resistance." Cold Spring Harb Perspect Med **6**(4): a025387.

Hanahan, D. (1983). "Studies on transformation of *Escherichia coli* with plasmids." J Mol Biol **166**(4): 557-580.

Hazards, E. P. o. B., K. Koutsoumanis, A. Allende, A. Alvarez-Ordóñez, D. Bolton, S. Bover-Cid, M. Chemaly, R. Davies, F. Hilbert, R. Lindqvist, M. Nauta, L. Peixe, G. Ru, M. Simmons, P. Skandamis, E. Suffredini, P. S. Cocconcelli, P. S. Fernandez Escamez, M. P. Maradona, A. Querol, J. E. Suarez, I. Sundh, J. Vlak, F. Barizzzone, S. Correia and L. Herman (2019). "Update of the list of QPS-recommended biological agents intentionally added to food or feed as notified to EFSA 9: suitability of taxonomic units notified to EFSA until September 2018." EFSA J **17**(1): e05555.

Hoff, B., J. Plassmeier, M. Blankschien, A. C. Letzel, L. Kourtz, H. Schroder, W. Koch and O. Zelder (2021). "Unlocking Nature's Biosynthetic Power-Metabolic Engineering for the Fermentative Production of Chemicals." Angew Chem Int Ed Engl **60**(5): 2258-2278.

Horbal, L., F. Marques, S. Nadmid, M. V. Mendes and A. Luzhetskyy (2018). "Secondary metabolites overproduction through transcriptional gene cluster refactoring." Metabolic engineering **49**: 299-315.

Huo, L., J. J. Hug, C. Fu, X. Bian, Y. Zhang and R. Muller (2019). "Heterologous expression of bacterial natural product biosynthetic pathways." Nat Prod Rep **36**(10): 1412-1436.

Huo, L., S. Rachid, M. Stadler, S. C. Wenzel and R. Muller (2012). "Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis." Chem Biol **19**(10): 1278-1287.

ICBAS. (2020). "ICBAS - Biolaboratório - Projeto de Experimentação Multidisciplinar." Retrieved 18/10/2022, 2022, from <https://s.up.pt/kw4l>.

Kahan, J. S., F. M. Kahan, R. Goegelman, S. A. Currie, M. Jackson, E. O. Stapley, T. W. Miller, A. K. Miller, D. Hendlin, S. Mochales, S. Hernandez, H. B. Woodruff and J. Birnbaum

(1979). "Thienamycin, a new beta-lactam antibiotic. I. Discovery, taxonomy, isolation and physical properties." J Antibiot (Tokyo) **32**(1): 1-12.

Kalinowski, J., B. Bathe, D. Bartels, N. Bischoff, M. Bott, A. Burkovski, N. Dusch, L. Eggeling, B. J. Eikmanns, L. Gaigalat, A. Goesmann, M. Hartmann, K. Huthmacher, R. Kramer, B. Linke, A. C. McHardy, F. Meyer, B. Mockel, W. Pfefferle, A. Puhler, D. A. Rey, C. Ruckert, O. Rupp, H. Sahm, V. F. Wendisch, I. Wiegrabe and A. Tauch (2003). "The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins." J Biotechnol **104**(1-3): 5-25.

Kallscheuer, N., H. Kage, L. Milke, M. Nett and J. Marienhagen (2019). "Microbial synthesis of the type I polyketide 6-methylsalicylate with *Corynebacterium glutamicum*." Appl Microbiol Biotechnol **103**(23-24): 9619-9631.

Kao, W. C., C. Ortmann de Percin Northumberland, T. C. Cheng, J. Ortiz, A. Durand, O. von Loeffelholz, O. Schilling, M. L. Biniossek, B. P. Klaholz and C. Hunte (2022). "Structural basis for safe and efficient energy conversion in a respiratory supercomplex." Nat Commun **13**(1): 545.

Karady, S., J. S. Amato, R. A. Reamer and L. M. Weinstock (2002). "Stereospecific conversion of penicillin to thienamycin." Journal of the American Chemical Society **103**(22): 6765-6767.

Katz, L. and R. H. Baltz (2016). "Natural product discovery: past, present, and future." J Ind Microbiol Biotechnol **43**(2-3): 155-176.

Ke, J. and Y. Yoshikuni (2020). "Multi-chassis engineering for heterologous production of microbial natural products." Curr Opin Biotechnol **62**: 88-97.

Keilhauer, C., L. Eggeling and H. Sahm (1993). "Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB-ilvN-ilvC* operon." J Bacteriol **175**(17): 5595-5603.

Keseru, G. M. and G. M. Makara (2009). "The influence of lead discovery strategies on the properties of drug candidates." Nat Rev Drug Discov **8**(3): 203-212.

Kiefer, D., M. Merkel, L. Lilge, R. Hausmann and M. Henkel (2021). "High cell density cultivation of *Corynebacterium glutamicum* on bio-based lignocellulosic acetate using pH-coupled online feeding control." Bioresour Technol **340**: 125666.

Kieser, T., M. J. Bibb, M. J. Buttner, K. F. Chater and D. A. Hopwood (2000). Practical streptomyces genetics, John Innes Foundation Norwich.

Kinoshita, S., K. Nakayama and S. Akita (2014). "Taxonomical Study of Glutamic Acid Accumulating Bacteria, *Micrococcus glutamicus* nov. sp." Bulletin of the Agricultural Chemical Society of Japan **22**(3): 176-185.

Kobayashi, S., H. Kawaguchi, T. Shirai, K. Ninomiya, K. Takahashi, A. Kondo and Y. Tsuge (2020). "Automatic Redirection of Carbon Flux between Glycolysis and Pentose Phosphate Pathway Using an Oxygen-Responsive Metabolic Switch in *Corynebacterium glutamicum*." ACS Synth Biol **9**(4): 814-826.

Kobayashi, Y., M. Ichioka, T. Hirose, K. Nagai, A. Matsumoto, H. Matsui, H. Hanaki, R. Masuma, Y. Takahashi, S. Omura and T. Sunazuka (2010). "Bottromycin derivatives: efficient chemical modifications of the ester moiety and evaluation of anti-MRSA and anti-VRE activities." Bioorg Med Chem Lett **20**(20): 6116-6120.

Kogure, T. and M. Inui (2018). "Recent advances in metabolic engineering of *Corynebacterium glutamicum* for bioproduction of value-added aromatic chemicals and natural products." Appl Microbiol Biotechnol **102**(20): 8685-8705.

Kuhl, M., C. Ruckert, L. Glaser, S. Beganovic, A. Luzhetskyy, J. Kalinowski and C. Wittmann (2021). "Microparticles enhance the formation of seven major classes of natural products in native and metabolically engineered actinobacteria through accelerated morphological development." Biotechnol Bioeng **118**(8): 3076-3093.

Kuska, J. and E. O'Reilly (2020). "Engineered biosynthetic pathways and biocatalytic cascades for sustainable synthesis." Curr Opin Chem Biol **58**: 146-154.

Lennox, E. S. (1955). "Transduction of linked genetic characters of the host by bacteriophage P1." Virology **1**(2): 190-206.

Lerchen, H.-G., G. Schiffer, H. BrÖtz-Österheld, A. Mayer-Bartschmid, S. Eckermann, C. Freiberg, R. Endermann, J. Schuhmacher, H. Meier, N. Svenstrup, S. Seip, M. Gehling and D. Häbich (2006). Cyclische iminopeptid-derivate. O. World Intellectual Property, Aicuris Gmbh & Co. Kg.

Li, C., C. A. Swofford, C. Ruckert and A. J. Sinskey (2021). "Optimizing recombineering in *Corynebacterium glutamicum*." Biotechnol Bioeng **118**(6): 2255-2264.

Li, R., A. Stapon, J. T. Blanchfield and C. A. Townsend (2000). "Three Unusual Reactions Mediate Carbapenem and Carbapenam Biosynthesis." Journal of the American Chemical Society **122**(38): 9296-9297.

Lima, L. M., B. Silva, G. Barbosa and E. J. Barreiro (2020). "beta-lactam antibiotics: An overview from a medicinal chemistry perspective." Eur J Med Chem **208**: 112829.

Linder, M., M. Haak, A. Botes, J. Kalinowski and C. Ruckert (2021). "Construction of an IS-Free *Corynebacterium glutamicum* ATCC 13 032 Chassis Strain and Random Mutagenesis Using the Endogenous ISCg1 Transposase." Front Bioeng Biotechnol **9**: 751334.

Liu, H., F. Shi, S. Tan, X. Yu, W. Lai and Y. Li (2021). "Engineering a Bifunctional ComQXPA-PsrfA Quorum-Sensing Circuit for Dynamic Control of Gene Expression in *Corynebacterium glutamicum*." ACS Synth Biol **10**(7): 1761-1774.

Lobanova, J. S., N. V. Gorshkova, A. A. Krylov, N. V. Stoyanova and S. V. Mashko (2022). "Genome engineering of the *Corynebacterium glutamicum* chromosome by the Extended Dual-In/Out strategy." J Microbiol Methods **200**: 106555.

Marques, F., A. Luzhetskyy and M. V. Mendes (2020). "Engineering *Corynebacterium glutamicum* with a comprehensive genomic library and phage-based vectors." Metab Eng **62**: 221-234.

Marques, F., M. V. Mendes and M. S. Almeida (2022). A microbiology outreach project within training in responsible research and innovation. Lisbon.

Marques, F., P. D. Ramos, M. M. Lopes, J. B. Santos, F. Tavares, I. A. S. Olsson and M. S. Almeida (2021). BioLab – um espaço de experimentação e cruzamento disciplinar. . Livro de Atas do 7.º Congresso Nacional de Práticas Pedagógicas no Ensino Superior. S. C. Soares, F. Remião, A. V. Martins and S. Nunes. Online, UA Editora – Universidade de Aveiro: 295-301.

McGowan, S. J., M. Sebahia, S. O'Leary, K. R. Hardie, P. Williams, G. S. Stewart, B. W. Bycroft and G. P. Salmond (1997). "Analysis of the carbapenem gene cluster of *Erwinia carotovora*: definition of the antibiotic biosynthetic genes and evidence for a novel beta-lactam resistance mechanism." Mol Microbiol **26**(3): 545-556.

McGowan, S. J., M. Sebahia, L. E. Porter, G. Stewart, P. Williams, B. W. Bycroft and G. P. C. Salmond (1996). "Analysis of bacterial carbapenem antibiotic production genes reveals a novel beta-lactam biosynthesis pathway." Mol Microbiol **22**(3): 415-426.

McLaughlin, K. J., C. M. Strain-Damerell, K. Xie, D. Brekasis, A. S. Soares, M. S. Paget and C. L. Kielkopf (2010). "Structural basis for NADH/NAD⁺ redox sensing by a Rex family repressor." Mol Cell **38**(4): 563-575.

Miethke, M., M. Pieroni, T. Weber, M. Bronstrup, P. Hammann, L. Halby, P. B. Arimondo, P. Glaser, B. Aigle, H. B. Bode, R. Moreira, Y. Li, A. Luzhetskyy, M. H. Medema, J. L. Pernodet, M. Stadler, J. R. Tormo, O. Genilloud, A. W. Truman, K. J. Weissman, E. Takano, S. Sabatini, E. Stegmann, H. Brotz-Oesterhelt, W. Wohlleben, M. Seemann, M. Empting, A. K. H. Hirsch, B. Loretz, C. M. Lehr, A. Titz, J. Herrmann, T. Jaeger, S. Alt, T. Hesterkamp, M. Winterhalter, A. Schiefer, K. Pfarr, A. Hoerauf, H. Graz, M. Graz, M. Lindvall, S. Ramurthy, A. Karlen, M. van Dongen, H. Petkovic, A. Keller, F. Peyrane, S. Donadio, L. Fraisse, L. J. V. Piddock, I. H. Gilbert, H. E. Moser and R. Muller (2021). "Towards the sustainable discovery and development of new antibiotics." Nat Rev Chem **5**(10): 726-749.

Milke, L., N. Kallscheuer, J. Kappelmann and J. Marienhagen (2019). "Tailoring *Corynebacterium glutamicum* towards increased malonyl-CoA availability for efficient synthesis of the plant pentaketide noreugenin." Microb Cell Fact **18**(1): 71.

Milke, L., M. Mutz and J. Marienhagen (2020). "Synthesis of the character impact compound raspberry ketone and additional flavoring phenylbutanoids of biotechnological interest with *Corynebacterium glutamicum*." Microb Cell Fact **19**(1): 92.

Miller, W. J., L. Chalet, G. Rasmussen, B. Christensen, J. Hannah, A. K. Miller and F. J. Wolf (1968). "Bottromycin. Separation of biologically active compounds and preparation and testing of amide derivatives." J Med Chem **11**(4): 746-749.

Miyashita, M., N. Chida and A. Yoshikohsi (1982). "Synthesis of the precursor of (+)-thienamycin utilizing D-glucosamine." Journal of the Chemical Society, Chemical Communications(23): 1354-1356.

Mochales, S. (1994). "Forty years of screening programmes for antibiotics." Microbiologia **10**(4): 331-342.

Monod, J. (1970). O acaso e a necessidade. Mem Martins, Europa-América.

Monson, R. E., K. Jones and G. P. C. Salmond (2018). "Draft Genome Sequence of *Pectobacterium carotovorum* subsp. *carotovorum* ATCC 39048, a Carbapenem-Producing Phytopathogen." Microbiol Resour Announc **7**(5): e00825-00818.

Mora-Lugo, R., J. Stegmüller and M. Mack (2019). "Metabolic engineering of roseoflavin-overproducing microorganisms." Microb Cell Fact **18**(1): 146.

Musiol-Kroll, E. M., A. Tocchetti, M. Sosio and E. Stegmann (2019). "Challenges and advances in genetic manipulation of filamentous actinomycetes - the remarkable producers of specialized metabolites." Nat Prod Rep **36**(9): 1351-1369.

Myronovskyi, M. and A. Luzhetskyy (2019). "Heterologous production of small molecules in the optimized *Streptomyces* hosts." Nat Prod Rep **36**(9): 1281-1294.

Myronovskyi, M., E. Welle, V. Fedorenko and A. Luzhetskyy (2011). "Beta-glucuronidase as a sensitive and versatile reporter in actinomycetes." Appl Environ Microbiol **77**(15): 5370-5383.

Nakamura, Y., T. Gojobori and T. Ikemura (2000). "Codon usage tabulated from international DNA sequence databases: status for the year 2000." Nucleic Acids Res **28**(1): 292.

Nett, M., H. Ikeda and B. S. Moore (2009). "Genomic basis for natural product biosynthetic diversity in the actinomycetes." Nat Prod Rep **26**(11): 1362-1384.

Newman, D. J. and G. M. Cragg (2020). "Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019." J Nat Prod **83**(3): 770-803.

Olimpo, J. T., L. A. Diaz-Martinez, J. M. Bhatt and C. E. D'Arcy (2017). "Integration of RCR and Ethics Education into Course-Based Undergraduate Research Experiences in the Biological Sciences: A Needed Discussion." J Microbiol Biol Educ **18**(2): 18.12.30.

Otaka, T. and A. Kaji (1976). "Mode of action of bottromycin A2. Release of aminoacyl- or peptidyl-tRNA from ribosomes." J Biol Chem **251**(8): 2299-2306.

Parker, W. L., M. L. Rathnum, J. S. Wells, Jr., W. H. Trejo, P. A. Principe and R. B. Sykes (1982). "SQ 27,860, a simple carbapenem produced by species of *Serratia* and *Erwinia*." J Antibiot (Tokyo) **35**(6): 653-660.

Parte, A., W. B. Whitman, M. Goodfellow, P. Kämpfer, H.-J. Busse, M. E. Trujillo, W. Ludwig and K.-i. Suzuki (2012). Bergey's manual of systematic bacteriology: volume 5: the Actinobacteria, Springer Science & Business Media.

Pawar, S. V., J. C. Ho, G. D. Yadav and V. G. Yadav (2017). "The Impending Renaissance in Discovery & Development of Natural Products." Curr Top Med Chem **17**(2): 251-267.

Pedwell, R. K., J. A. Fraser, J. T. H. Wang, J. K. Clegg, J. D. Chartres and S. L. Rowland (2018). "The beer and biofuels laboratory: A report on implementing and supporting a large, interdisciplinary, yeast-focused course-based undergraduate research experience." Biochem Mol Biol Educ **46**(3): 213-222.

Petrie, K. L. (2020). "There're CRISPRs in My Yogurt: A Discovery-Based CURE at the Intersection of Industrial Food Production and the Human Microbiome." Front Microbiol **11**: 578737.

Pfeifer-Sancar, K., A. Mentz, C. Ruckert and J. Kalinowski (2013). "Comprehensive analysis of the *Corynebacterium glutamicum* transcriptome using an improved RNAseq technique." BMC Genomics **14**(1): 888.

Portevin, D., C. De Sousa-D'Auria, C. Houssin, C. Grimaldi, M. Chami, M. Daffe and C. Guilhot (2004). "A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms." Proc Natl Acad Sci U S A **101**(1): 314-319.

Rahmat, E. and Y. Kang (2020). "Yeast metabolic engineering for the production of pharmaceutically important secondary metabolites." Appl Microbiol Biotechnol **104**(11): 4659-4674.

Ramos, P. D., F. Marques, M. M. Lopes, J. B. Santos, F. Tavares, I. A. S. Olsson and M. S. Almeida (2022). BioLab – uma proposta para o desenvolvimento de projetos em ambiente multidisciplinar. Cadernos de Inovação Pedagógica, n.º 2 Unidades Curriculares INOVPED da Universidade do Porto. Transversalidade e criatividade. Porto, U.Porto Press: 125-139.

Rapinel, V., O. Claux, M. Abert-Vian, C. McAlinden, M. Bartier, N. Patouillard, L. Jacques and F. Chemat (2020). "2-Methyloxolane (2-MeOx) as Sustainable Lipophilic Solvent to Substitute Hexane for Green Extraction of Natural Products. Properties, Applications, and Perspectives." Molecules **25**(15): 3417.

Reis, T. (2021, 2021/09/27). "A Noite em que a Ciência da Universidade do Porto se abriu à cidade." Notícias U.Porto, from <https://noticias.up.pt/a-noite-em-que-a-ciencia-da-asprela-se-abriu-a-cidade/>.

Richter, M., L. Vieira and V. Sieber (2021). "Sustainable Chemistry - An Interdisciplinary Matrix Approach." ChemSusChem **14**(1): 251-265.

Salazar, A. (1933). A socialização da ciência. Lisboa, Editorial Liberdade.

Salzmann, T. N., R. W. Ratcliffe, F. A. Bouffard and B. G. Christensen (1980). "A stereocontrolled, enantiomerically specific total synthesis of thienamycin." Philos Trans R Soc Lond B Biol Sci **289**(1036): 191-195.

Sambrook, J. and D. W. Russell (2001). Molecular cloning : a laboratory manual. Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press.

Schafer, A., J. Kalinowski, R. Simon, A. H. Seep-Feldhaus and A. Puhler (1990). "High-frequency conjugal plasmid transfer from gram-negative Escherichia coli to various gram-positive coryneform bacteria." J Bacteriol **172**(3): 1663-1666.

Scheifele, L. Z. and T. Burkett (2016). "The First Three Years of a Community Lab: Lessons Learned and Ways Forward." J Microbiol Biol Educ **17**(1): 81-85.

Sharp, P. M. and W. H. Li (1987). "The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications." Nucleic Acids Res **15**(3): 1281-1295.

Shomar, H., S. Gontier, N. J. F. van den Broek, H. Tejeda Mora, M. J. Noga, P. L. Hagedoorn and G. Bokinsky (2018). "Metabolic engineering of a carbapenem antibiotic synthesis pathway in Escherichia coli." Nat Chem Biol **14**(8): 794-800.

Shortlidge, E. E., G. Banger and S. E. Brownell (2017). "Each to Their Own CURE: Faculty Who Teach Course-Based Undergraduate Research Experiences Report Why You Too Should Teach a CURE." J Microbiol Biol Educ **18**(2): 18.12.10.

Siebert, D., J. Altenbuchner and B. Blombach (2021). "A Timed Off-Switch for Dynamic Control of Gene Expression in Corynebacterium Glutamicum." Front Bioeng Biotechnol **9**: 704681.

Silver, S. (2020). "Patenting a living microbial cell: 40th anniversary of US Supreme Court decision Diamond versus Chakrabarty." FEMS Microbiol Lett **367**(13): fnaa091.

Singh, S. B. and J. F. Barrett (2006). "Empirical antibacterial drug discovery--foundation in natural products." Biochem Pharmacol **71**(7): 1006-1015.

Son, J., J. H. Jang, I. H. Choi, C. G. Lim, E. J. Jeon, H. Bae Bang and K. J. Jeong (2021). "Production of trans-cinnamic acid by whole-cell bioconversion from L-phenylalanine in engineered Corynebacterium glutamicum." Microb Cell Fact **20**(1): 145.

Southgate, R. (1994). "The synthesis of natural β -lactam antibiotics." Contemp. Org. Synth. **1**(6): 417-431.

Stanziani, A. (2008). "Defining "natural product" between public health and business, 17th to 21st centuries." Appetite **51**(1): 15-17.

Sun, H., D. Zhao, B. Xiong, C. Zhang and C. Bi (2016). "Engineering *Corynebacterium glutamicum* for violacein hyper production." Microb Cell Fact **15**(1): 148.

Suntar, I., S. Cetinkaya, U. S. Haydaroglu and S. Habtemariam (2021). "Bioproduction process of natural products and biopharmaceuticals: Biotechnological aspects." Biotechnol Adv **50**: 107768.

Tanaka, N., T. Nishimura, S. Nakamura and H. Umezawa (1966). "Biological studies on bottromycin A and its hydrazide." J Antibiot (Tokyo) **19**(4): 149-154.

Taton, A., A. Ecker, B. Diaz, N. A. Moss, B. Anderson, R. Reher, T. F. Leao, R. Simkovsky, P. C. Dorrestein, L. Gerwick, W. H. Gerwick and J. W. Golden (2020). "Heterologous Expression of Cryptomaldamide in a Cyanobacterial Host." ACS Synth Biol **9**(12): 3364-3376.

Tsuge, Y. and A. Yamaguchi (2021). "Physiological characteristics of *Corynebacterium glutamicum* as a cell factory under anaerobic conditions." Appl Microbiol Biotechnol **105**(16-17): 6173-6181.

Unthan, S., M. Baumgart, A. Radek, M. Herbst, D. Siebert, N. Bruhl, A. Bartsch, M. Bott, W. Wiechert, K. Marin, S. Hans, R. Kramer, G. Seibold, J. Frunzke, J. Kalinowski, C. Ruckert, V. F. Wendisch and S. Noack (2015). "Chassis organism from *Corynebacterium glutamicum*--a top-down approach to identify and delete irrelevant gene clusters." Biotechnol J **10**(2): 290-301.

Vertes, A. A., M. Inui and H. Yukawa (2012). "Postgenomic approaches to using corynebacteria as biocatalysts." Annu Rev Microbiol **66**(1): 521-550.

Vertès, A. A., M. Inui and H. Yukawa (2013). The Biotechnological Potential of *Corynebacterium glutamicum*, from Umami to Chemurgy. *Corynebacterium glutamicum*. H. Yukawa and M. Inui. Berlin, Heidelberg, Springer Berlin Heidelberg: 1-49.

Waisvisz, J. M., M. G. van der Hoeven, J. van Peppen and W. C. M. Zwennis (2002). "Bottromycin. I. A New Sulfur-containing Antibiotic." Journal of the American Chemical Society **79**(16): 4520-4521.

Wang, G., Z. Zhao, J. Ke, Y. Engel, Y. M. Shi, D. Robinson, K. Bingol, Z. Zhang, B. Bowen, K. Louie, B. Wang, R. Evans, Y. Miyamoto, K. Cheng, S. Kosina, M. De Raad, L. Silva, A. Luhrs, A. Lubbe, D. W. Hoyt, C. Francavilla, H. Otani, S. Deutsch, N. M. Washton, E. M. Rubin, N. J. Mouncey, A. Visel, T. Northen, J. F. Cheng, H. B. Bode and Y. Yoshikuni (2019). "CRAGE enables rapid activation of biosynthetic gene clusters in undomesticated bacteria." Nat Microbiol **4**(12): 2498-2510.

Wang, Q., J. Zhang, N. H. Al Makishah, X. Sun, Z. Wen, Y. Jiang and S. Yang (2021). "Advances and Perspectives for Genome Editing Tools of *Corynebacterium glutamicum*." Front Microbiol **12**: 654058.

Wang, T., Y. Li, J. Li, D. Zhang, N. Cai, G. Zhao, H. Ma, C. Shang, Q. Ma, Q. Xu and N. Chen (2019). "An update of the suicide plasmid-mediated genome editing system in *Corynebacterium glutamicum*." Microb Biotechnol **12**(5): 907-919.

Wei, L., J. Zhao, Y. Wang, J. Gao, M. Du, Y. Zhang, N. Xu, H. Du, J. Ju, Q. Liu and J. Liu (2022). "Engineering of *Corynebacterium glutamicum* for high-level gamma-aminobutyric acid production from glycerol by dynamic metabolic control." Metab Eng **69**: 134-146.

Weixler, D., M. Berghoff, K. V. Ovchinnikov, S. Reich, O. Goldbeck, G. M. Seibold, C. Wittmann, N. S. Bar, B. J. Eikmanns, D. B. Diep and C. U. Riedel (2022). "Recombinant production of the lantibiotic nisin using *Corynebacterium glutamicum* in a two-step process." Microb Cell Fact **21**(1): 11.

Wendisch, V. F., K. M. Nampoothiri and J. H. Lee (2022). "Metabolic Engineering for Valorization of Agri- and Aqua-Culture Sidestreams for Production of Nitrogenous Compounds by *Corynebacterium glutamicum*." Front Microbiol **13**: 835131.

Wenzel, S. C. and R. Muller (2009). "Myxobacteria--'microbial factories' for the production of bioactive secondary metabolites." Mol Biosyst **5**(6): 567-574.

Wiechert, J., C. Gatgens, A. Wirtz and J. Frunzke (2020). "Inducible Expression Systems Based on Xenogeneic Silencing and Counter-Silencing and Design of a Metabolic Toggle Switch." ACS Synth Biol **9**(8): 2023-2038.

Williams, D. H., M. J. Stone, P. R. Hauck and S. K. Rahman (1989). "Why are secondary metabolites (natural products) biosynthesized?" J Nat Prod **52**(6): 1189-1208.

Wolf, S., J. Becker, Y. Tsuge, H. Kawaguchi, A. Kondo, J. Marienhagen, M. Bott, V. F. Wendisch and C. Wittmann (2021). "Advances in metabolic engineering of *Corynebacterium glutamicum* to produce high-value active ingredients for food, feed, human health, and well-being." Essays Biochem **65**(2): 197-212.

World Health Organization (2017). Prioritization of pathogens to guide discovery, research and development of new antibiotics for drug-resistant bacterial infections, including tuberculosis, World Health Organization.

Wright, F. (1990). "The 'effective number of codons' used in a gene." Gene **87**(1): 23-29.

Wu, J., A. Deng, Q. Sun, H. Bai, Z. Sun, X. Shang, Y. Zhang, Q. Liu, Y. Liang, S. Liu, Y. Che and T. Wen (2018). "Bacterial Genome Editing via a Designed Toxin-Antitoxin Cassette." ACS Synth Biol **7**(3): 822-831.

Wu, X., J. Liu, D. Liu, M. Yuwen, M. A. G. Koffas and J. Zha (2022). "Biosynthesis of eriodictyol from tyrosine by *Corynebacterium glutamicum*." Microb Cell Fact **21**(1): 86.

Xu, X., Y. Liu, G. Du, R. Ledesma-Amaro and L. Liu (2020). "Microbial Chassis Development for Natural Product Biosynthesis." Trends Biotechnol **38**(7): 779-796.

Yang, D., W. J. Kim, S. M. Yoo, J. H. Choi, S. H. Ha, M. H. Lee and S. Y. Lee (2018). "Repurposing type III polyketide synthase as a malonyl-CoA biosensor for metabolic engineering in bacteria." Proc Natl Acad Sci U S A **115**(40): 9835-9844.

Yang, D., C. P. S. Prabowo, H. Eun, S. Y. Park, I. J. Cho, S. Jiao and S. Y. Lee (2021). "Escherichia coli as a platform microbial host for systems metabolic engineering." Essays Biochem **65**(2): 225-246.

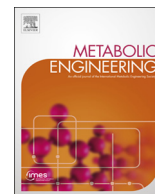
Zhang, C. and H. P. Too (2019). "Revalorizing Lignocellulose for the Production of Natural Pharmaceuticals and Other High Value Bioproducts." Curr Med Chem **26**(14): 2475-2484.

Zhang, S., D. Liu, Z. Mao, Y. Mao, H. Ma, T. Chen, X. Zhao and Z. Wang (2018). "Model-based reconstruction of synthetic promoter library in *Corynebacterium glutamicum*." Biotechnol Lett **40**(5): 819-827.

ANNEXES

COMPLEMENTARY PUBLICATION I

Horbal, L., **Marques, F.**, Nadmid, S., Mendes, M. V., & Luzhetskyy, A. (2018). Secondary metabolites overproduction through transcriptional gene cluster refactoring. *Metabolic Engineering*, 49, 299-315.



Secondary metabolites overproduction through transcriptional gene cluster refactoring



Liliya Horbal^a, Filipe Marques^{a,c,d}, Suvd Nadmid^a, Marta V. Mendes^{c,d}, Andriy Luzhetskyy^{a,b,*}

^a Department Microbial Natural Products, Actinobacteria Metabolic Engineering Group, Saarland University, Campus C2.3, 66123 Saarbrücken, Germany

^b Helmholtz-Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Center for Infectious Research (HZI), Germany

^c i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal

^d IBMC - Instituto de Biologia Molecular e Celular, Universidade do Porto, Porto, Portugal

ARTICLE INFO

Keywords:

Random rational engineering
Overproducer
Bottromycin
Streptomyces
Synthetic promoters

ABSTRACT

We present a random rational approach enabling the construction of overproducing strains in two steps. The approach first involves creating a library of clusters of interest, in which native promoters are substituted with randomly generated constitutive synthetic promoters, and then expressing this library in an appropriate host strain. This strategy is fast, easy to use, accounts for the architecture of a cluster and completely decouples the expression of a gene cluster from complex native regulatory networks. The strategy was applied to improve the production of a macrocyclic peptide, bottromycin, which possesses antibacterial activity against multidrug-resistant bacteria and is a blueprint for a new class of antibacterials. We successfully optimized the expression of genes in operons and created several variants of the bottromycin gene cluster that provide 5–50 fold higher titres of bottromycin than the natural one, thus resulting in the identification of several new bottromycin derivatives not previously described. Moreover, due to the higher bottromycin yield, bottromycin derivatization was performed via the biosynthetic engineering of the gene cluster. The abovementioned features make this generic strategy a promising tool for the overproduction of known secondary metabolites and the activation of silent secondary metabolites in *Actinobacteria*.

1. Introduction

More than 50% of FDA-approved drugs are natural products (NPs) or NP derivatives (Cragg and Newman, 2013). Bacteria, mainly of the genus *Streptomyces*, produce approximately one-fourth of all NPs (Cragg and Newman, 2013; Katz and Baltz, 2016; Patridge et al., 2016). Moreover, these bacteria produce more than half of FDA-approved antibacterial NPs (Patridge et al., 2016). For the reasons mentioned above, microorganisms are an important source of pharmacologically active and industrially relevant secondary metabolites. The period from the 1950s to the 1970s is considered ‘the Golden Era’ of microbial NP discovery since most compounds with antibacterial activities were discovered during this time (Brown and Wright, 2016). The end of the 20th century was characterized by a severe decline in NP identification due to frequent re-discovery of known compounds and the notion that one strain is unable to produce more than 3–4 molecular entities. However, recent genome sequencing data revealed that it may be possible to obtain many more compounds from even the most

comprehensively studied strains, as one strain was found to possess the ability to produce more than 20 chemical entities (Bentley et al., 2002; Ikeda et al., 2003; Rebets et al., 2014b). Currently, there are thousands of prokaryotic genomes in sequence databases containing a plethora of secondary metabolite gene clusters that comprise a treasure trove of potentially new chemical scaffolds and biological activities awaiting discovery (Watve et al., 2001; Rinke et al., 2013; Cimermanic et al., 2014; Roberts et al., 2017). This realization has reinvigorated the field and may give rise to the second ‘Golden Era’ of microbial NP discovery. However, the majority of secondary metabolite gene clusters within genomes are silent or poorly expressed and therefore need to be ‘re-factored’ to yield the corresponding NPs (Baltz, 2011; Bachmann et al., 2014; Katz and Baltz, 2016; Weber et al., 2015; Fedorenko et al., 2015; Zarins-Tutt et al., 2016; Bilyk and Luzhetskyy, 2016). The accessibility of NP-derived lead candidates has always been a critical issue. Even though the 2D chemical structure of a given new compound can now be elucidated based on one digit or even sub-milligram amounts (Janssen et al., 2007), several milligrams are needed to evaluate the activity of a

* Corresponding author at: Department Microbial Natural Products, Actinobacteria Metabolic Engineering Group, Saarland University, Campus C2.3, 66123 Saarbrücken, Germany.

E-mail addresses: a.luzhetskyy@mx.uni-saarland.de, andriy.luzhetskyy@helmholtz-hzi.de (A. Luzhetskyy).

<https://doi.org/10.1016/j.ymben.2018.09.010>

Received 13 July 2018; Received in revised form 17 September 2018; Accepted 17 September 2018

Available online 18 September 2018

1096-7176/ © 2018 Published by Elsevier Inc. on behalf of International Metabolic Engineering Society.

Table 1
Strains and plasmids used in this study.

Bacterial strains and plasmids	Description	Source or reference
<i>S. sp.</i> BC16019	Bottomycin producer	Huo et al. (2012)
<i>S. lividans</i> TK24	Derivative of <i>S. lividans</i> TK21 that contains mutation in the <i>rpsL</i> gene and is resistant to spectinomycin	Kieser et al. (2000)
<i>S. albus</i> J1074	Isoleucine and valine auxotrophic derivative of <i>S. albus</i> G (DSM 40313) lacking <i>Sall</i> -restriction activity	Salas J., Oviedo, Spain
<i>S. sp.</i> Δ <i>botA</i>	Derivative of BC16019 with deleted <i>botA</i> gene	This work
<i>S. lividans</i> BottR-2	Derivative of <i>S. lividans</i> TK24, resistant to bottomycin	This work
<i>E. coli</i> ET12567 (pUB307)	conjugative transfer of DNA	Kieser et al. (2000)
<i>E. coli</i> GB05-red	Derivative of GB2005 containing integration of PBAD-ETgA operon	Zhang et al. (2000)
pUCstrongP	Derivative of pUC57 containing hygromycin resistance gene surrounded with two outwards oriented strong promoters	Manderscheid (2016)
pUC57	Apr, general cloning vector	Thermo Fisher Scientific
pSET152	Am ^r ; ϕC31-based integrative vector	Kieser et al. (2000)
pGUS	Promoter probe vector containing promoterless <i>gusA</i>	Myronovskiy et al. (2011)
pOJ436	ϕC31-based integrative cosmid vector, Am ^r	Kieser et al. (2000)
DG2-Km	Derivative of the DG2 cosmid with the left shoulder substituted to Km ^r marker, contains bottomycin gene cluster	Huo et al. (2012)
DG2-Km-P21RBSHyg	Derivative of the DG2-Km cosmid with strong promoters integrated in between two operons	This work
DG2-KmgusA	Derivative of DG2-Km with integrase gene substituted to <i>gusA</i>	This work
DG2-KmRandom	Derivative of DG2-KmgusA with random pair of promoters integrated in between two operons	This work
DG2-Km-P41hyg	Derivative of DG2-Km with P41 pair of promoters integrated in between two operons	This work
DG2-KmgusA-P41hyg	Derivative of DG2-KmgusA with P41 pair of promoters integrated in between two operons	This work
DG2-KmgusAΔ <i>botA</i>	Derivative of DG2-KmgusA with inactivated <i>botA</i> gene	This work
DG2-KmBotAMeth	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution Valine 1 to Methionine	This work
DG2-KmBotAlso	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution Valine 1 to Isoleucine	This work
DG2-KmBotATyr	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution Valine 1 to Tyrosine	This work
DG2-KmBotACys	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution Valine 1 to Cysteine	This work
DG2-KmBotAGlu	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution Valine 1 to Glutamic acid	This work
DG2-KmBotAVal3-Tyr	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution Valine 3 to Tyrosine	This work
DG2-KmBotAAsp-Meth	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution Aspartic acid to Methionine	This work
DG2-KmBotAAsp-Tyr	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution Aspartic acid to Tyrosine	This work
DG2-KmBotAAsp-Glu	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution Aspartic acid to Glutamic acid	This work
DG2-KmBotACys-Pro	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution Cysteine to Proline	This work
DG2-KmBotAP41hygMeth	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution valine 1 to methionine and P41 pair of promoters	This work
DG2-KmBotAVal1-TAG	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution Valine 1 to stop codon	This work
DG2-KmBotAVal3-TAG	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution Valine 3 to stop codon	This work
DG2-KmBotAAsp-TAG	Derivative of DG2-KmgusA with the <i>botA</i> gene coding for the propeptide with substitution Aspartic acid to stop codon	This work

compound in biological systems as well as to elucidate the absolute stereochemistry of a novel compound. For structure-activity relationship (SAR) studies involving micro-derivatization and pharmacological experiments, at least 100 mg are normally required, and no compound has ever entered into the late preclinical stage without being made available on a multi-gram scale. Classical strain improvement methods (e.g.: non-directed mutagenesis, single colony productivity screening, and media optimization) are rather time-consuming and very laborious. Therefore, synthetic biotechnology, which is currently under further development, is becoming increasingly important as a genome-based approach to engineer NPs for improved properties and optimized production.

Tight and complicated regulatory networks controlling the biosynthesis of secondary metabolites are among the main hurdles to the high-level production of metabolites of interest in native and heterologous hosts (Bibb, 2005; Ostash et al., 2013; van Wezel and McDowall, 2011; Liu et al., 2013). Furthermore, the complicated architecture of clusters, which often contain dozens of genes arranged into operons facing different directions and are governed by promoters that differ in strength and result in differing expression levels, add another level of complexity to ‘refactoring’ clusters (Li et al., 2004; Sosio et al., 2004; Mast et al., 2011; Horbal et al., 2016). Substitution of native promoters with strong orthogonal constitutive and/or synthetic promoters was long regarded as a key strategy for the activation and/or increase of gene transcription and to overcome the regulatory barriers of the cell (Siegl and Luzhetskyy, 2013; Rebets et al., 2014a, 2014b; Myronovskiy and Luzhetskyy, 2016; Zhou et al., 2017). However, a growing body of

data indicates that such a ‘blind’ approach, which assumes that modifications are additive (a linear behaviour), often does not lead to the overproduction of a target metabolite and does lead to a distinct metabolic burden (Keasling, 2012; Smanski et al., 2016; Sun et al., 2014). This is related to the fact that a blind approach ignores nonspecific effects, such as cell fitness, resistance and the existence of genes encoding enzymes that constitute bottlenecks in the biosynthesis of a secondary metabolite, and leads to the accumulation of metabolic intermediates or even toxic products (Keasling, 2012). Thus, a prerequisite for NP overproduction is not ‘hyper’ transcription but rather a balanced expression of genes in operons that lead to the biosynthesis of enzymes in catalytic amounts sufficient for the efficient transformation of an intermediate into a desired product (Ajikumar et al., 2010; Woolston et al., 2013). To this end, there is a need for strategies enabling the accurate control and balanced expression of native and heterologous genes within the cell. One of the latest approaches in synthetic biology is called multivariate modular metabolic engineering, an approach that groups enzymes in metabolic pathways into a series of modules based upon pathway branching, compound chemistry, and enzyme turnover; different combinations of well-defined controlling elements (promoters, tunable intergenic elements, RBSs, etc.) are tested for the modules (Biggs et al., 2014). This engineering approach also takes into account enzymes kinetics, the order of reactions or pathway branching architecture; however, such a rational approach requires a great deal of effort, time and additional investigation.

Herein, we report an alternative, a so-called random rational strategy for the expression and activation of gene clusters. The

approach is based on the generation of a library of clusters with a random combination of synthetic promoters and the further screening of those clusters demonstrating the highest production of a metabolite of interest. The strategy is fast and easy to use, takes into account cluster architecture and does not require detailed knowledge of gene expression and enzymes kinetics. Furthermore, combining this approach with bioactivity tests or reporter systems allows significant levels of metabolite production to be achieved in a relatively short amount of time. As a case study, we refactored the gene cluster coding for bottromycin, a promising macrocyclic ribosomally synthesized and post-translationally modified peptide (RiPP). This peptide is active against methicillin-resistant *Staphylococcus aureus* (Kobayashi et al., 2010), possesses a unique chemical structure (Crone et al., 2012), and may be a blueprint for a new class of antibacterials; however, it is produced at extremely low yields (Huo et al., 2012; Crone et al., 2016). Using the abovementioned approach, we successfully optimized the expression of genes in operons and created a library of modified gene clusters that produce 5–50 fold higher yields of bottromycin than natural cluster. The significantly increased production yields enabled the generation of several new bottromycins. We anticipate that this generic strategy will facilitate the overproduction of known NPs and the discovery of new potentially active NPs in *Actinobacteria*.

2. Materials and methods

2.1. Bacterial strains and culture conditions

The bacterial strains used in this study are listed in Table 1. *E. coli* strains were grown in Lysogeny Broth (LB) medium. When required, antibiotics (Roth, Germany; Sigma, USA) were added to cultures at the following concentrations: 75 µg ml⁻¹ ampicillin, 50 µg ml⁻¹ kanamycin, 50 or 120 µg ml⁻¹ hygromycin, and 50 µg ml⁻¹ apramycin. *E. coli* GB05-red (Fu et al., 2008) was employed in Red/ET recombineering experiments (Zhang et al., 2000).

For conjugation, *Streptomyces* sp. BC16019 and *Streptomyces lividans* strains were grown on oatmeal or mannitol soy (MS) agar (Kieser et al., 2000) for sporulation. For glucuronidase activity measurements, strains were grown in liquid tryptic soy broth (TSB).

2.2. Recombinant DNA techniques

Chromosomal DNA from *Streptomyces* strains and plasmid DNA from *E. coli* were isolated using standard protocols (Kieser et al., 2000; Sambrook and Russell, 2001). Restriction enzymes and molecular biology reagents were used per the manufacturer's protocol (NEB, England; Thermo Scientific, Germany).

2.3. Construction of the DG2-KmgusA cosmid

A 1.9 kb DNA fragment containing the *gusA* gene was amplified from the pGUS plasmid with the use of the GusAHindIIIForw/GusAHindIIIRev primer pair (Table 2). The fragment was then digested with *HindIII* and cloned into respective site of DG2-Km cosmid. As a result a DG2-KmgusA cosmid (Table 1) was obtained that contains the *gusA* gene cloned instead of ϕ C31 integrase.

2.4. Construction the pUCStrongP plasmid

A hygromycin cassette surrounded with outwards oriented strong promoters (P21 (Siegl et al., 2013) and Virolle (Seghezzi et al., 2011)) was amplified using the P21EcoRIForw and P21HindIIIRev primers (Table 2). The obtained fragment was then digested with *EcoRI*/*HindIII* and cloned into respective sites of pUC57, yielding the pUCStrongP plasmid (Table 1).

2.5. Modification of the DG2-Km and DG2-KmgusA cosmids by Red/ET recombination

In all Red/ET recombination experiments linear DNA fragments flanked by suitable homology arms were generated by PCR. The PCR reactions were carried out with Phusion DNA polymerase (Thermo Fisher Scientific) according to the manufacturer's protocol. PCR products were concentrated by ethanol precipitation prior to further use. In general, 300 µl overnight culture of *E. coli* GB05-red (Table 1) cells harboring the parent vector to be modified were inoculated into 15 ml of LB, and cultivated on a shaker at 37 °C and 200 rpm for 2 h. Afterwards, 400 µl 10% L-arabinose was added to the culture to induce expression of the recombinases. Cultivation was then continued for 45 min. The cells were then harvested by centrifugation, washed twice with ice-cold distilled H₂O, and resuspended in 600 µl of 10% ice-cold glycerol. PCR product was mixed with the electrocompetent cells and transferred into an ice-cold electroporation cuvette (1 mm). The mixture was then electroporated at 1800 V (Eppendorf electroporator), followed by addition of 750 µl LB. The cells were cultivated at 37 °C for 60–90 min, before the culture was transferred to LB agar plates with suitable antibiotics. The plates were incubated at 37 °C overnight. Correct transformants were verified by restriction analysis and/or sequencing of the isolated cosmid DNA.

2.6. Construction DG2-Km-P21RBSHyg

A hygromycin cassette surrounded with outwards oriented strong promoters was amplified from pUCStrongP (Table 1) using the BotromP21RedForw1/BotromP21RedRev primer pair (Table 2), containing 59 bp homologous regions to the *botOMT* and *botRMT* genes, respectively. The fragment was used for performing Red/ET recombination with DG2-Km. After recombination, colonies were selected on LB agar plates with 120 µg/ml hygromycin. Correct transformants were verified by sequencing. As a result the DG2-Km-P21RBSHyg cosmid was obtained.

2.7. Construction the library of the DG2-Km cosmids with random promoters

A hygromycin cassette surrounded with randomly generated outwards oriented promoters was amplified using degenerate BotRandomPFor and BotRandomPRev primers, containing 50 bp homologous regions to *botOMT* and *botRMT* genes, respectively. The PCR fragment was used to perform recombination with the DG2-Km cosmid. After recombination, colonies were selected on LB agar plates with 120 µg/ml hygromycin. As a result, library of the DG2-Km-Prandom cosmids was obtained.

2.8. Construction the DG2-KmgusA-P41hyg cosmid

A hygromycin cassette surrounded with the P41 pair of promoters was amplified using the 41PromSubstFor and 41PromSubsRev primers, containing 50 bp homology arms to the *botOMT* and *botRMT* genes, respectively. The fragment then was used to perform recombination with the DG2-KmgusA cosmid. After recombination, colonies were selected on LB agar plates with 120 µg/ml hygromycin. Correct transformants were verified by sequencing. As a result, the DG2-KmgusA-P41hyg cosmid (Table 1) was constructed.

2.9. Construction of the DG2-KmgusAdelBotA cosmid

A hygromycin cassette was amplified using the BotAHygrRForw and BotAHygrRRev primers (Table 2) containing 57 bp homology regions up to start and stop codons of the *botA* gene, respectively. The fragment then was used to perform recombination with the DG2-KmgusA cosmid. After recombination, colonies were selected on LB agar plates with

Table 2
Primers used in this work.

Primers	Sequence 5'-3'	Purpose
GusAHindIIIForw	AAAAAAAAAGCTTACGAGGCGCTTTCGTGcagcag	Integrase gene substitution with <i>gusA</i>
GusAHindIIIRev	AAAAAAAAAGCTTTTATCAGTGTCTCCGCCCTGCTG	Integration of strong promoters in between two operons
BottomP21RedForw11	GTAGAGGGGTGCGAGCAGGGTCTCCGGCGCGCGGTACAGGAAATCTTACGGGACTTCCTCCTGTAGGCTGGAGCTGCTTCGG	
BottomP21RedRev	TGACGCTGTTCAGGCAATTGATACAAGAGAAAAAGTTTTTGGGGCGCAGGCACGAAACAACCTCTTCGCGGGGATCCGTCGACCCCT	Check the integration of the hygromycin cassette in between two operons
BotOMITForw	AAGTCGTAGTCGATCGCGTC	
BotRMTRev	TCTACTTGCACGCCCTTCACTC	Construction of library cosmid with random promoters
BotRandomPFor	TGCGACGAGGTCTCCGGCGCGCGGTACAGGAAATCTTACGGGACTTCCTCCTTCANNNNNNNAGCCNNNNNTGGGGCCTACGATGTATCAG	
BotRandomPRev	AAGGCAATTGATACAAAGAGAAACGTTTTTTTGGGGCGCAGGCACGAAACAACCTCCGATANNNNNNNATCCATANNNNNNNNNNAGCCNNNNN TGTAATACTTTGACATATCAC	The <i>botA</i> gene substitution
BotAHygRForw	TGAGGCACGGAACCGAAATCGAAACGCAAGCCAGACGGAAGGATCGAGATATGCTACGCCCCCACTGAGAGAAC	
BotAHygRRev	GGTGTCCGTCCTCGCCGACTTCACACCCGCCGGCTACTCACCGCGCGGCTCAGTCCGTACCTCGGTTGCTCGTGGATGATCCTAATAGTTATTTG	Amplification of the fragment containing the <i>botA</i> gene
BotAForw11	ACCAAGAGTCCCGCCGCTTCAG	
BotAREv22	ACAGCCGTCCGGCGGGTCCGTAAC	Amplification of the <i>botA</i> gene coding for propeptide with substitution Cysteine to Proline
BotACysVSPProRev11	TGCTGTTGAGGAAATCCGGGTCAATTGGTTCGAATACGACTACGGGTCC	
BotACysVSPProForw22	TGGGACCCGTAGTCGTATTCGACCCAATGACGGCGGATTTCTCAACGAC	Amplification of the <i>botA</i> gene coding for propeptide with substitution Aspartic acid to Methionine
BotAAspVSMethionRev11	GACGAAAGGATCGAGATATGGAACCCGTAGTCGTATTCATGTGCAATGACCGCGGATTTCC	
BotAAspVSTyrForw22	GGAAATCCCGGGTCATGCAGTAAATACGACTACGGGTCCCATATCTCGATCCTTTGTC	Amplification of the <i>botA</i> gene coding for propeptide with substitution Aspartic acid to Tyrosine
BotAAspVSTyrRev11	GACGAAAGGATCGAGATATGGAACCCGTAGTCGTATTCGAGTGCATGACCGCGGATTTCC	
BotAAspVSGluForw22	GGAAATCCCGGGTCATGCAGTAAATACGACTACGGGTCCCATATCTCGATCCTTTGTC	Amplification of the <i>botA</i> gene coding for propeptide with substitution Aspartic acid to Glutamic acid
BotAAspVSGluRev11	GACGAAAGGATCGAGATATGGAACCCGTAGTCGTATTCGAGTGCATGACCGCGGATTTCC	
BotAVAl3VSTyrForw22	GGAAATCCCGGGTTCATGCACCTCGAATACGACTACGGGTCCCATATCTCGATCCTTTGTC	Amplification of the <i>botA</i> gene coding for propeptide with substitution Valine 3 to Tyrosine
BotAVAl3VSTyrRev11	GACGAAAGGATCGAGATATGGAACCCGTAGTCGTATTCGAGTGCATGACCGCGGATTTCC	
BotAVAl3VSTyrRev11	GGAAATCCCGGGTTCATGCAGTGCAGTAGACTAGAGCTACGGGTCCCATATCTCGATCCTTTGTC	

(continued on next page)

Table 2 (continued)

Primers	Sequence 5'-3'	Purpose
BotAV/Al11VSCysForw22	GACGAAAGGATCGAGATATGGGACCCCTGGTCTGCTATTTCGACTGCATGACCGGGGATTTC	Amplification of the <i>botA</i> gene coding for propeptide with substitution Valine 1 to Cysteine
BotAV/Al11VSCysRev11 BotAV/Al11VSIsoForw22	GGAATCCGGGGTCATGCAGTCGAAATACGACGACGAGGTCCCATATCTCGATCCTTTGGTC GACGAAAGGATCGAGATATGGGACCCATCGTCGTATTTCGACTGCATGACCGGGGATTTC	Amplification of the <i>botA</i> gene coding for propeptide with substitution Valine 1 to Isoleucine
BotAV/Al11VSIsoRev11 BotAV/Al11VSMethow22	GGAATCCGGGGTCATGCAGTCGAAATACGACGATGGGTCCCATATCTCGATCCTTTGGTC GACGAAAGGATCGAGATATGGGACCCATGGTCGTATTTCGACTGCATGACCGGGGATTTC	Amplification of the <i>botA</i> gene coding for propeptide with substitution Valine 1 to Methionine
BotAV/Al11VSMethRev11 BotAV/Al11VSTyrForw22	GGAATCCGGGGTCATGCAGTCGAAATACGACCATGGGTCCCATATCTCGATCCTTTGGTC GACGAAAGGATCGAGATATGGGACCCATCGTCGTATTTCGACTGCATGACCGGGGATTTC	Amplification of the <i>botA</i> gene coding for propeptide with substitution Valine 1 to Tyrosine
BotAV/Al11VSTyrRev11 BotAV/Al11VSGluForw22	GGAATCCGGGGTCATGCAGTCGAAATACGACGTAGGGTCCCATATCTCGATCCTTTGGTC GACGAAAGGATCGAGATATGGGACCCGAGGTCTGTATTTCGACTGCATGACCGGGGATTTC	Amplification of the <i>botA</i> gene coding for propeptide with substitution Valine 1 to Glutamic acid
BotAV/Al11VSGluRev11 BotAV/Al11VSTAGorw22	GGAATCCGGGGTCATGCAGTCGAAATACGACCTCGGGTCCCATATCTCGATCCTTTGGTC GACGAAAGGATCGAGATATGGGACCCCTAGGTCTGTATTTCGACTGCATGACCGGGGATTTC	Amplification of the <i>botA</i> gene coding for propeptide with substitution Valine 1 to Pyrolysine
BotAV/Al11VSTAGRev BotAV/Al13VSTAGorw22	GGAATCCGGGGTCATGCAGTCGAAATACGACCTAGGGTCCCATATCTCGATCCTTTGGTC GACGAAAGGATCGAGATATGGGACCCGTAGTCTAGTTCGACTGCATGACCGGGGATTTC	Amplification of the <i>botA</i> gene coding for propeptide with substitution Valine 3 to Pyrolysine
BotAV/Al13VSTAGRev BotAAspVSTAGorw22	GGAATCCGGGGTCATGCAGTCGAACTAGACTACGGGTCCCATATCTCGATCCTTTGGTC GACGAAAGGATCGAGATATGGGACCCGTAGTCTGTATTCTAGTGCATGACCGGGGATTTC	Amplification of the <i>botA</i> gene coding for propeptide with substitution Aspartic acid to Pyrolysine
BotAAspVSTAGRev P21EcoRIForw	GGAATCCGGGGTCATGCAGTACGACTACGGGTCCCATATCTCGATCCTTTGGTC TTTTTGAATTCCTGCTCATCTAGCATAGTGGACCCGTAGAGCCGGCACACTAGCCCCCAACTGAGAG	Cloning hygromycin cassette surrounded with outwards oriented promoters
P21HindIIIRev	AAAAAAGCTTCGCATATCCTAACACTCCCAAAATACATCGTCAATAGACCCACCTGCAGGGGAGTAC	

(continued on next page)

Table 2 (continued)

Primers	Sequence 5'-3'	Purpose
41PromSubstFor	TGGAGCAGGGTCTCCGCCGCCGCGGTCAAGGAAATCTTTCACGGGACTTCCTCCTTCATATGATATCCCTAGCATGTATGTGTGTAACAATCAGCCGAGATTGGGGCAGCTACGATGTATCAG	Amplification hygromycin cassette surrounded with the p41 pair of outwards oriented promoters
41PromSubstRev	AAGGCATTGATACAGAGAGAAACGTTTTTGGCGGCGAGGCACGAAACAAACCTCCGATAACTTTCTCTACTTGGCATACACGGCTTGTAGCCACAGCT	RT-PCR analysis of the <i>botI</i> gene expression
botI_F	CGATACCGATCGTGGAGGTCA	RT-PCR analysis of the <i>botRMT1</i> gene expression
botI_R	CCTCGTGGCAAGGACTCA	RT-PCR analysis of the 16S rRNA gene express
RMT1_F	AAGCTCAACACGGACCACTC	
RMT1_R	GTGGCACCCCTTCTCTGAGATG	
16S_F	CGCAACGCCGAGAACCTTAC	
16S_R	CCCAACATCTCAGGACACGA	

120 µg/ml hygromycin. Correct transformants were verified by sequencing. As a result, the DG2-KmgusAdelBotA cosmid (Table 1) was constructed.

2.10. Construction of the series of cosmids with *botA* coding for propeptides with aa substitutions

In all cases 4 different primers were used. Two primers, namely BotAForw11 and BotARev22 (Table 2), were always the same. All the cosmids were constructed using the same principle and primers listed in Table 2. As an example, the construction of the DG2-KmBotAMeth (Table 1) cosmid is described. A fragment 1–0.53 kb long, containing homology region to the part of the *botRMT* and *botA* genes was amplified using the BotAForw11 and BotAVal1VSMethRev11 primers (Table 2). The latter primer contained the *botA* gene carrying the substitution of a triplet coding for valine in position one with methionine coding trinucleotide. A fragment 2–0.465 kb long, containing homology region to the part of the *botA* and *botC* genes, was amplified using the BotAVal1VSTyrForw22 and BotARev22 primers (Table 2). The former primer contained part of the *botA* gene carrying the substitution of a triplet coding for valine in position one with methionine coding trinucleotide. Afterwards, both fragments, 1 and 2, were used as a template in the PCR with the BotAForw11 and BotARev22 primers. Obtained 0.996 kb fragment was used together with *SwaI* digested DG2-KmgusAdelBotA cosmid to electroporate *E. coli* GB05-red cells. After recombination, colonies were selected on LB agar plates with 50 µg/ml apramycin. In case of correct recombination hygromycin resistance gene should be substituted with the *botA* gene coding for the propeptide with substitution valine amino acid in position 1 to methionine. Correct transformants were verified by sequencing. As a result, the DG2-KmBotAMeth cosmid (Table 1) was constructed.

2.11. Construction of the DG2-KmBotAP41hygMeth cosmid

A hygromycin cassette surrounded with the P41 promoter pair was amplified using the 41PromSubstFor and 41PromSubstRev primers (Table 2), containing 50 bp homology arms to the *botOMT* and *botRMT* genes, respectively. The fragment then was used to perform recombination with the DG2-KmBotAMeth cosmid. After recombination, colonies were selected on LB agar plates with 120 µg/ml hygromycin. Correct transformants were verified by sequencing. As a result, the DG2-KmBotAP41hygMeth cosmid (Table 1) was constructed.

2.12. Construction of the *S. sp.* $\Delta botA$ mutant

The DG2-KmgusAdelBotA (Table 1) gene disruption cosmid was transferred from *E. coli* ET12567/pUB307 into *S. sp.* BC16019 by means of conjugation (Kieser et al., 2000). Transconjugants were selected for resistance to apramycin and hygromycin (50 µg ml⁻¹). Since above mentioned cosmid does not contain origin of replication for *Streptomyces*, all the obtained exconjugants are result of a first crossover event. For the generation of the *S. sp.* $\Delta botA$ mutant single-crossover apramycin and hygromycin resistant mutants were screened for loss of apramycin resistance that is the result of a double-crossover event. Replacement of the *botA* gene was confirmed by PCR using the BotA-Forw11/BotARev22 primer pair (Table 2).

2.13. Construction of the *S. sp.* *botA* mutants producing “unnatural” natural derivatives of bottromycin

To create mutants producing different bottromycins, cosmids with certain substitutions in the *botA* gene were used. As a particular example we describe construction of the *S. sp.* BotA-Meth strain producing bottromycin derivative with methionine instead of valine 1.

The DG2-KmBotAMeth cosmid was transferred from *E. coli* ET12567/pUB307 into *S. sp.* $\Delta botA$ by means of conjugation (Kieser

et al., 2000). Transconjugants were selected for resistance to apramycin and hygromycin ($50 \mu\text{g ml}^{-1}$). Since above mentioned cosmid does not contain origin of replication for *Streptomyces*, all the obtained exconjugants are result of a first crossover event. For the generation of the *S. sp.* BotA-Meth mutant single-crossover apramycin and hygromycin resistant mutants were screened for loss of hygromycin resistance that is the result of a double-crossover event. Replacement of the hygromycin gene with the *botA-meth* allele was confirmed by PCR using the BotAForw11/BotARev22 primer pair (Table 2). The same procedure was used to obtain any other bottromycin derivative.

2.14. Construction of the *S. sp.* BotA-MethP41 strain

The DG2-KmBotAP41hygMeth cosmid (Table 1) was transferred from *E. coli* ET12567/pUB307 into *S. sp.* BotA-Meth by means of conjugation (Kieser et al., 2000). Transconjugants were selected for resistance to apramycin and hygromycin ($50 \mu\text{g ml}^{-1}$). Since above mentioned cosmid does not contain origin of replication for *Streptomyces*, all the obtained exconjugants are result of a first crossover event. For the generation of the *S. sp.* BotA-Meth-P41 mutant single-crossover apramycin and hygromycin resistant mutants were screened for loss of apramycin resistance that is the result of a double-crossover event. Insertion of the hygromycin gene with outwards oriented P41 promoters in between *botOMT* and *botRMT* was confirmed by PCR using the BotOMTForw/BotRMTRev primer pair (Table 2).

2.15. RT-qPCR analysis of gene expression in *S. lividans* recombinant strains

The expression levels of the original and modified versions of bottromycin gene cluster DG2-Km and DG2-KmP41hygR in *S. lividans* TK24 was assessed by reverse-transcription quantitative polymerase chain reaction (RT-qPCR). One microgram of total RNA treated with DNaseI was reverse transcribed to cDNA using the RevertAid RT Kit (Thermo Fisher Scientific) in a final volume of $20 \mu\text{l}$, following manufacturer's instructions. qPCRs were assembled using $0.2 \mu\text{M}$ of specific primers (Table 2), $10 \mu\text{l}$ of QuantiNova SYBR Green PCR Master Mix (QIAGEN) and $2 \mu\text{l}$ of template cDNA (dilution 1/32). qPCRs were carried out in the Rotor-Gene™ 3000 Real Time Thermal Cycler (Corbett Research) with the following conditions: 95°C for 10 min, 40 cycles of 95°C for 10 s and 60°C for 10 s. Surveys on the efficiency of qPCR and the purity of the samples were conducted using five serial dilutions of the cDNA and negative controls (no-template and no-reverse transcriptase), respectively. The resulting amplicons were analysed by agarose gel electrophoresis to exclude the formation of nonspecific products. Rotor-Gene software (version 6.1) was employed to retrieve the quantification cycle (C_q) values based on fluorescence readings. The C_q values of 16S rRNA were used for normalization. The relative transcriptional level of each gene was calculated using the method described by Pfaffl (2001). Mean values of two biological replicates per experiment are shown.

2.16. Analysis of bottromycin production and novel derivatives thereof

For bottromycin production, 2 ml of a 2-day-old pre-culture grown in 50 ml of TSB media (Sigma-Aldrich, USA) was inoculated into 50 ml of SG media and grown for 7 days at 28°C with agitation at 200 rpm. Bottromycin was extracted from biomass with ethyl acetate. Samples were evaporated, dissolved in $450 \mu\text{l}$ of methanol and subjected to LC-MS or MS² analysis. The LC-HRMS analysis of the crude extracts was done on a Dionex Ultimate 3000 RSLC system using a BEH C18, $100 \times 2.1 \text{ mm}$, $1.7 \mu\text{m}$ dp column (Waters, Germany). Separation of $1 \mu\text{l}$ sample was achieved by a linear gradient with a mobile phase of water/acetonitrile each containing 0.1% formic acid, using a gradient from 5% to 95% acetonitrile over 9 min at a flow rate of $600 \mu\text{l/min}$ and 45°C . The mass-spec data was collected on a maXis 4Gh-ToF ultrahigh resolution mass spectrometer (Bruker Daltonics, Germany) using the

Apollo II ESI source. Mass spectra were acquired in centroid mode ranging from 200 to $2500 m/z$ at a 2 Hz scan rate. Data were collected and analysed with the Bruker Compass Data Analysis software, version 4.1 (Bruker, Billerica, USA). High-resolution mass spectrometry was performed on an Accela UPLC-system (Thermo-Fisher) coupled to a linear trap-FT-Orbitrap combination (LTQ-Orbitrap), operating in positive ionization mode. Separation was achieved using a BEH RP-C18 column ($50 \times 2 \text{ mm}$; 1.7 m particle diameter; flow rate 0.6 ml/min , Waters), with a mobile phase of water/acetonitrile each containing 0.1% formic acid over 9 min. The UPLC-system was coupled to the LTQ-Orbitrap by a Triversa Nanomate (Advion), a chip based nano-ESI interface. The yield of bottromycin A2 was determined from the peak area (extracted ion chromatogram at m/z $[M+H]^+ = 823.4528$) by reference to a standard curve generated with pure bottromycin-Meth. For quantitative measurement of bottromycin A2 and other bottromycins production, a standard curve was established by preparing a concentration gradient (from 0.01 mg/ml to 3.2 ng/ml in a dilution factor 1:10) of pure bottromycin as reference substance.

Two ml of biomass from culture broth was dried and used to assess growth of the strains. In all cases bottromycin production was normalized to biomass. All the data were calculated as mean values of three independent analyses.

2.17. Isolation and purification of bottromycin-Meth

The methanol extract obtained after ethyl acetate extraction was used for purification of the compound on a 1-m-long column packed with cross-linked dextran polymer beads. Methanol was used as a solvent for elution. Fractions were collected every 12 min with a flow rate of 0.6 ml/min . The fractions containing bottromycins were pooled together, evaporated, and dissolved in methanol. Afterward, the pre-purified methanol extract was further purified on an Agilent Technologies 1260 Infinity semi-preparative HPLC using a Synergi column ($250 \times 10 \text{ mm}$) and isocratic gradient at 35% acetonitrile. The fractions with bottromycins were again evaporated and dissolved in methanol. After these two purification steps, the bottromycin-Meth was used for NMR analysis.

2.18. NMR analysis of new bottromycin with methionine substitution

NMR spectra were acquired on a Bruker Ascend 500 MHz NMR spectrometer equipped with a 5 mm TXI cryoprobe. As solvent, deuterated DMSO- d_6 were used and HSQC, HMBC, 1–1H COSY spectra were recorded using standard pulse programs.

2.19. Selection of bottromycin resistant *S. lividans* mutants

A bottromycin enriched crude extract (on average 33 mg/ml) from *S. lividans* TK24 DG2-Km-P41hyg, was used for the selection of bottromycin resistance mutants. Spontaneous bottromycin-resistant (BotR) mutants were obtained from colonies that grew within 7 days after spores (usually 10^9 spores) were spread on MS agar containing various concentrations of bottromycin extract (20, 100, 200, 400, $800 \mu\text{l}$). In each step, one colony resistant to certain amount of bottromycin was used for the next round of screening. To be sure obtained resistant colonies are mutants, they were cultivated for three passages in non-selective conditions and afterwards plated on MS with certain concentration of bottromycin.

3. Results

3.1. Heterologous expression of the bottromycin gene cluster

Bottromycin (Fig. 1A) is a ribosomally synthesized macrocyclic peptide and is active against methicillin-resistant *Staphylococcus* and vancomycin-resistant *Enterococcus* strains (Shimamura, 2009). This

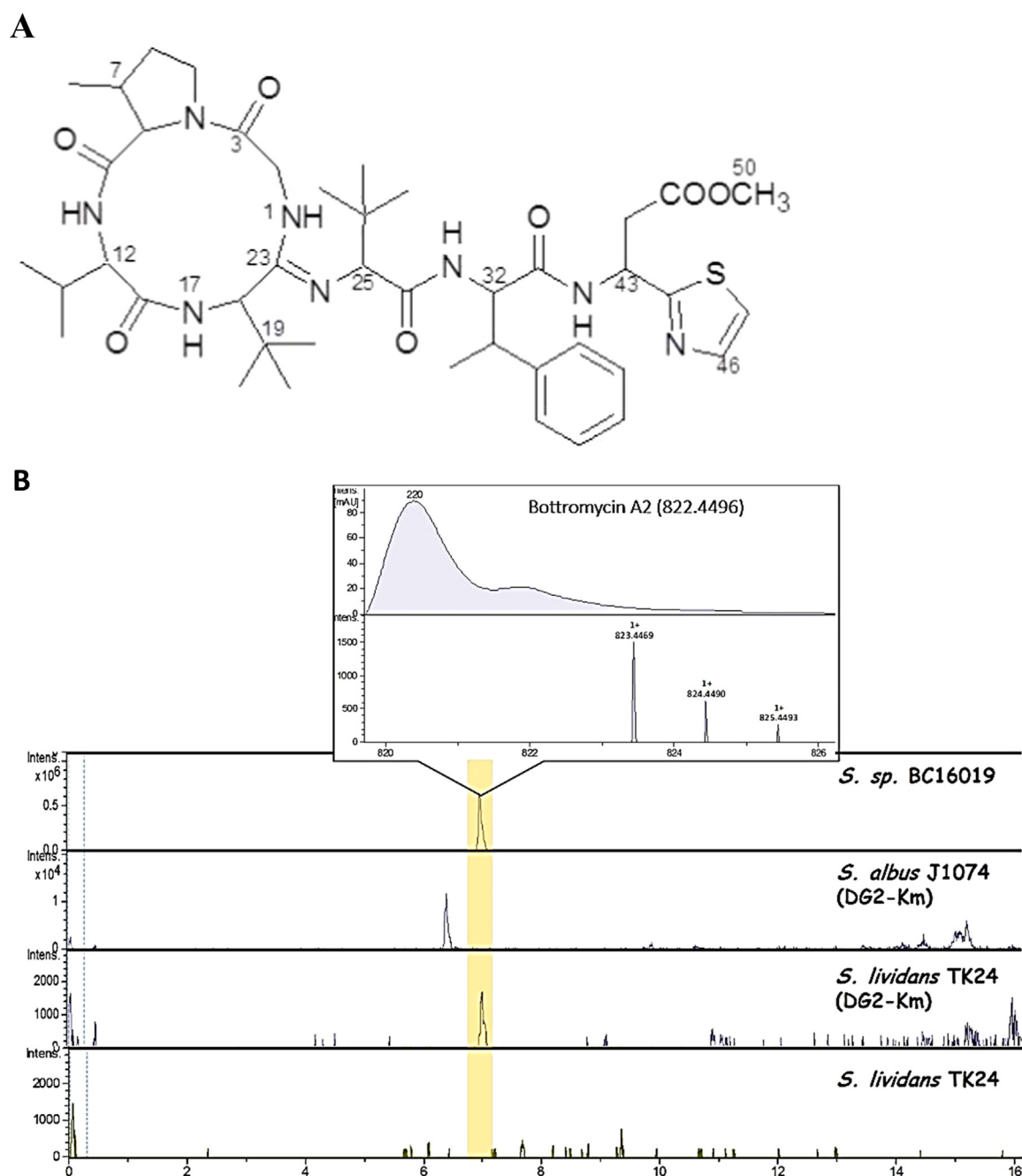


Fig. 1. Heterologous bottromycin production. (A) Chemical structure of bottromycin A2. (B) Comparison of bottromycin A2 production in different *Streptomyces* strains.

peptide targets therapeutically unexploited A-sites of the bacterial ribosome and thus is a highly promising blueprint for a new class of drugs (Lin and Tanaka, 1968; Otaka and Kaji, 1976). However, the production yield of bottromycin in native or heterologous hosts is extremely low, hampering further development of this compound. Chemical synthesis of the antibiotic is challenging due to its unique structure with a macrocyclic amidine architecture and β -methylated amino acids. Therefore, bioengineering approaches are the only feasible way to obtain high yields of new bottromycins with improved activities.

There are several known producers of bottromycin (Hou et al., 2012; Crone et al., 2012; Gomez-Escribano et al., 2012); however, the production levels in these strains are low, and genetic manipulations are very challenging. Thus, we decided to use a well-defined heterologous host to optimize and boost bottromycin production. Recently

published work on the heterologous expression of the bottromycin gene cluster (*bot*-cluster) together with pathway engineering in *Streptomyces coelicolor* A3(2) reported a 20 $\mu\text{g/L}$ product yield (Huo et al., 2012). Taking this knowledge into account, we decided to assay two other known and widely used *Streptomyces*, namely, *S. albus* J1074 and *S. lividans* TK24 (Table 1), as potential host strains for bottromycin production. We first transferred the DG2-Km cosmid (Table 1), which contains the entire *bot*-cluster, into the abovementioned strains via conjugation. Two recombinant strains, *S. albus* DG2-Km⁺ and *S. lividans* TK24 DG2-Km⁺, were generated. The strains were grown in SG production medium, and then bottromycin was extracted with ethyl acetate from the culture broth, dissolved in methanol and analysed by HPLC-MS. *S. albus* J1074 and *S. lividans* TK24 were used as negative controls, and *Streptomyces* sp. BC16019 (Table 1) was used as a positive

control. High-resolution mass spectrometry HRMS analysis revealed the presence of bottromycin A2 (mass – 822.43 Da) in the extracts of *S. lividans* TK24 DG2-Km⁺ and *S. sp.* BC16019 (Fig. 1B). Furthermore, MS/MS analysis of the extracts revealed a characteristic fragmentation pattern identical to that of bottromycin A2 (data not shown). However, we were not able to detect any bottromycins in *S. albus* J1074 (Fig. 1B) that might be related to the strain's sensitivity to the drug or inability to synthesize the final molecule. Thus, *S. lividans* TK24 was used for further investigations.

3.2. Rational approach for increased bottromycin production

S. lividans TK24 DG2-Km⁺ is capable of producing bottromycin, meaning that native promoters in the gene cluster are functional in this strain. However, the production rate was very low and averaged only 0.23 mg/L. One possible reason for this low production may be the low expression levels achieved with native promoters in the *S. lividans* strain. There are numerous examples of native promoter substitutions in secondary metabolite gene clusters with well-defined strong constitutive promoters leading to compound overproduction in native or heterologous hosts (Stratigopoulos et al., 2004; Horbal et al., 2014; Myronovskiy et al., 2016). For our first approach, we decided to use the

abovementioned rational strategy to increase bottromycin production. The *bot*-cluster (Fig. 2A) comprises 13 genes (Huo et al., 2012) that are most likely organized into only two operons, as genes are transcribed from two promoters according to real-time reverse transcription-PCR (RT-PCR) data (Truman A., personal communication). Furthermore, the BotR regulator within the *bot*-cluster has no influence on bottromycin biosynthesis (Crone et al., 2012). Thus, we chose two strong promoters with similar strengths from the library (Seghezzi et al., 2011; Siegl et al., 2013), fused them with the hygromycin resistance marker and introduced them upstream of *botOMT* and *botRMT1*, which are the first genes in the operons, using Red/ET technology (Fig. 2A, for details see Materials and Methods). The obtained DG2-Km-P21RBSHyg cosmid (Table 1) was transferred into *S. lividans* TK24 via conjugation. *S. lividans* TK24 DG2-Km-P21RBSHyg⁺ exconjugants, which were very small in size, were obtained. However, these exconjugants did not grow after plating on MS medium. This lack of growth may be related to a toxicity effect; however, other reasons cannot be excluded.

3.3. Increased resistance of a host strain leading to a further increase in bottromycin production

S. lividans DG2-Km-P21RBSHyg exconjugants containing cosmids

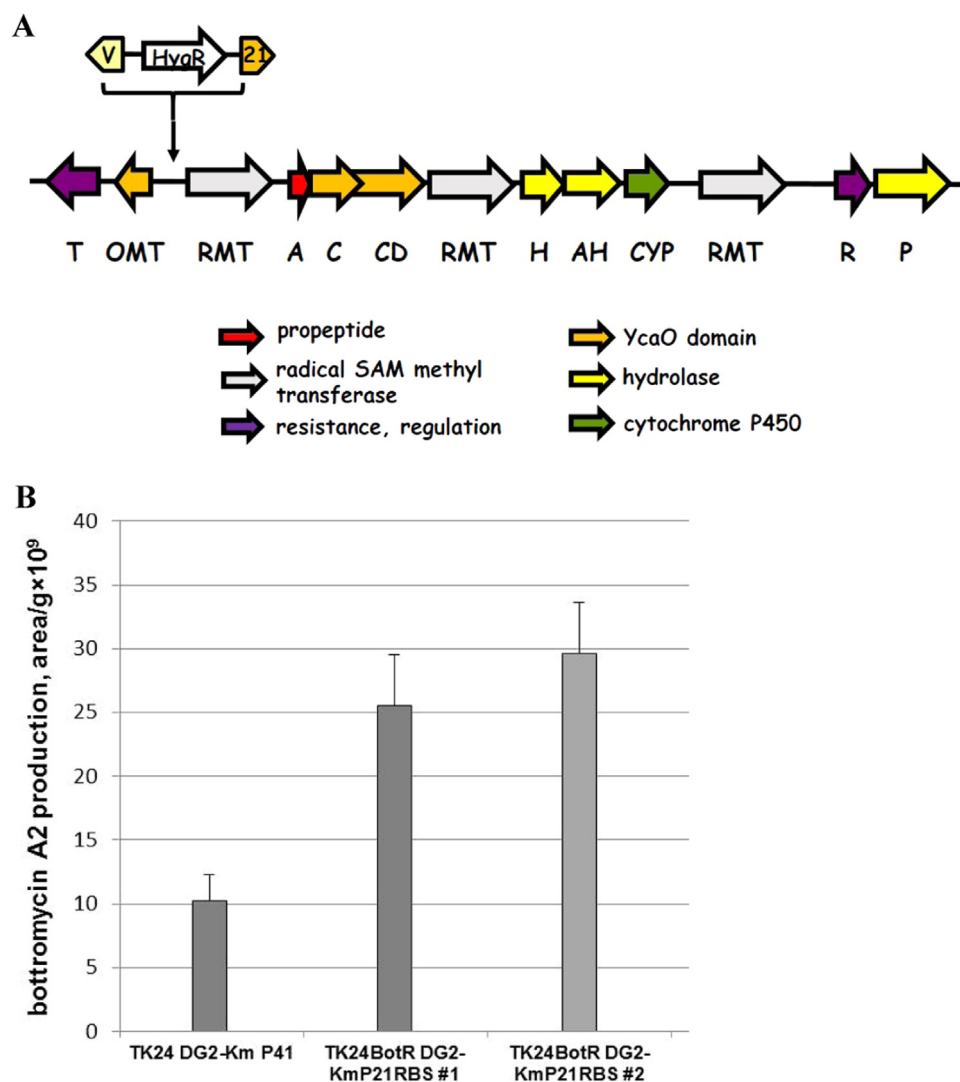


Fig. 2. Bottromycin production in *S. lividans* strains. (A) Schematic depiction of the *bot*-cluster and the cassette with strong promoters. P21 and V – semi-synthetic promoters, HygR – hygromycin-resistance gene. (B) Bottromycin production analysis in *S. lividans* recombinant strains. All results are equilibrated to the dry biomass of the strains, thus on the Y axis peak area divided per dry biomass (in g) is represented.

with strong constitutive promoters were unable to grow on MS medium. This effect, as mentioned earlier, may be attributable to the toxicity of the produced compound and/or the low resistance levels of the producer strain. To determine whether increased resistance improves the production rate, *S. lividans* TK24 bottromycin resistance mutants were selected. Bottromycin enriched crude extract was used to select bottromycin-resistant *S. lividans* strains (for details, see Materials and Methods). Two mutants, Bot-R2 and BotR-3, were obtained and used for expression of the DG2-Km-P21RBShyg cosmid, which contains strong promoters introduced between two operons. Exconjugants were obtained only with the *S. lividans* BotR-2 strain. All exconjugants were viable on MS and other tested media and did not show any significant differences in growth in comparison to the wild-type strain. LC-MS analysis of the extracts from this strain and the *S. lividans* DG2-Km-P41hyg control strain (for details see 3.4) revealed the presence of bottromycin A2 (mass – 822.43 Da) in both extracts. Furthermore, the production level of bottromycin A2 in the recombinant strain was an average of 1.5–3-fold higher than that of the control strain (Fig. 2B).

3.4. Random rational strategy for increasing bottromycin production in *S. lividans* TK24

In most cases, secondary metabolite gene clusters carry a number of operons in which expression is governed at various levels, and the products of some genes may constitute bottlenecks; thus, it is complicated to predict the combination of promoters leading to increased production. Therefore, different combinations of various promoters should often be tested. As the ideal expression levels of the genes/operons within the *bot*-cluster are unknown, we generated a library of *bot*-gene clusters carrying random synthetic promoters of unknown strengths inserted between two operons (Fig. 3A). Random promoters were created based on the consensus –35 and –10 sequences of the *ermEp1* promoter (Strohl, 1992). The sequences located upstream, in between and downstream of these consensus sequences were randomized using degenerate primers (Fig. 3A). The use of degenerate primers and an amplified hygromycin resistance gene facilitated the rapid and easy Red/ET cloning of a variety of different semi-synthetic promoters between two operons in the *bot*-cluster (Fig. 3A, for details see Materials and Methods). An average of 100 GB05-red DG2-KmRandom⁺ mutant colonies were randomly selected and used for direct conjugation into *S. lividans* TK24 using triparental mating. One hundred *S. lividans* TK24 DG2-KmRandom⁺ recombinant strains containing the library of bottromycin cosmids were constructed and used for production assays. In all assays the biosynthesis was expressed as μg of product per mg of dry biomass. In general, most of the strains had the same biomass yield as a control *S. lividans* TK24 DG2-Km⁺ strain. Liquid chromatography–mass spectrometry (LC-MS) analysis of the extracts revealed 10 strains producing 5–50 fold more bottromycin A2 than the control strain (Fig. 3B). *S. lividans* TK24 DG2-Km-P41hyg⁺ proved to be the best producer resulting from this applied strategy. The biomass yield in this strain was in most cases 1.2–1.5 folds lower than in the case of the control strain. This recombinant strain produced an average of 3.3 fold more bottromycin A2 than the native *S. sp.* BC16019 strain that has the averaged production rate 3–5 mg/L (Fig. 3B). Out of 100 analysed strains, 35 did not produce bottromycin at all, 12 were characterized by very low production and 43 had production levels comparable to or slightly higher than those of the control strain (*S. lividans* TK24 DG2-Km⁺).

3.5. Quantitative RT-PCR analysis of gene expression in *S. lividans* recombinant strains

To compare gene expression levels in the best *S. lividans* TK24 DG2-Km-P41hyg⁺ producer and the control *S. lividans* TK24 DG2-Km⁺ strain and to determine the factors leading to bottromycin overproduction, quantitative RT-PCR (RT-qPCR) experiments were

performed. Both strains were cultivated for 2 days at 28 °C in TSB media. RNA was isolated using standard protocols with Kirby mix (Kieser et al., 2000). The isolated RNA was used to synthesize cDNA, which was then utilized as a template for RT-qPCR with primers annealing to internal parts of the *botT* and *botRMT1* genes. These genes were chosen since they are localized in two different operons and provide an opportunity to compare expression levels in the left and right operons, respectively, of the recombinant and control strains (Fig. 3A). Transcriptional analysis showed that the expression of the *botT* and *botRMT1* genes was an average of 14 and 175 fold higher, respectively, in the *S. lividans* TK24 DG2-Km-P41hyg⁺ overproducer than that in the control strain (Table S1). Furthermore, the ratio between gene transcription in the left and right operons was approximately 1–4.8 in the *S. lividans* TK24 DG2-Km⁺ control strain and 1–59 in the overproducer. These data confirm that not all the genes in the cluster are expressed or need to be expressed at the same level to achieve high production. In summary, both significantly increased as well as somewhat balanced gene expression levels in the *S. lividans* TK24 DG2-Km-P41hyg⁺ recombinant strain led to bottromycin overproduction.

3.6. Improved bottromycin biosynthesis in a native producer using the identified promoter pair and cluster amplification

Because the random rational strategy we applied to increase bottromycin production appeared to be efficient in a heterologous host, we decided to use the pair of promoters obtained from the best producer (TK24 DG2-Km-P41hyg⁺) to tune production in the native strain, *S. sp.* BC16019. For this purpose, the integrase gene in the DG2-Km cosmid was substituted with the *gusA* reporter (for details, see Materials and Methods), resulting in construction of the DG2-KmgusA cosmid (Table 1). To obtain the sequence of the P41 promoter pair (Table S2), DG2-Km-P41hyg was partially sequenced. A hygromycin cassette flanked with the outward-oriented P41 promoter pair and containing a region of homology to the *bot*-cluster was amplified. The obtained fragment was incorporated into the DG2-KmgusA cosmid via Red/ET recombination to generate DG2-KmgusA-P41hyg (Table 1). To exclude potential mutations from the cloning procedure, the expression construct was verified via partial sequencing. The obtained cosmid was used to introduce a P41 promoter pair into the *bot*-cluster in the *S. sp.* BC16019 strain. Since the cosmid is suicide, because it does not contain any origin of replication for *Streptomyces*, all colonies obtained after its transfer into the *S. sp.* BC16019 strain resulted from the first crossover event. Single-crossover apramycin- and hygromycin-resistant mutants were screened for the loss of apramycin resistance and a blue colour resulting from a double-crossover event. After genotypic verification, the obtained *S. sp.* P41hyg⁺ mutant and the control wild-type strain were cultivated in SG medium at 28 °C for 5 days. The cultures were extracted with ethyl acetate and analysed for the presence of bottromycin. LC-MS analysis revealed that the *S. sp.* P41hyg⁺ mutant produced an average of 12 fold more bottromycin A2 than the wild-type strain (Fig. 4). The data obtained proves the efficiency of the P41 promoter pair in both heterologous and native producer strains for the genetic engineering of bottromycin production. Thus, using a random rational strategy, we succeeded in creating a promoter pair that boosts expression and biosynthesis in both producers.

Inspired by the abovementioned results, we aimed to further increase bottromycin production in the *S. sp.* BC16019 strain by introducing a second copy or amplifying the gene cluster (Yanai et al., 2006; Jiang, 2013). To this end, the cosmid DG2-Km-P41hyg containing the bottromycin gene cluster, the P41 promoters and the integrase gene was transferred into *S. sp.* BC16019 via conjugation, resulting in a *S. sp.* DG2-KmP41hyg strain carrying two copies of the *bot*-cluster, one native and one modified with the P41 promoter pair. Constructed strains were cultivated in SG medium for 5 days, extracted and analysed. LC-MS analysis of extracts revealed an expected array of

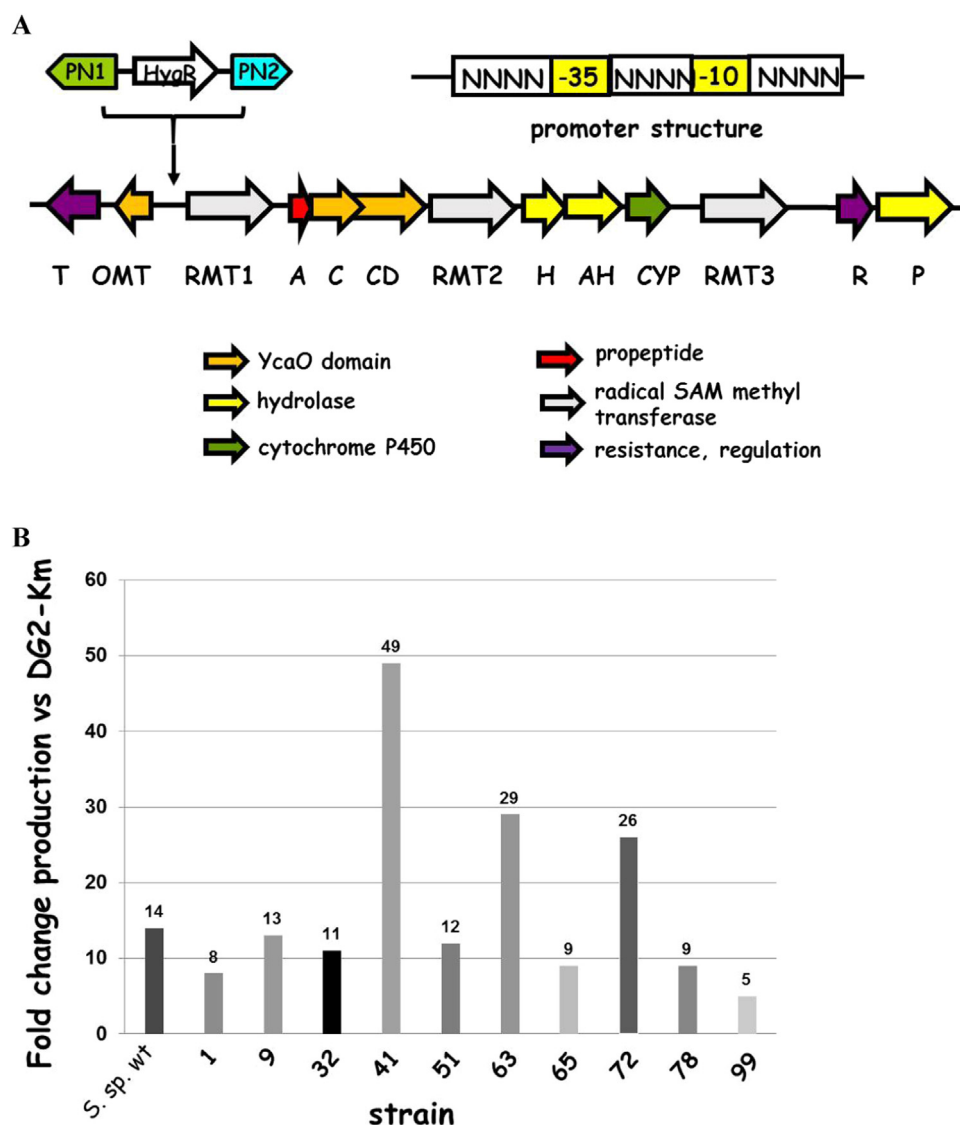


Fig. 3. Engineering and characterization of bottromycin production levels. (A) Schematic depiction of the *bot*-cluster, the cassette with random promoters and the structure of a random promoter. (B) Bottromycin production in different *S. lividans* engineered strains. PN1 and PN2 – randomly generated promoters, – 10 and – 35 – promoter consensus regions.

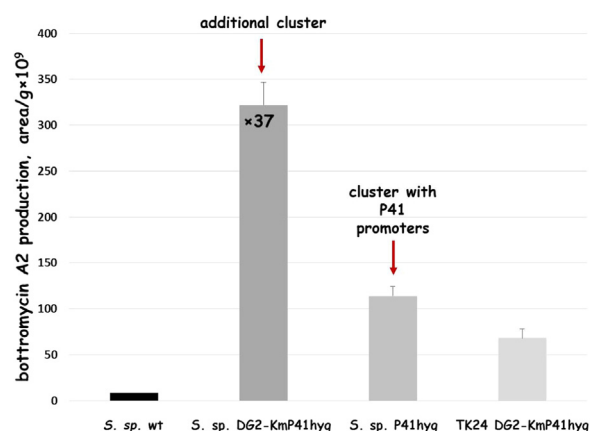


Fig. 4. Bottromycin A2 production in *Streptomyces* sp. BC16019 and *S. lividans* recombinant strains. All results are equilibrated to the dry biomass of the strains, thus on the Y axis peak area divided per dry biomass (in g) is represented.

bottromycins (A2 and B2). The recombinant strain containing two copies of the cluster produced an average of 37 and 2.8 fold more bottromycin A2 than the wild-type strain or the strain with one copy of the modified *bot*-cluster (Fig. 4), respectively. Thus, introduction of the additional copy of the modified cluster into the genome of the *S. sp.* BC16019 strain led to the further improvement of bottromycin production.

3.7. New methylated bottromycin in an overproducing strain

After careful analysis of the metabolite profile of the *S. sp.* strain carrying two copies of the engineered *bot*-cluster, a new peak with a mass-to-charge ratio of 837.4 was revealed (Fig. 5A). The peak was not detectable on the base chromatogram of extracts from the wild-type strain (Fig. 5A). MS/MS analysis confirmed a characteristic fragmentation pattern similar to that of bottromycin A2; however, in the case of some fragments, there was a mass difference of 14 Da (Fig. S1). We suggest that this difference of 14 Da indicated a methyl group. Based on the MS/MS data, we predicted that the probable methyl group was located in the macrocycle. The *bot*-gene cluster contains a set of four methyltransferases that perform methylation in different positions of

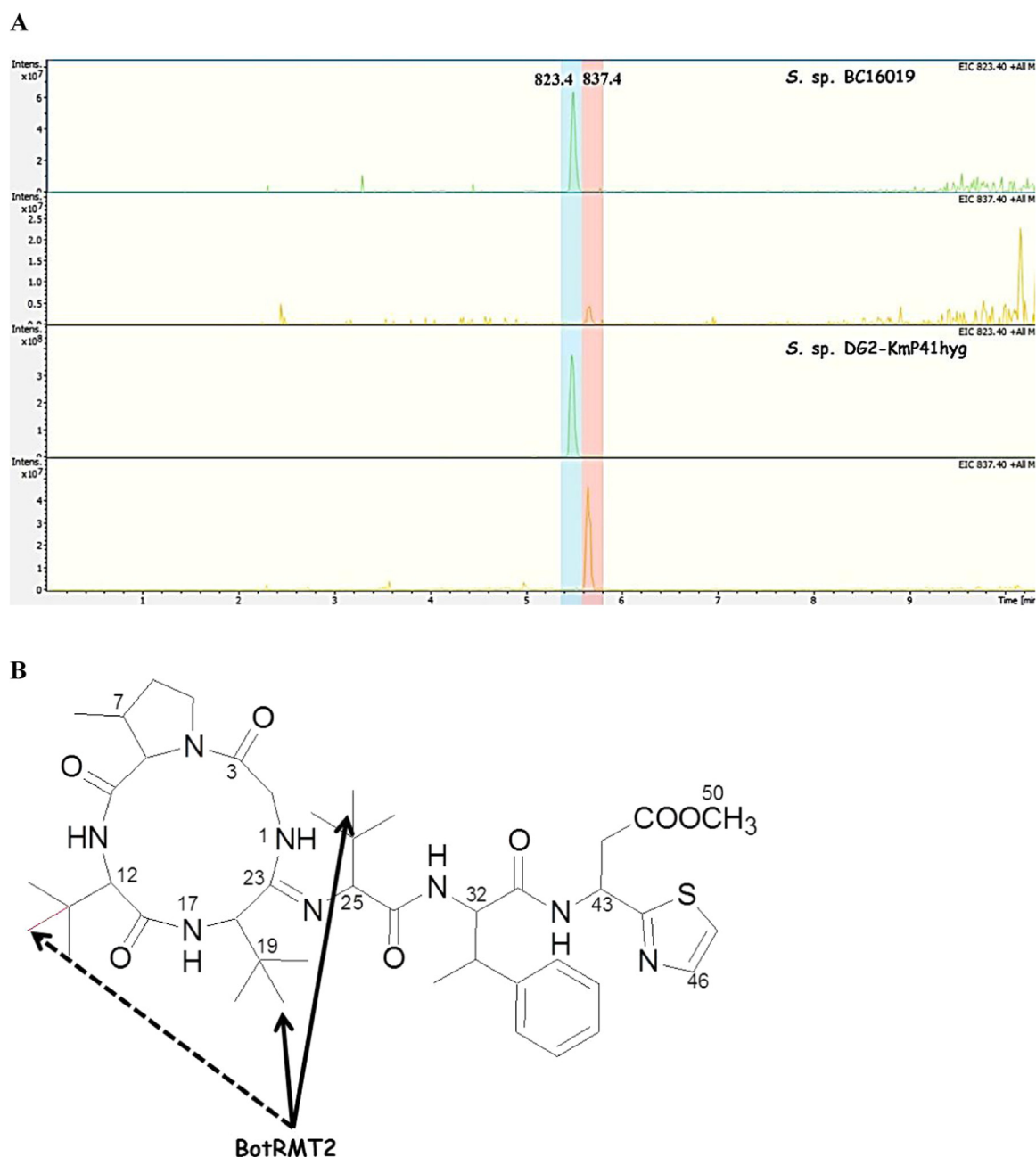


Fig. 5. New bottromycins in *S. sp.* recombinant strains. (A) Bottromycin production profile in *S. sp.* strains. (B) Proposed chemical structure of new bottromycin with methylation at valine in third position.

the bottromycin molecule (Huo et al., 2012). However, only two methyltransferases, *botRMT2* and *botRMT3*, carry out methylations in the macrocycle. There are few other positions that are able to be methylated in the amidine ring. The BotRMT2 methyltransferase performs β -methylation of valine residues in the fourth and fifth positions (Fig. 5B) (Huo et al., 2012). Taking this into account and noting that the valine in the third position of the macrocycle is not methylated, we suggested that BotRMT2 is also responsible for the methylation in that position (Fig. 5B). Thus, the new bottromycin is likely a derivative of BotA2 with a methylated valine residue in position 3; however, additional experiments are required to confirm proposed planar structure.

3.8. Generation of ‘unnatural’ natural bottromycin derivatives

Since bottromycin is a ribosomally synthesized peptide, the *botA*

gene coding for its precursor provides an opportunity to introduce altered amino acids, including non-canonical ones, through nucleotide substitutions. To implement this strategy, a mutant strain with a deleted gene encoding a propeptide is created, and then new alleles with codons coding for amino acid substitutions are introduced. To achieve this objective, a DG2-KmgusAdelBotA cosmid (Table 1) carrying a *botA* gene deletion with the hygromycin cassette was constructed using Red/ET technology (for details, see Materials and Methods). The cosmid was transferred into *S. sp.* BC16019 strain via conjugation. Obtained transconjugants were selected based on resistance to hygromycin and apramycin. Single-crossover mutants were screened for the loss of apramycin resistance and a blue colour resulting from a double-crossover event. A *S. sp.* Δ *botA* strain (Table 1) carrying a *botA* gene deletion was created. The identity of the strain was confirmed by PCR analysis (data not shown). LC-MS analysis of extracts from Δ *botA* did not detect

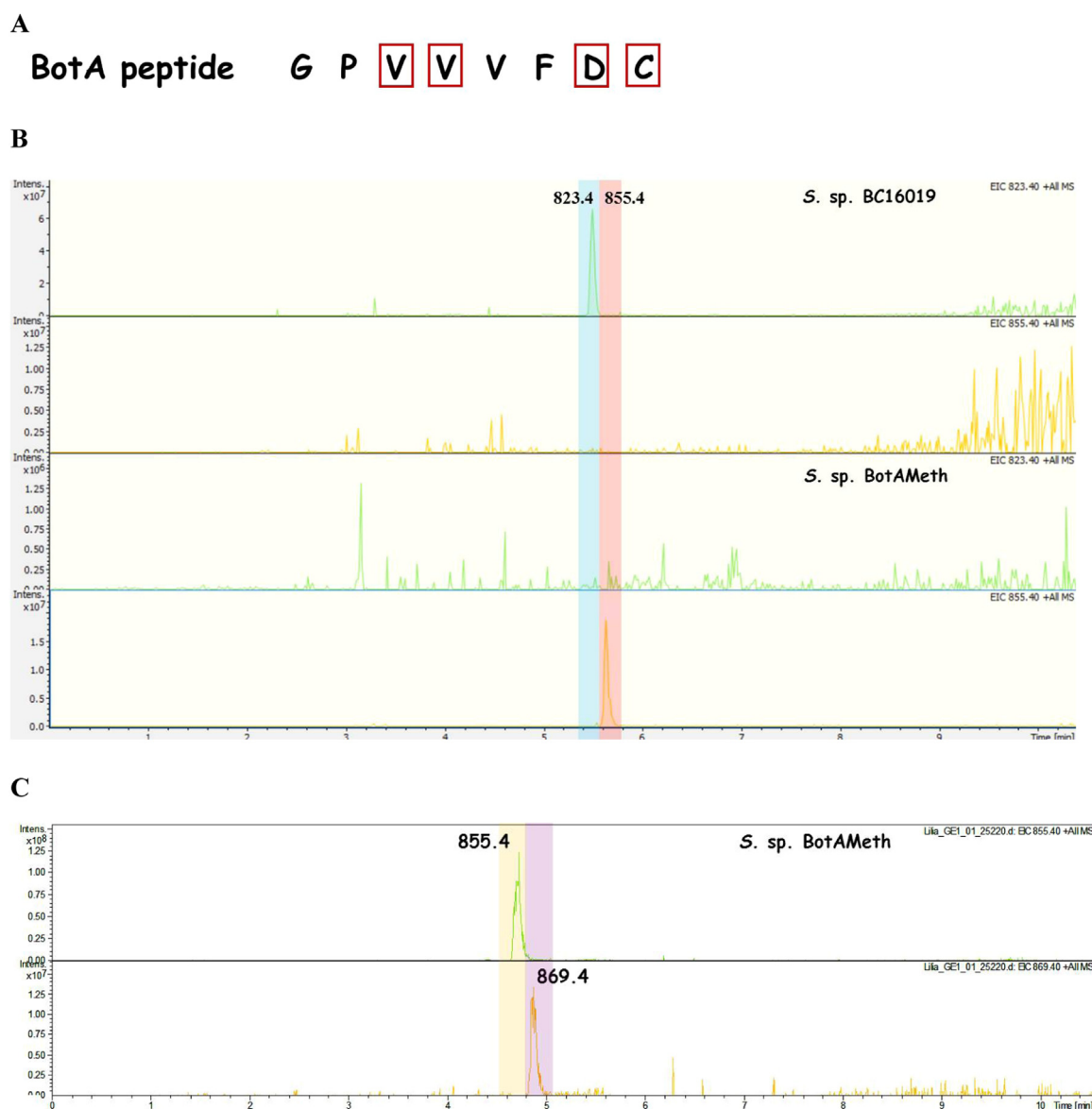


Fig. 6. New unnatural natural bottromycin derivatives. (A) Amino acids that were going to be substituted in bottromycin peptide. (B) Bottromycin production profile in the *S. sp.* strains. (C) Bottromycin production profile in the *S. sp.* BotAMeth strain.

any bottromycin biosynthesis, confirming the deletion of the key biosynthetic gene. The *S. sp.* Δ botA strain was further used as a chassis to introduce new engineered *botA* alleles and create ‘unnatural’ natural bottromycins.

We created a set of cosmids (Table 1) carrying mutant *botA* alleles with different nucleotide substitutions including stop codons for the incorporation of non-canonical amino acids (for details, see Materials and Methods). Amino acids in four positions of the BotA mature peptide, namely, valine in position 3, valine 5, aspartic acid 7 and cysteine in position 8, were mutated (Fig. 6A, Table S3). Methionine, isoleucine, tyrosine, cysteine, glutamic acid and pyrrolysine were incorporated instead of valine in position 3. Tyrosine was incorporated instead of valine in position 4. Cysteine was substituted with proline or pyrrolysine. Methionine, tyrosine, glutamic acid and pyrrolysine were introduced in place of aspartic acid. The idea was to develop a method that will allow the easy introduction of different amino acids in bottromycin molecule in order to expand a chemical space of this molecule and test the activities of the obtained derivatives in the future. Valines in positions 3 and 4 were targeted, since substitutions in these positions should not severely influence the structure and thus the maturation of

the molecule, at least in case of introduction of similar amino acids, and therefore were expected to give positive results. This was also important for validation of the method. Aspartic acid was chosen since methyl group in this position is unstable and is critical for the activity (Kobayashi et al., 2010). Thus, it is of utmost importance to get derivatives with substitutions in this position that are more stable or contain groups which can be chemically modified later.

The constructed cosmids were used to generate *S. sp.* mutants producing bottromycin derivatives. The same scheme was used to obtain all mutants. As an example, the construction of a strain producing bottromycin containing methionine instead of valine in position 3 is described. The DG2-KmBotAMeth cosmid (Table 1) carrying the *botA* allele with methionine substituting for the valine in position 3 was transferred into the *Streptomyces sp.* Δ botA strain via conjugation. Since the construct is a suicide cosmid, all colonies obtained after its transfer resulted from the first crossover event. Single-crossover apramycin- and hygromycin-resistant mutants were screened for the loss of apramycin and hygromycin resistance and a blue colour resulting from a double-crossover event. After genotypic verification, the *S. sp.* BotAMeth strain was obtained. LC-MS analysis of the extracts from this strain revealed a

new peak with a mass-to-charge ratio of 855.4, which corresponds to the predicted mass of the bottromycin derivative with the incorporated methionine in position 3 (Fig. 6B). This peak was not present in the Δ botA strain or the wild-type strain. MS/MS analysis revealed a characteristic fragmentation pattern similar to that of bottromycin A2, but in the case of some fragments, the mass difference was 32 Da, corresponding to a sulphur atom (Fig. S3). Based on these results, we assumed that the newly identified peak corresponds to a new bottromycin with methionine substitution (bottromycin M). Careful analysis of the metabolic profile of the *S. sp.* BotAMeth strain and comparison to the wild-type strain revealed the presence of an additional new peak with a mass-to-charge ratio of 869.4 Da (Fig. 6C). MS/MS analysis revealed a characteristic fragmentation pattern similar to that of bottromycin A2 and a putative derivative containing methionine; however, for some fragments, the mass difference was 46 or 14 Da in comparison to bottromycin A2 or bottromycin M (data not shown). Based on these data, we suspect that this peak corresponds to bottromycin with an S-methylmethionine residue.

The *S. sp.* BotAlso mutant expresses a *botA* mutant allele that encodes a bottromycin peptide with isoleucine instead of valine in position 3. LC-MS analysis of the extracts of this strain revealed the presence of a new peak with a mass-to-charge ratio of 837.4 Da, corresponding to the predicted mass of the bottromycin derivative with incorporated isoleucine instead of valine in position 3 (Fig. 7A). This peak was not present in the Δ botA strain. MS/MS analysis revealed a characteristic fragmentation pattern similar to that of bottromycin A2, but some fragments had a mass difference of 14 Da corresponding to a methyl group (Fig. S3). Based on these data, we conclude that the new peak is a bottromycin derivative with isoleucine (Fig. 7B).

Analysis of other *S. sp.* mutants with substitutions in the BotA propeptide did not reveal the biosynthesis of any bottromycins, including expected derivatives (Table S3). This observation might be related to production levels below the detection range or inability of post-translational modification machinery to process the altered propeptide into a mature, bottromycin-like structure. Thus, additional genetic modifications are most likely required to obtain additional new derivatives.

3.9. Structural elucidation of a new bottromycin with a methionine substitution

To confirm the structure and position of methionine in the new bottromycin molecule, NMR analysis was performed. To increase the production of the new bottromycin derivative and simplify the purification procedure, the P41 pair of promoters was introduced into the *S. sp.* BotAMeth strain (for details, see Materials and Methods). We obtained the *S. sp.* BotAP41Meth strain, an overproducer of bottromycin M, ideal for compound purification. This strain was cultivated in SG media at 28 °C for 7 days. The obtained culture broth was extracted with ethyl acetate, evaporated and dissolved in methanol. The peak with a mass of 854.4 Da was purified using size-exclusion chromatography on a dextran polymer bead LH20 column and reverse-phase chromatography on a preparative C18 column. Then, 0.8 mg of the compound was used in NMR studies.

Careful analysis of 2D NMR data, such as HSQC, HMBC and COSY spectra, revealed the 'common' constitution of the following amino acid residues: Gly; *cis*-3Me-Pro; and two each of t-Leu, Me-Phe and Thia- β -Ala, as observed in other bottromycin analogues. The remaining spin system comprising an alpha proton at δ 3.85 adjacent to two methylene signals between δ 1.84–2.56 and a methyl singlet at δ 2.02 was identified as methionine. Subsequently, the HMBC spectrum was further used to confirm the amino acid sequence and finalize the planar structure of bottromycin M (Fig. 7C, Table S4).

4. Discussion

A growing body of genome sequencing data confirms that micro-organisms are a treasure trove of bioactive compounds and new chemical scaffolds (Rigali, 2018). However, the low yield of NPs represents a serious obstacle on the way to their discovery. The ability to quickly turn on or up-regulate biosynthetic pathways (Baltz, 2011; Rigali, 2018) will 'open the door' to new chemical entities. Thus, strategies that provide an opportunity to rewire and overcome the genetic circuitry controlling low yields are in great demand. In this article, we have presented a new rational random approach that was used, for the first time, for cluster 'refactoring' and the construction of balanced pathways with high yields. The key idea underlying this strategy is a construction of a library of gene clusters of interest with random constitutive promoters introduced in different combinations into operons followed by their expression in host strains under specific conditions. Using this approach, we successfully balanced the expression of genes organized in two operons and created modified gene clusters that provide 5–50 fold higher yields of bottromycins than the initial cosmid. This allowed for the identification of several new derivatives that have not previously been described. Furthermore, we demonstrated, for the first time, that an increase in the resistance of the heterologous producer strain led to a further increase in production.

In this work, we clearly showed that the developed strategy is rapid, easy to use and successful, as it resulted in strains producing high levels of bottromycin. The simple incorporation of randomly generated promoters upstream of operons, based on the well-defined constitutive *ermEp* promoter, resulted in the construction of 10 overproducing strains (on average 10% of all tested strains were overproducers). Moreover, this approach led to the identification of new derivatives that are most likely produced in minute amounts in the wild-type strain and have not been previously detected or described. Thus, a simple increase in bottromycin production levels allowed us to access a new molecule from a well-studied and previously described pathway that would otherwise be overlooked (Crone et al., 2016). Based on the obtained MS-MS data, we suggested that this molecule is a derivative of bottromycin A2, with a methylated valine residue at position 3. The BotRMT2 enzyme in the pathway methylates two other valine residues in the molecule (Huo et al., 2012) and is predicted to be responsible for the methylation at the abovementioned position as well. There are no other candidates that might be responsible for this reaction since BotOMT is an O-methyltransferase and two others, BotRMT1 and BotRMT3, methylate proline and phenylalanine, which are completely different substrates (Huo et al., 2012). Additional experiments are required to verify this hypothesis.

Application of the random rational strategy does not require the extensive preliminary investigation of biosynthetic pathways, enzyme activities and kinetics. The only requirement is verification of the genetic organization of a cluster. Moreover, in the case of complicated clusters, the architecture might be simplified by Red/ET cloning prior to promoter engineering. Our strategy also provides an opportunity to introduce different combinations of not only promoters, but also other controlling parts, such as ribosomal binding sites and variable tunable intergenic elements that should be incorporated into PCR primers; however, in this case much more recombinant strains should be analysed, since the probability of successful combinations will go down. Combining this approach with bioactivity tests, reporter systems or biosensors allows to quickly estimate the biosynthesis of molecules, resulting in an opportunity to achieve significant levels of metabolite production in a relatively short time. Once a library of clusters is created, it can be kept and used for other purposes, e.g., selection of the best cluster in other strains and conditions.

Random rational strategy allowed us to create a combination of promoters in just two steps that led to a 50-fold increase in bottromycin production in the heterologous host *S. lividans*. Application of the same P41 activating cassette to modulate biosynthesis in the native *S. sp.*

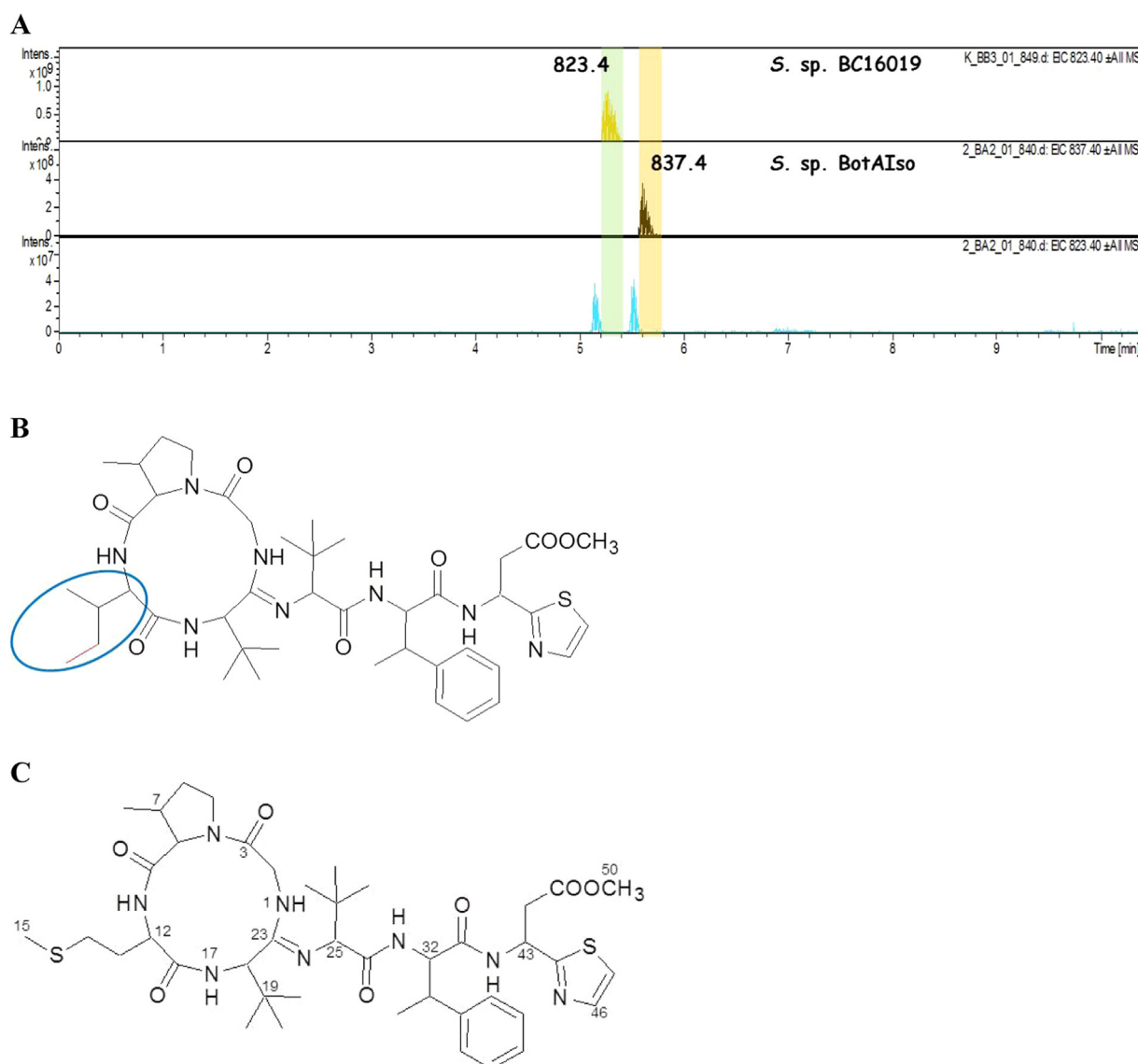


Fig. 7. Bottromycin derivatives. (A) Bottromycin production profile in the *S. sp.* BC16019 and *S. sp.* BotAIso strains. (B) Proposed chemical structure of new bottromycin with isoleucine instead of valine in third position. (C) Chemical structure of new bottromycin with methionine instead of valine in position 3.

BC16019 strain also increased production, although only by 12-fold. The presence of the same trend in both native and heterologous hosts provides evidence that synthetically created promoters are orthogonal, do not undergo intracellular regulation and do provide stable and constitutive expression in different streptomycetes strains. Therefore, utilization of the orthogonal controlling elements that increase biosynthesis to certain levels in one strain can in some cases be directly extrapolated to regulate the same biosynthesis in another strain.

Analysis of the expression of genes in the operons in the best bottromycin producer and comparison with the control strain revealed stark differences. Transcription was increased 13 fold by the left operon and 179 fold by the right operon. Furthermore, the ratio of gene expression in the operons was 1–13, indicating that transcription was modulated to a different extent in the overproducer. The data obtained once again confirmed the growing body of evidence indicating that increment in the expression of biosynthetic genes should be balanced to certain and differing levels so as not to create bottlenecks and metabolic burden on the cell. Thus, in some but definitely not all cases, genes should be expressed as strongly as possible. However, the enhancement of bottromycin production is not strictly proportional to the huge increase in the transcriptional activity reported. This might be due to

limitation in the translation of abundant mRNA. Therefore, a prerequisite for NP overproduction is indeed not “hyper” transcription that can lead to metabolic burden, but rather an efficient coordination between transcriptional and translational activity together with improved resistance. Taking into account our results, it is obvious that it is complicated to rationally predict the best expression levels for the genes in operons that will lead to increased production. Therefore, our strategy provides a straightforward mechanism to identify the optimal combination of regulatory elements for certain pathways. Moreover, we clearly demonstrated that coupling random rational approaches with other widely used strategies for metabolite overproduction, such as cluster amplification (Manderscheid, 2016), paves the way for further strain optimization. This assumption was confirmed by the additional increase in bottromycin production after the integration of an additional copy of the modified cluster into the chromosome of the native producer.

In addition, in this study we developed methods for genetic engineering the native producer strain. Using these molecular tools, we succeeded for the first time in creating producers of ‘unnatural’ natural bottromycin derivatives. Three novel bottromycin molecules were obtained and verified by MS/MS analysis. The structure of a novel

bottromycin containing methionine was confirmed by NMR analysis.

Detailed analysis of the metabolic profile of the bottromycin producer with methionine substitution revealed a novel peak with a mass-to-charge ratio of 869.4. Based on MS/MS data, we predict that this compound is a derivative of bottromycin A2 with methylated methionine. Taking this prediction into account, we assume that BotRMT2 methyltransferase possess a certain level of promiscuity and thus is able to methylate methionine as well as valine at position 3.

For the first time, it was shown that an increase in the resistance of the heterologous host strain to the produced metabolite paves the way for the increased production of this metabolite. A bottromycin cosmid with strong promoters, which was not efficiently transferred into the *S. lividans* wild-type strain, was easily expressed in the *S. lividans* BotR-2 resistant mutant. Furthermore, the obtained recombinant strains produced 1.5–2.5 fold more bottromycin than the control strain with the best pair of randomly generated promoters. Thus, low resistance to metabolites of interest is a severe obstacle standing in the way of overproduction.

In summary, an alternative, fast and easy strategy was developed to balance gene expression and construct strains that overproduce metabolites of interest. As a result of the study, several overproducing strains were constructed, and four new bottromycin derivatives were produced. Taking into account the data described here, we anticipate that this generic strategy will facilitate the overproduction of known NPs and the discovery of new potentially active NPs in *Actinobacteria*.

Acknowledgments

This work was supported by the European Commission under the 7th Framework Program through the “Collaborative Project” action “STREPSYNTH” Grant no. 613877 and through the ERC starting Grant EXPLOGEN no. 281623 to AL. A part of this work was also supported by the Deutsche Forschungsgemeinschaft grant LU 1524/4-1 to AL. MVM was supported by Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa fellowship SFRH/BPD/95683/2013. FM was supported by the Erasmus program of the European Union..

Conflict of interest

None.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ymben.2018.09.010

References

- Ajikumar, P.K., et al., 2010. Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science* 330, 70–74.
- Bachmann, B.O., Van Lanen, S.G., Baltz, R.H., 2014. Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? *J. Ind. Microbiol. Biotechnol.* 41, 175–184.
- Baltz, R.H., 2011. Strain improvement in actinomycetes in the postgenomic era. *J. Ind. Microbiol. Biotechnol.* 38, 657–666.
- Bentley, S.D., et al., 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417, 141–147.
- Bibb, M.J., 2005. Regulation of secondary metabolism in streptomycetes. *Curr. Opin. Microbiol.* 8, 208–215.
- Biggs, B.W., et al., 2014. Multivariate modular metabolic engineering for pathway and strain optimization. *Curr. Opin. Biotechnol.* 29, 156–162.
- Bilyk, O., Luzhetskyy, A., 2016. Metabolic engineering of natural product biosynthesis in actinobacteria. *Curr. Opin. Biotechnol.* 42, 98–107.
- Brown, E.D., Wright, G.D., 2016. Antibacterial drug discovery in the resistance era. *Nature* 529, 336–343.
- Cimermancic, P., et al., 2014. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* 158, 412–421.
- Cragg, G.M., Newman, D.J., 2013. Natural products: a continuing source of novel drug leads. *Biochim. Biophys. Acta* 1830, 3670–3695.
- Crone, W.J., et al., 2016. Dissecting bottromycin biosynthesis using comparative untargeted metabolomics. *Angew. Chem. Int. Ed. Engl.* 55, 9639–9643.
- Crone, W.J.K., Leeper, F.J., Truman, A.W., 2012. Identification and characterisation of the gene cluster for the anti-MRSA antibiotic bottromycin: expanding the biosynthetic diversity of ribosomal peptides. *Chem. Sci.* 3, 3516–3521.
- Fedorenko, V., et al., 2015. Antibacterial discovery and development: from gene to product and back. *Biomed. Res. Int.* 591349. <https://doi.org/10.1155/2015/591349>.
- Fu, J., et al., 2008. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition. *Nucleic Acids Res.* 36, e113.
- Gomez-Escribano, J.P., Song, L., Bibb, M.J., 2012. Posttranslational β -methylation and macrolactamidation in the biosynthesis of the bottromycin complex of ribosomal peptide antibiotics. *Chem. Sci.* 3, 3522–3525.
- Horbal, L., et al., 2014. The pathway-specific regulatory genes, *tei15** and *tei16**, are the master switches of teicoplanin production in *Actinoplanes teichomyceticus*. *Appl. Microbiol. Biotechnol.* 98, 9295–9309.
- Horbal, L., et al., 2016. A gene cluster for the biosynthesis of moenomycin family antibiotics in the genome of teicoplanin producer *Actinoplanes teichomyceticus*. *Appl. Microbiol. Biotechnol.* 100, 7629–7638.
- Hou, Y., et al., 2012. Structure and biosynthesis of the antibiotic bottromycin D. *Org. Lett.* 14, 5050–5053.
- Huo, L., et al., 2012. Synthetic biotechnology to study and engineer ribosomal bottromycin biosynthesis. *Chem. Biol.* 19, 1278–1287.
- Ikeda, H., et al., 2003. Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat. Biotechnol.* 21, 526–531.
- Janssen, D., et al., 2007. Chivosazole A - Elucidation of the absolute and relative configuration. *Angew. Chem. Int. Ed.* 46, 4898–4901.
- Jiang, L., 2013. Combined gene cluster engineering and precursor feeding to improve gougouerin production in *Streptomyces gramineus*. *Appl. Microbiol. Biotechnol.* 97, 10469–10477.
- Katz, L., Baltz, R.H., 2016. Natural product discovery: past, present, and future. *J. Ind. Microbiol. Biotechnol.* 43, 155–176.
- Kieser, T., et al., 2000. *Practical Streptomyces Genetics*, 2nd ed. John Innes Foundation, Norwich.
- Keasling, J.D., 2012. Synthetic biology and the development of tools for metabolic engineering. *Metab. Eng.* 14, 189–195.
- Kobayashi, Y., et al., 2010. Bottromycin derivatives: efficient chemical modifications of the ester moiety and evaluation of anti-MRSA and anti-VRE activities. *Bioorg. Med. Chem. Lett.* 20, 6116–6120.
- Li, T.L., et al., 2004. Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: characterization of two glycosyltransferases and the key acyltransferase. *Chem. Biol.* 11, 107–119.
- Lin, Y.C., Tanaka, N., 1968. Mechanism of action of bottromycin in polypeptide biosynthesis. *J. Biochem.* 63, 1–7.
- Liu, G., et al., 2013. Molecular regulation of antibiotic biosynthesis in *Streptomyces*. *Microbiol. Mol. Biol. Rev.* 77, 112–143.
- Manderscheid, N., 2016. An influence of the copy number of biosynthetic gene clusters on the production level of antibiotics in a heterologous host. *J. Biotechnol.* 232, 110–117.
- Mast, Y., et al., 2011. Characterization of the ‘pristinamycin supercluster’ of *Streptomyces pristinaespiralis*. *Microb. Biotechnol.* 4, 192–206.
- Myronovskiy, M., et al., 2016. Generation of new compounds through unbalanced transcription of landomycin A cluster. *Appl. Microbiol. Biotechnol.* 100, 9175–9186.
- Myronovskiy, M., Luzhetskyy, A., 2016. Native and engineered promoters in natural product discovery. *Nat. Prod. Rep.* 33, 1006–1019.
- Myronovskiy, M., Welle, E., Fedorenko, V., Luzhetskyy, A., 2011. Beta-glucuronidase as a sensitive and versatile reporter in actinomycetes. *Appl. Environ. Microbiol.* 77, 5370–5383.
- Ostasz, B., et al., 2013. Exploring and exploiting gene networks that regulate natural products biosynthesis in actinobacteria. *Nat. Prod. J.* 3, 189–198.
- Otake, T., Kaji, A., 1976. Mode of action of bottromycin A2. Release of aminoacyl- or peptidyl-tRNA from ribosomes. *J. Biol. Chem.* 251, 2299–2306.
- Patridge, E., et al., 2016. An analysis of FDA-approved drugs: natural products and their derivatives. *Drug Discov. Today* 21, 204–207.
- Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29 (e45–e45).
- Rebets, Y., et al., 2014a. Actinomycetes biosynthetic potential: how to bridge in silico and in vivo? *J. Ind. Microbiol. Biotechnol.* 41, 387–402.
- Rebets, Y., et al., 2014b. Complete genome sequence of producer of the glycopeptide antibiotic *Aculeximycin Kutzneria albidia* DSM 43870T, a representative of minor genus of Pseudonocardiaceae. *BMC Genom.* 15, 885.
- Rigali, S., 2018. Cracking the regulatory code of biosynthetic gene clusters as a strategy for natural product discovery. *Biochem. Pharmacol.* 153, 24–34.
- Rinke, C., et al., 2013. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499, 431–437.
- Robertson, H.L., 2017. Toward systems metabolic engineering of *Streptomyces* for secondary metabolites production. *Biotechnol. J.* 13 (1). <https://doi.org/10.1002/biot.201700465>.
- Sambrook, J., Russell, D., 2001. *Molecular Cloning: A Laboratory Manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Seghezzi, N., et al., 2011. The construction of a library of synthetic promoters revealed some specific features of strong *Streptomyces* promoters. *Appl. Microbiol. Biotechnol.* 90, 615–623.
- Shimamura, H., 2009. Structure determination and total synthesis of bottromycin A2: a potent antibiotic against MRSA and VRE. *Angew. Chem. Int. Ed. Engl.* 48, 914–917.
- Siegl, T., et al., 2013. Design, construction and characterization of a synthetic promoter library for fine-tuned gene expression in actinomycetes. *Metabol. Eng.* 19, 98–106.
- Smanski, M.J., et al., 2016. Synthetic biology to access and expand nature’s chemical

- diversity. *Nat. Rev. Microbiol.* 14, 135–149.
- Sosio, M., et al., 2004. Organization of the teicoplanin gene cluster in *Actinoplanes teichomyceticus*. *Microbiology* 150, 95–102.
- Stratigopoulos, G., Bate, N., Cundliffe, E., 2004. Positive control of tylosin biosynthesis: pivotal role of TylR. *Mol. Microbiol.* 54, 1326–1334.
- Strohl, W.R., 1992. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Res.* 20, 961–974.
- Sun, J.G., et al., 2014. Distinct promoters affect pyrroloquinoline quinone production in recombinant *Escherichia coli* and *Klebsiella pneumoniae*. *Curr. Microbiol.* 69, 451–456.
- van Wezel, G.P., McDowall, K.J., 2011. The regulation of the secondary metabolism of *Streptomyces*: new links and experimental advances. *Nat. Prod. Rep.* 28 (1311–1333).
- Watve, M.G., et al., 2001. How many antibiotics are produced by the genus *Streptomyces*? *Arch. Microbiol.* 176, 386–390.
- Weber, T., et al., 2015. Metabolic engineering of antibiotic factories: new tools for antibiotic production in actinomycetes. *Trends Biotechnol.* 33, 15–26.
- Woolston, B.M., Edgar, S., Stephanopoulos, G., 2013. Metabolic engineering: past and future. *Annu. Rev. Chem. Biomol. Eng.* 4, 259–288.
- Yanai, K., Murakami, T., Bibb, M., 2006. Amplification of the entire kanamycin biosynthetic gene cluster during empirical strain improvement of *Streptomyces kanamyceticus*. *Proc. Natl. Acad. Sci. USA* 103, 9661–9666.
- Zarins-Tutt, J.S., et al., 2016. Prospecting for new bacterial metabolites: a glossary of approaches for inducing, activating and upregulating the biosynthesis of bacterial cryptic or silent natural products. *Nat. Prod. Rep.* 33, 54–72.
- Zhang, Y., Muylers, J.P.P., Testa, G., Stewart, A.F., 2000. DNA cloning by homologous recombination in *Escherichia coli*. *Nat. Biotechnol.* 18, 1314–1317.
- Zhou, S., et al., 2017. The application of powerful promoters to enhance gene expression in industrial microorganisms. *World J. Microbiol. Biotechnol.* 33, 23.

COMPLEMENTARY PUBLICATION II

Marques, F., Ramos, P. D., Lopes, M. M., Santos, J. B., Tavares, F., Olsson, I. A. S., & Almeida, M. S. (2022). BioLab – um espaço de experimentação e cruzamento disciplinar. In S. C. Soares, F. Remião, A. V. Martins & S. Nunes (Coords.), *Livro de Atas do 7.º Congresso Nacional de Práticas Pedagógicas no Ensino Superior* (pp. 295-301). UA Editora – Universidade de Aveiro.

BioLab – um espaço de experimentação e cruzamento disciplinar

Filipe Marques ‡,*
Pedro Dias Ramos ‡,*
Maria Manuela Lopes ‡
Júlio Borlido Santos ‡
Fernando Tavares †
I. Anna S. Olsson ‡,*
Maria Strecht Almeida *

‡ i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto

filipe.marques@ibmc.up.pt
pedro.ramos@ibmc.up.pt
maria.lopes@i3s.up.pt
borlido.santos@i3s.up.pt
olsson@ibmc.up.pt

† FCUP – Faculdade de Ciências, Universidade do Porto

ftavares@fc.up.pt

* ICBAS – Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto

msalmeida@icbas.up.pt

Resumo

Os grandes desafios atuais, como saúde, produção de alimentos, alterações climáticas, ambiente e biodiversidade, exigem um trabalho colaborativo no qual a investigação dialoga com a sociedade sobre prioridades e formas de ação. Nesta comunicação reportamos sobre a unidade curricular UC InovPed *Biolaboratório – Projeto de Experimentação Multidisciplinar* (BioLab), um espaço de formação onde estudantes de diferentes áreas desenvolvem trabalho em projetos. No processo, os estudantes dão resposta a preocupações atuais da sociedade, identificando questões científicas pertinentes enquadradas no conhecimento em biologia. Envolvendo uma equipa docente multidisciplinar (ciências de vida, sociologia, comunicação e arte), implementou-se uma estratégia centrada no estudante, recorrendo a metodologias de aprendizagem ativa, focada no diálogo biologia-sociedade. Em 2021 integrou sessões remotas (apresentações formais, discussões de grupo, *world cafés*, *role-plays*) e sessões laboratoriais. A formulação de projetos incluiu: chamada para ideias-resposta a um tema integrador; chamada para composição de equipas. Os grupos mantiveram um diário aberto aos membros do BioLab. A par de sessões laboratoriais tuteladas, o desenvolvimento dos projetos foi regularmente aferido em apresentações seguidas de discussão coletiva. Considera-se que o BioLab é uma iniciativa bem-sucedida – o envolvimento e interesse demonstrado pelos estudantes são disso indicadores – e que poderá ser usada noutros contextos de formação.

Palavras-Chave: aprendizagem ativa, colaboração multidisciplinar, diálogo biologia-sociedade.

1. Contextualização

É reconhecido que os grandes desafios atuais – como saúde, produção de alimentos, alterações climáticas, ambiente e biodiversidade – exigem um trabalho colaborativo e diálogo com a sociedade sobre prioridades e formas de ação. É ainda reconhecido que esses desafios deverão ser perspectivados em termos da Agenda 2030 da Nações Unidas que define os objetivos do desenvolvimento sustentável (United Nations, 2015). Adicionalmente, considera-se hoje que a aliança das áreas STEAM (*Science, Technology, Engineering, Arts and Math*) na formação contribui para uma melhor resposta a desafios sociais (Comissão Europeia, 2017). A criação da unidade de formação contínua cuja implementação aqui se reporta, *Biolaboratório – Projeto de Experimentação Multidisciplinar* (BioLab), foi motivada por esse tipo de considerações.

BioLab é uma das mais recentes unidades curriculares de inovação pedagógica (UC InovPed) da Universidade do Porto (U.Porto). Estas UC InovPed, formalmente unidades de formação contínua, têm participação de docentes de diferentes unidades orgânicas da universidade, podendo ter ainda a participação de investigadores e técnicos superiores (das unidades orgânicas, de centros de investigação ou ainda dos serviços centrais) (Universidade do Porto, 2017). De acordo com o seu regulamento, as UC InovPed são pensadas para objetivos de aprendizagem transversais, não focados em áreas específicas e podem ser frequentadas também como Opção UP – unidade curricular a selecionar dentro da oferta formativa da universidade – nos ciclos de estudo que incluem esse tipo de unidade curricular no seu plano de estudos. Assim, a unidade curricular BioLab integra no corpo docente elementos de duas unidades orgânicas da U.Porto (Instituto de Ciências Biomédicas Abel Salazar, sede administrativa do curso, e Faculdade de Ciências) e de um centro de investigação (Instituto de Investigação e Inovação em Saúde), assim cobrindo diferentes áreas do conhecimento: ciências da vida, sociologia, comunicação e arte.

Nesta unidade curricular procurou implementar-se uma metodologia de aprendizagem orientada para a investigação (*inquiry-based learning*), seguindo práticas reconhecidas internacionalmente (Fung, 2017). De facto, o BioLab desenrola-se como um CURE (Course-based Undergraduate Research Experience) (Bell *et al.*, 2016), proporcionando aos estudantes o desenvolvimento de um pequeno projeto que, neste caso, integra contribuições a partir de diferentes áreas disciplinares.

Apresentam-se de seguida detalhes da prática desenvolvida, incluindo elementos de avaliação, bem como seus resultados e, finalmente, algumas notas conclusivas.

2. Descrição da prática pedagógica

A unidade curricular BioLab foi concebida no sentido de proporcionar aos estudantes oportunidade de explorar cruzamentos entre ciência, arte e sociedade, tendo funcionado pela primeira vez no segundo semestre do ano letivo 2020-21. Creditada em 6 ECTS, prevê 42 horas de contacto. A tipologia de aulas, incluindo práticas laboratoriais, significou um desafio acrescido pelas restrições da situação pandémica, tendo implicado adaptações relativamente ao planeado aquando da proposta de criação. A descrição que se apresenta integra essas condicionantes.

2.1. Objetivos e público-alvo

Como enfatiza a designação do curso, com a unidade curricular BioLab procura-se promover um ambiente multidisciplinar de aprendizagem que contribua para integrar a aquisição de novos conhecimentos, o desenvolvimento de aptidões laboratoriais, e a experimentação (não só científica como também artística e social) enquanto resposta à resolução de problemas específicos. Neste entendimento, como se apresenta na ficha de unidade curricular institucional (<https://s.up.pt/kw4l>), podem elencar-se como objetivos:

- reconhecer a relevância das abordagens multidisciplinares na produção de conhecimento científico e no diálogo ciência-sociedade;
- promover capacidades de diálogo e colaboração em equipa multidisciplinar;
- desenvolver a capacidade de conceber, planear e implementar um trabalho interdisciplinar de projeto, promovendo o trabalho colaborativo que potencie a complementaridade de conhecimentos e competências;
- adquirir (e/ou aprofundar) competências de trabalho laboratorial no contexto das ciências da vida e das questões em estudo;
- reconhecer a contribuição de práticas artísticas na reflexão em torno de problemas associados às ciências da vida;
- estimular competências que promovam a comunicação do projeto de grupo perante diferentes públicos, pelos seus pares e outros agentes sociais.

Pretende-se proporcionar uma aprendizagem que mobilize as diferentes áreas STEAM, de forma a potenciar o desenvolvimento de competências múltiplas, tais como a autonomia, o pensamento crítico, a inovação, e a capacidade de resolução de problemas, que possa resultar na formação de estudantes com elevada polivalência e capacidade de adaptação. Desse modo, no que respeita a resultados de aprendizagem e competências, ainda conforme a já mencionada ficha de unidade curricular institucional, espera-se que resulte:

- na aquisição de novos conhecimentos, particularmente em áreas académicas fora da formação base de cada estudante, potenciando a complementaridade de conhecimentos individuais;
- no reconhecimento da importância de recrutar conhecimentos multidisciplinares, para identificar questões ou problemas atuais e de interesse transversal que possam ser abordados numa perspetiva biológica e promovendo um olhar de perspetivas de arte/humanidades e das ciências sociais sobre questões de biologia/biotecnologia;
- no reforço de competências de trabalho em grupo no âmbito do projeto, nomeadamente na obtenção de consensos sobre um fluxograma de trabalho que identifique e hierarquize tarefas, estabeleça marcas, aponte responsabilidades individuais, valorize sinergias dentro do grupo, permita diagnosticar constrangimentos, avaliar riscos e proponha soluções alternativas;
- num aumento de competências de comunicação do saber científico que valorize as novas tecnologias de informação e vá de encontro às exigências crescentes de uma sociedade digital;
- num aumento de capacidade de pensamento crítico e criatividade que a arte/humanidades e as ciências sociais podem trazer.

Em 2020-21 o BioLab captou o interesse de 13 estudantes, alguns dos quais frequentaram o curso como Opção UP no âmbito do seu ciclo de estudos; a maior parte fê-lo fora desse âmbito, mesmo nos casos em que o estudante frequentava algum ciclo de estudos (1.º ciclo (nove estudantes), 2.º ciclo (um estudante) ou 3.º ciclo (um estudante)) na U.Porto. O grupo de estudantes foi relativamente homogêneo em termos de área, uma vez que 12 eram ou tinham sido estudantes no campo alargado das ciências da vida.

2.2. Metodologia

A unidade de formação foi concebida para articular formação teórico-prática partindo de sessões expositivas, formação laboratorial supervisionada e o desenvolvimento, em forma autónoma e em grupo, de uma atividade exploratória de projeto num tema que é construído em resposta a um desafio colocado, conforme se detalha adiante.

A formação teórico-prática integra a apresentação de estudos de caso (incluindo projetos artísticos interconectados com as ciências da vida) com desconstrução dos mesmos nas diferentes componentes do método científico (observação; questão; problema; hipóteses;

experiências; conclusões) ou, no caso de projetos artísticos, do enquadramento conceptual, processual e da sua materialização.

As aulas laboratoriais preveem a aplicação/exemplificação de algumas das técnicas laboratoriais abordadas nas sessões teórico-práticas, contemplam oficinas de comunicação e de experimentação artística, mas grande parte deste tempo é destinado ao desenvolvimento de projetos. Tendo em conta a formação dos estudantes, que nesta primeira edição tinham, maioritariamente, alguma formação nas ciências da vida, em 2020-21 não se dedicou tempo à exemplificação de técnicas laboratoriais básicas.

Para partilha de informação e interação entre docente e estudantes, recorreu-se às plataformas de *e-learning* Moodle e Teams, no primeiro caso apenas de forma assíncrona, recursos disponibilizados pela universidade. Este último recurso foi fundamental para o acompanhamento dos projetos e para o envolvimento de toda a equipa de docentes.

Para implementar o processo de trabalho de projeto pelos grupos (tentativamente heterogêneos relativamente à área de formação dos seus elementos) foi apresentado um tema integrador e lançada uma chamada para ideias-respostas; seguiu-se uma chamada para projetos e finalmente para equipas. Após identificação dos projetos de trabalho, os grupos trabalharam de uma forma autónoma, mas orientada no sentido de estimular os conhecimentos multidisciplinares e o pensamento crítico. Organizaram-se tutoriais de grupo regulares. Os diferentes grupos mantiveram um diário e repositório de ideias (na plataforma Teams) acessíveis aos membros do BioLab para discussão. Promoveram-se apresentações regulares pelos grupos (em formato *pitch*) e discussão coletiva. Os passos para o desenvolvimento dos projetos encontram-se esquematizados na figura 1.

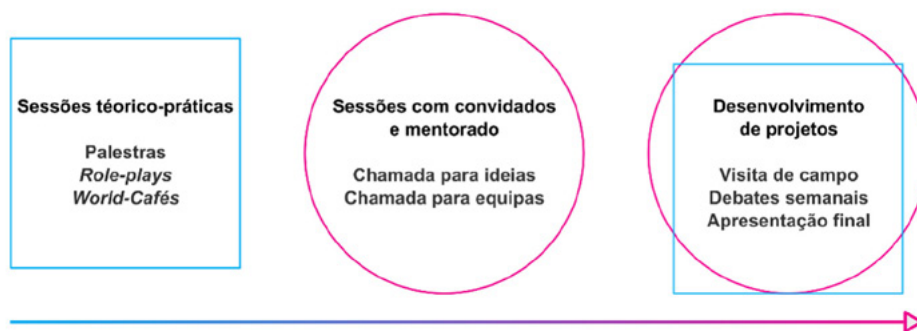


Figura 1: Etapas do BioLab com vista ao desenvolvimento dos projetos

Na edição de 2020-21, o tema apresentado foi: “Vestuário bio-positivo”, tendo o desafio sido colocado em termos de explorar possíveis alterações nos modelos de fabrico de vestuário evitando processos poluentes, que esgotem recursos naturais ou impliquem sacrifício de animais para a sua produção. Assim e em resultado, nesta primeira edição, os estudantes organizaram-se em quatro grupos para desenvolver quatro projetos co-desenhados no BioLab. Estes projetos focaram, designadamente: i) produção de celulose bacteriana a partir de Kombucha e fermentados de frutos; ii) biodegradação de pigmentos de origem biológica incomumente utilizados no têxtil; iii) produção de biopeles/biocouros com base em resíduos ou excedentes alimentares frutícolas; iv) sensibilização para a produção e uso de bioplásticos a partir de materiais baseados em tapioca ou alginato.

Tendo em conta a metodologia de ensino, as componentes de avaliação dos estudantes incluem a participação presencial, um trabalho escrito individual bem como o trabalho de projeto, de grupo, e sua apresentação.

2.3. Avaliação

No sentido de melhor avaliar a prática pedagógica implementada, procurou-se obter *feedback* dos estudantes recorrendo a um inquérito por questionário. O estudo, com a referência 2021/CE/P028(P368/CETI/ICBAS), obteve parecer favorável da Comissão de Ética CHUPorto/ICBAS em 22 de setembro, tendo sido realizado cerca de três meses após conclusão do BioLab. O questionário foi aplicado utilizando a plataforma Google Forms, tendo sido obtido consentimento dos estudantes-participantes à recolha anónima de dados e ao tratamento de forma agregada dos mesmos, para este fim. As perguntas estão organizadas em quatro secções: i) experiência pessoal do BioLab; ii) competências prévias importantes para a frequência; iii) relevância de diferentes componentes pedagógicas; iv) probabilidade de recomendação do curso. A apresentação de resultados segue essa organização. Dada uma população-alvo de 13 participantes, os números que se apresentam, resultantes de um pequeno número de respostas, devem ser lidos com reserva. Para minorar o efeito de respostas discrepantes, sempre que aplicável, privilegiou-se o cálculo de medianas. Foram submetidos 7 questionários, correspondendo a uma taxa de resposta de 54%.

3. Resultados, implicações e recomendações

Considera-se que a unidade curricular BioLab é uma iniciativa bem-sucedida – a percepção por parte da equipa docente, o envolvimento dos estudantes e os resultados do questionário são disso indicadores. Apresenta-se de seguida uma sumarização desses resultados.

Experiência pessoal do BioLab

Questionados a fornecer termos ou expressões que pudessem descrever o BioLab, para além da palavra multidisciplinar (a mais frequente), surgem termos como enriquecedor, inovador, prático, criativo. A caracterização da experiência incluiu uma apreciação do modo de funcionamento balizado entre unidade curricular formal e *biohacking space*, no que diz respeito à expectativa e à experiência tida de facto. As respostas mostraram o desaparecimento da percepção unidade curricular formal 100% com aparecimento da percepção *biohacking space* 100%, entre o antes e o depois da frequência do curso. Adicionalmente, a mediana desta percepção passou de 50% para 75% *biohacking space*.

Competências prévias importantes para a frequência

A capacidade de trabalho em grupo foi ressaltada como muito importante. Tidas como importantes surgem a capacidade de criação de registo de atividades, de investigação autónoma, de discussão e argumentação bem como a de escrita de ensaios. Identificados como pouco importantes em termos de competência prévias surgem os conhecimentos de artes plásticas e de filosofia da ciência. Algo importantes surgem os conhecimentos de fundamentos de química e de biologia e ainda a destreza no trabalho laboratorial. É mencionado que eventuais lacunas individuais são facilmente colmatadas graças ao trabalho em grupo. Por exemplo afirmando que o “trabalho em equipa num grupo multidisciplinar [...] permite mitigar a necessidade de competências prévias nas áreas de biologia e química”. Por outro lado, as competências prévias de trabalho em grupo e as *soft skills* correspondentes são consideradas muito importantes e vistas como podendo ser adquiridas no decurso da participação no BioLab. Como ilustração desta ideia, refere-se: “o biolaboratório treina muito bem essas competências [capacidade de síntese e à-vontade em apresentações orais], permitindo ao aluno melhorá-las”.

Relevância de diferentes componentes pedagógicas

As componentes percecionadas como muito importantes referem-se ao desenvolvimento prático dos projetos, nomeadamente, a assistência dos docentes (de áreas profissionais diversas), o acesso a espaço laboratorial e a própria componente de trabalho em grupo. Foi valorizado o apoio docente no desenvolvimento das ideias, assim como a aprendizagem de “técnicas laboratoriais novas, [de] como seguir um trabalho experimental, [e de] como

trabalhar da melhor forma em equipa”. Também tidas como muito relevantes foram as iniciativas com a participação de pessoas e entidades externas através de palestras e visitas de campo realizadas. De destacar ainda como importantes as palestras sobre abordagens multidisciplinares em biologia e o trabalho autónomo dos estudantes desde a escolha do tema de trabalho até à produção e discussão do mesmo. Foi mencionado que as sessões expositivas permitiram uma “interação contínua”. Foi ainda referido que a interação entre os diferentes participantes permitiu “pensar na comunicação em ciência de uma forma menos tradicional”. Foi sugerido que o período de *brainstorming* na resposta ao desafio (chamada para ideias-resposta) fosse alargado e pudesse acontecer em grupo.

Probabilidade de recomendação do curso

As respostas mostram que os participantes muito provavelmente recomendarão a participação no BioLab aos seus colegas de curso/profissão.

Do exposto, parece claro que o modelo seguido foi bem-sucedido e poderá ser usado noutros contextos de formação onde seja aplicável proporcionar uma aprendizagem baseada em investigação (*inquiry-based learning*), seja em unidades curriculares de ciclos de estudo seja em formações complementares para estudantes, investigadores ou docentes.

4. Conclusões

Procurando concluir, a prática pedagógica cuja implementação aqui se reporta pautou-se por um modelo inclusivo quanto à área disciplinar, tendo proporcionado oportunidade de aprendizagem ativa, para a qual, como mostra o *feedback* pelos estudantes, foram importantes o acompanhamento próximo e regular durante os trabalhos em grupo, o acesso facilitado e acompanhado a espaço laboratorial, o desenvolvimento de *soft skills* associadas à produção e comunicação de conhecimento científico. De alguma forma, estes são sucessos da primeira edição do Biolab. Encontramos ainda alguns constrangimentos. Será pertinente notar, por exemplo, que o tipo de abordagem seguido envolve um rácio docentes-estudantes elevado, o que poderá condicionar a aplicação em contextos que signifiquem grupos de maior dimensão. A situação pandémica teve impacto no modo de funcionamento do curso, impondo sessões remotas no período inicial. Subsiste o desafio de uma apresentação final dos projetos perante uma audiência diversificada e, eventualmente fora do contexto académico. Apesar de não ter sido possível realizar esse plano na primeira edição do BioLab, uma experiência posterior ao curso integrada em evento da Noite Europeia dos Investigadores 2021 poderá ter sido um bom ensaio neste sentido.

5. Referências

Bell, J. K., Eckdahl, T. T., Hecht, D. A., Killion, P. J., Latzer, J., Mans, T. L., . . . Bell, J. E. (2016). CUREs in biochemistry – where we are and where we should go. *Biochemistry and Molecular Biology Education*, 45(1), 7-12. <https://doi.org/10.1002/bmb.20989>

Comissão Europeia. (2017). *Comunicação da Comissão ao Parlamento Europeu, ao Conselho, ao Comité Económico e Social Europeu e ao Comité das Regiões sobre uma nova agenda da UE em prol do ensino superior (COM/2017/0247)*. Consultado em 17.10.2021 em <https://eur-lex.europa.eu/legal-content/PT/ALL/?uri=CELEX:52017DC0247>

Fung, D. (2017). *A connected curriculum for higher education*. UCL Press.

United Nations. (2015). *Transforming our world: the 2030 agenda for sustainable development* (A/RES/70/1). Consultado em 05.08.2021 em <https://sustainabledevelopment.un.org/post2015/transformingourworld/publication>

Universidade do Porto. (2017). *Regulamento do Programa de “Unidade Curricular InovPed” Unidade de Formação Contínua*. Consultado em 17.10.2021 em <https://s.up.pt/n7rx>

COMPLEMENTARY PUBLICATION III

Ramos, P. D., **Marques, F.**, Lopes, M. M., Santos, J. B., Tavares, F., Olsson, I. A. S., & Almeida, M. S. (2022). BioLab – uma proposta para o desenvolvimento de projetos em ambiente multidisciplinar. In João Veloso (Dir.), *Cadernos de Inovação Pedagógica, n.º 2 Unidades Curriculares INOVPED da Universidade do Porto. Transversalidade e criatividade* (pp. 125-139), U.Porto Press

BioLab – uma proposta para o desenvolvimento de projetos em ambiente multidisciplinar

Pedro Dias Ramos(*)§,
Filipe Marques(**)§,
Maria Manuela Lopes(***)
Júlio Borlido Santos(****),
Fernando Tavares(*****),
I. Anna S. Olsson(*****),
& Maria Strecht Almeida(*****)

(*)i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto; ICBAS – Instituto de Ciências Biomédicas
Abel Salazar, Universidade do Porto
(*)pedro.ramos@ibmc.up.pt

(**)i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto; ICBAS – Instituto de Ciências Biomédicas
Abel Salazar, Universidade do Porto
(**) filipe.marques@ibmc.up.pt

(***)i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto; ESE.IPP – Escola Superior de Educação,
Instituto Politécnico do Porto
(***) maria.lopes@i3s.up.pt

(****)i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto
(****) borlido.santos@i3s.up.pt

(*****)FCUP – Faculdade de Ciências, Universidade do Porto
(*****) ftavares@fc.up.pt

(*****)i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto; ICBAS – Instituto de Ciências
Biomédicas Abel Salazar, Universidade do Porto
(*****) olsson@ibmc.up.pt

(*****)ICBAS – Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto
(*****) msalmeida@icbas.up.pt

§ – estes autores contribuíram igualmente para o trabalho

Resumo

Este capítulo descreve a implementação da unidade curricular de inovação pedagógica *Biolaboratório – Projeto de Experimentação Multidisciplinar* (BioLab). A unidade de formação contínua, que teve a sua primeira edição em 2020-21, procura explorar cruzamentos ciência/arte/sociedade. Na prática, desenvolve-se como investigação integrada em unidade curricular e procura tirar partido do ambiente informal que existe em laboratórios comunitários. Pelos projetos desenvolvidos e pelo entusiasmo de todos os envolvidos, e principalmente dos estudantes, considera-se uma iniciativa bem-sucedida.

Abstract

This chapter describes the implementation of the curricular unit of pedagogical innovation *Biolaboratory – Multidisciplinary Experimentation Project* (BioLab). The continuing training unit, launched in 2020-21, aims to explore science/arts/society intersections by applying course-based research experiences inspired by community

laboratories. Awe-inspiring projects and positive feedback from all players involved, notably the students, suggest BioLab was a successful initiative.

Palavras-Chave: STEAM; CURE; laboratório comunitário; trabalho colaborativo.

Keywords: STEAM; CURE; community lab; collaborative work.

1 Introdução

Tem sido defendido que a educação científica deve focar-se hoje numa mudança de STEM (*Science, Technology, Engineering and Mathematics*) para STEAM, que acrescenta as artes a STEM e assim liga a ciência todas as outras disciplinas (European Commission 2015). A UC InovPed *Biolaboratório – Projeto de Experimentação Multidisciplinar* (BioLab) cuja implementação aqui se reporta foi concebida para um contexto STEAM, explorando cruzamentos ciência/arte/sociedade.

Enfatizar a ligação entre investigação e ensino proporcionando uma aprendizagem orientada para a investigação (*inquiry-based learning*) é hoje uma prática reconhecida internacionalmente (Fung 2017). A proposta de currículo conectado (*ibidem*) considera essa ligação como uma das suas vertentes. E a essa acrescenta a das ligações entre disciplinas e entre a academia e o mundo exterior. Na prática, o BioLab desenrola-se como um CURE (*Course-based Undergraduate Research Experience*), promovendo uma aprendizagem pela investigação (Bell, Eckdahl *et al.* 2016), proporcionando aos estudantes o desenvolvimento de um pequeno projeto que, neste caso, tem a particularidade de integrar contribuições a partir de diferentes áreas disciplinares. Nesta UC InovPed procurou ainda integrar-se a informalidade (regulada) de um laboratório comunitário (Scheifele & Burkett 2016).

A educação STEAM visa trazer o “mundo real”, com a sua complexidade, para a sala de aula, ligando diferentes currículos de uma forma que se relacionem tanto com o mundo como entre si. Neste sentido, os processos de pensamento crítico, e métodos artísticos e de design foram introduzidos num processo de aprendizagem ativa centrada no estudante esperando aumentar a motivação, a autoeficácia das aprendizagens e a capacidade de resolução de problemas. Assente na lógica apresentada por Martinez (2017), o sucesso da proposta de unidade curricular dependia do carácter multidisciplinar da equipa (tanto de docentes quanto de estudantes) em criar uma experiência de aprendizagem conjunta. Assim sendo, os papéis não estavam predefinidos *à priori*, havendo lugar a um ajuste permanente na função de cada um no decorrer do desenvolvimento do ambiente de aprendizagem.

No que se segue, a unidade curricular é apresentada em detalhe, incluindo a motivação para a sua criação, a proposta de inovação pedagógica que representa e como funcionou de facto na edição de 2020-21. Uma análise crítica da implementação do curso reflete sobre a prática, discutindo lições aprendidas em termos de desafios e oportunidades, tentando-se, finalmente, algumas notas conclusivas.

2 A unidade curricular *Biolaboratório – Projeto de Experimentação Multidisciplinar*

2.1 Motivação

Proporcionar uma experiência de investigação que integrasse alguma da informalidade dos laboratórios comunitários foi a ideia subjacente à proposta desta UC InovPed. A iniciativa desenvolveu-se a partir de alguns eventos prévios entre os quais se destaca um workshop sobre CRISPR & *Biohacking* com estudantes frequentando o 1.º ano dos programas doutorais BiotechHealth (ICBAS/FFUP), MCBiology (ICBAS/FCUP) e Neurociências (FMUP/ICBAS). O *workshop* que ocorreu em janeiro de 2020 pretendeu associar o tema da edição de genoma e da sua utilização por parte da comunidade de *biohacking* em exercícios de reflexão contando para isso com a elaboração de diferentes perguntas que pretendiam levar à discussão entre os estudantes. Debateram-se a utilização de microrganismos geneticamente modificados, a que classe de segurança estariam associados no caso do seu uso e ainda, quais os mecanismos de regulação e transparência necessários para que estes pudessem ser utilizados numa lógica de laboratório aberto ou *OpenBiolab*. Após 3 horas de discussão, foi ainda possível uma participação remota de um *biohacker* convidado do espaço OLGA (Open Lab Graz Austria) num momento interativo que proporcionou aos estudantes a oportunidade de colocarem questões sobre a atividade desenvolvida pelo mesmo, bem como as condições em que o *hackerspace* se pode organizar. O sucesso do *workshop* que contou com a presença de quase todos os elementos docentes que mais tarde acabariam por integrar o projeto da UC InovPed evidenciou a questão que todos se colocavam: seria possível a fundação de um laboratório aberto de biologia, vulgo *OpenBiolab*, em contexto académico na nossa Universidade? E se sim, estariam estudantes de diferentes áreas do conhecimento abertos à sua criação e integração? O mote estava lançado e a candidatura da InovPed esboçou-se a partir desse momento acrescentando-se à base biológica da UC e à ideia de investigação integrada num curso (internacionalmente descrita como CURE) as componentes de responsabilidade social e expressão artística, bem como o diálogo entre ciência e sociedade na definição dos objetivos para a concretização da UC.

2.2 Proposta pedagógica

No BioLab proporciona-se uma aprendizagem que integra a aquisição de conhecimentos teóricos e práticos no decurso de um processo de experimentação científica, artística e social, desenvolvido num ambiente multidisciplinar, no sentido da resolução de um problema específico. Nesse âmbito, como se apresenta na ficha de unidade curricular institucional (<https://s.up.pt/kw4l>) podem considerar-se como objetivos;

- reconhecer a relevância das abordagens multidisciplinares na produção de conhecimento científico e no diálogo ciência-sociedade;
- promover capacidades de diálogo e colaboração em equipa multidisciplinar;

- desenvolver a capacidade de conceber, planejar e implementar um trabalho interdisciplinar de projeto, promovendo o trabalho colaborativo que potencie a complementaridade de conhecimentos e competências;
- adquirir (e/ou aprofundar) competências de trabalho laboratorial no contexto das ciências da vida e das questões em estudo;
- reconhecer a contribuição de práticas artísticas na reflexão em torno de problemas associados às ciências da vida;
- estimular competências que promovam a comunicação do projeto de grupo perante diferentes públicos, pelos seus pares e outros agentes sociais.

A capacitação dos estudantes enquanto cidadãos dotados de autonomia, pensamento crítico, capacidade de inovação e resolução de problemas é proporcionada por uma experimentação consciente e engajada em problemas sociais mobilizando conceitos e ferramentas das diferentes áreas STEAM. Assim, de forma sucinta, os resultados de aprendizagem e competências que o BioLab propõe alcançar, segundo a já mencionada ficha de unidade curricular institucional, traduzem-se:

- na aquisição de novos conhecimentos, particularmente em áreas académicas fora da formação base de cada estudante, potenciando a complementaridade de conhecimentos individuais;
- no reconhecimento da importância de recrutar conhecimentos multidisciplinares, para identificar questões ou problemas atuais e de interesse transversal que possam ser abordados numa perspetiva biológica e promovendo um olhar de perspetivas de arte/humanidades e das ciências sociais sobre questões de biologia/biotecnologia;
- no reforço de competências de trabalho em grupo no âmbito do projeto, nomeadamente na obtenção de consensos sobre um fluxograma de trabalho que identifique e hierarquize tarefas, estabeleça marcas, aponte responsabilidades individuais, valorize sinergias dentro do grupo, permita diagnosticar constrangimentos, avaliar riscos e proponha soluções alternativas;
- num aumento de competências de comunicação do saber científico que valorize as novas tecnologias de informação e vá de encontro às exigências crescentes de uma sociedade digital;
- num aumento de capacidade de pensamento crítico e criatividade que a arte/humanidades e as ciências sociais podem trazer.

Para implementar o processo de trabalho de projeto pelos grupos, a metodologia baseou-se em diferentes etapas, conforme ilustra o diagrama da figura 1. Destacam-se três vetores principais: i) capacitação dos estudantes para trabalho de projeto multidisciplinar, através de aulas teórico-práticas (que em 2020-21 decorreram *online* dadas as restrições para contenção da pandemia COVID-19) com diferentes temáticas cobertas; ii) envolvimento dos estudantes num tema integrador, incluindo participação de *stakeholders* externos, uma chamada para ideias de cada estudante naquilo que poderia ser um provável projeto dentro do contexto Biolab, e uma chamada para grupos em que cada estudante teve a oportunidade

de integrar os projetos com que mais se identificasse dentro dos selecionados pelos docentes após proposta de cada estudante; iii) ação, isto é, desenvolvimento dos projetos selecionados no contexto de prática laboratorial tutorada, com apresentações regulares semanais para discussão de ideias, uma apresentação final de cada grupo sobre o seu projeto, incluindo quais os sucessos e insucessos bem como reflexão crítica, em formato de ensaio individual, da utilidade do seu projeto a nível científico, artístico e/ou social.

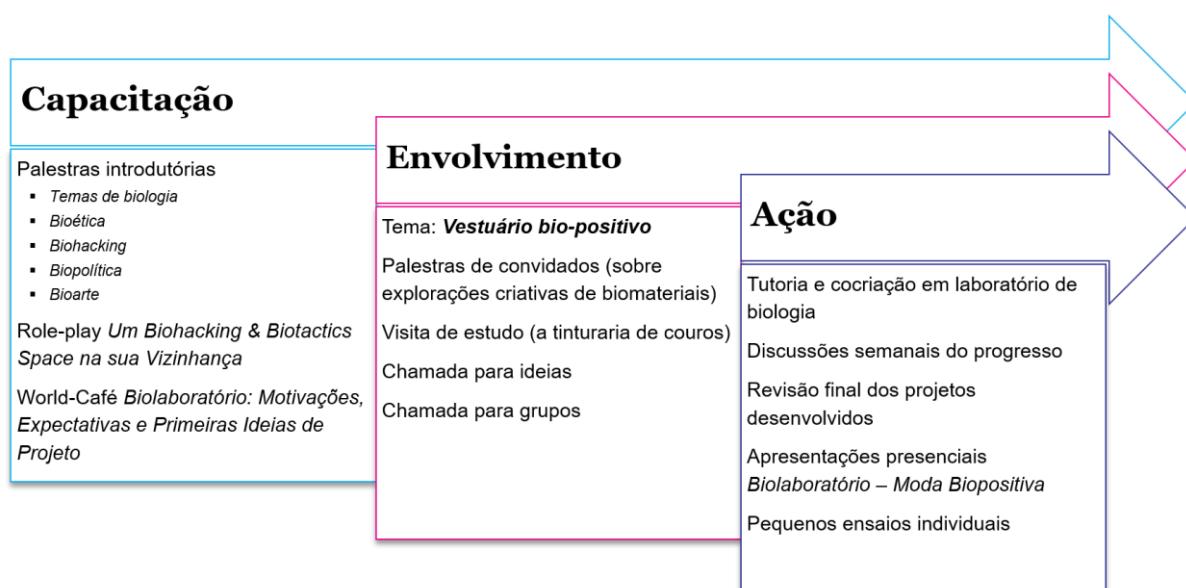


Figura 1. Vetores pedagógicos da unidade curricular BioLab com vista ao desenvolvimento dos projetos pelos grupos de estudantes

As componentes de avaliação dos estudantes incluem a participação presencial, um trabalho escrito individual bem como o trabalho de projeto, de grupo, e sua apresentação.

2.3 Edição de 2020-21

Na sua primeira edição, o BioLab captou o interesse de treze estudantes. Ainda que a alguns estudantes tenham selecionado o Biolab como Opção UP no âmbito do seu ciclo de estudos, a maior parte fê-lo fora desse âmbito. Em termos de distribuição por ciclos de estudo, a grande maioria frequentava uma licenciatura (9), com dois estudantes a frequentarem cursos de pós-graduação (1 mestrado, 1 doutoramento) na U.Porto. A quase totalidade dos estudantes (12) frequentava ou tinha frequentado um ciclo de estudos na área das ciências da vida, contribuindo assim para uma certa homogeneidade do grupo.

“Vestuário bio-positivo” foi o tema proposto aos estudantes, tendo o desafio sido colocado em termos de explorar possíveis alterações nos modelos de fabrico de vestuário evitando processos poluentes, que esgotem recursos naturais ou impliquem sacrifício de animais para a sua produção. Os estudantes organizaram-se em quatro grupos para desenvolver quatro projetos colaborativos que os próprios desenharam no BioLab em resposta ao tema proposto. Estes projetos focaram, designadamente: i) produção de celulose bacteriana, na perspetiva de obtenção de matéria-prima mais sustentável do que o algodão; ii)

biodegradação de pigmentos de origem biológica potencialmente utilizáveis em vestuário, com vista à redução da poluição causada pela indústria têxtil; iii) produção de biopeles/biocouros com base em resíduos ou excedentes alimentares frutícolas, procurando a utilização de subprodutos em matéria-prima para a indústria de moda; iv) sensibilização para a produção e uso de bioplásticos, incorporando a sua utilização em iniciativas de consciencialização sobre o impacto ambiental da moda.

3 Lições aprendidas

A experiência do BioLab foi, em grande medida, uma aventura em que todos os elementos da equipa docente embarcaram com entusiasmo. O trabalho colaborativo, de projeto e com contribuição de diferentes áreas do conhecimento constituiu simultaneamente um estímulo e um desafio. Na perceção dos docentes, foi assim também para os estudantes e isso faz com que esta seja considerada uma experiência bem-sucedida.

Mas vejamos em maior detalhe o que podemos retirar da experiência. Para pensar as lições aprendidas na sua globalidade, haveria que considerar três fontes principais de informação acerca da pertinência da abordagem pedagógica: os estudantes, outras pessoas que colaboraram no curso e ainda *stakeholders* nos projetos desenvolvidos. Centremo-nos em primeiro lugar nos estudantes. Fez-se uma avaliação da primeira edição do BioLab através de um inquérito por questionário implementado após a conclusão do curso. O estudo obteve parecer favorável da Comissão de Ética CHUPorto/ICBAS (referência do projeto: 2021/CE/P28(P368/CETI/ICBAS)). Uma análise detalhada dos resultados desse questionário foi incluída no artigo submetido às Atas do 7.º Congresso Nacional de Práticas Pedagógicas no Ensino Superior, referente à comunicação que tinha sido apresentada na reunião (Marques, Ramos *et al.* 2021). As questões propostas aos estudantes incidiram sobre a experiência subjetiva dos atributos e forma do curso, a classificação da importância relativa das competências prévias e das componentes pedagógicas do curso, e a probabilidade de recomendação do curso. Sete dos treze estudantes responderam ao questionário. Os estudantes sublinharam o carácter multidisciplinar, prático e criativo do curso, o qual evoluiu num sentido que consideraram ser de aproximação a um espaço de experimentação autónoma. De entre as competências prévias sugeridas, os estudantes consideraram muito importante a capacidade de trabalho em grupo e menos importantes, quer o contacto prévio com fundamentos de química e biologia, quer a destreza laboratorial. A este respeito, a instrução entre pares foi apontada como fator mitigante das disparidades de preparação teórica e técnica entre estudantes. Em termos das componentes do curso, os estudantes avaliaram como muito importantes, entre outras, a assistência dos docentes no desenvolvimento dos projetos e o acesso a um laboratório de biologia para realização de experiências. Por outro lado, os estudantes consideraram que haveria margem para melhorar as palestras sobre abordagens multidisciplinares, a discussão semanal do progresso dos projetos e a escolha da abordagem ao tema proposto. Pode ainda fazer-se referência a outras impressões positivas recolhidas informalmente. Entre as pessoas que externamente colaboraram no curso, contaram-se duas designers (uma nacional e outra internacional) que colaboraram pela apresentação em

seminário breve do seu trabalho relacionado com o tema “Vestuário bio-positivo”, bem como um representante de empresa de tinturaria do Grande Porto que proporcionou uma visita de estudo às instalações da empresa e que participou no momento das apresentações dos projetos finais. De alguma forma, este último é um *stakeholder* nos projetos desenvolvidos no curso que se mostrou entusiasta e com interesse em desenvolver colaboração.

Em 2020-21, pelas restrições da pandemia, não foi possível concretizar a planeada exposição pública final dos projetos desenvolvidos no curso. Já após a conclusão do curso, tendo sido lançado o repto para participação na *Noite Europeia dos Investigadores*, no i3S, em setembro de 2021, nove estudantes mobilizaram-se para ajudar a dinamizar uma pequena banca de demonstração dos projetos dos seus grupos. O sucesso da representação do BioLab neste evento ilustra bem a qualidade dos resultados do curso, sendo que a participação voluntária dos estudantes após conclusão do curso também parece indicar que a experiência terá sido positiva.

4 Considerações finais

O Biolab apresentou-se como unidade curricular formal pioneira na abordagem multidisciplinar das ciências da vida na U.Porto. Com uma abrangência no conjunto alargado das áreas STEAM, o BioLab procurou transpor as (boas) práticas de experimentação em laboratórios comunitários para um ambiente académico, multidisciplinar, de investigação baseada em projetos de índole científica e artística, respondendo a preocupações societais. O BioLab envolveu um conjunto alargado de docentes, estudantes, técnicos e convidados, representando uma diversidade considerável de áreas do conhecimento, que cocriaram o próprio processo pedagógico, através das contribuições no sentido de uma capacitação, envolvimento ou desenvolvimento efetivo dos projetos.

A elevada satisfação de todos os intervenientes, quer estudantes quer docentes leva-nos a considerar que o BioLab é uma experiência bem-sucedida. Apesar dessa perceção, reconhece-se que existe espaço para evoluir, nomeadamente através da integração de estudantes de áreas do conhecimento ainda mais diversas além das ciências da vida, num plano multidisciplinar, e o desejável envolvimento de um conjunto mais alargado de possíveis *stakeholders* no contexto do Biolab. A exploração científica e artística de temas com impacto social continuará a ser o *motto* do Biolab. Esta UC InovPed foi essencialmente uma experiência de cocriação entre estudantes e docentes, carácter que se pretende enfatizar em próximas edições.

Agradecimentos. Os autores agradecem a Carla Oliveira (ICBAS) o apoio técnico na edição 2020-21.

Referências

Bell, J. K., Eckdahl, T. T., Hecht, D. A., Killion, P. J., Latzer, J., Mans, T. L., . . . Bell, J. E. (2016). CUREs in biochemistry – where we are and where we should go. *Biochemistry and Molecular Biology Education*, 45(1), 7-12. <https://doi.org/10.1002/bmb.20989>

- European Commission. (2015). *Science education for responsible citizenship: report to the European Commission of the expert group on science education*. Publications Office. <https://doi.org/doi/10.2777/13004>
- Fung, D. (2017). *A connected curriculum for higher education*. UCL Press. <https://doi.org/10.14324/111.9781911576358>
- Marques, F., Ramos, P. D., Lopes, M. M., Santos, J. B., Tavares, F., Olsson, I. A. S., & Almeida, M. S. (2021). *BioLab – um espaço de experimentação e cruzamento disciplinar [Resumo de apresentação em conferência]* 7.º Congresso Nacional de Práticas Pedagógicas no Ensino Superior, Aveiro. https://cnappes.org/files/2021/08/Livro_Resumos_CNaPPES.pdf
- Martinez, J. E. (2017). *The Search for Method in STEAM Education*. Palgrave Macmillan, Cham. <https://doi.org/10.1007/978-3-319-55822>
- Scheifele, L. Z., & Burkett, T. (2016). The first three years of a community lab: lessons learned and ways forward. *Journal of Microbiology & Biology Education*, 17(1), 81-85. <https://doi.org/10.1128/jmbe.v17i1.1013>

SEDE ADMINISTRATIVA INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
FACULDADE DE FARMÁCIA

