UMinho | 2016 Márcia Linhares Couto **Construction and validation of *Escherichia coli* mutants to improve the curcumin production by an engineered strain**

Universidade do Minho
Escola de Engenharia

Márcia Regina Linhares Couto

Construction and validation of *Escherichia coli* mutants to improve the curcumin production by an engineered strain

Outubro de 2016

Universidade do Minho
Escola de Engenharia

Márcia Regina Linhares Couto

Construction and validation of *Escherichia coli* mutants to improve the curcumin production by an engineered strain

Dissertação de Mestrado
Mestrado em Bioengenharia

Trabalho efetuado sob a orientação de
Professora Doutora Lígia Raquel Marona Rodrigues
Doutora Joana Lúcia de Lima Correia Rodrigues

Outubro de 2016

Universidade do Minho
Escola de Engenharia

Márcia Regina Linhares Couto

Construction and validation of *Escherichia coli* mutants to improve the curcumin production by an engineered strain

Dissertação de Mestrado
Mestrado em Bioengenharia

Trabalho efetuado sob a orientação de
Professora Doutora Lígia Raquel Marona Rodrigues
Doutora Joana Lúcia de Lima Correia Rodrigues

Outubro de 2016

DECLARAÇÃO

Nome: Márcia Regina Linhares Couto

Endereço eletrónico: marciacouto93@gmail.pt

Cartão de Cidadão: 14381053

Título da dissertação: Construction and validation of *Escherichia coli* mutants to improve the curcumin production by an engineered strain

Orientadores:

Professora Doutora Lígia Raquel Marona Rodrigues

Doutora Joana Lúcia de Lima Correia Rodrigues

Ano de conclusão: 2016

Mestrado em Bioengenharia

É AUTORIZADA A REPRODUÇÃO INTEGRAL DESTA DISSERTAÇÃO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE.

Universidade do Minho, ______/______/___________

Assinatura:

“*Be an enzyme - a catalyst for change.*”

Pierre Omidyar

AGRADECIMENTOS

No final de mais uma etapa queria agradecer a todos os que de alguma forma contribuíram para a realização da presente tese, aos que me acompanharam, e aos que me ampararam assim que precisava.

Este trabalho foi orientado pela professora doutora Lígia Rodrigues e pela doutora Joana Rodrigues às quais quero expressar o meu maior apreço pelos sábios conselhos e pela dedicação. A sua supervisão, as suas sugestões e também o seu encorajamento foram o que mais contribuiu para o sucesso deste trabalho porque me fizeram sempre dar o melhor de mim.

Agradeço imenso também à Aline Barros, pela ajuda e pelos muitos ensinamentos no HPLC, que me deu de forma carinhosa e prestável.

Aos meus companheiros do laboratório Plataforma de Biologia Molecular e Sintética, que no dia-a-dia no laboratório me apoiaram, ensinaram e suportaram as minhas cantorias, Rita, Débora, Carla, Vanessa Gonçalves, Fernando, Ana Silva, Adriana, Rodrigo, Diana, e à Joana novamente, obrigada!

Muito obrigada às companheiras do laboratório vizinho de Bioprocessos, em especial à Joana Azevedo, pela simpatia com que sempre me auxiliou quando preparei soluções, mas também à Sofia que pacientemente discutiu várias vezes os protocolos de deleção comigo.

Um profundo agradecimento ao Hélder Gonçalves, à Raquel Gonçalves, à Ana Brandão, e ao Henrique Ribeiro, pelas suas amizades verdadeiras e incondicionais, porque todos os dias me prestaram momentos únicos de alegria mesmo depois de um dia cansativo.

Não posso deixar de agradecer às pessoas que fizeram tanto esforço para que eu pudesse chegar aqui. Estou-vos eternamente grata pais Maria Amélia Carvalho e José Couto porque sempre acreditaram em mim.

ABSTRACT

Curcumin has been reported for its beneficial therapeutic properties including as anti-cancer agent. However, it has poor bioavailability and it is quickly metabolized in the human body, implying a repetitive oral administration if a therapeutic effect is envisaged. Besides, its extraction from plants is very expensive. For these reasons, the use of microorganisms to produce it on large scale and with greater yields constitutes an interesting alternative. With this aim, *Escherichia coli* K-12 MG1655 (DE3) was previously engineered with three enzymatic steps (4-coumarate-CoA ligase, diketide-CoA synthase and curcumin synthase 1) that catalyze the production of curcumin from ferulic acid. In the present study, the optimal strain, operational conditions and media composition for the production of curcumin by *E. coli* harboring the artificial biosynthetic pathway were established. Previously, a standard two-step fermentation strategy (LB+M9 minimal medium) was used. Although feasible at the laboratory scale, the biomass separation is much more difficult, laborious and expensive in large-scale fermentations. Therefore, herein a single medium formulation more suitable for the production of curcumin at an industrial set-up was implemented. MOPS minimal medium, TB and LB were evaluated. Using the optimized conditions, the curcumin concentration obtained in this study was the highest reported to be produced by a heterologous organism, 686.7±59.7 µM in TB (43 h) and 822.6±28.1 µM in LB+M9 (63 h). These results were obtained using *E. coli* BL21 (DE3) that was identified as the best producer since it produced 3.7 times more curcumin than *E. coli* K-12 MG1655 (DE3). Moreover, curcumin toxicity against *E. coli* cells was evaluated. The tests performed showed that curcumin concentrations above 400 µM influence negatively the *E. coli* cells growth. Furthermore, one of the purposes of the current work was to construct and validate several *E. coli* mutants (e.g. Δ*fumA,fumB,fumC*) previously identified by an *in silico* approach as the most promising towards an increased production of curcumin from ferulic acid. The deletion of *fumB* gene from *E. coli* K-12 MG1655 (DE3) genome was accomplished and it resulted in a faster curcumin production in the initial 21 h, but after 63 h, the curcumin production by this mutant was 2.6 times lower as compared to the 'original' strain (i.e. the strain harboring the curcuminoids biosynthetic pathway but with no gene knockout). The same deletion in *E. coli* BL21 (DE3) genome resulted in a more significant decrease in curcumin production. In the future, the triple knock-out (Δ*fumA,fumB,fumC*) should be constructed to evaluate if curcumin production can indeed be improved as predicted *in silico*.

Keywords: curcumin production; genetic engineering; *E. coli* mutants; gene deletion; CRISPR-Cas9 system

RESUMO

A curcumina tem sido reportada pelas suas propriedades benéficas incluindo como agente anticancerígeno. Apesar da curcumina apresentar um alto potencial terapêutico, tem uma baixa biodisponibilidade e é rapidamente metabolizada no organismo humano, o que implica uma repetitiva administração oral para atingir o efeito terapêutico pretendido. Além disso, a sua extração a partir das plantas é muito dispendiosa. Por estas razões, o uso de microrganismos para a produzir em larga escala e com melhores rendimentos constitui uma alternativa atraente. Neste sentido, *Escherichia coli* K-12 MG1655 (DE3) foi previamente geneticamente modificada adicionando três reações enzimáticas (4-cumarato-CoA ligase, dicetídeo-CoA sintase e curcumina sintase 1) que catalisam a produção de curcumina a partir do ácido ferúlico. No presente trabalho foi estabelecida a estirpe*,* as condições operacionais e a composição de meio ótimas para a produção de curcumina por *E. coli* contendo a via biossintética artificial. Anteriormente, utilizou-se uma estratégia comum de dois passos (LB+meio mínimo M9). Apesar de ser praticável numa escala laboratorial, a separação de biomassa é muito mais difícil, trabalhosa e dispendiosa numa fermentação em grande escala. Assim, foi implementada uma única formulação de meio mais adequada para a produção de curcumina a nível industrial. O meio mínimo MOPS, TB e LB foram avaliados. Usando as condições otimizadas, a concentração de curcumina produzida neste estudo foi mais elevada do que as previamente descritas na literatura, 686,7±59,7 μM em TB (43 h) e 822,6±28,1 μM em LB+M9 (63 h). Estes resultados foram obtidos usando a estirpe *E. coli* BL21 (DE3) que foi a identificada como melhor produtora após produzir 3,7 vezes mais curcumina do que a *E. coli* K-12 MG1655 (DE3). Além disso, a toxicidade da curcumina para células de *E. coli* foi avaliada. Os resultados mostraram que concentrações de curcumina acima de 400 μM influenciam negativamente o crescimento de *E. coli*. Adicionalmente, um dos objetivos do presente trabalho foi construir e validar vários mutantes de *E. coli* (p. ex., Δ*fumA,fumB,fumC*) que foram previamente identificados *in silico* como os mais promissores no sentido de aumentar a produção de curcumina a partir do ácido ferúlico. A deleção do gene *fumB* do genoma de *E. coli* K-12 MG1655 (DE3) foi efetuada e resultou numa produção mais rápida de curcumina nas 21 h iniciais, mas após 63 h, a produção de curcumina usando este mutante foi 2,6 vezes mais baixa do que a da estirpe que lhe deu origem. A mesma deleção no genoma de *E. coli* BL21 (DE3) resultou num decréscimo ainda mais significativo na produção de curcumina. No futuro, é necessário ainda construir o mutante triplo (Δ*fumA,fumB,fumC*) para avaliar se a produção de curcumina é efetivamente aumentada como previsto *in silico*.

Palavras chave: produção de curcumina; engenharia genética; mutantes de *E. coli*; deleção de genes; sistema CRISPR-Cas9

SCIENTIFIC OUTPUT

PAPERS SUBMITTED TO PEER REVIEWED JOURNALS:

Joana L Rodrigues, **Márcia R Couto**, Rafael G Araújo, Kristala LJ Prather, Leon Kluskens, Lígia R Rodrigues. Hydroxycinnamic Acids and Curcumin Production in *Escherichia coli* using heat shock promoters. *(Submitted in October 2016)*

PAPERS IN PREPARATION FOR SUBMISSION TO PEER REVIEWED JOURNALS:

Márcia R Couto, Joana L Rodrigues, Lígia R Rodrigues. Optimization of heterologous curcumin production titers by *Escherichia coli* regarding the fermentation conditions.

INTERNATIONAL CONFERENCE POSTERS:

Márcia R Couto, Joana L Rodrigues and Lígia R Rodrigues. Enhancing curcuminoids production using *E. coli* engineered strains. International Congress “Bio.Iberoamérica 2016”, Salamanca, June 2016.

CONTENTS

Abstract vii
Resumo ix
Scientific Output xi
List of Figures xv
List of Tables xvii
Abbreviations xix
Scope xxi
1. State-of-the-art 1
  1.1 Curcuminoids 1
    1.1.1 Curcuminoids relevance and therapeutic use 1
    1.1.2 Cancer and bacterial therapies 2
    1.1.3 Biosynthetic pathway of curcuminoids in *Curcuma longa* 4
  1.2 Heterologous production of curcumin 6
    1.2.1 Possible hosts – advantages and drawbacks 7
    1.2.2 Curcumin production in *Escherichia coli* 9
    1.2.3 Optimization of curcumin production 12
2. Objectives 17
3. Materials and Methods 19
  3.1 Strains and plasmids 19
  3.2 Chemicals and culture media 19
  3.3 Preparation of electrocompetent cells and transformation step 21
  3.4 Preparation of chemical competent cells and transformation step 21
  3.5 Optimization of curcumin production in *E. coli* using an artificial biosynthetic pathway 22
    3.5.1 Curcumin production 22
    3.5.2 Curcumin extraction and samples preparation 24
    3.5.3 Substrate and product quantification by HPLC 24
  3.6 Curcumin toxicity assay 25
  3.7 Gene deletion in *E. coli* 26

4. Results and Discussion 33
4.1 Curcumin production optimization 33
4.1.1 Evaluation of the performance of different *E. coli* strains 33
4.1.2 Establishment of the optimal optical density for induction of BL21 (DE3) in the process using two culture media 35
4.1.3 Evaluation of different IPTG concentrations for induction of the protein expression 37
4.1.4 Optimization of operational conditions 38
4.1.5 Assessment of different optical densities for the induction in TB 42
4.1.6 Improvement of curcumin production in TB by adding extra carbon sources 42
4.2 Curcumin toxicity assay 44
4.3 Confirmation of the *gnd* deletion for *E. coli* mutant validation 46
4.4 Construction of *E. coli* mutants using the CRISPR-Cas9 system 47
4.5 Curcumin production by engineered *E. coli* mutants 51
5. Conclusions 53
6. Future perspectives 55
7. Bibliography 57
Appendix I – Gene deletion by Lambda Red–Mediated, Linear DNA–Based Deletion Method 63
Primers 63
Deletion cassette construction 64
Deletion protocol 65
Appendix II – Gene deletion by the Landing Pad Technology 69
Primers 70
Landing pad construction 70
Deletion protocol 71
Appendix III– Results from recombineering strategies 73
AIII.1 *fumA*, *fumB* and *fumC* deletion through Cm cassette recombineering 73
AIII.1 *fumA* and *fumC* deletion through recombineering of a landing pad 77

LIST OF FIGURES

Figure 1. Curcuminoid biosynthesis mechanism in *Curcuma longa* 5

Figure 2. Possible curcumin biosynthesis pathways with plant genes to construct engineered *E. coli* strains 8

Figure 3. Part of the metabolic model of *E. coli* K-12 MG1655 indicating the reactions altered by *gnd*, *fumA*, *fumB* and *fumC* deletions 15

Figure 4. General scheme of the *in vivo* curcumin production fermentative process 23

Figure 5. Curcumin extraction with ethyl acetate. 24

Figure 6. 96-wells plate design for curcumin toxicity assay 26

Figure 7. Curcumin production over time using *E. coli* K-12 MG1655 (DE3), *E. coli* BL21 (DE3) and *E. coli* JM109 (DE3) harboring the heterologous biosynthetic pathway 34

Figure 8. Effect of IPTG (isopropyl β-D-1-thiogalactopyranoside) induction at different optical density (600 nm) values of *E. coli* BL21 (DE3) in the production of curcumin 36

Figure 9. Curcumin production by *E. coli* BL21 (DE3) using LB (lysogeny broth) and M9 media and different IPTG (Isopropyl β-D-1-thiogalactopyranoside) concentrations 37

Figure 10. Curcumin production over time in different media compositions by *E. coli* BL21 (DE3) 39

Figure 11. Curcumin production in lysogeny broth (LB) agar supplemented with the required antibiotics and 2 mM of ferulic acid 41

Figure 12. Curcumin production over time by *E. coli* BL21 (DE3) in TB (terrific broth), induced at different optical densities (0.4, 0.6 and 0.8) 42

Figure 13. Curcumin production over time by *E. coli* BL21 (DE3) in terrific broth (TB) supplemented with different carbon sources 43

Figure 14. Wild-type *E. coli* BL21 (DE3) (A) and the one harboring the artificial biosynthetic pathway (B) growth during 8 h in the presence of different curcumin concentrations 45

Figure 15. Agarose gel (0.7% (w/v)) to confirm the deletion of *gnd* from *E. coli* K-12 MG1655 (DE3) .. 46

Figure 16. Agarose gels (1% (w/v)) showing the amplification of the part 1 (2848 bp) and part 2 (4414 bp) of the pKDsgRNA to construct pKDsgRNA_*fumB*, pKDsgRNA_*fumAC*, pKDsgRNA_*ccmA* and pKDsgRNA_*argO* by CPEC (circular polymerase extension cloning) 47

Figure 17. Agarose gel (1% (w/v)) with colony PCR results to confirm the *fumB* deletion in *E. coli* K-12 MG1655 DE3 48

Figure 18. Agarose gel 1% with colony PCR results to confirm the *fumB* deletion in *E. coli* BL21 (DE3). 49
Figure 19. Agarose gel 1% (w/v) showing the colony PCR results using the primers fumA_D and fumAC_E to confirm the deletion of *fumA* and *fumC*........ 50
Figure 20. Curcumin production by *E. coli* K-12 MG1655 (DE3) and its derived mutants, *E. coli* K-12 MG1655 (DE3) Δ*fumB* and *E. coli* K-12 MG1655 (DE3) Δ*gnd* 51
Figure 21. Curcumin production by *E. coli* BL21 (DE3) and *E. coli* BL21 (DE3) Δ*fumB*. LB (lysogeny broth) and M9 media combination was used 52

Figure A 1. Agarose gel (0.7% (w/v)) to confirm the presence of the deletion cassette with Cm resistance from pSG76-CS plasmid with primers FumA_AB and FumA_C (A) FumB_AB, FumB_C (B) and FumC_AB, FumC_C (C) 73
Figure A 2. Agarose gel (0.7% (w/v)) to confirm the insertion of the *fumA* deletion cassete with Cm resistance after electroporation 74
Figure A 3. Agarose gel (0.7% (w/v)) to confirm the correct insertion of the Cm resistance fragment on *E. coli* K-12 MG1655 (DE3) through colony PCR using primers fumA_D and Cm_Rv 74
Figure A 4. Agarose gel (1.5% (w/v)) to confirm the elimination of the Cm resistance fragment from the genome of *E. coli* K-12 MG1655 (DE3) using a PCR with primers FumA_D and FumA_E (Up) and using primers Cm_Fw and Cm_Rv (Down) 75
Figure A 5. Agarose gel 0.7% showing pKD46 digestion with *Bam*HI and *Nco*I (1) and pSTKST digestion with *Nde*I and *Sph*I (2 and 3) 76
Figure A 6. Alignment between the expected pSG76-CS plasmid sequence (pSG76-CS) and the sequences obtained by sequencing using two different primers 76
Figure A 7. Agarose gel (1% (w/v)) to visualize the constructed landing pad for *fumA* and *fumC* deletion. 77
Figure A 8. Agarose gel (1% (w/v)) from colony PCR used to confirm if the deletion cassette was integrated into *E. coli* K-12 MG1655 (DE3) genome 77

LIST OF TABLES

Table 1. Biological systems and conditions used for curcumin production in literature 11

Table 2. Strains used in this study and relevant genotype information 19

Table 3. Primers used for scarless deletion using the CRISPR-Cas9 system 27

Table 4. Oligonucleotides used for scarless deletion using the CRISPR-Cas9 system 28

Table 5. PCR conditions to amplify the two parts of the pKDsgRNA plasmid for *fumB* deletion using Phusion HiFi DNA polymerase 29

Table 6. Setting up PCR conditions for the construction of the pKDsgRNA_fumB using CPEC 30

Table 7. PCR conditions for colony PCR to verify the deletion of *fumB* using primers fumB_D and fumB_E and KAPA *Taq* DNA Polymerase 31

Table A 1. Primers used in the construction and confirmation of the deletion cassette with Cm resistance 64

Table A 2. PCR conditions for the generation of the AB primer, using A and B primers and KAPA HiFi DNA Polymerase 65

Table A 3. Second PCR conditions for the deletion cassette construction with primers AB, A and C and KAPA HiFi DNA Polymerase 65

Table A 4. PCR conditions for colony PCR to confirm the insertion of the cassette Cm using primers D and E and KAPA *Taq* DNA polymerase 66

Table A 5. PCR conditions for colony PCR to confirm the insertion of the cassette Cm using primers D and Cm_Rv and KAPA *Taq* DNA polymerase 67

Table A 6. PCR conditions for colony PCR to confirm the insertion of the cassette Cm using primers Cm_Fw and E and KAPA *Taq* DNA polymerase 67

Table A 7. Primers used for fumA and fumC genes deletion using the landing pad system. 70

Table A 8. PCR conditions for primer AB formation to delete *fumA* and *fumC*..................................... 70

Table A 9. PCR conditions for landing pad construction to delete *fumA* and *fumC*............................... 71

Table A 10. Colony PCR conditions for confirming landing pad integration using KAPA *Taq* Polymerase. 72

ABBREVIATIONS

4CL	4-Coumarate-CoA ligase
6PDH	Glucose-6-phosphate dehydrogenase
6PGDH	6-Phosphogluconate dehydrogenase)
ACA	Acetyl-CoA
ACC	Acetyl-CoA carboxylase
Amp	Ampicillin
*At*4CL1	4CL from *Arabidopsis thaliana*
aTc	Anhydrotetracycline (hydrochloride)
C3H	4-Coumarate 3-hydroxylase
C4H	Cinnamate-4-hydroxylase
CCoAOMT	Caffeoyl-CoA*O*-methyltransferase
CCR	Carbon catabolite repression
Cm	Chloramphenicol
CoA	Coenzyme A
COMT	Caffeic acid 3-O-methyltransferase
CPEC	Circular Polymerase Extension Cloning
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CS3H	*p*-Coumaroyl 5-*O*-shikimate 3-hydroxylase
CST	*p*-Coumaroyl shikimate transferase
CURS	Curcumin synthase
CUS	Curcuminoid synthase
DCS	Diketide-CoA synthase
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
GRAS	Generally Recognized as Safe
HPLC	High-performance liquid chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
Kan	Kanamycin
KDPG	2-Keto-3-deoxy 6-phosphogluconate

LB	Lysogeny broth
*Le*4CL1	4CL from *Lithospermum erythrorhizon*
MOPS	Morpholinepropanesulfonic acid
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
OD	Optical density
PAL	Phenylalanine ammonia-lyase
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
Phe	Phenylalanine
PKS	Polyketide synthase
PPP	Pentose phosphate pathway
Rpm	Revolutions per minute
SOC	Super optimal broth with catabolite repression
Spec	Spectinomycin
STS	Steroid sulfatase
TAL	Tyrosine ammonia-lyase
TB	Terrific broth
TCA	Tricarboxylic acid
TFA	Trifluoroacetic acid
Tyr	Tyrosine
*Zo*CURS	CURS from *Zingiber officinale*

SCOPE

Breast cancer is a major health problem nowadays, being the most common cancer type among women worldwide and the second most common cancer in both women and men after lung cancer. Statistics show that breast cancer is the first leading cause of cancer death in women, being responsible for 521,900 deaths worldwide in 2012 (Torre *et al.*, 2015). Along with the mortality, cancer in general has several consequences on human life and this particular kind of cancer has a great impact on women's self-esteem.

With the aim to reduce the significant numbers of mortality, great efforts through the search of new compounds and ways to produce them have been conducted. Curcumin has been exhaustively reported for its several interesting properties, including as a chemopreventive and anti-angiogenesis agent (Sun *et al.*, 2004), which makes this compound a potential weapon against cancer that is worth exploring. Indeed, this possibility is being studied, however curcumin extraction from plants is inefficient and it is hard to synthetize chemically, thus greatly limiting its use for this purpose. Hence, using microorganisms to produce curcumin became an attractive and innovative solution. The current work was developed under this scope, aiming at an optimization of the approaches developed by our research group (Rodrigues *et al.*, 2015c).

1. STATE-OF-THE-ART

1.1 Curcuminoids

Curcuminoids are hydrophobic polyphenols naturally found on the turmeric rhizome (*Curcuma longa* Linn) that are responsible for its yellow color. Specifically, these compounds are diarylheptanoids (C6-C7-C6) and may represent 2 to 9% (w/w) of turmeric, depending on the varieties of *C. longa* (Jayaprakasha *et al.*, 2002). Curcuminoids have also been isolated from other species like *Curcuma aromatica*, *Curcuma phaeocaulis*, *Curcuma mangga, Curcuma zedoaria, Costus speciosus, Curcuma xanthorrhiza, Etlingera elatior* and *Zingiber cassumunar* (Tohda *et al.*, 2006; Aggarwal *et al.*, 2007), but *C. longa* is the most studied and the main source from which these compounds are extracted.

Curcumin (diferuloylmethane) is the most abundant curcuminoid on the rhizome of *Curcuma* followed by demethoxycurcumin and bisdemethoxycurcumin. Curcumin was isolated for the first time in 1815 (Vogel A, 1815). Its structure as diferuloylmethane was confirmed and synthesized in 1913 by Lampe and Milobedzka (Lampe & Milobedzka, 1913; Goel *et al.*, 2008). Other turmeric curcuminoids include dicinnamoylmethane (Katsuyama *et al.*, 2007), 1-hydroxy-1,7-bis(4-hydroxy-3-methoxyphenyl)-6-heptene-3,5-dione, 1,7-bis(4-hydroxyphenyl)-1-heptene-3,5-dione, 1,7-bis(4-hydroxyphenyl)-1,4,6-heptatrien-3-one and 1,5-bis(4-hydroxy-3-methoxyphenyl)-1,4-pentadien-3-one (Park & Kim, 2002). The commercial curcumin is a semi-purified extract containing a mixture of the three main curcuminoids, namely curcumin (75–81%), demethoxycurcumin (15–19%) and bisdemethoxycurcumin (2.2–6.6%) (Jayaprakasha *et al.*, 2005; Bagchi, 2012).

1.1.1 Curcuminoids relevance and therapeutic use

Although it is clear that *Curcuma* has an industrial relevance as a condiment and as a dye, and besides being used since ancient times on Asian medicine, it was only on the middle of the 20th century that its use was reported for the treatment of human diseases (Oppenheimer, 1937; Gupta *et al.*, 2012).

Indeed, in the past decades, curcuminoids have been extensively reported for their benefits as anti-cancer, anti-inflammatory, antioxidant, anti-septic, anti-amyloid agents, anti-HIV, anti-depressive and chemopreventive (Aggarwal *et al.*, 2003; Chattopadhyay *et al.*, 2004; Goel *et al.*, 2008; Gupta *et al.*, 2012; Park & Kim, 2002; Ramírez *et al.*, 2005; Ringman *et al.*, 2005; Rodrigues *et al.*, 2015b; Wilken *et al.*, 2011). Several studies suggest their ability to suppress the proliferation of a broad variety of tumor cells (Kuo *et al.*, 1996; Odot *et al.*, 2004; Goel *et al.*, 2008). The targets and mechanisms underlying the curcumin activities have been widely discussed (Aggarwal *et al.*, 2007; Ravindran *et al.*, 2009; Zhang *et al.*, 2014; Jordan *et al.*, 2016; Kumar *et al.*, 2016; Neutzling *et al.*, 2016). Furthermore, to date there is no evidence of curcumin's toxicity on clinical trials (Aggarwal *et al.*, 2003; Lao *et al.*, 2006; Anand *et al.*, 2007; Gupta *et al.*, 2012) and an oral dosage of 10 to 12 g per day of curcumin was found to be safe (Lao *et al.*, 2006; Aggarwal *et al.*, 2007). Besides, United States Food and Drug Administration (FDA) has classified the turmeric among substances Generally Recognized as Safe (GRAS). Outside the gastrointestinal tract, the concentration of curcumin found after repetitive oral doses is on the nanomolar range due to the extensive metabolism in liver and intestine, thus toxic doses are not achieved (Burgos-Morón *et al.*, 2010). However, the low natural abundance of curcuminoids and their fast elimination from the body upon ingestion represents a major limitation for their use as oral therapeutic agents. Indeed, the poor absorption and rapid elimination are the main reasons contributing to the observed low plasma and tissues levels of curcumin (Ahmad, 2007; Anand *et al.*, 2007).

For this reason, several approaches have been proposed towards the enhancement of curcumin bioavailability, such as the administration of heat-solubilized curcumin (Kurien & Scofield, 2009); the use of nanoparticles (Mancarella *et al.*, 2015); micelles; phospholipid complexes or adjuvants (like piperine which increases 2000% the curcumin bioavailability) (Anand *et al.*, 2007; Prasad *et al.*, 2014); as well as its microencapsulation in yeast cells (Paramera *et al.*, 2011). Also, the use of resveratrol along with curcumin has been found to create synergistic effects to inhibit colon cancer, suggesting that despite being known as anticancer, it may act as a stabilizing agent for curcumin (Majumdar *et al.*, 2009; Shindikar *et al.*, 2016). Nevertheless, the results obtained so far were not sufficiently promising to enable the use of these strategies in therapeutics (Anand *et al.*, 2007). Therefore, alternatives are ought to be developed in order to overcome such limitations, including the heterologous production of curcumin and the administration of live recombinant bacteria.

1.1.2 Cancer and bacterial therapies

Cancer is a growing worldwide problem and therefore its treatment is a global priority. Breast cancer is the one which causes more mortality among women presenting the highest incidence (International Agency for Research on Cancer, 2012). Several strategies have been pursued to find new biomarkers, drugs, treatments and to overcome the well-known side effects of common therapies. One of the most recently discussed approaches is using bacteria as a specific drug delivery system.

Targeting cancer cells with bacteria has been widely studied (Fox *et al.*, 1996; Lemmon *et al.*, 1997; Zhao *et al.*, 2005; Stritzker *et al.*, 2007; Gardlik & Fruehauf, 2010; Massa *et al.*, 2013; Nuyts *et al.*, 2014; Felgner *et al.*, 2016). The idea is not new, but it was forgotten for a while because of the inability to control bacterial infections or therapy side effects. The exact mode of action of the bacteria in tumors is not fully understood but they might enhance the immune response against the cancer cells by, for example, activation of natural killer cells (Felgner *et al.*, 2016). The use of bacteria systemically as anticancer agents is now undergoing renaissance and the enhanced attention on bacterial tumor therapy is due to the progress in biomedical research, as well as the increasing knowledge on bacterial behavior and genetic engineering (Felgner *et al.*, 2016). The main advantage of using bacteria as drug vectors is its specificity to its preferred colonization area, obviously avoiding the undesired systemic effects of the delivered drugs (Rubinstein, 1990). Results from the latest pre-clinical and clinical studies on bacterial tumor therapy were recently reviewed by Felgner *et al.* (2016). Several Gram-positive and Gram-negative bacteria namely *Escherichia coli*, *Clostridia sp., Salmonella sp.*, *Bifidobacteria sp.*, *Lactobacillus sp*. and *Listeria sp.* have already been evaluated regarding their potential for cancer therapy by using animal models. Strains from the genera *Listeria*, *Clostridia*, and *Salmonella* have made it into clinical trials (Toso *et al.*, 2002; Nemunaitis *et al.*, 2003; Le *et al.*, 2012, 2015; Roberts *et al.*, 2014).

E. coli cells, after being ingested, move through the body and colonize preferentially the lower intestine (Freter *et al.*, 1983). Hence, *E. coli* may comprise a specific drug delivery vector to the gut (Rubinstein, 1990). The potential use of *E. coli* mutants targeting other colonization sites or even the use of other microorganisms may promote the drug adsorption on distinct sites (Paton *et al.*, 2012). *Salmonella sp.* and *Clostridia* spores became the most studied because these bacteria are able to successfully target neoplastic tissue without seriously harming the host. Also, *Clostridia* spores are able to pass the blood brain barrier under certain conditions (Felgner *et al.*, 2016). Therefore, they might represent a suitable carrier in bacteria mediated tumor therapy. However, the restriction to anaerobic regions might also be a great disadvantage because they need to be injected in or close to the tumor, thus metastases or small tumors without necrotic areas as colonizing niche may not be targeted as easily by this therapeutic approach (Felgner *et al.*, 2016).

Since curcumin has a poor bioavailability, in this particular case, the bacteria ability to move to a specific site may result in a more effective drug absorbance where it is actually needed, thus resulting in a more effective treatment of solid tumors.

In the breast cancer treatment, ultrasound is often used to treat solid tumors. Nonetheless, this procedure is not totally successful, as sometimes it just heats the tumor without destroying it and its continuous use may promote the occurrence of metastasis (Maxwell, 1995). Therefore, combining this treatment with the release of a therapeutic agent may represent a more effective treatment approach. Some studies have been developed towards this goal (Rodrigues, 2014). Administrating live bacteria able to produce curcuminoids triggered by heat may be more efficient than the ultrasound itself.

1.1.3 Biosynthetic pathway of curcuminoids in *Curcuma longa*

Curcuminoids are chemically bis-α,β-unsaturated β-diketones and consist of two phenylpropanoid units chemically derived from the amino acid phenylalanine (Phe) connected by a central carbon unit derived from malonyl-CoA (Katsuyama *et al.*, 2009a). The biosynthesis of curcuminoids in plants is accomplished by type III polyketide synthases (PKS), structurally simple enzymes consisting of a homodimer of ketosynthase. The understanding of the biosynthetic pathway that leads to the production of curcuminoids contributed to the advance of the biotechnological knowledge on this field.

The main contribution to the identification of the enzymes from the curcuminoid biosynthetic pathway was given by Ramirez-Ahumada and co-workers (2006). By that time, they confirmed the involvement of the phenylpropanoid pathway in the production of these compounds in plants. The first enzyme identified was phenylalanine ammonia-lyase (PAL), which is involved in both amino acid and plant secondary metabolisms. They also identified *p*-coumaroyl shikimate transferase (CST) and caffeoyl-CoA *O*-methyltransferase (CCoAOMT), among others.

Later, Katsuyama *et al.* (2009a) proposed a pathway for the curcuminoid biosynthesis in the herb *C. longa*, which involved two PKS: diketide-CoA synthase (DCS) and curcumin synthase 1 (CURS1). DCS catalyzes the synthesis of feruloyl-diketide-CoA from feruloyl-CoA and malonyl-CoA and CURS1 catalyzes the hydrolysis of feruloyl-diketide-CoA in an β-keto acid and subsequently, uses the β-keto acid and other molecule of feruloyl-CoA as substrates to catalyze the formation of curcumin (Katsuyama *et al.*, 2007, 2011) (Figure 1). CURS1 enzyme is responsible for the last steps in the production of curcuminoids. DCS and CURS1 are very similar, sharing 62% amino acid sequence identity and a conserved Cys-His-Asn catalytic triad (Katsuyama *et al.*, 2009a) and both are able to synthetize bisdemethoxycurcumin but at low efficiency. Demethoxycurcumin, an asymmetric curcuminoid, can be produced from *p*-coumaroyl-diketide-CoA and feruloyl-CoA or from feruloyl-diketide-CoA and *p*-coumaroyl-CoA (Figure 1). Katsuyama *et al.* (2009b) identified and characterized two other type III PKS from turmeric, CURS2 and CURS3. The presence of three curcumin synthases, CURS1, CURS2 and CURS3, with different substrate specificities might explain the distribution of the main curcuminoids (curcumin, demethoxycurcumin, and bisdemethoxycurcumin) in *C. longa* (Rodrigues *et al.*, 2015b).

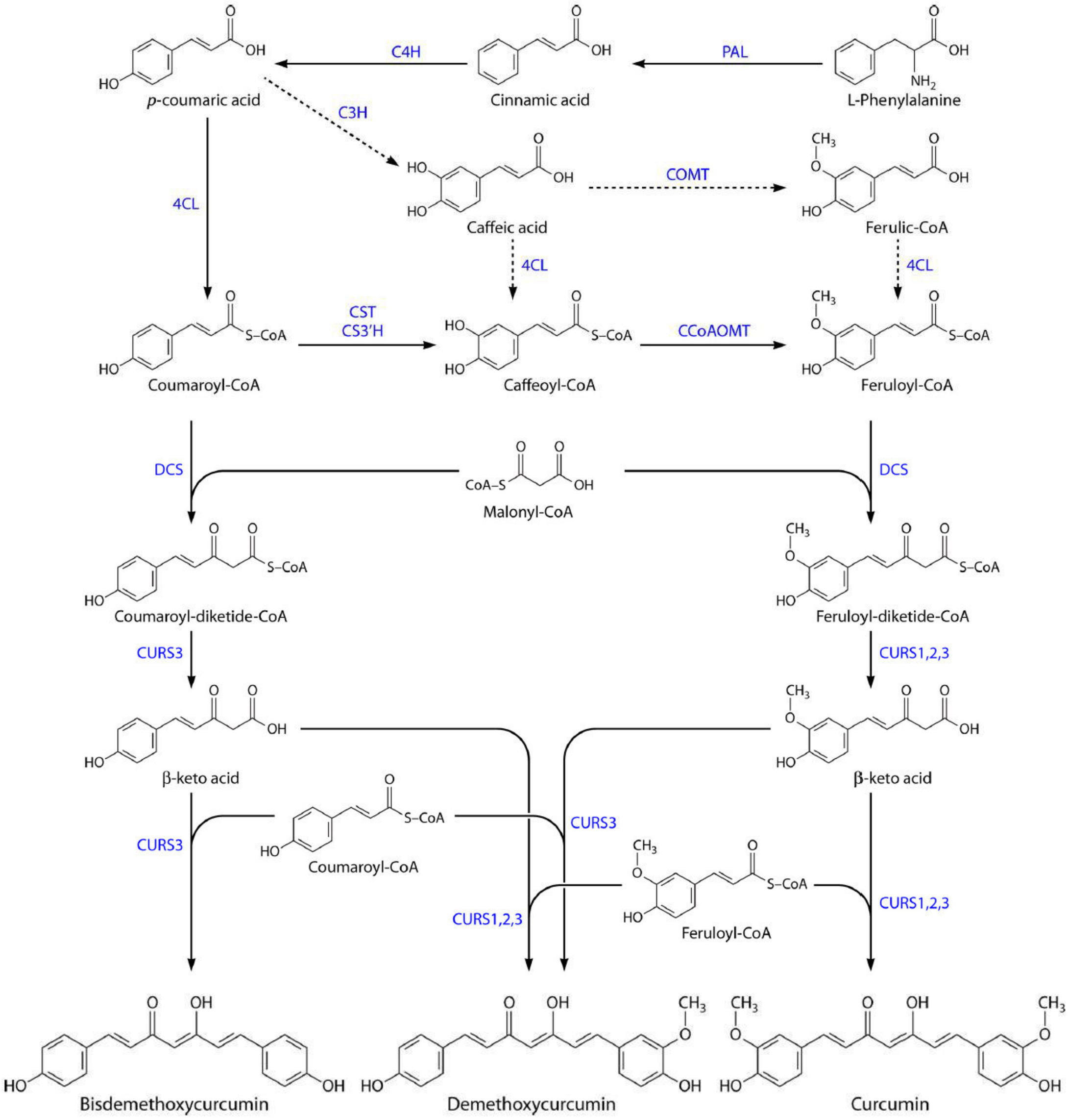

p-coumaric acid
C4H
Cinnamic acid
PAL
L-Phenylalanine
C3H
4CL
Caffeic acid
COMT
Ferulic-CoA
4CL
4CL
Coumaroyl-CoA
CST
CS3'H
Caffeoyl-CoA
CCoAOMT
Feruloyl-CoA
DCS
Malonyl-CoA
DCS
Coumaroyl-diketide-CoA
Feruloyl-diketide-CoA
CURS3
CURS1,2,3
β-keto acid
β-keto acid
CURS3
Coumaroyl-CoA
CURS3
Feruloyl-CoA
CURS1,2,3
CURS1,2,3
Bisdemethoxycurcumin
Demethoxycurcumin
Curcumin

Figure 1. Curcuminoid biosynthesis mechanism in *Curcuma longa*. PAL: Phenylalanine ammonia lyase; C4H: cinnamate-4-hydroxylase; 4CL: 4-coumarate-CoA ligase; CST*p*-coumaroyl shikimate transferase; CS3H: *p*-coumaroyl 5-*O*-shikimate 3-hydroxylase; CCoAOMT: caffeoyl-CoA*O*-methyltransferase; DCS: diketide-CoA synthase; CURSs: curcumin synthases. Taken from Rodrigues *et al.* (2015b).

Additionally to the most recognized curcuminoids, other diarylheptanoids can also be found in turmeric, for example, 3'-hydroxy-bisdemethoxycurcumin and 3'-hydroxy-demethoxycurcumin (Xie *et al.*, 2009). Besides, derivatives from curcuminoids as dihydrocurcuminoids (Resmi & Soniya, 2012) have also been identified in *C. longa*, for example dihydrodemethoxycurcumin, dihydrobisdemethoxycurcumin, and dihydrocurcumin (Kita *et al.*, 2009).

More studies in this field are required since understanding the biosynthetic pathway of curcuminoids may contribute to find enzymes with improved characteristics for the heterologous production of curcumin. Also, curcuminoids derivatives might be novel drug candidates for the treatment of several diseases, including breast cancer.

Figure 1 represents the biosynthetic pathways present in *C. longa* that leads to the production of curcuminoids. The enzymes responsible for these steps have been the basis of several bioengineering constructions.

1.2 Heterologous production of curcumin

As previously mentioned, curcumin has been the subject of hundreds of scientific papers for more than four decades due to its interesting properties. Plants store very low quantities of curcuminoids, as many other plant secondary metabolites, so its extraction from *C. longa* is insufficient and unprofitable. Besides, their chemical synthesis is difficult and complex. As a result, there have been increasing efforts to produce it in heterologous microorganisms (Katsuyama *et al.*, 2008; Lussier *et al.*, 2012; Rodrigues *et al.*, 2015b,c).

Synthetic biology approaches or combinatorial biosynthesis can be used to develop those heterologous curcumin producers. These approaches consist in combining genes from different species/organisms and designing a new biosynthetic pathway to produce the desired compounds in a heterologous host (Horinouchi, 2008, 2009).

The heterologous production of curcuminoids in microorganisms holds several advantages comparing to the native producers, including their fast growth on inexpensive substrates and their ease of manipulation compared to plants (Lussier *et al.*, 2012), thus curcuminoids can be produced faster and most probably in larger amounts. Moreover, metabolites produced by microbes can be more easily purified since, in principle, they do not have competing pathways to the recombinant one (Chemler & Koffas, 2008).

1.2.1 Possible hosts – advantages and drawbacks

Currently, the heterologous production of curcumin has only been accomplished in *E. coli*. This bacteria is the most well-known and used host for the most varied applications since it has a rapid and easy growth, and there is a great knowledge of its genome and manipulation techniques (Demain & Vaishnav, 2009). Although it is recognized that there are some problems of expressing eukaryotic proteins on prokaryotic hosts, several strategies have been developed to overcome these issues (Weickert *et al.*, 1996).

Saccharomyces cerevisiae, which is also easy to grow and manipulate and is well characterized, has been used only to produce other polyketides, such as resveratrol, naringenin, pinocembrin, among others (Katsuyama *et al.*, 2010b; Koopman *et al.*, 2012; Lussier *et al.*, 2012). What makes *S. cerevisiae* distinctively advantageous over *E. coli* for the design and construction of complex biosynthetic pathways, as the one of curcuminoids, is the fact that, as an eukaryote it has the post-translational machinery, thus possessing the intracellular compartments similar to those of plant cells (Becker *et al.*, 2003). Additionally, this yeast permits an easier enzyme expression than in prokaryotes, which generally results on much better yields. Besides, *S. cerevisiae* has a food-grade status (GRAS organism) which allows its use in human nutrition and pharmaceuticals. In addition, as membrane proteins - cytochrome P450 enzymes (C4H and C3H) – are needed for curcuminoid production from phenylalanine and/or tyrosine (Tyr) (Figure 2), such enzymes would be more adequately expressed in a eukaryotic organism. In addition, the heterologous production of a similar compound exhibited higher yields using *S. cerevisiae* than *E. coli* (Watts *et al.*, 2006). The major drawback of the use of *S. cerevisiae* as a heterologous host relies not only on the difficulty of the DNA manipulation compared to *E. coli*, but also this yeast over-glycosylates N-linked sites leading to reduction in both activity and receptor-binding (Demain & Vaishnav, 2009).

Pichia pastoris could also be an interesting host for the recombinant production of curcuminoids, as it generally has strong and inducible promotors (Potvin *et al.*, 2012), so its use usually allows greater yields on recombinant proteins synthesis (Demain & Vaishnav, 2009). Therefore, the synthesis of curcuminoids could be greater on this microbe, however further studies are needed. Some disadvantages of *P. pastoris* cultivation include the high protease expression levels, nutrient-deficiency when grown on defined media (Potvin *et al.*, 2012) and the difficulty to manipulate its genome.

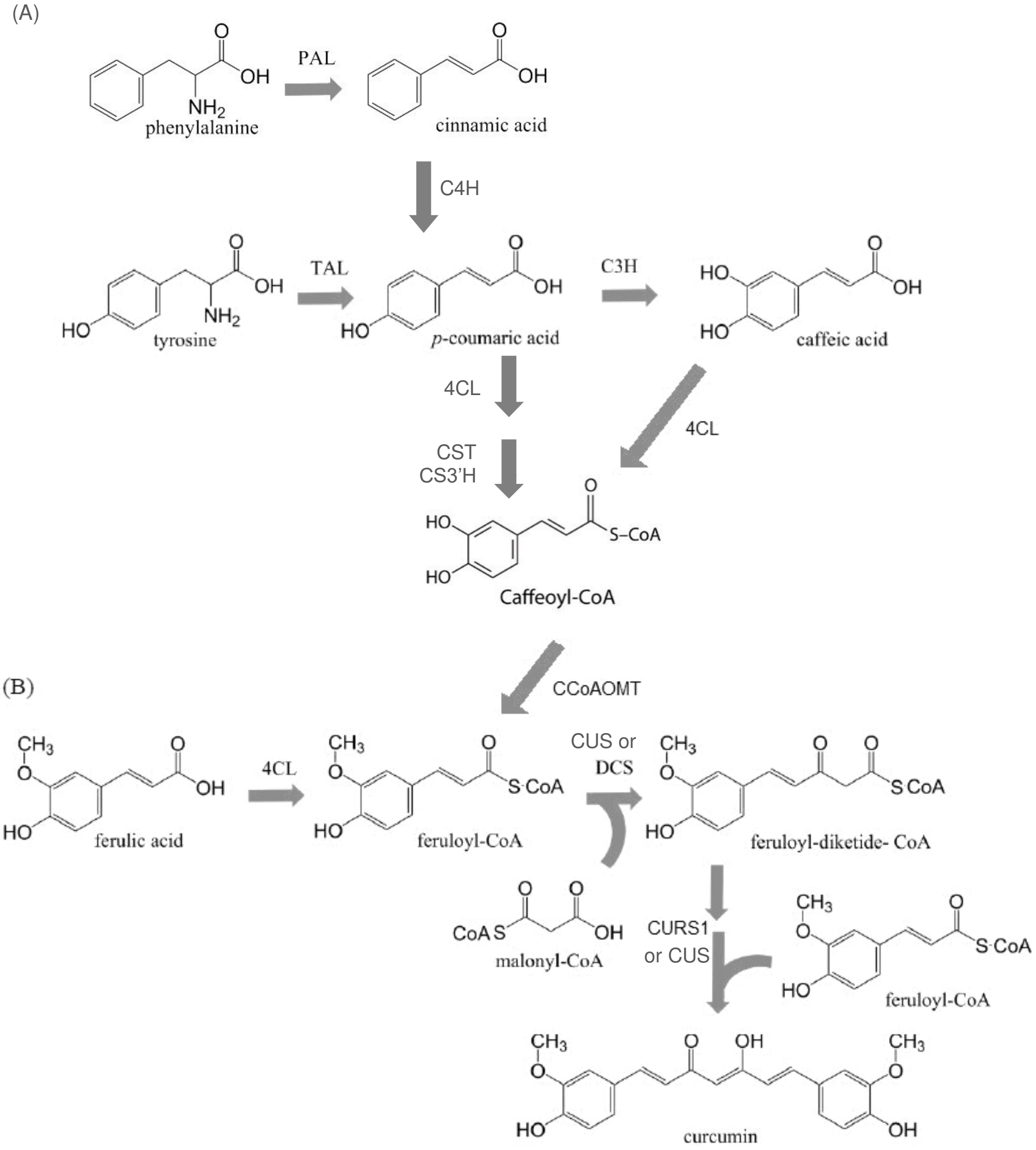

(A)
PAL
phenylalanine
cinnamic acid
C4H
TAL
C3H
tyrosine
p-coumaric acid
caffeic acid
4CL
4CL
CST
CS3'H
Caffeoyl-CoA
(B)
CCoAOMT
4CL
ferulic acid
feruloyl-CoA
CUS or
DCS
feruloyl-diketide- CoA
malonyl-CoA
CURS1
or CUS
feruloyl-CoA
curcumin

Figure 2. Possible curcumin biosynthesis pathways with plant genes to construct engineered *E. coli* strains. The production can occur from amino acids (tyrosine or phenylalanine) (A) or directly from phenylpropanoid acids such as ferulic acid (B). PAL: phenylalanine ammonia lyase; C4H: cinnamate-4-hydroxylase; TAL: tyrosine ammonia lyase; C3H: 4-coumarate 3-hydroxylase; 4CL: 4-coumarate-CoA ligase; CST: *p*-coumaroyl shikimate transferase; CS3'H: *p*-coumaroyl 5-*O*-shikimate 3'-hydroxylase; CCoAOMT: caffeoyl-CoA *O*-methyltransferase; DCS: diketide-CoA synthase; CURS1: curcumin synthase 1; CUS: curcuminoid synthase. Adapted from Rodrigues *et al.* (2015a,c).

The heterologous synthesis of curcuminoids in plants has the only advantage that it is not required a lot of genetic modifications because the genes from phenylpropanoid pathway already belong to plants genome (Jeandet *et al.*, 2012). However, products derived from genetically modified plants are difficult to accept by the society and the regulations for these kind of crops are extremely complex (Halls & Yu, 2008). Hence, the heterologous production of curcuminoids by microorganisms is much more attractive.

Mammalian cells could also be an interesting approach to produce curcuminoids for therapeutic uses. These cells have already been engineered to produce other polyketides such as resveratrol (Zhang *et al.*, 2006) to explore possible medical and biotechnology applications. In that work, TAL (tyrosine ammonia-lyase from *Rhodobacter sphaeroides*) and a 4CL::STS (4-coumarate-CoA ligase from *Arabidopsis thaliana* and steroid sulfatase from *Vitis Vinifera*) fusion enzyme were expressed in a human kidney cell line (HEK293) and *de novo* biosynthesis of resveratrol was obtained. *In situ* studies were not included, but the authors believe that one possible application of engineered resveratrol production might be the improvement of cell replacement therapies by prolonging the lifetime of donor cells. For example, β-cell replacement therapy for type I diabetes using inlet transplantation, that seems to be a promising therapy, is limited by the short life span of the transplanted β-cells (Stock & Bluestone, 2004), hence engineering donor β-cells to produce resveratrol may provide longer survival, improving the efficacy and efficiency of the treatment. Another strategy is the production of curcumin by bacteria and its delivery on mammalian cells. In fact, the *E. coli* secretion system has been recently studied aiming at exploring the secretion of proteins directly into mammalian cells cytoplasm (Reeves *et al.*, 2015). This same approach could be explored for the treatment of cancer, for example expressing curcumin *in situ*.

Despite the recent research efforts, more studies on the heterologous production of curcuminoids are crucial for the development of their technological application. As discussed above, alternative organisms may lead to higher yields, while the use of animal cells may enable a more specific therapeutic application.

1.2.2 Curcumin production in *Escherichia coli*

Curcuminoids artificial production by *E. coli* was reported for the first time by Katsuyama *et al.* (2008) even before they could clarify the enzymatic steps involved in the natural biosynthetic pathway on *C. longa* (Katsuyama *et al.*, 2009a). At that time, the authors identified a type III PKS from rice (*Oryza sativa*) that they called curcuminoid synthase (CUS). This was the first identified type III PKS able of catalyzing the synthesis of bisdemethoxycurcumin from two molecules of *p*-coumaroyl-CoA (starter substrate) and one molecule of malonyl-CoA (extender substrate) through only one-step. CUS itself catalyzes both steps that in *C. longa* are catalyzed by DCS and CURS.

The pathway constructed in *E. coli* by Katsuyama *et al.* (2008) started with the conversion of phenylalanine into cinnamic acid by PAL from the yeast *Rhodotorula rubra* (Figure 2). This enzyme was previously shown to have TAL activity (Miyahisa *et al.*, 2005). Thus, using this enzyme, tyrosine can be used as a precursor and can be further converted to *p*-coumaric acid. The phenylpropanoid acids are converted to CoA esters by 4CL from *Lithospermum erythrorhizon* and then to curcuminoids by CUS from *O. sativa*. To increase the intracellular pool of malonyl-CoA in *E. coli*, acetyl-CoA carboxylase (ACC) from *Corynebacterium glutamicum* was also overexpressed. The cultivation of the recombinant *E. coli* in M9 medium with the supplementation of tyrosine or/and phenylalanine resulted on the production of bisdemethoxycurcumin (favored by supplementation of tyrosine), dicinnamoylmethane (enhanced when Phe is supplemented), and cinnamoyl-*p*-coumaroylmethane (better on M9+Tyr+Phe medium). In addition to curcuminoids, CUS produces triketide pyrones as by-products, by condensing two malonyl-CoAs to *p*-coumaroyl-CoA and its synthesis was greater on the M9+Tyr medium (Katsuyama *et al.*, 2008).

Phenylpropanoid acids (*p*-coumaric acid, cinnamic acid and ferulic acid) can also be directly supplied, which allows obtaining higher productivities of curcuminoids (Katsuyama *et al.*, 2008) (Figure 2). When *p*-coumaric acid, cinnamic acid or ferulic acid were supplied, the yield of bisdemethoxycurcumin, dicinnamoylmethane or curcumin, respectively, was about 100 mg/L (Katsuyama *et al.*, 2008). Rice bran pitch was also used for the production of curcumin since it is an industrial waste, rich in ferulic acid. The *E. coli* strain engineered by Katsuyama *et al.* (2008), harboring 4CL, CUS, and ACC genes, produced 57±21 mg/L of curcumin from 11±1.4 mg of ferulic acid extracted from 500 mg of rice bran pitch.

Additionally, the supplementation of unnatural carboxylic acids to the recombinant *E. coli* cells carrying the artificial pathway (4CL, CUS and ACC) resulted in the production of unnatural compounds and this way of producing unnatural curcuminoids may provide new drug candidates with improved or novel activities. Proving this possibility, Katsuyama *et al.* (2010a) produced 15 asymmetric curcuminoids (9 of these compounds are not found in nature) by adding two different carboxylic acids simultaneously (analogues of *p*-coumaric acid).

Another genes combination (PALs from *Trifolium pratense*, *At*4CL1 from *Arabidopsis thaliana* and CUS from *O. sativa*) in *E. coli* resulted in better results (Wang *et al.*, 2013) than the ones obtained by Katsuyama *et al.* (2008) using PAL from *R. rubra* and 4CL from *L. erythrorhizon*. Wang *et al.* (2013) used different PAL enzymes from *T. pretense* and PAL1 was the one that allowed greater dicinnamoylmethane production among PALs tested.

More recently, *At*4CL1 from *A. thaliana*, DCS and CURS1 from *C. longa* were used to produce curcuminoids in *E. coli* and 70 mg/L of curcumin was obtained from ferulic acid (Rodrigues *et al.*, 2015c). By feeding *p*-coumaric acid or a mixture of *p*-coumaric acid and ferulic acid, bisdemethoxycurcumin and demethoxycurcumin were also produced respectively, but in lower concentrations. Moreover, the same authors (Rodrigues *et al.*, 2015a,c) produced curcuminoids from tyrosine through an alternative caffeic acid pathway. To produce caffeic acid the authors used TAL from *Rhodotorula glutinis* and 4-coumarate 3-hydroxylase (C3H) from *Saccharothrix espanaensis.* Then, caffeoyl-CoA 3-*O*-methyltransferase (CCoAOMT) from *Medicago sativa* was used to convert caffeoyl-CoA to feruloyl-CoA. Using caffeic acid, *p*-coumaric acid or tyrosine as a substrate, 3.9 mg/L, 0.3 mg/L and 0.2 mg/L of curcumin were produced, respectively. This was the first time DCS and CURS1 were used for the heterologous production of curcuminoids and that curcuminoids were produced by feeding tyrosine, which was a great improvement on the heterologous production of curcumin using *E. coli*.

Very recently, Zhang *et al.* (2016) identified a new curcuminoid synthase from ginger *Zingiber officinale* (*Zo*CURS) and proved that using an unnatural fusion protein DCS::CURS using DCS from *C. longa* and the newly identified *Zo*CURS, the *in vitro* curcuminoid production was improved compared to the experiments in which co-incubation of DCS and *Zo*CURS was used.

In Table 1, a summary of the results obtained in several reported works for curcumin production is presented.

Table 1. Biological systems and conditions used for curcumin production in literature. *Le*4CL1: 4-coumarate-CoA ligase from *Lithospermum erythrorhizon*; ACC: Acetyl-CoA carboxylase (to improve the intracellular pool of malonyl-CoA) from *Corynebacterium glutamicum*; CUS: curcuminoid synthase from *Oryza sativa*; *At*4CL1: 4-coumarate-CoA ligase from *Arabidopsis thaliana*; DCS: diketide-CoA synthase from *Curcuma longa*; CURS1: Curcumin synthase from *C. longa*; *Rg*TAL from *Rhodotorula glutinis*; *Se*TAL: tyrosine ammonia-lyase from *Saccharothrix espanaensis*; C3H: 4-coumarate 3-hydroxylase from *S. espanaensis*; CCoAOMT: caffeoyl-CoA 3-O-methyltransferase from *Medicago sativa*; COMT: caffeic acid 3-O-methyltransferase from *M. sativa*. *Genes codon-optimized to *E. coli* system.

Organism	Plasmids_Genes	Substrate (Concentration)	Conditions	Curcumin production	Reference
E. coli BLR (DE3)	pCDF_*Le*4CL1 pRSF_ACC pET_CUS	Ferulic acid (1 mM)	LB+M9 26°C 60 h	113±22 mg/L	(Katsuyama *et al.*, 2008)
		Ferulic acid in rice bran pitch (1.1±0.15 mM)		57±21 mg/L	
E. coli K-12 MG1655 (DE3)	pAC_*At*4CL1 pRSFDuet_CUS*	Ferulic acid (2 mM)	LB+M9 26°C 63 h	2.47±0.17 mg/L	(Rodrigues *et al.*, 2015c)
	pAC_*At*4CL1 pCDFDuet_DCS* pRSFDuet_CURS1*			69.2±3.0 mg/L	
	pAC_*At*4CL pCDFDuet_DCS* pRSFDuet_CURS1*_CCoAOMT*	Caffeic acid (1 mM)		1.44 mg/L	
	pKVS45_C3H* pAC_*At*4CL pRSFDuet_CURS1*_CCoAOMT* pCDFDuet_DCS*	*p*-Coumaric acid (2 mM)		0.11 mg/L	
	pETDuet_*Rg*TAL*_C3H* pRSFDuet_CURS1*_CCoAOMT* pAC_*At*4CL1 pCDFDuet_DCS*	L-Tyrosine (3 mM)		0.07 mg/L	
E. coli BL21 (DE3)	pSW97_*Se*TAL_C3H pSW99_COMT_*At*4CL pSW24_CUS	L-Tyrosine (3 mM)	LB 28°C 48 h	0.67±0.1 mg/L	(Wang *et al.*, 2015)

In summary, the artificial biosynthetic pathway in *E. coli* can start with the supplementation of the amino acids tyrosine or phenylalanine, or of the carboxylic acids and lead to the production of several natural curcuminoids. Unnatural carboxylic acids can also be supplemented as precursors and lead to the production of unnatural compounds with possibly novel therapeutic properties. For the development of the production strategy, the next step should be to engineer *E. coli* mutants that overproduce tyrosine and are able of converting simple carbon sources as glucose to curcuminoids.

1.2.3 Optimization of curcumin production

In the last decade, noteworthy advances in the understanding of the biosynthetic pathway of curcuminoid production in *C. longa* and its heterologous production in *E. coli* have been reported. However, continuous efforts towards exploiting new heterologous hosts and finding the most adequate and compatible synthetic enzymes and plasmids are needed.

Evaluating the most productive strain is of great interest for the process optimization. It is important to notice that all the studies reported on curcumin production using different strains are not comparable as these have not been conducted under the same conditions. *E. coli* K-12 MG1655 (DE3) was used in the previous studies from Rodrigues *et al.* (Rodrigues, 2014; Rodrigues *et al.*, 2015c), but B strains such as BL21 (DE3) have been described as more suitable for protein expression in some cases. Gene expression levels related to the acetate metabolism and glyoxylate shunt are higher in the BL21 compared to K–12 strains and there is less acetate formation during cultivations of BL21 relative to K–12 strains (Shiloach *et al.*, 1996). Furthermore, *E. coli* BL21 can achieve greater amounts of biomass and higher growth rates than *E. coli* K–12 strains, possibly because of differentially regulated glucose transport, as recently suggested (Shiloach *et al.*, 2010). Moreover, the BL21 strain is deficient in two proteases (*lon*, *ompT*) (Studier *et al.*, 2009) and is therefore more suitable for expression of non-toxic genes. Besides, most part of the studies use the B strains for protein expression and in particular, for plant metabolites production (Miyahisa *et al.*, 2005; Leonard *et al.*, 2008) and its growth is faster.

The enzymatic pathway has obviously an important role. Choosing enzymes and catalyzed-steps that represent the best candidates for the curcuminoid production pathway construction is required to adequately transform the metabolic intermediates into the desired products (Rodrigues *et al.*, 2015b). Enzymes from different organism's genome must be found in order to find the one(s) that lead to higher productivities.

Accumulation of toxic intermediates to the heterologous host should also be avoided (Rodrigues *et al.*, 2015b), as their presence represents a limitation of the process and lower is the desired final product production. The addition of genes that produce curcumin from tyrosine (Figure 2) (addition of TAL, C3H, and CCoAOMT) was made by Rodrigues *et al.* (2015c), which successfully avoided feeding the expensive substrate ferulic acid. Further optimizations can include the introduction of enzymes to convert phenylalanine into tyrosine in order for bacteria to use more cellular resources. Moreover, undesired by-products should be decreased or even eliminated (Rodrigues *et al.*, 2015b). As in the Katsuyama approach (Katsuyama *et al.*, 2008), ACC can also be overexpressed to improve malonyl-CoA cellular pool (Miyahisa *et al.*, 2005; Leonard *et al.*, 2007). However, it is necessary to study if malonyl-CoA is really a limiting substrate for the production of curcuminoids as comparing the results from Katsuyama *et al.* (2008) (that overexpressed ACC) and Wang *et al.* (2013) it is perceptible that a high production of curcuminoids can also be obtained without ACC overexpression.

Additionally, there is still a great amount of optimization of the operational conditions that need to be conducted in order to obtain greater curcumin productions and ideally, higher *E. coli* growth. Usually, for the production of curcuminoids, *E. coli* strains are first cultured at 37°C on LB (lysogeny broth) medium to favor biomass growth and reach a suitable protein production level (Santos *et al.*, 2011; Rodrigues *et al.*, 2015b). Afterwards, when the exponential phase is reached, the cells are harvested and transferred to minimal medium, such as M9 medium, where the substrates are supplemented and the curcuminoids are produced. This medium exchange is inconvenient for the production of curcumin, especially on a large scale. Terrific broth (TB) can be tested since it was already used on heterologous expression and production of a similar compound, the flavonoid (2S)-pinocembrin (Wu *et al.*, 2013). The use of MOPS (morpholinepropanesulfonic acid) minimal medium (Neidhardt *et al.*, 1974) may also be a good alternative since its use on the production of flavonoids in an one-step fermentation did not seem to affect the culture growth (Santos *et al.*, 2011; Wu *et al.*, 2013).

The delay of IPTG (isopropyl-β-D-thiogalactopyranoside) induction can also alleviate the metabolic burden. Induction should be performed at the exponential phase, since after that point, although bacteria are still growing, the production titers decrease (Santos *et al.*, 2011; Wu *et al.*, 2013).

The compounds used as substrates (caffeic acid, *p*-coumaric acid and ferulic acid) present some toxicity to the cells at high concentrations (Watts *et al.*, 2006; Zhang & Stephanopoulos, 2013), so when using them as precursors, a lower concentration added at the beginning of the experiment, as well as fed-batch fermentations should be considered.

Using metabolic engineering approaches, non-essential genes can be deleted to enhance curcumin production by *E. coli*, directing the carbon fluxes towards the production of the recombinant enzymes (Berry, 1996; Matsuoka & Shimizu, 2012). On this approach, global metabolic studies are needed to predict physiological responses of the mutants. Using *E. coli* metabolic models containing the artificial curcumin biosynthetic pathway, it was possible to simulate and predict *in silico* which gene knockouts would increase curcumin production from ferulic acid (Machado, 2012). These deletions included one single knockout (*gnd* gene) and multiple knockouts of five non-essential genes for aerobic growth (*fumA, fumB, fumC, ccmA* and *argO*). After optimizing the genes expression, the several deletions can be implemented in the heterologous curcumin producer to confirm the improvement of its production.

Gnd encodes for 6PGDH (6-Phosphogluconate dehydrogenase) (EC 1.1.1.44), an enzyme responsible for a key reaction of the pentose phosphate pathway (PPP) that is an alternative route for the carbohydrate metabolism, besides glycolysis (Kobayashi, 2004) (Figure 3). The obstruction of the PPP by *gnd* knockout leads to the decreasing of reducing power (NADPH) for anabolism. The Δ*gnd* mutant is well described to decrease the carbon flux through glucose-6-phosphate dehydrogenase (6PDH) (Jiao *et al.*, 2003) which means that the carbon can be directed for the pyruvate formation using 2-keto-3-deoxy 6-phosphogluconate (KDPG) as intermediate, an activity that exists when *gnd* is knocked-out (Jiao *et al.*, 2003). This means that acetyl-CoA (ACA) formation can be higher and ACC converts it into malonyl-CoA (fatty acid synthesis pathway), a precursor for curcumin synthesis.

Fumarase A, fumarase B and fumarase C (encoded from *fumA*, *fumB* and *fumC*, respectively) are the three fumarase isozymes (EC 4.2.1.2) participating in the TCA (tricarboxylic acid) cycle (Figure 3). Their function is to convert fumarate into malate and their knockout leads to higher fumarate production (Tseng *et al.*, 2001; Song & Lee, 2013) and higher malonyl-CoA concentration (Xu *et al.*, 2011), which has been pointed as essential to the curcuminoids production as previously discussed.

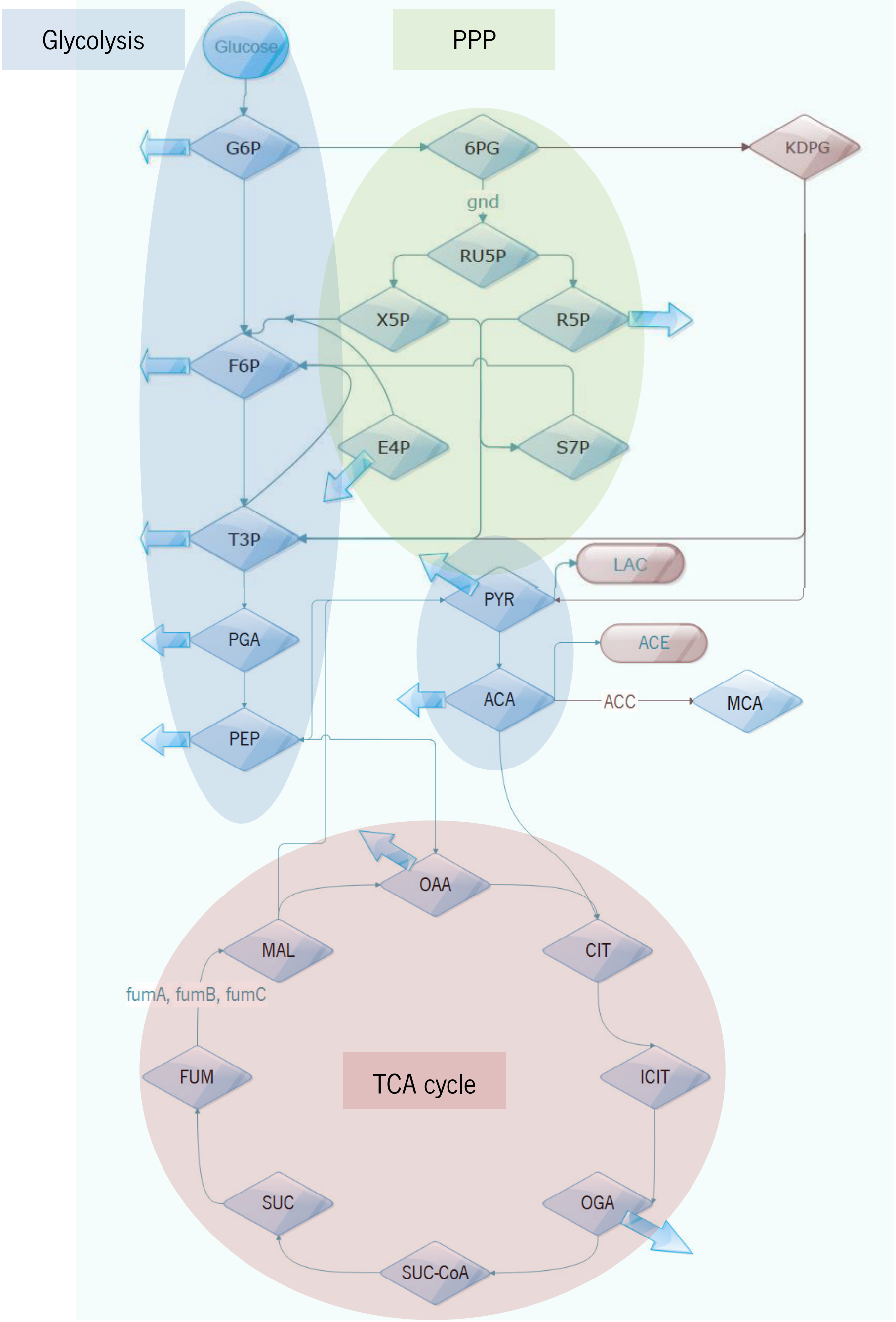

Glycolysis
PPP
Glucose
G6P
6PG
KDPG
gnd
RU5P
X5P
R5P
F6P
E4P
S7P
T3P
LAC
PYR
PGA
ACE
ACA
ACC
MCA
PEP
OAA
MAL
CIT
fumA, fumB, fumC
FUM
TCA cycle
ICIT
SUC
OGA
SUC-CoA

Figure 3. Part of the metabolic model of *E. coli* K-12 MG1655 indicating the reactions altered by *gnd*, *fumA*, *fumB* and *fumC* deletions. Based in the work from Jiao *et al.* (2003). PPP: pentose phosphate pathway; TCA cycle: tricarboxylic acids cycle; G6P: glucose 6-phosphate; F6P: fructose 6-phophate; T3P: glyceraldehydes-3-phosphate; PGA: 3-phosphoglycerate; PEP: phosphoenolpyruvate; 6PG: 6-phosphogluconate; RU5P: ribulose-5-phosphate; X5P: xylulose-5-phosphate; E4P: erythrose-4-phosphate; R5P: ribose-5-phosphate; S7P: sedoheptulose-7-phosphate; KDPG: 2-keto-3-deoxy 6PG; PYR: pyruvate; LAC: lactate; ACA: Acetyl-CoA; ACE: acetate; OAA: oxaloacetate; CIT: citrate; ICIT: isocitrate; GLYO: glyoxilate; OGA: oxoglutarate; SUC-CoA: succinyl-CoA SUC: succinate; FUM: fumarate; MAL: malate. MCA: malonyl-CoA; ACC: acetyl-CoA carboxylase.

In *E. coli*, CcmA is the ATP-binding component of the ABC complex required for the formation of cytochrome c, so this protein is not synthetized in *ΔccmA* mutant (Fischer *et al.*, 1995). ArgO (formerly YggA) is a member of the LysE family of lysine efflux transporters in *E. coli* K-12 and it is responsible for the export of arginine. These genes are not important for aerobic growth and the prediction that *ccmA* and *argO* knock-outs would improve curcumin production can be explained by the redirection of the resources that cells would spend in the synthesis of this biomolecules to the desired reactions.

The non-essentiality of the *gnd*, *fumA*, *fumB*, *fumC* and *ccmA* genes was confirmed by genetic footprinting under conditions of aerobic logarithmic cell growth in rich medium (Gerdes *et al.*, 2003). According to the same study, *argO*/*yggA* is essential for *E. coli* growth. Unlike the findings of Gerdes *et al.* (2003), Baba *et al.* (2006) verified that *E. coli* K-12 MG1655 Δ*argO* is able to grow in different media. *E. coli* growth can be blocked or delayed by some of the deletions. Jiao *et al.* (2003) studied the growth of *E. coli* BW25113 with and without *gnd* knock-out and they demonstrate that the mutant growth is slightly lower, so this can also influence curcumin production.

2. OBJECTIVES

The main goal of this thesis is to improve the production of curcumin by an engineered *E. coli* holding an artificial biosynthetic pathway constructed with 4CL, DCS and CURS1 codifying genes derived from plants. The specific aims are the following:

- Evaluation of alternative strains with higher protein expression than *E. coli* K-12 MG1655 (DE3) and consequently improved curcumin production.
- Evaluation of different media composition and operational conditions to optimize the production of curcumin.
- Construction and validation of the *E. coli* mutants previously identified through an *in silico* approach (e.g. *ΔfumA,fumB,fumC*) using an engineered strain able to produce curcumin from ferulic acid.
- Validation of the *E. coli* mutants with or without gene deletions for the production of curcumin using ferulic acid as substrate.

3. MATERIALS AND METHODS

3.1 Strains and plasmids

E. coli NZY5α competent cells (NZYTech, Lisbon, Portugal) were used for molecular cloning and vector propagation. *E. coli* K-12 MG1655 (DE3), *E. coli* K-12 JM109 (DE3) and *E. coli* BL21 (DE3) strains (relevant genotype information is presented in Table 2) were used as possible hosts to hold the artificial curcuminoid production pathway. The plasmids pCDFDuet_DCS and pRSFDuet_CURS1 herein used were previously constructed by Rodrigues *et al.* (2015c). pAC_*At*4CL1 was previously purchased from Addgene (#35947) (Cambridge, MA, USA).

For the CRISPR-Cas9 based technique, the plasmids used were the pCas9-CR4 for Cas9 expression that harbors a Cm resistance gene, and the pKDsgRNA-p15 plasmid, which serves as a PCR (polymerase chain reaction) template to construct a plasmid containing the protospacer and conferring Spec resistance. pCas9-CR4 (#62655) and pKDsgRNA-p15 (#62656) were purchased from Addgene.

Table 2. Strains used in this study and relevant genotype information.

Strain	Relevant genotype	Source
E. coli NZY5α	*fhuA2*Δ*(argF-lacZ)*U169 *phoA glnV44* Φ80 Δ*(lacZ)*M15 *gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	NZYTech
E. coli K-12 MG1655 (DE3)	F^- λ^- *ilvG*$^-$ *rfb-50 rph-1,* λDE3	(Nielsen *et al.*, 2010)
E. coli K-12 JM109 (DE3)	*end*A1, *rec*A1, *gyr*A96, *thi, hsd*R17 (r_k^-,m_k^+), *rel*A1, *sup*E44, λ–, Δ(*lac-pro*AB), [F', *tra*D36, *pro*AB, *lacI*^{q}ZΔM15], λDE3	Promega (Madison, USA)
E. coli BL21 (DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS*	New England Biolabs (Whitby, Canada)

3.2 Chemicals and culture media

Standard reagents used include LB (NZYTech), agar (Liofilchem, Roseto degli Abruzzi, Italia), Super Optimal broth with catabolite repression (SOC) (NZYTech), glucose (Acros Organics, New Jersey, USA) and glycerol (HiMedia, Mumbai, India), $MgCl_2$ (VWR, Radnor, USA), $CaCl_2.2H_2O$ (Panreac, Barcelona, Spain).

One liter of TB was prepared by mixing 12 g of tryptone (Oxoid, Basingstoke, United Kingdom), 24 g of yeast extract (Oxoid), and 4 mL of a 10% (v/v) glycerol solution with deionized water up to 900 mL. Also, 9.4 g of K_2HPO_4 (Panreac) and 2.2 g of KH_2PO_4 (Panreac) were weight and added to water up to 100 mL. Both solutions were mixed after sterilization.

MOPS minimal medium contained (per liter): 40 g of glucose (Acros Organics), 10 mL of 0.132 M K_2HPO_4 (Panreac) and 100 mL of 10x MOPS mixture. The 10x MOPS mixture contained (per liter) 83.72 g of MOPS (Fisher Scientific, Waltham, USA), 7.17 g of Tricine (ChemCruz, Dallas, USA), 0.028 g of $FeSO_4$ (Sigma, St. Louis, USA), 50 mL of 1.9 M NH_4Cl (Panreac), 10 mL of 0.276 M K_2SO_4 (Panreac), 0.25 mL of 0.02 M $CaCl_2.2H_2O$ (Panreac), 4.2 mL of 1.25 M $MgCl_2$ (VWR), 100 mL of 5 M NaCl (NZYTech), 0.2 mL of micronutrient solution [containing per 50 mL: 9 mg of $(NH_4)_6Mo_7O_{24}.4H_2O$ (Fluka, Buchs, Switzerland), 62 mg of H_3BO_3 (Merck, Kenilworth, USA), 18 mg of $CoCl_2$ (Riedel-de Haën, Seelze, Germany), 6 mg of $CuSO_4$ (Sigma), 40 mg of $MnCl_2$ (Merck) and 7 mg of $ZnSO_4$ (Sigma)].

Modified M9 minimal salt medium contained (per liter): 40 g of glucose (Acros Organics), 6 g of Na_2HPO_4 (Chem-Lab NV, Zedelgem, Belgium), 3 g of KH_2PO_4 (Panreac), 1 g of NH_4Cl (Panreac), 0.5 g of NaCl (NZYTech), 15 mg of $CaCl_2$ (Panreac), 110 mg of $MgSO_4$ (Panreac), 340 mg of thiamine (Fisher Scientific), and 5 g of $CaCO_3$ (Panreac) (to control the pH), trace elements [54 mg of $FeCl_3$ (Panreac), 4 mg of $ZnCl_2$ (Riedel-de Haën), 4 mg of $CoCl_2$ (Riedel-de Haën), 4 mg of $NaMoO_4$ (Merck), 2 mg of $CaCl_2$, 2 mg of $CuCl_2$ (Riedel-de Haën), 1 mg of H_3BO_3 (Merck)] and vitamins [0.84 mg of riboflavin (Riedel-de Haën), 10.8 mg of pantothenic acid (Fluka), 12.2 mg of nicotinic acid (Riedel-de Haën), 0.12 mg of biotin (Fisher Bioreagents), 84 μg of folic acid (Merck) and 2.8 mg of pyridoxine (Fisher Scientific)].

The antibiotics were added according to the plasmids present in the strain at the following final concentrations: 100 μg/mL of ampicillin (Amp) (NZYTech), 100 μg/mL of spectinomycin (Spec) (AppliChem Panreac), 25 μg/mL of chloramphenicol (Cm) (NZYTech), and/or 50 μg/mL of kanamycin (Kan) (NZYTech or Fisher Scientific). Differently, in the CRISPR-Cas9 method used for gene deletion, a final concentration of 34 μg/mL of Cm and 50 μg/mL of Spec were used when the pCas9-CR4 and pKDsgRNA plasmids were present, respectively.

IPTG (NZYTech), L-arabinose (Sigma-Aldrich) or anhydrotetracycline hydrochloride (aTc) (Fluka) were used to induce the protein expression. Ferulic acid (Acros Organics) was supplemented as substrate for the production of curcuminoids.

Acetonitrile (Fisher Chemical), ethyl acetate (Fisher Chemical) and trifluoroacetic acid (TFA) (Fluka) were used for the high-performance liquid chromatography (HPLC) analysis.

DMSO (dimethyl sulfoxide) (Fisher) and PBS (phosphate-buffered saline) 1x were used to dissolve and dilute curcumin, respectively, to be further used in the curcumin toxicity evaluation tests. PBS 1x was prepared by mixing 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.47 mM KH_2PO_4 and adjusting to a final pH of 7.4.

3.3 Preparation of electrocompetent cells and transformation step

In order to prepare electrocompetent cells, *E. coli* cells were grown overnight at 30 or 37°C (depending on temperature sensitivity of plasmids) and 200 rpm in 5 mL of LB medium. Then, 50 mL of LB medium were inoculated with 500 µL of the overnight culture and grown at 30 or 37°C. At an OD_{600nm} (optical density at 600 nm) ≈ 0.35 read in the microplate spectrophotometer (about 2-3 h from start), the culture was transferred to falcons and placed on ice. The cells were centrifuged at 3000 g for 15 min at 4°C. The centrifuged culture was placed on ice and the supernatant was discarded. The pellet was gently resuspended in 40 mL of ice-cold 10% (v/v) glycerol. Afterwards, the suspension was centrifuged using the same conditions. The supernatant was discarded and the pellet was slowly resuspended in 20 mL of ice-cold 10% glycerol. The solution was centrifuged again in the same conditions. Cells were placed on ice and the supernatant was discarded. The pellet was gently resuspended in 10 mL of ice-cold 10% glycerol. Finally, a fourth centrifugation was done using the same conditions. The cells were placed on ice, the supernatant was removed and the pellet was resuspended in 80 µL of ice-cold 10% glycerol. Aliquots of 20 µL of electrocompetent cells were frozen at -80°C.

For transformation, 0.5-2 µL of the targeting DNA fragment or plasmid were mixed with 20 µL of *E. coli* electrocompetent cells and the mixture was electroporated at 1.8 kV and 25 pF in a refrigerated electroporation cuvette. 1 mL of SOC was immediately added to the shocked cells and the culture was incubated 1 to 2 h at 30 or 37°C (depending on temperature sensitivity of plasmids) and 200 rpm. Afterwards, the cells were plated in LB Agar with the appropriate antibiotics and incubated overnight at the required temperature.

3.4 Preparation of chemical competent cells and transformation step

E. coli cells were grown overnight at 30 or 37°C (depending on temperature sensitivity of plasmids) and 200 rpm in 5 mL LB medium. Then, 50 mL of LB medium were inoculated with 500 µL of the overnight culture and grown at 30 or 37°C. At an OD_{600nm} ≈ 0.35 read in the microplate spectrophotometer (about 2-3 h from start) the culture was placed on ice and were kept there for 20 min. The cells were harvested by centrifugation at 3000 g for 15 min at 4°C. The supernatant was decanted and each pellet was gently resuspended in about 5 mL of ice cold $MgCl_2$ 100 mM. Next, the cells were again harvested by centrifugation at 2000 g for 15 min at 4°C. The supernatant was then removed and the pellet was resuspended in about 10 mL of ice cold $CaCl_2$. In this step, the solution was kept on ice for at least 20 min. Then, a centrifugation at 2000 g for 15 min at 4°C was made to harvest the cells. The resulting supernatant was decanted and the pellet was resuspended in about 2.5 mL of ice cold 85 mM $CaCl_2$ solution containing 15% (v/v) glycerol. Cells were then centrifuged at 1000 g for 15 min at 4°C. The supernatant was removed and the pellet was resuspended in about 100 μL of ice cold 85 mM $CaCl_2$ containing 15% glycerol. Aliquots of 50 μL were then stored at -80°C.

Heat shock was used for the cells transformation. Around 0.5-2 μL of the targeting DNA were added to the *E. coli* chemical competent cells and the mixture was incubated on ice for 30 min. Afterwards, the *E. coli* cells were heat-shocked in a water bath at 42°C for 40 seconds. The tubes were placed on ice for 2 min to reduce damage to the *E. coli* cells. Then, 1 ml of SOC was added and the tubes were incubated for 1-2 h at 30 or 37°C. Finally, cells were spread on LB agar plates with appropriate antibiotic and were grown overnight.

3.5 Optimization of curcumin production in *E. coli* using an artificial biosynthetic pathway

3.5.1 Curcumin production

To produce curcumin, the plasmids pCDFDuet_DCS and pRSFDuet_CURS1 available in our group (Rodrigues *et al.*, 2015c), and pAC_*At*4CL1 were simultaneously electroporated into the electrocompetent *E. coli* hosts (K-12 MG1655 (DE3), BL21 (DE3) and K-12 JM109 (DE3)), that were prepared and transformed as described in **Section 3.3**. The resulting three strains were tested for curcumin production.

Usually, a combination of LB and M9 was used for the production of curcumin. A representative scheme of the general fermentation process is illustrated in Figure 4. However, different medium compositions were evaluated, as well as different media combinations.

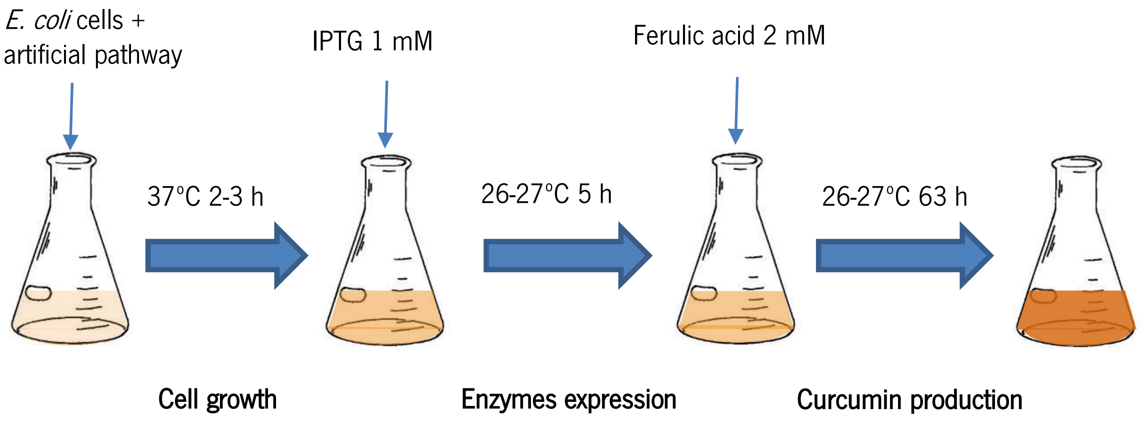


Figure 4. General scheme of the *in vivo* curcumin production fermentative process. In the first phase it occurs the biomass production. When the exponential phase is reached, the protein expression is induced by adding IPTG (isopropyl-β-D-thiogalactopyranoside) in a final concentration of 1 mM and the temperature is decreased. After 5 h of induction the substrate is added to the same or fresh media, depending on the approach. Curcumin starts to be produced while the culture is maintained at 26°C-27°C for 63 h.

The cultures were grown in shake flasks at 37°C in 50 mL of LB until an OD_{600nm} reached 0.4 (in the case of *E. coli* K-12 MG1655 (DE3) strain) or 0.6-0.7 (when using *E. coli* JM109 (DE3) or BL21 (DE3)). IPTG (generally at 1 mM) was added to the medium to induce protein expression and the culture was incubated for more 5 h at 26°C-27°C. Next, the cells were harvested by centrifugation and suspended in M9 minimal medium supplemented with the substrate, ferulic acid as well as more IPTG (in the same concentration as before). The resulting culture was incubated at 26°C-27°C for 63 h. Samples were taken at different times (0 h, 20 h, 43 h and 63 h).

Additionally to this standard LB+M9 combination, LB+LB, LB+MOPS and TB+MOPS combinations were tested. LB, TB, M9 and MOPS were the media tested for curcumin production in a single medium (see **Section 3.2** for media composition description) in which ferulic acid was directly added after protein expression time. In some tests, after 5 h of protein expression and before adding the substrate, the culture was centrifuged and resuspended in the same fresh medium.

When using M9, different concentrations of IPTG (0.1 mM, 0.5 mM, 1 mM or 1.5 mM) were tested to trigger protein expression. The $CaCO_3$ was not added to the M9 minimal medium in one experiment to study its relevance in the production of curcumin.

Additionally, some curcumin production experiments were also performed for the constructed (as described in **Section 3.7**) mutants (*E. coli* K-12 MG1655 (DE3) Δ*fumB*, *E. coli* BL21 (DE3) Δ*fumB* and *E. coli* K-12 MG1655 (DE3) Δ*gnd*) in the optimal production conditions, i.e. the ones that favor curcumin production by the wild-type strains.

All the experiments were performed in triplicate unless otherwise specified.

3.5.2 Curcumin extraction and samples preparation

During fermentation, samples were taken (1 mL) at the time the substrate was added (time 0) to confirm the initial concentration of substrate (ferulic acid). The treatment of this sample included a centrifugation (14000 g for 3 min) to remove most of the cells followed by filtration of the remaining supernatant with a 0.2 μm filter.

After 20 h, 43 h and 63 h, 500 μL were taken from the flasks in which curcumin was being produced by *E. coli* cells. To each sample, 10 μL of HCl 6M were added for acidification. Then, to extract curcumin from cells an equal volume (500 μL) of ethyl acetate was roughly mixed with each sample by vortexing. Afterwards, the mixture was centrifuged for 3 min at 14000 g. Two phases were formed (as can be seen in Figure 5) and the organic upper phase, containing curcumin, was collected. This procedure was repeated while the cells were still yellow. The samples were concentrated by evaporation of the ethyl acetate in a fume hood. Afterwards, the dried curcumin was resuspended in acetonitrile. The extracts (and the substrate) were further analyzed by HPLC.

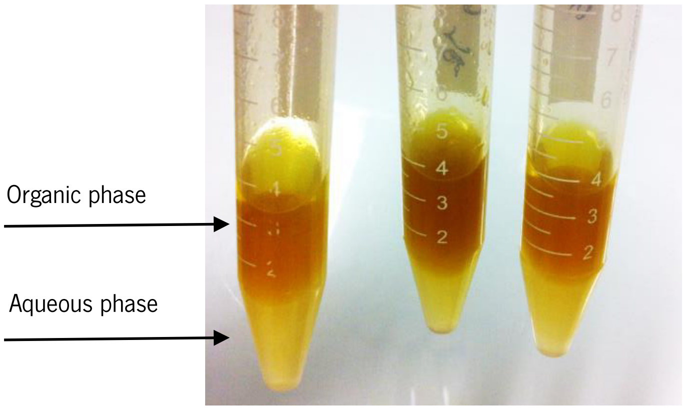


Figure 5. Curcumin extraction with ethyl acetate.

3.5.3 Substrate and product quantification by HPLC

HPLC was used to quantify the substrate and the curcumin produced. The chromatographic system used was the Shimadzu Nexera-X2 (Shimadzu Corporation, Kyoto, Japan) (CBM-20A system controller, LC-30AD pump unit, DGU-20A 5R degasser unit, SPD-M20A detector unit, SIL-30AC autosampler unit, CTO-20AC column oven) and the column was a Grace Alltech Platinum EPS C18 (3 μm, 150 mm × 4.6 mm) (Grace, Columbia, MD, USA). Mobile phase A was composed of water with 0.1% of TFA and the mobile phase B was acetonitrile. For ferulic acid quantification, the following gradient was used at a flow rate of 1 mL/min: 10 - 20% of acetonitrile (mobile phase B) for 17 min.

Quantification was based on the peak area at 310 nm for ferulic acid. In these conditions, the retention time of ferulic acid was 9.7 min. For curcumin quantification, a gradient of 40 - 43% of acetonitrile (mobile phase B) for 15 min and 43% of acetonitrile for more 5 min was used. Curcumin was detected at an absorbance of 425 nm at a retention time of 17.5 min. Ferulic acid and curcumin concentrations were determined using a calibration curve with known concentrations.

3.6 Curcumin toxicity assay

The toxicity of the commercially available curcumin (Acros Organics) against *E. coli* BL21 (DE3) was determined to identify the concentrations at which curcumin becomes toxic. The killing assays were conducted according to the work of Gudiña *et al.* (2010a, 2010b). A pre-culture of *E. coli* BL21 (DE3) (with or without the heterologous pathway) was grown overnight. Afterwards, the pre-culture was diluted 10x in LB and the OD_{600nm} was measured. The OD_{600nm} was then adjusted to 0.6 in LB 2x concentrated. A curcumin stock of 10 mM was prepared in DMSO and then diluted to 4000 µM in PBS 1x to minimize DMSO toxicity to the cells. A 96-well plate assay was prepared with different concentrations of curcumin in triplicate and several controls. In each well, 100 µL of the culture in LB 2x was added (except for the controls with only curcumin). Another 100 µL composed of curcumin and PBS were added to each well to a volume up to 200 µL. Curcumin at 4000 µM and PBS 1x were added in different amounts to get the intended curcumin concentration. Figure 6 shows the distribution of conditions through the 96-well plate: each column with a different curcumin concentration (1000 µM, 800 µM, 700 µM, 600 µM, 500 µM, 400 µM, 300 µM, 200 µM, 100 µM and finally the control with bacteria without curcumin). Moreover, a control was prepared only with 100 µL LB 2x + 100 µL PBS 1x in the last column. The indicated lines contained the triplicates for both wild-type and curcumin producing strain. The plate was incubated at 37°C and 180 rpm for bacteria growth and the OD_{600nm} was measured every hour for 8 h.

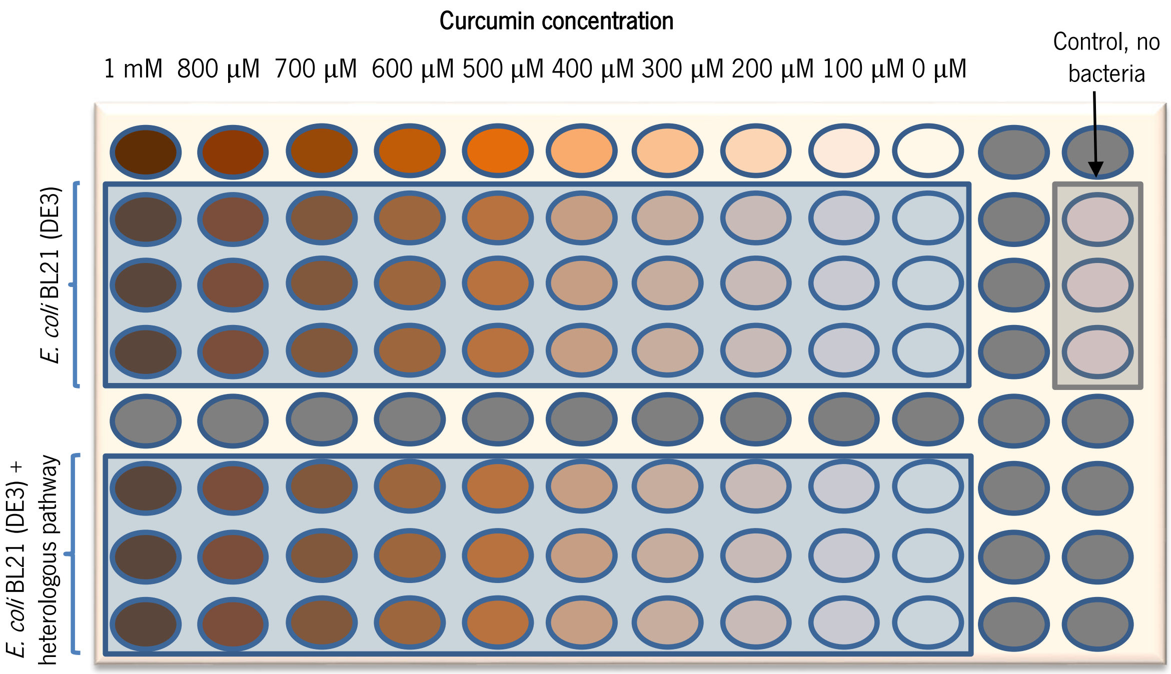


Figure 6. 96-wells plate design for curcumin toxicity assay. Wells from each column contain decreasing curcumin concentrations, being 0 µM curcumin the negative control in which cells were grown without curcumin. In the top line, only curcumin and no cells were added to the wells in order to define the blank. The three lines indicated for both wild-type and curcumin producing strain contained triplicates for the conditions. The control in the last column contained 100 µL of the medium LB 2x and 100 µL of PBS 1x to check for contamination.

3.7 Gene deletion in *E. coli*

Three methodologies were tested for the construction of the mutants predicted to favor curcumin production (e.g. Δ*gnd*, Δ*fumA,fumB,fumC*). One of the recombineering strategies used consisted in an optimized procedure of the homologous recombination-mediated genetic engineering, a Lambda Red–mediated, linear DNA–based deletion method, well described by Kolisnychenko *et al.* (2002) and Fehér *et al.* (2007). The procedure and the primers designed for this strategy are described in **Appendix I**. A second recombineering approach was implemented to delete *fumA* and *fumC* (simultaneously) from *E. coli* K-12 MG1655 (DE3) genome involving a Tet^R cassette integration (landing pad), as described more recently in the literature (Kuhlman & Cox, 2010; Tas *et al.*, 2015). The procedure and the primers designed for this strategy are described in **Appendix II**.

Finally, the gene deletion method using CRISPR-Cas9 system (Reisch & Prather (2015)) was used. All the primers required for this strategy are summarized in Table 3.

Table 3. Primers used for scarless deletion using the CRISPR-Cas9 system

Gene to be deleted	Primer name	Primer Sequence 5'-3'
fumB	K_fumB_A	GGCGCTGGACATGTACAAAGgttttagagctagaaatagcaag
	K_fumB_B	CTTTGTACATGTCCAGCGCCgtgctcagtatctctatcactga
	fumB_D	CATCAACCACAGCTTTATTCTG
	fumB_E	GAATTAACGGCTTGAGTGTG
fumA and fumC	K_fumAC_A	GTACGCAGCGAAAAAGATTCGAgttttagagctagaaatagcaag
	K_fumAC_B	TCGAATCTTTTTCGCTGCGTACgtgctcagtatctctatcactga
	fumA_D	GCAAAAAGTCGTACTAGTCTC
	fumAC_E	TACACCACTGATTGCACTGAT
ccmA	ccmA_A	TATTCCTGTAAATCAGCTCTCGGCgttttagagctagaaatagcaag
	ccmA_B	GCCGAGAGCTGATTTACAGGAATAgtgctcagtatctctatcactga
	ccmA_D	GCCGTAACTGTCACAACTTC
	ccmA_E	CAGATAATGCCCGGTGCAATAC
argO	argO_A	GTTGTACTGGGCAGCCTTGGgttttagagctagaaatagcaag
	argO_B	CCAAGGCTGCCCAGTACAACgtgctcagtatctctatcactga
	argO_D	CTGCGTAATCAACGCTGAGG
	argO_E	CGGAGGTGACAATATGCGGT
gnd	K_gnd_A	CAAAATCGTTTCTTACGCCCgttttagagctagaaatagcaag
	K_gnd_B	GGGCGTAAGAAACGATTTTGgtgctcagtatctctatcactga
	gnd_D	TCAGCATCCCGGTAGGCTTCA
	gnd_E	CGGTTGTTGATTGGTGCGCAGG
fumA	K_fumA_A	AAAGGGCAGCGTGTATGGACgttttagagctagaaatagcaag
	K_fumA_B	GTCCATACACGCTGCCCTTTgtgctcagtatctctatcactga
	fumA_D	GCAAAAAGTCGTACTAGTCTC
	fumA_E	GAATCTTTTTCGCTGCGTAC
fumC	K_fumC_A	GACCTGTGATGCCCTGGTTCgttttagagctagaaatagcaag
	K_fumC_B	GAACCAGGGCATCACAGGTCgtgctcagtatctctatcactga
	fumA_D	GCAAAAAGTCGTACTAGTCTC
	fumC_E	GATCAAAAACAAGTCCAACACG
-	pKD_Seq5	CAGTGAATGGGGGTAAATGG
	K_Reverse	TTTATAACCTCCTTAGAGCTCGA
	K_Forward	CCAATTGTCCATATTGCATCA

The lowercase letters represent the parts of the sequence with homology with the pKDsgRNA-p15 (template) plasmid.

This was the only technique that allowed the gene deletion in *E. coli* K-12 MG1655 (DE3) and *E. coli* BL21 (DE3), generating the mutants *E. coli* K-12 MG1655 (DE3) Δ*fumB* and *E. coli* BL21 (DE3)

Δ*fumB*. The mutant *E. coli* K-12 MG1655 (DE3) Δ*gnd* was already constructed through recombineering in a previous work from our group (Araújo, 2014). Thus, in this work, this deletion was confirmed and the mutant was tested for curcumin production.

Primers D and E are the forward and reverse primers, respectively, used to verify the deletion by colony PCR. The oligonucleotides (Oligo_gene) used to insert the desired mutation contained 37 bp from the region before the target gene plus 36 bp from the region after. The same criterion was used in the design of all the oligonucleotides that are presented in Table 4.

Table 4. Oligonucleotides used for scarless deletion using the CRISPR-Cas9 system

Gene to be deleted	Oligonucleotide name	Primer Sequence 5'-3'
fumB	Oligo_fumB	TTTCGAATAACAAATACAGAGTTACAGGCTGGAAGCT**CCTCTTCGGCCCAGCGCCTGGCAGCATGCTGCCAGG**
fumA and fumC	Oligo_fumAC	CAGAGCATAACCAAACCAGGCAGTAAGTGAGAGAACA**TCTGCAACATACAGGTGCAGCCGTGGAATGATCAAA**
ccmA	Oligo_ccmA	GCGTGAAGTCGAGCCAGGTTTTTAACAGGGTTATTGC**TGTTCTGGCGCATTTTCCGTCTTGAGCTGCGTGTAG**
argO	Oligo_argO	AGTCTCTGCATCAGATACTTAATTCGGAATATCCAAC**GCGCTATCTGATGGAAAAATAAAACAGAGGCGCTAA**
gnd	Oligo_gnd	CATTCAGCGCGGTGATCACACCTGACAGGAGTATGTA**TCTGATTTAACCAACAATAAAATTGAGGCCCGGCGT**
fumA	Oligo_fumA	CAGAGCATAACCAAACCAGGCAGTAAGTGAGAGAACA**ACAGAGCCGCCCTTCGGGGCGGTTTTTTTACATGGC**
fumC	Oligo_fumC	AACAGAAAGAAAAAATTAATCAGGTGAGGAGCAGGTC**TCTGCAACATACAGGTGCAGCCGTGGAATGATCAAA**

The nucleotides from the genome immediately before and after the target gene are underlined or at **bold,** respectively.

The first step of the protocol for gene deletion using the CRISPR-Cas9 system is the cloning of the protospacer into the pKDsgRNA plasmid, and this was done using CPEC (Circular Polymerase Extension Cloning) (Tian & Quan, 2009; Quan & Tian, 2014). The protospacer consists in a 20 bp region of the gene that has to be deleted and it was designed to be specific (this was confirmed by blastn (https://blast.ncbi.nlm.nih.gov)). To clone the protospacer into the pKD plasmid through CPEC, four primers were used. Two PCR reactions were done to amplify the 7 kb plasmid in two parts, and then the two parts were mixed and the CPEC reaction was set. pKDsgRNA_*fumB*, pKDsgRNA_*fumAC*, pKDsgRNA_*ccmA*, pKDsgRNA_*argO*, pKDsgRNA_*fumA* and pKDsgRNA_*fumC* were constructed through this approach using the primers described in Table 3. All plasmids were sequenced using primer pKD_seq5 (GATC Biotech, Konstanz, Germany).

For example, for *fumB* gene deletion, K_fumB_A and K_reverse were firstly used to amplify the first part (2848 bp) of the pKDsgRNA-p15 plasmid. The amplification conditions are described in Table 5. In order to amplify the second part (4414 bp), the primers used were K_fumB_B and K_forward and the conditions are also specified in Table 5.

During the amplification of the two parts of the plasmid by Phusion High-Fidelity DNA Polymerase (Thermo Scientific, Waltham, USA), 10 ng of pKDsg-p15 were used as template in a 25 µL reaction.

Table 5. PCR conditions to amplify the two parts of the pKDsgRNA plasmid for *fumB* deletion using Phusion HiFi DNA polymerase

Cycling step		Part 1 (2848 bp)	Part 2 (4414 bp)
Initial denaturation		98°C, 30 s	
Denaturation	15 x	98°C, 10 s	
Annealing		59°C, 30 s	61°C, 30 s
Extension		72°C, 86 s	
Final extension		72°C, 5 min	

After the amplification of the two parts of the plasmid, both PCR products were digested with 0.5 µL of *Dpn*I restriction enzyme (Thermo Scientific) for each 25 µL reaction for at least 2 h at 37°C. Afterwards, the DNA was mixed with 5 µL of 6x loading dye and run in 1% (w/v) agarose gel. Afterwards it was purified from a gel using NucleoSpin® Gel and PCR clean-up (Macherey-Nagel, Düren, Germany). The CPEC reaction (50 µL) was set with Phusion High-Fidelity DNA Polymerase, 129 ng of the Part 1 and 200 ng of the Part 2 due to their size difference, therefore the number of molecules were the same, and no primers were added to this reaction. The PCR program, in which a slow ramp annealing was used, was set as described in Table 6.

Table 6. Setting up PCR conditions for the construction of the pKDsgRNA_fumB using CPEC. Phusion Hifi DNA polymerase was used in this reaction

Cycling step		Condition
Initial denaturation		98°C, 30 s
Denaturation	15 x	98°C, 10 s
Slow ramp annealing		70 to 55°C, 0.1°C/s
Annealing		55°C, 30 s
Extension		72°C, 140 s
Final extension		72°C, 5 min

After PCR amplification, 10 μL of the resulting product were run in a 1% agarose gel to confirm the size (6962 bp). Then, 100 μL of *E. coli* NZY5α competent cells were transformed with 10 μL of the CPEC-resulting plasmid. The colonies were plated in LB agar supplemented with Spec (50 μg/L), and were grown at 30°C for 20 h. Afterwards, a colony was patched into 10 mL of LB supplemented with Spec and the plasmid was extracted using the kit NucleoSpin® Plasmid (NoLid) (Macherey-Nagel) and sent for sequencing (GATC Biotech) with the primer pKD_Seq5 to confirm the presence of the designed protospacer.

The desired host strains (*E. coli* K-12 MG1655 (DE3) and *E. coli* BL21 (DE3)) were transformed with the pCas9-CR4 plasmid as described in **Section 3.3**. After transformation, cells were spread in LB plates with Cm and grown at 37°C. Next, a colony from each strain was grown at 37°C in LB with Cm. These cells were then made electrocompetent and transformed (as described in **Section 3.3**) with the new pKDsgRNA_*gene* plasmid which contains the protospacer. The cells recovered in SOC at 30°C for 2 h and were then spread on LB+Spec+Cm plates. Around 1 to 3 big colonies were taken from the plate and grown in liquid LB+Spec+Cm at 30°C. At an OD_{600nm} 0.4–0.5, L-arabinose (1.2% (w/v) final concentration) was added to induce λ-Red recombinase and the culture was incubated for more 15-20 min. Then, the cells were made electrocompetent and were further transformed with about 1000 ng of the oligo_fumB. After electroporation, 1 mL of SOC was immediately added, the cells recovered at 30°C for 2 h, before plating serial dilutions on LB with Cm, Spec, and 100 ng/mL aTc, that induces Caspase 9 expression. Afterwards, the plates were incubated at 30°C overnight. The correct mutation was verified by colony PCR with KAPA *Taq* DNA Polymerase (Kapa Biosystems, Boston, USA) using the primers fumB_D and fumB_E (PCR conditions in Table 7) and it was also confirmed by sequencing (GATC Biotech) with primer fumB_D.

Table 7. PCR conditions for colony PCR to verify the deletion of *fumB* using primers fumB_D and fumB_E and KAPA *Taq* DNA Polymerase

Cycling step		Condition
Initial denaturation		95°C, 10 min
Denaturation	35 x	95°C, 30 s
Annealing		47°C, 30 s
Extension		72°C, 115 s
Final extension		72°C, 5 min

After mutation, it was necessary to cure the transformed plasmids. The pKDsgRNA_*fumB* plasmid was cured first by growing in LB at 37°C for 4-24 h. Then, the culture was plated and incubated at 37°C. Spec sensitivity was checked by platting also in LB supplemented with Spec. In the cases where double or triple knock outs were envisaged, the cells were transformed with the correspondent pKDsgRNA_*gene* plasmid and the process was repeated. To cure the pCas9-CR4 plasmid, cells were transformed (as described in **Section 3.3**) with pKD-kill plasmid (pKDsgRNA-p15), which is Spec resistant and temperature sensitive. Afterwards, a colony was induced with aTC in liquid culture and plated on Spec + aTC at 30°C. Finally, individual colonies were patched on LB Cm+ / Cm- at 37°C twice to make sure that the pKD plasmid was also cured.

4. RESULTS AND DISCUSSION

4.1 Curcumin production optimization

The optimization of the fermentation conditions for the production of curcumin by *E. coli* strains harboring the artificial biosynthetic pathway was herein conducted. Additionally to the standard LB+M9 combination, LB+LB, LB+MOPS and TB+MOPS combinations were tested. LB, TB and MOPS were the media tested for curcumin production in a single medium. In some cases, the medium was changed to a fresh medium with the same composition at the substrate addition step to evaluate the influence of the nutritive content. Different strains were also tested, as well as different IPTG concentrations.

After pCDFDuet_DCS, pRSFDuet_CURS1 and pAC_*At*4CL1 simultaneous transformation in *E. coli* cells, the mixture of plasmids was extracted and digested with *Hind*III and their presence and correct sizes were confirmed (*data not shown*).

During the fermentation, the production of curcuminoids by the *E. coli* cells was visually suggested by the yellow color of the culture, when grown in liquid medium, as well as by the colonies when grown on solid medium.

4.1.1 Evaluation of the performance of different *E. coli* strains

Firstly, different strains were tested for the production of curcumin. In addition to the *E. coli* K-12 MG1655 (DE3), that was used in the previous works from Rodrigues *et al.* (2015a;c), *E. coli* BL21 (DE3) and *E. coli* JM109 (DE3) strains were also evaluated. The results for these three *E. coli* strains are gathered in Figure 7. Different OD_{600nm} were used depending on the strain based on preliminary tests, but the optimal OD for the best strain was afterwards studied with more detail (**Section 4.1.2**).

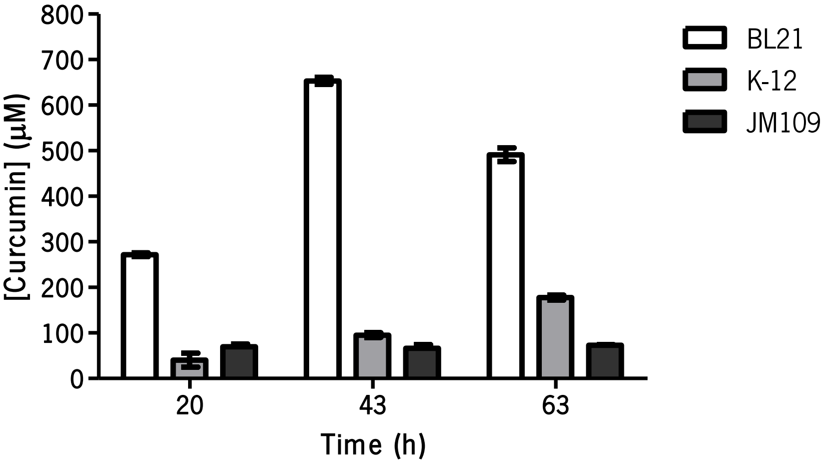


Figure 7. Curcumin production over time using *E. coli* K-12 MG1655 (DE3), *E. coli* BL21 (DE3) and *E. coli* JM109 (DE3) harboring the heterologous biosynthetic pathway. The medium lysogeny broth (LB) was used for cell growth and then cells produced curcumin in M9 minimal medium. The optical density 600 nm at the moment of induction was 0.6 for BL21 and JM109 and 0.4 for K-12. IPTG (Isopropyl β -D-1-thiogalactopyranoside) concentration was 1 mM in both LB and M9 media. Standard deviations are indicated by error bars.

The curcumin production from ferulic acid (2 mM) by *E. coli* K-12 MG1655 (DE3) in this work (177.9±4.3 μM) was similar to the previously obtained in the work of Rodrigues *et al.* (2015c) in the same conditions (187.9 μM=70 mg/L), as expected. On the other hand, *E. coli* BL21 DE3 was able to produce a curcumin maximum of 653.2±6.4 μM at 43 h. The difference between curcumin production by *E. coli* K-12 MG1655 (DE3) and *E. coli* BL21 (DE3) is statistically relevant in the three time points studied (*p*-value <0.001). *E. coli* JM109 (DE3) was also used for the production of curcumin. JM109 is a K strain that is *recA*– and *endA*– to minimize recombination and improve the quality of plasmid DNA (Casali, 2003). However, the curcumin concentration obtained by this strain was even lower (73.2±1.1 μM) than the one obtained with the K-12 strain. Katsuyama *et al.* (2008) reported the production of 113±22 mg/L (306.8 μM) of curcumin after 60 h from ferulic acid using *E. coli* BLR (DE3), a BL21 related strain (Table 1), carrying *Le*4CL, CUS and ACC as producer strain and the LB and M9 media combination. Therefore, herein the *E. coli* BL21 (DE3) produced 2.1 times more curcumin in only 43 h than the obtained by Katsuyama *et al.* (2008).

Hence, curcumin production can be highly influenced not only by the genes used for the heterologous pathway construction but also by the host strain. B strains, such as *E. coli* BL21 and BLR, as mentioned before, have higher gene expression levels related to the acetate metabolism and glyoxylate shunt, so there is less acetate formation, triggered by excess glucose supply during batch cultivations compared to the K–12 strains, like JM109 or K-12 MG1655 (Shiloach *et al.*, 1996; Marisch *et al.*, 2013a; Marisch *et al.*, 2013b). Besides, most of B strains cells are deficient in the Lon protease, which degrades many foreign proteins (Gottesman, 1996). Also, their genome misses one gene coding for the outer membrane protease OmpT, whose function is to degrade extracellular proteins and is critical because after cell lysis the recombinant proteins would be degraded (Grodberg & Dunn, 1988). In addition, plasmid loss is prevented thanks to the *hsd*SB mutation in BL21 (DE3) (Rosano & Ceccarelli, 2014), a mutation that disrupts the system of methylation and restriction that allow *E. coli* to recognize DNA as foreign.

All of these can justify the higher performance of *E. coli* BL21 (DE3) as protein expression host and curcumin producer compared to the *E. coli* K-12 MG1655 (DE3) or JM109 (DE3). As *E. coli* BL21 (DE3) was identified as the best curcumin producing strain, the following optimizations were done using this strain.

From the time point of 20 h to 43 h and 63 h, curcumin production was not increased when using *E. coli* JM109, at least at a statistically relevant level (*p*-value >0.05). The opposite happened when *E. coli* K-12 MG1655 (DE3) was used, since curcumin concentration was increased over time with differences statistically significant (*p*-value <0.001). *E. coli* BL21 (DE3) produced a maximum curcumin concentration at 43 h but it decreased at 63 h, with a determined statistically significant difference (*p*-value <0.001). However, when using this last strain in different assays (see **Sections 4.1.2** to **4.1.6**), the curcumin concentration did not usually change significantly from 43 h to 63 h. Therefore, in this particular case, the difference on curcumin production between 43 and 63 h may be due to curcumin extraction errors. However, the curcumin degradation can also be considered since it is reported to be highly instable and sensitive to light and pH (Nugroho *et al.*, 2009).

4.1.2 Establishment of the optimal optical density for induction of BL21 (DE3) in the process using two culture media

The optimal OD_{600nm} for IPTG induction in LB and M9 combination using *E. coli* BL21 (DE3) harboring the artificial biosynthetic pathway was also studied to improve the production of curcumin. Optimal induction OD_{600nm} for *E. coli* K-12 MG1655 (DE3) was previously studied in the work of Rodrigues *et al.* (2015) and it was concluded that for that strain, the optimal OD_{600nm} for IPTG induction was 0.4. Herein, the effect of different OD_{600nm} at the induction time in the production of curcumin by *E. coli* BL21 (DE3) is shown in Figure 8.

Also, for one of the conditions tested (induction at OD_{600nm} 0.8) the $CaCO_3$ (5 g/L) was not added to the M9 medium to evaluate its influence in the pH control and in the production of curcumin. The $CaCO_3$ is used in the medium to maintain a neutral pH since curcumin production acidifies the culture This particular parameter was evaluated because Katsuyama *et al.* (2008) used a very high amount of $CaCO_3$, 25 g/L, in their work during curcumin production, while Rodrigues *et al.* (2015a,c) used 5 g/L of $CaCO_3$. The necessity for this compound in such high level would represent a substantial cost in industrial facilities.

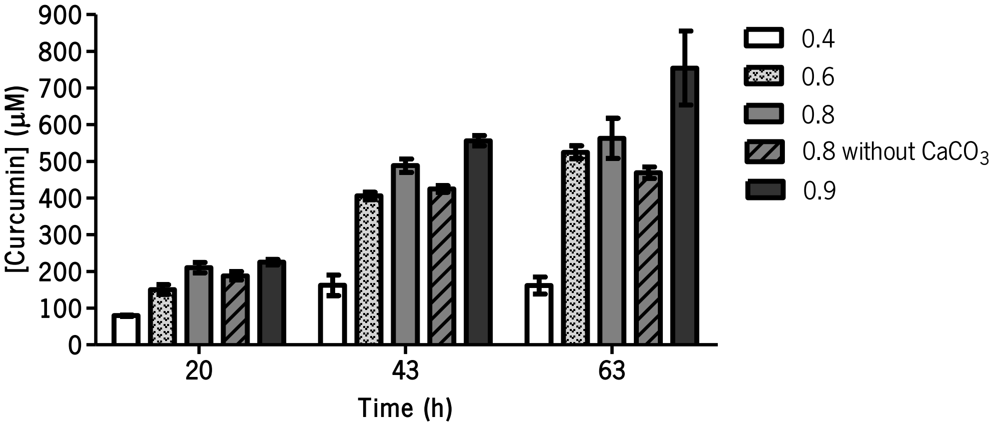


Figure 8. Effect of optical density (600 nm) at time of induction of protein expression in *E. coli* BL21 (DE3) in the production of curcumin. The LB and M9 combination was used. IPTG (Isopropyl β-D-1-thiogalactopyranoside) was used at a final concentration of 1 mM in both LB and M9 media. Standard deviations are indicated by error bars.

At 20 h from the beginning of the fermentation, only the difference between inducing at an OD_{600nm} of 0.4 and 0.9 was statistically significant (*p*-value <0.001). At 43 h and 63 h it was possible to conclude that the OD_{600nm} 0.9 is the best for the production of curcumin, resulting in statistically higher curcumin concentrations (*p*-value <0.001). Curcumin concentration obtained in this case was 754.3±82.3 µM. This can be due to the higher biomass formed at the time of induction. More cells should produce more curcumin. However, when inducing at the late-logarithmic or stationary phase, the metabolic state of the cells can be adverse for protein expression and stress-related proteins, such as proteases are produced and can reduce the yield of foreign proteins (Carneiro *et al.*, 2013). This influence of OD_{600nm} at which cells were induced in case of K-12 clearly showed different results (Rodrigues *et al.*, 2015c) probably because, differently to BL21 strain, K-12 expresses *lon* and *ompT* protease genes.

Comparing the cases where the OD_{600nm} of induction was the same (0.8), the $CaCO_3$ absence exhibited no significant difference in the production of curcumin. The difference becomes significant with a *p*-value <0.01 at 63 h. Therefore, these results suggest that the curcumin production by *E. coli* BL21 (DE3) can benefit from the addition of $CaCO_3$, which is in accordance with previous reports (Katsuyama *et al.*, 2008). However, in the previous work from Rafael (2014), a 15 g/L concentration of $CaCO_3$ was evaluated for curcumin production with *E. coli* K-12 MG1655 (DE3) and the curcumin concentration obtained was substantially lower than the achieved when using a 5 g/L concentration.

Therefore, the authors had already concluded that higher levels of $CaCO_3$ are probably not required for curcumin production.

4.1.3 Evaluation of different IPTG concentrations for induction of the protein expression

Generally, IPTG was added at a 1 mM final concentration in LB for the induction of protein expression. Also, 1 mM of IPTG was added to the culture medium when the cells were harvested and put to grow in the M9 medium. The optimization of IPTG concentrations (0.1, 0.5, 1 and 1.5 mM) to trigger protein expression in LB was conducted to determine which concentration allows optimal enzyme synthesis and consequently, optimal substrate conversion. Additionally, the effect of not adding IPTG to the fresh M9 medium (time 0 h) was studied (1-0 mM of IPTG). These results are shown in Figure 9.

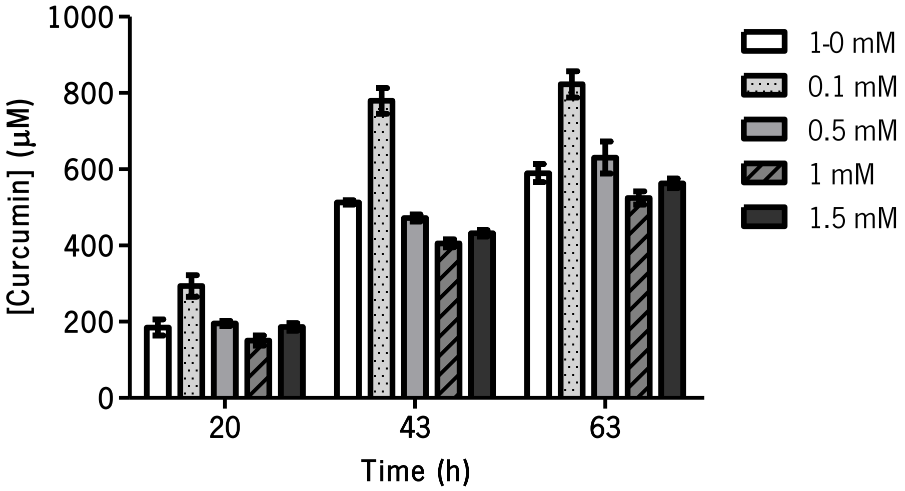


Figure 9. Curcumin production by *E. coli* BL21 (DE3) using LB (lysogeny broth) and M9 media and different IPTG (Isopropyl β-D-1-thiogalactopyranoside) concentrations. A final IPTG concentration was added to LB for protein expression: **0.1**, **0.5**, **1** or **1.5 mM** and the same correspondent concentrations (0.1, 0.5, 1 or 1.5 mM) were added to the M9 medium except in one of the assays where no IPTG was added to the M9 medium (**1-0 mM**). Protein expression was induced when cells reached an optical density at 600 nm of 0.6. Standard deviations are indicated by error bars.

The experiments represented in Figures 7, 8 and 9 were performed in parallel to assure similar physiological states of the cells and enable the comparison between the experiments. Differences observed between results from Figures 7, 8 and 9 using the same conditions, especially at 43 h, can be due to extraction and quantification errors, but also to slight differences in medium compositions or the overall cells physiological state as the different groups of experiments were performed in different weeks.

It was not expected to see a dose-dependent expression when using IPTG as inducer. Since IPTG can enter the cell by active transport (Fernández-Castané *et al.*, 2012) through Lac permease and the expression of Lac permease is not homogenous and the number of active permeases in each cell is highly variable, protein expression does not respond predictably to IPTG concentration (Rosano & Ceccarelli, 2014). However, it was notable that the highest curcumin production (822.6±28.1 µM at 63h in LB+M9 inducing protein expression with 0.1 mM of IPTG when cells reached an OD_{600nm} of 0.6) was obtained when using 10 times less IPTG than the usual 1 mM, and the improvement is statistically relevant comparing to any other conditions (*p*-value <0.001). Thus, it was possible to confirm that there is no dose-dependent effect of IPTG in curcumin production, but decreasing the IPTG concentration for protein expression can probably reduce the metabolic burden.

When comparing the case where IPTG was not added to the culture at time 0 h in M9 (1-0 mM IPTG) to the one where the IPTG was added to a 1 mM concentration (1 mM IPTG) the differences are not statistically significant in any of the time points. Probably, the enzyme expression is minimal during the curcumin production phase because the physiological condition of the cells is not so suitable for protein expression after the exponential phase (Carneiro *et al.*, 2013; Rodrigues *et al.*, 2015b), thus the addition of IPTG cannot improve significantly the enzyme expression.

The assays to seek for appropriate OD_{600nm} and the ones to find the optimal IPTG concentration were done simultaneously. In future work, new assays should be performed with both optimal OD_{600n} (0.9) and IPTG concentration (0.1 mM) to evaluate if curcumin production can be improved.

4.1.4 Optimization of operational conditions

Contrary to most of the works reported in the literature, a single medium was evaluated for the entire fermentative assay and different medium compositions were tested in order to avoid the medium exchange in the middle of the process. This would make the curcumin production easier and economically more interesting for an industrial scale-up. With this purpose, TB, LB and MOPS were tested since they have already been successfully used for the heterologous expression and production of (2S)-pinocembrin (Wu *et al.*, 2013) and other plant flavonoids (Leonard *et al.*, 2008). Besides, media combinations different from the typical LB and M9 combination were used to find a combination that provides higher production rates. In some cases, the exchange to fresh media (LB+LB and TB+TB) of the same composition was also performed to study the influence of the nutritive content during the production of curcumin. This allows evaluating if the fed-batch operation mode would be an interesting approach for the industrial production. The results of curcumin production with *E. coli* BL21 (DE3) in different media compositions are shown in Figure 10.

It is important to notice that, as preliminary optimizations with *E. coli* K-12 MG1655 (DE3) using LB medium from the start led to a residual curcumin production of 9.9±0.4 µM (*data not shown*), this condition was not evaluated in the experiments conducted with *E. coli* BL21 (DE3).

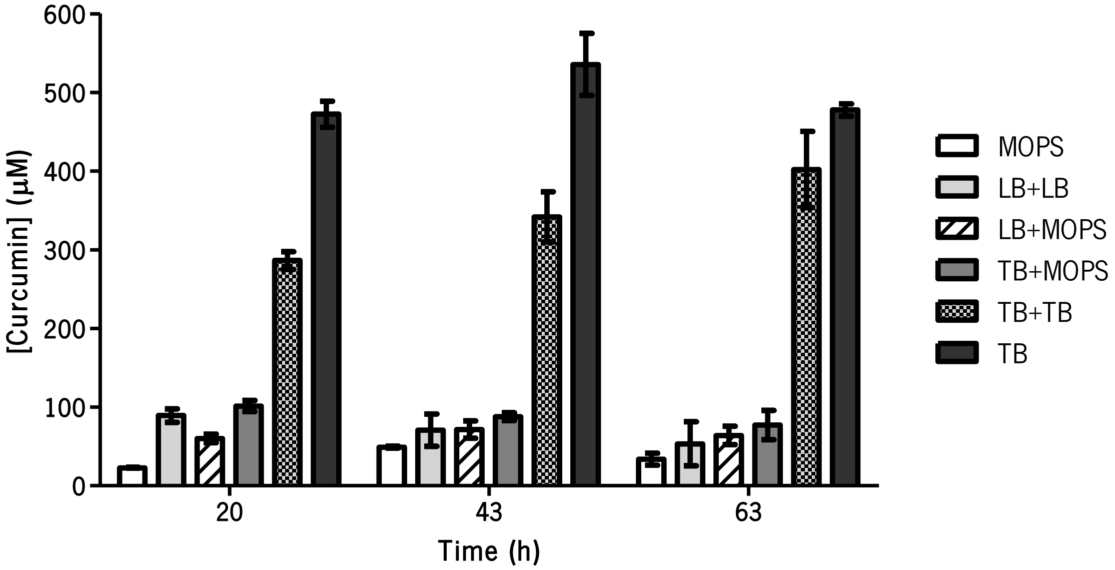


Figure 10. Curcumin production over time in different media compositions by *E. coli* BL21 (DE3). Cells were grown in: MOPS (morpholinopropane-1-sulfonic acid) minimal medium (**MOPS**); LB (lysogeny broth) broth with this medium being changed to new LB when the substrate was added (**LB+LB**); LB followed by MOPS medium (**LB+MOPS**). TB (terrific broth) and MOPS combination (**TB+MOPS**); TB refreshing to new TB when the substrate was added (**TB+TB**); and using only TB (**TB**). Protein expression was induced with 1 mM of IPTG (Isopropyl β-D-1-thiogalactopyranoside) when cells reached an optical density at 600 nm of 0.6. Standard deviations are indicated by error bars.

TB seems to be the best medium for the production of curcumin within the single media tested, reaching the highest curcumin production at 43 h (535.8±59.2 µM). Also, using different combinations of media did not result in any improvement for curcumin production. In fact, a two-way ANOVA test was done and it showed that the difference between TB and other media was always statistically significant (*p*-value < 0.001). Even when changing to fresh TB medium (TB+TB) the curcumin production was not significantly enhanced comparing to when using the entire time the same TB medium (TB), which was statistically better at 20h and 43h. Therefore, the higher nutritive content at the curcumin production phase seems to inhibit the reactions that produce curcumin in TB. Since after medium exchange more carbon sources (present in yeast extract) are present, cells growth and acetate synthesis is higher, which can lead to acetate over-accumulation and toxicity. Therefore, at this phase when protein expression is being induced, minimal carbon concentration should be available to minimize the use of cellular resources for cell growth and acetate formation as referred by several authors (Rosano & Ceccarelli, 2014; Wang *et al.*, 2014).

TB is a phosphate buffered rich medium. Compared to LB, TB has a slightly higher amino acids concentration (more tryptone) and 480% more yeast extract supplying large quantities of growth factors and nucleic acid precursors. TB also has 0.04% (v/v) glycerol as an extra carbon source, which is not metabolized into lactic acid (note that the usual concentration for this medium is 0.4%, but 0.04% was generally used in this study based on preliminary optimizations). TB also contains phosphate buffer to maintain a physiological pH during exponential growth. Since the pH is maintained bacteria grow more and it's lyse occurs more slowly, which allows the maintenance of plasmid DNA. Buffering tends to be important if the media contains a carbon source, such as glucose, which is converted into organic acids (acetate) resultant of aerobic metabolism. However, in TB, the carbon sources that generate these toxic compounds are limited. Glycerol is the main carbon source for bacteria growth in this medium, which is an advantage as it does not create carbon catabolite repression (CCR), a phenomenon that prevents consuming other carbon sources other than glucose. Thus, the carbon flux can be distributed between glycolysis and gluconeogenesis which avoids the accumulation of toxic TCA intermediates (Martínez-Gómez *et al.*, 2012). TB is particularly advantageous for curcumin production compared to the previously used LB and M9 media combination (Rodrigues *et al.*, 2015c) because the process becomes a lot easier, cheaper and more efficient in time and in energy consumption, since the use of two medium and separation of biomass in no longer needed. Besides, glycerol (a carbon source in TB) is a byproduct of biodiesel manufacture (10 %, w/w) (Pagliaro *et al.*, 2007; Silva *et al.*, 2009) so its availability has been increasing with the improvement of biodiesel production from triacylglycerol oils. As a waste from biodiesel industries, the utilization of glycerol as a carbon source constitutes a biotechnological advantage and a cheaper substrate.

In most of the conditions tested, curcumin production does not improve after the 20 h with a statistically significant difference (*p*-value is usually >0.05). Only in the case when MOPS medium was used, the curcumin production improved significantly between 20 h and 43 h. However, in general, even not at a significant difference, the curcumin seems to have its maximal concentration at 43 h. This was observed not in every but in most of the assays. In order to study what could happen to cells viability or the percentage of cells that express curcumin along the time, a representative assay was made in which cells from the flasks where curcumin was being produced were spread in LB Agar plates supplemented with Spec+Cm+Kan and 2 mM of ferulic acid. The plates contained no extra IPTG to verify if the cells were still able to produce curcumin along time under the conditions in which they were collected. Figure 11 shows the resulting aspect of the colonies removed from the flasks at time 0 h and 24 h after substrate addition in MOPS. Cells producing curcumin were yellow/orange.

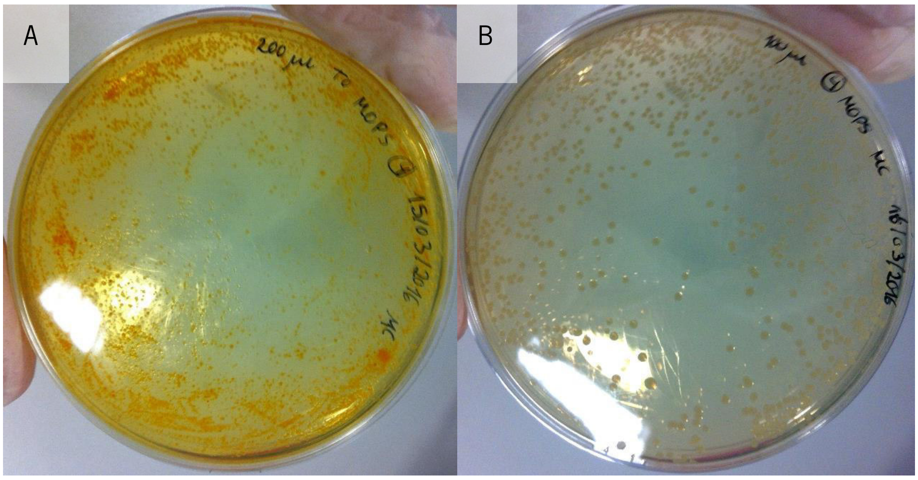


Figure 11. Curcumin production in lysogeny broth (LB) agar supplemented with the required antibiotics and 2 mM of ferulic acid. 200 μL from the flask immediately after substrate addition were spread in plates (**A**) and 100 μL were removed from the same flask after 24 h of incubation with substrate and spread (**B**). The flasks contained cultures growing in MOPS (morpholinopropane-1-sulfonic acid) minimal medium and the protein expression was induced at an optical density at 600 nm of 0.6 with 1 mM of IPTG (Isopropyl β-D-1-thiogalactopyranoside). The substrate ferulic acid was added to the flasks after the usual 5 h of induction.

It is visible that in the initial hours, the curcumin production is at maximum rate. All the cells appear to be producing curcumin (Figure 11A). However, the production becomes lower with increasing time (Figure 11B). This study allowed concluding that the decrease of curcumin production along time is not due to the antibiotic degradation and consequent loss of any of the plasmids because we observed that cells continued resistant to the three antibiotics.

After the 5 h of induction, cells growth is maintained in MOPS minimal medium. Nevertheless, when the exponential phase is finished, they are not expressing proteins anymore and the curcumin production is limited as suggested by the absence of yellow colonies. In fact, it is well-known that the protein expression induction is maximal during the mid-logarithmic growth phase when cells can provide sufficient levels of energy and metabolic precursors for the recombinant biosynthesis, since the cell is at its maximum catabolic capacity (Carneiro *et al.*, 2013; Rodrigues *et al.*, 2015b). However, if induced at the late-logarithmic or stationary phase there is a higher cell density for product formation, but the metabolic state of the cells is unfavorable and the presence of stress-related proteins like proteases can reduce the yield of foreign proteins (Carneiro *et al.*, 2013). For that reason, the induction of enzyme expression is not adequate after cells enter the stationary phase that corresponds to the end of the 5 hours of induction, so curcumin production might be residually improved probably because the enzymes already expressed still convert the ferulic acid into curcumin but no more enzymes are being produced. Thus, since after 5 hours the cells do not express enzymes at the same rate, the curcumin production is almost compared to an *in vitro* system.

4.1.5 Assessment of different optical densities for the induction in TB

Since the TB was the medium that allowed higher curcumin production by a single medium, several OD_{600nm} for IPTG-induced expression were tested with this medium in order to optimize the curcumin concentration. These results are gathered in Figure 12.

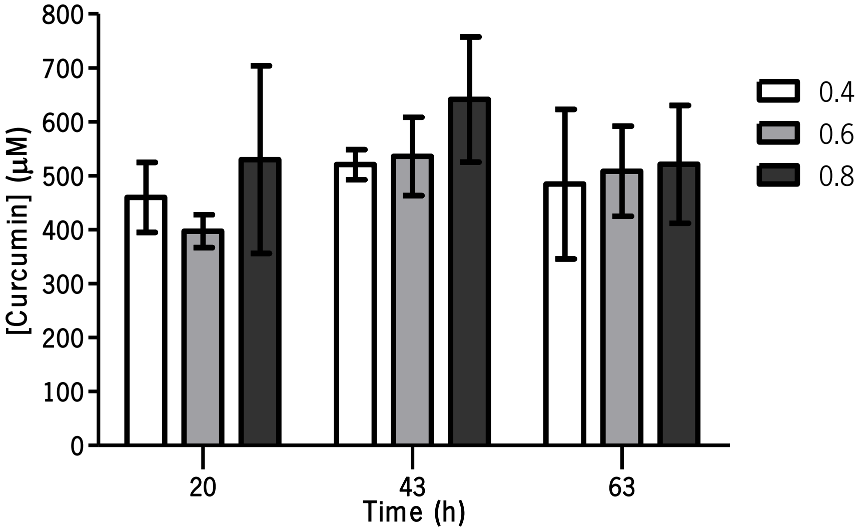


Figure 12. Curcumin production over time by *E. coli* BL21 (DE3) in TB (terrific broth), induced at different optical densities (0.4, 0.6 and 0.8). IPTG (Isopropyl β-D-1-thiogalactopyranoside) concentration for protein expression induction was 1 mM. Standard deviations are indicated by error bars.

When *E. coli* BL21 (DE3) cells grew in TB and the protein expression was induced with 1 mM IPTG at different OD_{600nm}, the differences in the production of curcumin were not statistically significant in none of the time points (*p*-value >0.05). In addition, the reproducibility of curcumin production in this rich medium was low (comparing to the LB and M9 combination), generating high standard deviations within the same conditions.

Through the analysis of the curcumin production over time it was possible to conclude that no difference between time points was statistically significant (*p*-value >0.05) in none of the tested conditions, due to the high standard deviations. However, the results suggest that the curcumin concentration is higher at 43 h, which is in accordance with the previous results obtained for other conditions.

4.1.6 Improvement of curcumin production in TB by adding extra carbon sources

In order to improve the production of curcumin in a single medium, the initial TB (0.04% (v/v) of glycerol) used was also supplemented with a higher concentration of carbon sources at the beginning of the fermentation or only when ferulic acid was added to the medium. Figure 13 summarizes the results obtained in this experiment.

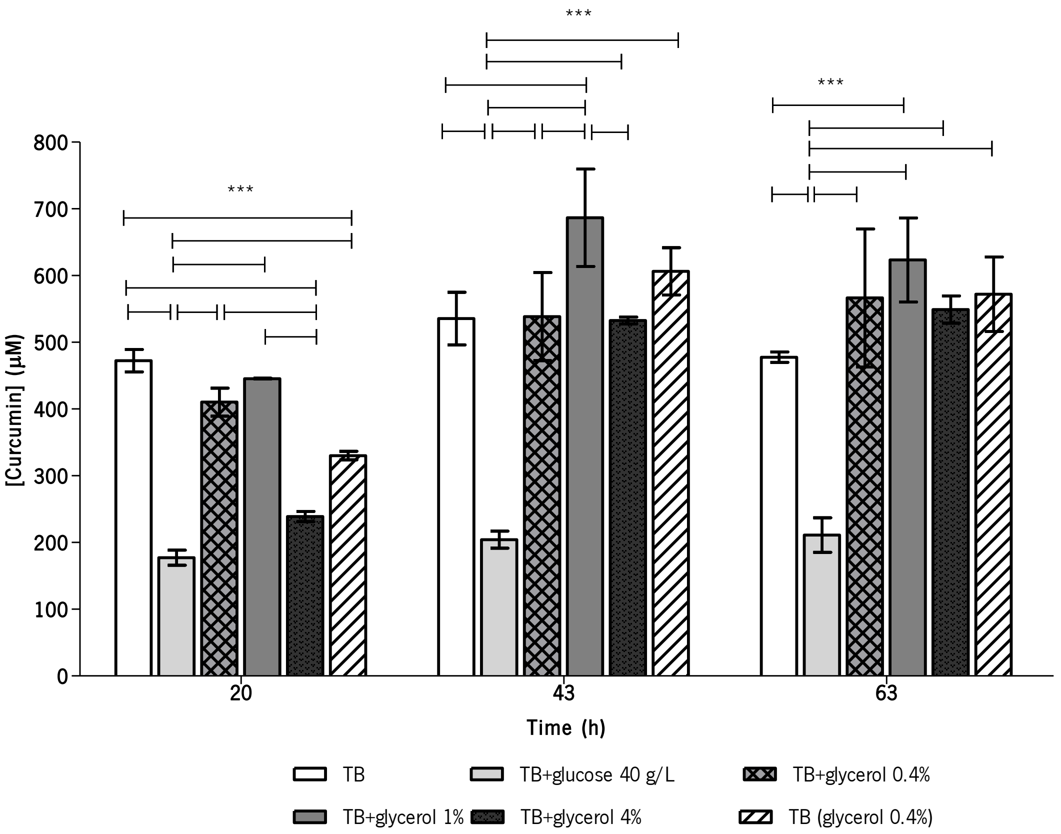


Figure 13. Curcumin production over time by *E. coli* BL21 (DE3) in terrific broth (TB) supplemented with different carbon sources. All TB contained glycerol in a final concentration of 0.04% (v/v) in the beginning of the fermentation except for TB (glycerol 0.4%) that contained 0.4% of glycerol. In the other cases, glucose or glycerol were supplemented when the substrate was added (time 0). All the significant differences indicated by the horizontal bars have a *p*-value <0.001. Protein expression was induced with 1 mM of IPTG (Isopropyl β-D-1-thiogalactopyranoside) when the optical density at 600 nm was 0.6. Standard deviations are indicated by error bars.

Comparing to the case when only the standard TB medium (0.04% (v/v) of glycerol) was used, the curcumin concentration obtained when supplementing the TB with extra glycerol at 1% final concentration after the 5 h of IPTG induction (*p*-value <0.001 at 43 h and 63 h time points) was statistically higher. At 63 h from the beginning, also TB medium with extra 0.4% glycerol represented a significant improvement comparing to the TB itself.

The addition of the same glucose concentration as in the M9 medium (when using the LB and M9 combination: 40 g/L of glucose) resulted in a clear inhibition of the curcumin production. This can be explained by the high concentration of carbon sources present in the culture that inhibited the curcumin production reactions. In fact, glucose uptake and respiration capacity become affected by the intense protein expression, as reported by Neubauer *et al.* (2003). As explained before, thanks to the CCR phenomena, glucose is the preferred *E. coli* carbon source and its presence prevents the use of other carbon sources such as glycerol (Bren *et al.*, 2016). CCR does not allow the coexistence of glycolytic and gluconeogenic pathways. Therefore, it leads to the accumulation of glucose in the fermentation medium which promotes the growth of plasmid-free cells (Neubauer *et al.*, 2003) that will not produce curcumin. The CCR effect can be eliminated by constructing strains lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system (PTS) which makes cells able to co-use glucose and other carbon sources and consequently improve their specific growth rate (Martínez *et al.*, 2008).

The initial glycerol concentration in TB was also increased to 0.4% [TB (glycerol 0.4%)], as the original TB recipe indicates, but it did not result in a statistically different curcumin production.

The highest curcumin concentration obtained in a single medium was 686.7±59.7 µM at 43 h from the beginning in TB (0.04%+1% glycerol (v/v)). This concentration is lower than the one obtained using LB + M9 medium (822.6±28.1 µM at 63h inducing protein expression with 0.1 mM of IPTG when cells reached an OD_{600nm} of 0.6). However, it is only 1.2 times lower and it was obtained using a single medium formulation, which represents a great economical advantage, as previously discussed. Besides, curcumin concentration obtained is higher than the previously reported in the literature by Katsuyama *et al.* (2008) (306.8 µM) or Rodrigues *et al.* (2015c) (187.9 µM).

4.2 Curcumin toxicity assay

A curcumin toxicity assay was done to determine the curcumin concentrations that are toxic to *E. coli* cells. The rationale of this assay was to find the maximum curcumin concentrations that the cells can tolerate in the medium without dying, as this will limit the optimization of the production. The assay was performed with *E. coli* BL21 (DE3), with and without the artificial biosynthetic pathway. The toxicity was evaluated by measuring the OD_{600nm} during 8 h. Figure 14 (A) and Figure 14 (B) present the results obtained for the wild-type *E. coli* BL21 (DE3) and for the *E. coli* carrying the biosynthetic pathway, respectively.

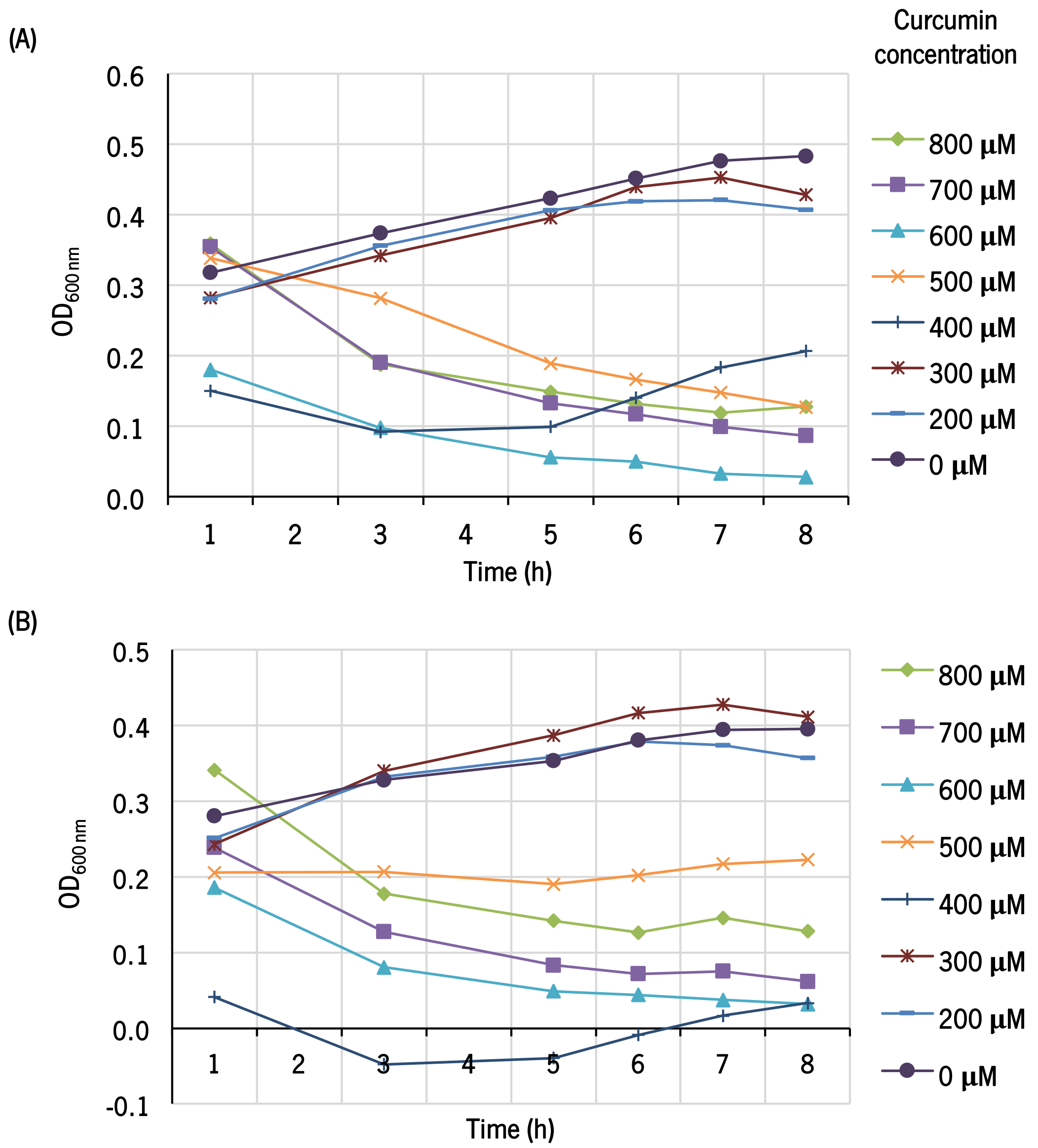


Figure 14. Wild-type *E. coli* BL21 (DE3) (A) and the one harboring the artificial biosynthetic pathway (B) growth during 8 h in the presence of different curcumin concentrations. In each curve it is possible to observe the OD_{600nm} variation over time of the bacterial culture exposed to different concentrations of curcumin. The medium values from triplicate wells are represented.

Comparing the Figure 14 (A) and (B), it is possible to conclude that this method for determining the curcumin toxicity was not reproducible. This might be caused by the curcumin instability, which has been reported to be strongly affected by factors such as pH and exposure to light (Nugroho *et al.*, 2009). Besides, the low solubility of curcumin in aqueous solution was visually observed, since it precipitates, influencing the measurement of the OD_{600nm}. However, since the goal of this experiment was to evaluate the concentration at which curcumin affects the *E. coli* growth in order to determine the maximum curcumin concentration that cells could produce, the results herein gathered are sufficient to draw some conclusions. In fact, it is notorious especially in Figure 14 (A) that from curcumin concentrations of 400 µM, the bacteria growth is affected. Curiously, in both images the growth curve of *E. coli* in the presence of 400 μM of curcumin, show that cells seem to recover from curcumin toxicity after 5 h but from curcumin concentrations above 500 μM, the cells are not able to recover.

Despite the proven curcumin toxicity at 400-500 μM, in the current study *E. coli* BL21 (DE3) cells were able to produce a higher curcumin concentration of 822.6±28.1 μM. This increased production is not due to cell growth on such a toxic concentration, but perhaps due to the enzymatic machinery activity of cells that are still converting ferulic acid to curcumin. Therefore, the fermentation can be acting as an *in vitro* reaction on the medium culture. Complementary assays are necessary to confirm these results, including platting cells after 20, 43 or 63 h of fermentation in the optimized medium to evaluate cells viability, when cells produced more than 500 μM of curcumin.

Curcumin toxicity can constitute a limitation for the heterologous curcumin production in *E. coli*. The curcumin toxicity can be minimized if a fed-batch fermentation mode is used for its heterologous production, leading to substrate and product concentrations reduction. Alternatively, a continuous approach where cells would remain in the bioreactor while the product (curcumin) would be removed over time could also be used.

4.3 Confirmation of the *gnd* deletion for *E. coli* mutant validation

The *gnd* gene was previously deleted from *E. coli* K-12 MG1655 (DE3) genome (Araújo, 2014) by using the recombineering technique (Kolisnychenko *et al.*, 2002; Fehér *et al.*, 2007). Herein, to confirm the absense of this gene for further studies, a colony PCR was performed using the primers gnd_D and gnd_E. As can be seen in Figure 15, the expected band of 917 bp was obtained. If the strain was not mutated, the band would correspond to 2524 bp because the deletion cassette would still be present.

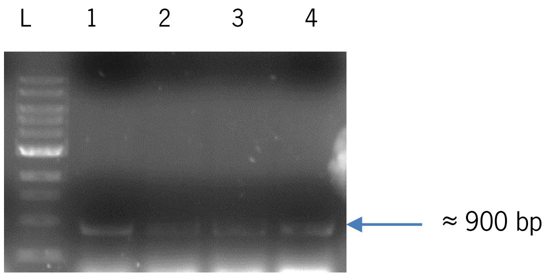


Figure 15. Agarose gel (0.7% (w/v)) to confirm the deletion of *gnd* from *E. coli* K-12 MG1655 (DE3). L represents the 1 kb molecular weight marker (NEB). Samples 1-4 were obtained using different colonies as template for colony PCR.

The samples (1-4) confirmed the successful *gnd* deletion. Afterwards, this gene deletion was also confirmed by sequencing.

4.4 Construction of *E. coli* mutants using the CRISPR-Cas9 system

In order to improve the heterologous production of curcumin by *E. coli*, the construction and validation of several mutants was herein attempted. For their construction, three different recombineering strategies were evaluated. However, only the CRISPR-Cas9 approach was successful and allowed the construction of the mutants *E. coli* K-12 MG1655 (DE3) Δ*fumB* and *E. coli* BL21 (DE3) Δ*fumB*. Therefore, the results obtained in the other two deletion strategies used in the current work can be found in **Appendix III**.

As explained in **Section 3.7** and described by Reisch & Prather (2015), the first step of the deletion protocol using CRISPR-Cas9 system is the construction of pKDsgRNA_*gene*, i.e. a plasmid containing a small sequence of the target gene (to be deleted) for Cas9 action. For each gene, two fragments were amplified from pKDsgRNA-p15 – part 1 (2848 bp) and part 2 (4414 bp). After *Dpn*I digestion of both PCR products, they were run in an agarose gel that can be visualized in Figure 16. *Dnp*I was used to make sure that no template plasmid remained in the solution.

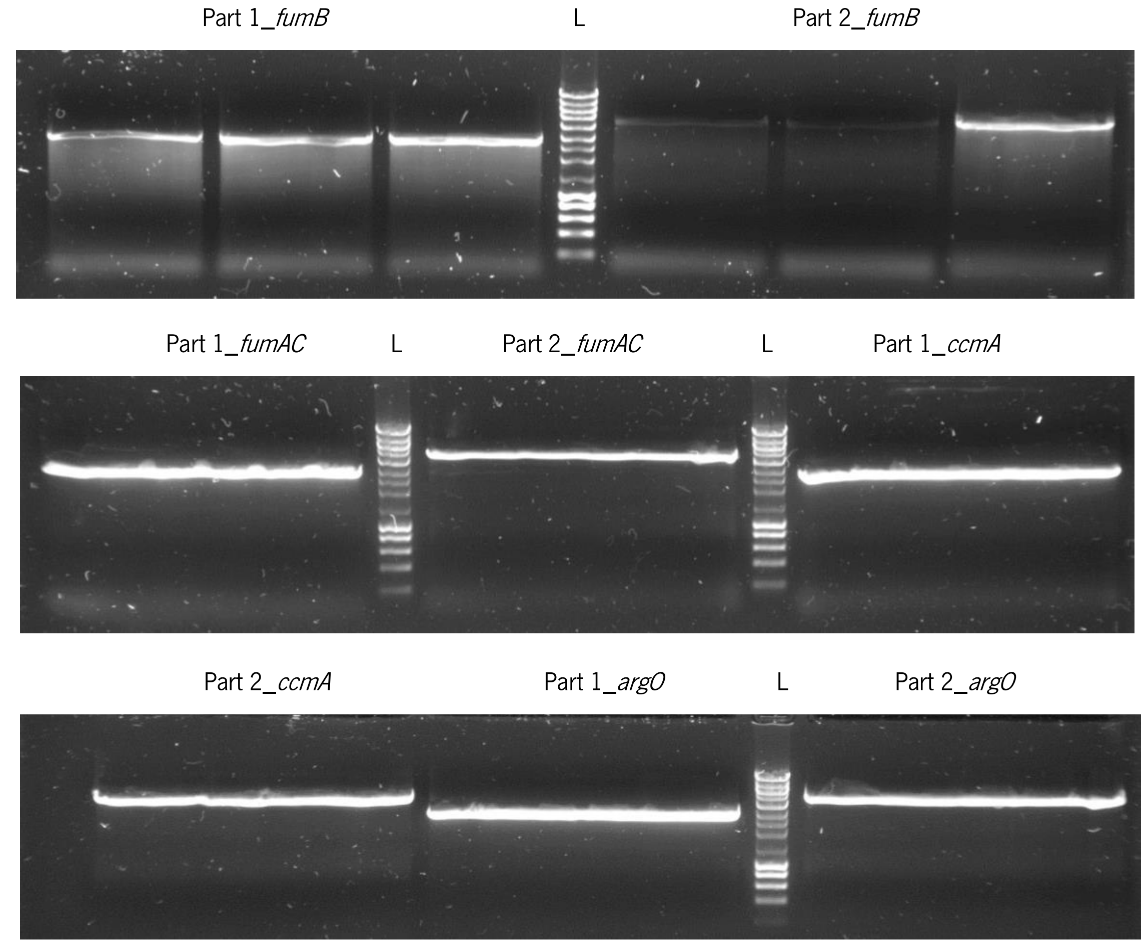


Figure 16. Agarose gels (1% (w/v)) showing the amplification of the part 1 (2848 bp) and part 2 (4414 bp) of the pKDsgRNA to construct pKDsgRNA_*fumB*, pKDsgRNA_*fumAC*, pKDsgRNA_*ccmA* and pKDsgRNA_*argO* by CPEC (circular polymerase extension cloning). L well contained the NZYDNALadder III (NZYTech).

All fragments were found to possess the expected size. Afterwards, the pKDsgRNA_*gene* plasmids were constructed by CPEC and the presence of the specific protospacer was confirmed by sequencing.

In order to continue the protocol for *fumB* gene deletion, *E. coli* K-12 MG1655 (DE3) and BL21 (DE3) were then transformed with the pCas9-CR4 plasmid. Afterwards, the plasmid was extracted and digested with *Hind*III-HF to confirm that pCas9-CR4 plasmid was inside the cells. The digested plasmid was run in an agarose gel and the bands from pCas9-CR4 were found to possess the correct sizes (*data not shown*).

Both strains harboring the pCas9-CR4 were then transformed with the constructed pKDsgRNA_*fumB* plasmid. Subsequently, the cells were also transformed with the corresponding oligonucleotide (Oligo_fumB) that replaces *fumB* gene in the *E. coli* genome. Afterwards, the colonies obtained were tested by colony PCR using the primers fumB_D and fumB_E. The expected negative band was 1967 bp and the positive one (i.e. in case of successful deletion of the gene) was 320 bp. These results are shown in Figure 17.

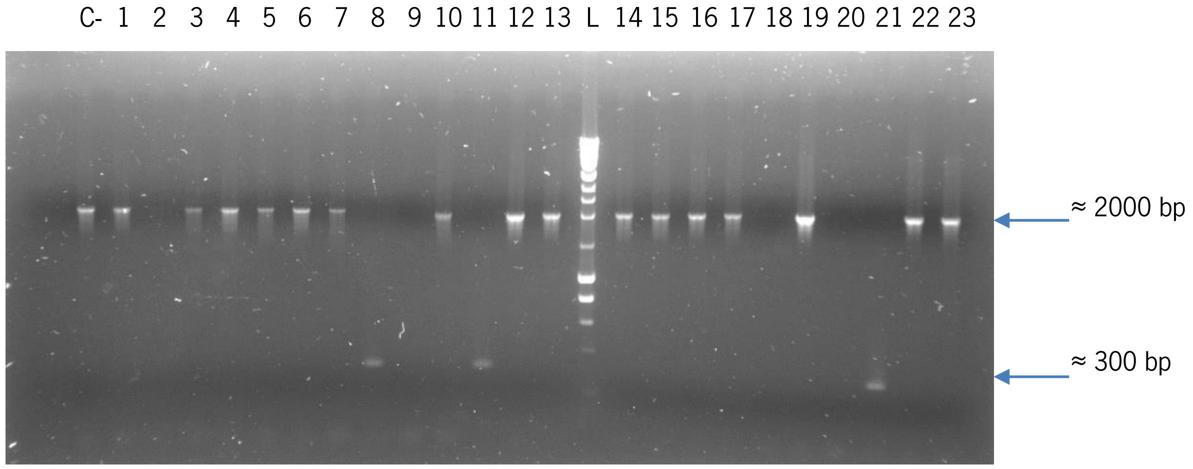

C- 1 2 3 4 5 6 7 8 9 10 11 12 13 L 14 15 16 17 18 19 20 21 22 23
≈ 2000 bp
≈ 300 bp

Figure 17. Agarose gel (1% (w/v)) with colony PCR results to confirm the *fumB* deletion in *E. coli* K-12 MG1655 DE3. C- corresponds to a wild-type colony for negative control. L well contains the NZYDNALadder III (NZYTech)

From the 23 *E. coli* K-12 MG1655 (DE3) colonies tested, three positive colonies were identified (8, 11 and 21). Thus, these colonies were sequenced with the primers fumB_D and fumB_E and the deletion was confirmed for all of them.

The results from the colony PCR of the *E. coli* BL21 colonies are shown in Figure 18.

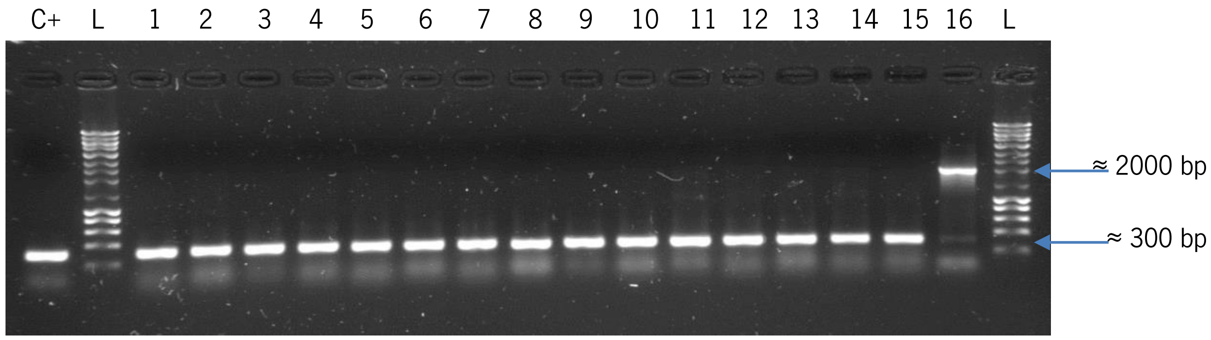


Figure 18. Agarose gel 1% with colony PCR results to confirm the *fumB* deletion in *E. coli* BL21 (DE3). C+ corresponds to a positive control (*E. coli* K-12 (DE3) Δ*fumB* colony). L well contains the NZYDNALadder III (NZYTech)

As can be seen in Figure 18, the colonies from 1 to 15 were positive. Interestingly, despite *E. coli* BL21 was more difficult to transform with the second plasmid (pKDsgRNA_*fumB*), the efficiency of deletion was very good, maybe because this strain is deficient in *lon* and *ompT* proteases which allows higher heterologous protein expression (Studier *et al.*, 2009) (in this case Cas9 and recombineering enzymes). The DNA from one of the positive colonies was sequenced with the fumB_D primer and the deletion was confirmed.

The same CRISPR-Cas9 strategy was used to attempt the simultaneous *fumA* and *fumC* deletion in *E. coli* K-12 MG1655 (DE3) Δ*fumB* and *E. coli* BL21 (DE3) Δ*fumB* mutants after curing the pKDsgRNA_*fumB* plasmid. After introduction of the pKDsgRNA_*fumAC*, the correspondent oligonucleotide (Oligo_AC) was transformed into cells. Figure 19 (A) presents some of the results of the colony PCR that was performed to confirm the *fumA* and *fumC* deletion. As all the bands were negative, the *fumA* and *fumC* deletion was attempted in the wild-types *E. coli* K-12 MG1655 (DE3) and *E. coli* BL21 (DE3) to ensure that the *fumB* deletion was not interfering biologically with the elimination of *fumA* and *fumC*. The results are shown in Figure 19 (B).

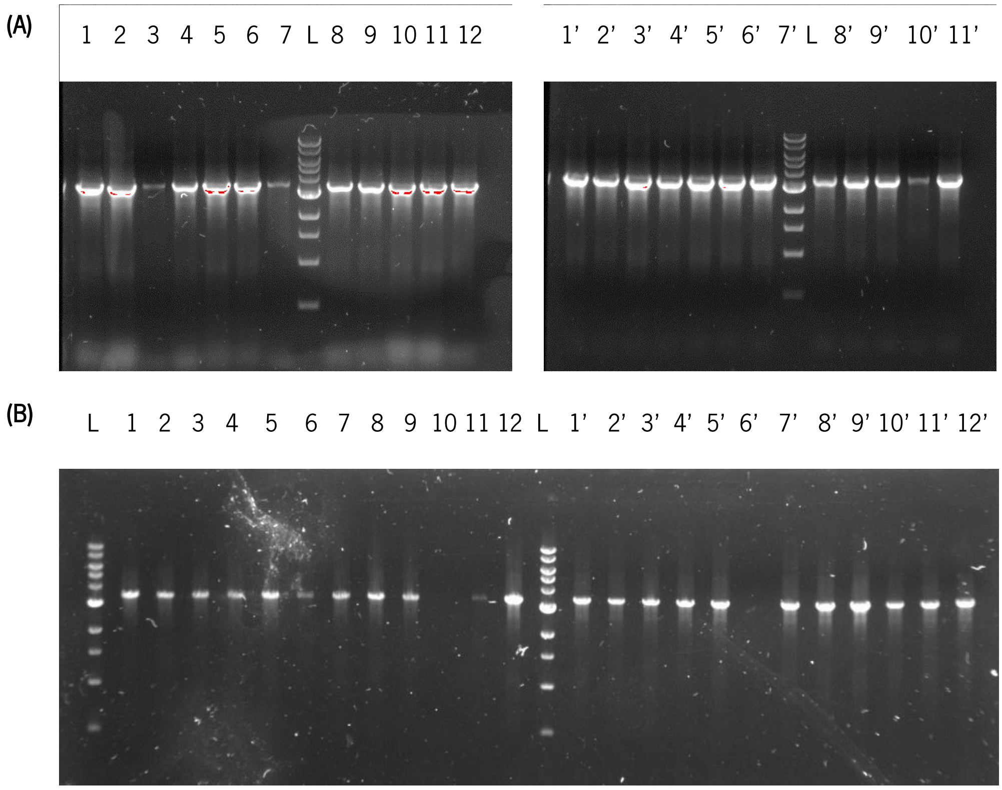

(A)
1 2 3 4 5 6 7 L 8 9 10 11 12
1’ 2’ 3’ 4’ 5’ 6’ 7’ L 8’ 9’ 10’ 11’
(B)
L 1 2 3 4 5 6 7 8 9 10 11 12 L 1’ 2’ 3’ 4’ 5’ 6’ 7’ 8’ 9’ 10’ 11’ 12’

Figure 19. Agarose gel 1% (w/v) showing the colony PCR results using the primers fumA_D and fumAC_E to confirm the deletion of *fumA* and *fumC.* The deletion was attempted in *E. coli* K-12 MG1655 (DE3) Δ*fumB* (1-12:**A**) and *E. coli* BL21 (DE3) Δ*fumB* (1’-12’:**A**) and in *E. coli* K-12 MG1655 (DE3) (1-12:**B**) and *E. coli* BL21 (DE3) (1’-12’:**B**).

The results clearly show that no positive band (≈ 300 bp) was present in any colony. Therefore, the simultaneous deletion of *fumA* and *fumC* could not be accomplished in none of the tested strains. The inability to delete these genes might be related to the larger size of the sequence to delete (≈ 3 kb) which reduces the recombination efficiency (Sharan *et al.*, 2009). Besides, even when genes are identified as non-essential, the deletion of these two genes at the same time can be harmful for the cell.

New primers for the individual deletion of *fumA* and *fumC* were designed, however these deletions are still being made, as well as the deletion of *gnd* in *E. coli* BL21 (DE3). Deletions of *argO* and *ccmA* in both *E. coli* K-12 MG1655 (DE3) and *E. coli* BL21 (DE3) strains are still being attempted. In the future, all the mutants constructed, whether with single or cumulative mutations, will be evaluated regarding the production of curcumin. The mutants herein successfully obtained (*E. coli* K-12 MG1655 (DE3) Δ*fumB* and *E. coli* BL21 (DE3) Δ*fumB*) were then used to evaluate the production of curcumin (**Section 4.5**). Despite the metabolic model has predicted that the triple mutation (Δ*fumA*, *fumB*, *fumC*) would enhance curcumin production, the single deletion of *fumB* was evaluated just to verify if it could result in an improvement *per se*.

4.5 Curcumin production by engineered *E. coli* mutants

The main goal of the present work was the enhancement of curcumin production using the previously designed *E. coli* mutants. *E. coli* K-12 MG1655 (DE3) Δ*gnd* was formerly constructed by Araújo (2014). Until now, another two mutants have been constructed, namely *E. coli* K-12 MG1655 (DE3) Δ*fumB* and *E. coli* BL21 (DE3) Δ*fumB*. The artificial biosynthetic pathway was inserted in the three mutants and a fermentation was conducted to compare their curcumin production against the 'original' strains (without genes deletion but harboring the artificial biosynthetic pathway). The comparison of curcumin production by *E. coli* K-12 MG1655 (DE3) and the above mentioned mutants is presented in Figure 20.

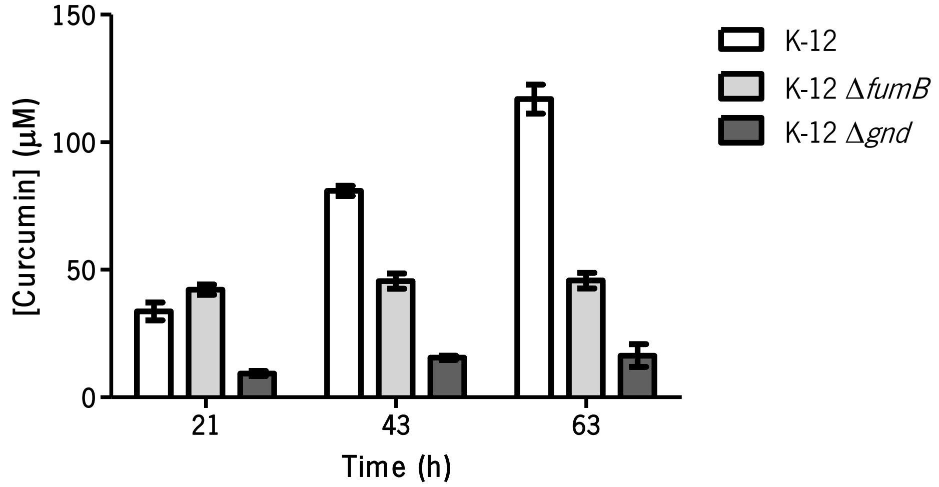


Figure 20. Curcumin production by *E. coli* K-12 MG1655 (DE3) and its derived mutants, *E. coli* K-12 MG1655 (DE3) Δ*fumB* and *E. coli* K-12 MG1655 (DE3) Δ*gnd*. LB (lysogeny broth) and M9 media combination was used. Optical density at 600 nm was 0.4 when protein expression was induced with 1 mM of IPTG (Isopropyl β-D-1-thiogalactopyranoside). Standard deviations are indicated by error bars.

Comparing the results from *E. coli* K-12 MG1655 (DE3) and *E. coli* K-12 MG1655 (DE3) Δ*fumB*, the strain without the *fumB* deletion produced higher amounts of curcumin at the end of the fermentation (116.9±4.6 µM). However, the curcumin production was, in the initial hours, higher when using the Δ*fumB* mutant, exhibiting a difference statistically relevant at 21 h (*p*-value <0.05). This result shows that the designed *E. coli* K-12 MG1655 (DE3) Δ*fumB* mutant is able to produce curcumin faster, but the curcumin concentration remained constant after the maximum curcumin production by this strain was achieved (45.8±2.5 µM of curcumin at 63 h).

E. coli K-12 MG1655 (DE3) Δ*gnd* was able to produce 16.4±3.6 µM. Araújo (2014) was able to produce in the same conditions 35.5 µM of curcumin with this mutant. This difference can be due to some loss of viability of the mutated strain over the time. However, this may also be due to some medium components as the result obtained for *E. coli* K-12 MG1655 (DE3) was also lower than the previous curcumin concentrations obtained in the current work (177.9±4.3 µM). The *gnd* deletion is being performed again and the curcumin production will then be evaluated to confirm this result.

Although the *fumB* deletion was previously predicted (*in silico*) to enhance the production of curcumin by *E. coli* K-12 MG1655 (DE3), the same deletion was attempted in *E. coli* BL21 (DE3) to see if the curcumin production could be improved also in this strain by using the same deletions. The results for *E. coli* BL21 (DE3) and *E. coli* BL21 (DE3) Δ*fumB* are presented in Figure 21.

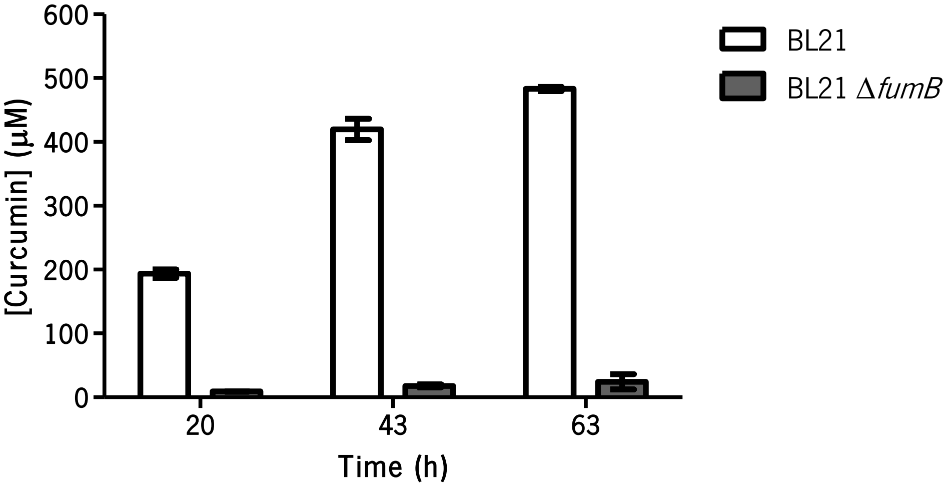


Figure 21. Curcumin production by *E. coli* BL21 (DE3) and *E. coli* BL21 (DE3) Δ*fumB*. LB (lysogeny broth) and M9 media combination was used. Protein expression was induced with 1 mM of IPTG (Isopropyl β-D-1-thiogalactopyranoside) when the optical density at 600 nm of the culture reached 0.6. Standard deviations are indicated by error bars.

It is well-known that the *in silico* predictions not always correspond to the results experimentally observed, due to unknown reactions and interactions between compounds from the cells. Besides, the prediction of those mutants was simulated with the *E. coli* K-12 MG1655 metabolic model, so the improvement of curcumin production by *fumB* deletion on *E. coli* BL21 has not been predicted and this was only an attempt to check if the same trend could be found.

Comparing the results from the *E. coli* BL21 (DE3) and the *E. coli* BL21 (DE3) Δ*fumB*, it is clear that the production of curcumin is not improved from the deletion of *fumB* in *E. coli* BL21 (DE3), contrary to what could be expected. The curcumin concentration at 63 h was 483.1±2.6 µM using *E. coli* BL21 (DE3) and 24.4±9.7 µM using *E. coli* BL21 (DE3) Δ*fumB*.

5. CONCLUSIONS

The main goal of this work was to optimize the production of curcumin in an engineered *E. coli*. This goal was accomplished in different ways. First, different strains were evaluated for the production of curcumin. It was proven that it is possible to obtain higher curcumin concentrations with *E. coli* BL21 (DE3) comparing to the previously used strain, *E. coli* K-12 MG1655 (DE3) under the same operational conditions. Since *E. coli* BL21 (DE3) presented the best performance within the tested strains, it was used for the following optimization tests. The optimal OD_{600nm} for IPTG induction in the LB and M9 medium combination generally used was assessed. The improvement of curcumin production was accomplished by inducing the cells at a higher OD_{600nm} (0.9). Also, the influence of the IPTG concentration was studied in LB+M9 and it was concluded that lower concentrations (0.1 mM) can indeed improve the curcumin concentrations. Additionally, given the goal of developing a sustainable and economic process, single culture media (to avoid the medium change step) were evaluated. For that purpose, LB, TB and MOPS were tested. Growing *E. coli* cells only in TB led to similar curcumin concentrations as those obtained using the LB and M9 combination. In order to improve curcumin production in TB, the optimal OD_{600nm} for IPTG induction in this medium was also assessed. However, it was not clearly demonstrated that the curcumin production is improved as a result of changing the OD_{600nm} for induction of protein expression. Extra carbon source (glucose or glycerol) were added to the TB medium to improve the curcumin production. This could be clearly observed when adding more 1% (v/v) of glycerol (final concentration) at time 0 h of substrate addition. Up to now, the highest curcumin concentration produced by a heterologous organism is the one reported in this study, 686.7±59.7 μM using the single medium TB (43 h) and 822.6±28.1 μM using LB+M9 (63 h).

Moreover, to assess what is the maximum curcumin concentration tolerated by the cells within this batch operation, the product toxicity was studied in a 96-well plate assay. The results showed that curcumin concentrations above 400 μM influence negatively the *E. coli* cells growth, thus this should be the limit of curcumin in the extracellular medium.

Finally, in order to optimize curcumin production in *E. coli* strains, the *fumB* gene was deleted from both *E. coli* K-12 MG1655 (DE3) and *E. coli* BL21 (DE3) strains. The production of curcumin by the *E. coli* K-12 MG1655 (DE3) mutant (*Δgnd)* previously constructed was also evaluated in this work. A fermentative assay was performed using the mutated strains and the wild-type as a control. After 63 h, *E. coli* K-12 MG1655 (DE3) Δ*fumB* was able to produce 45.8±2.5 μM of curcumin, *E. coli* K-12

MG1655 (DE3) Δ*gnd* produced 16.4±3.6 μM and BL21 (DE3) Δ*fumB* produced 24.4±9.7 μM of curcumin. These results showed that the mutants constructed based on the previous *in silico* analysis of possible gene knockouts to improve curcumin production did not enhance curcumin production as predicted. This can be due to several factors. First, the single knock-out of the gene *fumB* was not predicted to improve curcumin production. The mutant predicted had three mutations (*fumA*, *fumB* and *fumC*) but as *fumB* was the only gene deleted, the Δ*fumB* mutant was used hoping that the deletion of one of the fumarase genes would be enough for evidencing curcumin production enhancement. Also, as discussed before in **Section 4.5**, the metabolic model used for *in silico* analysis was from *E. coli* K-12 MG1655 (DE3). The same prediction was not made to *E. coli* BL21 (DE3) but it was assumed that the model would be similar so as the trend. In the case of *gnd* deletion, the mutant presented lower growth compared to the strain carrying the *gnd* gene. The altered bacteria growth can influence curcumin production. The *gnd* deletion in *E. coli* K-12 MG1655 (DE3) will be reattempted in future work to confirm if the cells viability was compromised due to the age of the stocks or to influence of the gene deletion (*data not shown*).

Overall, the results gathered in this work are very promising and suggest that further optimizations both at operational and genetic levels are still possible.

6. FUTURE PERSPECTIVES

Despite the interesting findings of the current work, a number of recommendations both at operational and genetic levels can be pointed towards a further improvement of the production of curcumin by *E. coli*.

Regarding the operational conditions, different culture media can be evaluated. For example, SB or Super Broth, that is similar to TB medium, could be evaluated for curcumin production, since it is an even richer culture medium in peptone and yeast extract and is also described as suitable for protein production giving high yields of plasmid DNA. Additionally, the best strains and optimized conditions for the production of curcumin should be tested in a bioreactor, in order to evaluate if the industrial scale-up could be implemented. Another optimization suggestion is raising the temperature after the 5 h of incubation with IPTG, since the results from this work suggests that only a residual protein expression is being induced during this time. The idea is to increase the conversion rate from ferulic acid to curcumin by using a temperature closer to the optimal temperature for the enzymes activity, which is 50°C for CURS1 from *C. longa*, 25-35°C for DCS from *C. longa*, and 30°C for 4CL (from *Streptomyces coelicolor*; since the optimal temperature for *A. thaliana*'s enzyme is not reported), instead of the 26°C used in this work.

The toxicity of curcumin to *E. coli* cells is clearly an inconvenient for the design of super-producing strains and optimized curcumin productivities using this heterologous host. Therefore, a fed-batch or continuous approach should be considered to minimize curcumin accumulation and toxicity. For these approaches, the strains should be engineered with secretion systems in order to secrete each one of the three curcumin producing enzymes and therefore, produce curcumin on the extracellular medium from ferulic acid. The usual approach for proteins exportation consists in sending proteins to the periplasm and some of them diffuse or leak into the extracellular medium. Routing to the periplasm or to the extracellular space is achieved by fusing the recombinant protein to a proper leader peptide (a signal sequence or fusion partner, such as peptides from proteins naturally secreted in *E. coli*). The signal peptides of the following proteins are widely used for secretion: Lpp, LamB, LTB, MalE, OmpA, OmpC, OmpF, OmpT, PelB, PhoA, PhoE, DsbA, SpA, endoxylanase, or StII. (Makrides, 1996; Jeong & Lee, 2002; Choi & Lee, 2004; Kotzsch *et al.*, 2011; Wong *et al.*, 2012; Rosano & Ceccarelli, 2014). Passive transport from periplasm to the extracellular culture can be stimulated by external or internal destabilization of the *E. coli* structural components (Sørensen & Mortensen, 2005). Suggestions for increasing the membrane permeability include the use of lytic proteins to permeabilize *E. coli* outer membrane integrity in a way that proteins present in the periplasm can pass through it. Examples include the co-expression of the *kil* gene (killing protein), Bacteriocin Release Protein (BRP) or TolAIII gene with the recombinant proteins as well as the deletion of Braun's lipoprotein (LPP) gene (Wong *et al.*, 2012). Usually, these genes can compromise the viability of the host cells, so it would not be helpful for a continuous approach unless genes are modified to reduce the damage of the membranes (Ignatova *et al.*, 2003; Choi & Lee, 2004). External destabilization includes the use of the gene III product of filamentous phage f1 to permeabilize the outer membrane and glycine or Triton X-100 to help disrupt cell membrane integrity (Wong *et al.*, 2012). Usually, more than one approach can be used simultaneously. For example, in the work of Ignatova *et al.*, 2003, the overproduction of SecA, SecB, and SecF improved the transport efficiency and enhanced up to threefold the level of expression of an OmpT-penicillin amidase fusion in the periplasm. Different organisms with improved resistance to curcumin can be also tested in order to obtain higher and industrially interesting curcumin concentrations.

Regarding the mutant construction for improved curcumin production, different *E. coli* mutants, which have been previously predicted to improve the production of curcumin from ferulic acid, are still to be constructed. *fumA*, *fumC*, *argO* and *ccmA* genes can be deleted and the influence of the single or cumulative mutations in curcumin production can be further evaluated.

Finally, as discussed before, the heterologous pathway could be used for therapeutic applications, e.g. in breast cancer treatment. This would require the construction of the complete pathway from natural resources, such as the aminoacids phenylalanine and/or tyrosine. Moreover, the strain can be engineered for the deletion of virulence factors to reduce health risks and increase therapeutic acceptance. Besides, the genes from the artificial pathway for curcumin production can be regulated by heat shock promoters, therefore curcumin production can, in principle, be triggered by heat, as for e.g. by the heat generated in the ultrasound treatment (Rodrigues, 2014).

7. BIBLIOGRAPHY

Aggarwal BB, Kumar A, Bharti AC, Anderson TMD (2003) Anticancer Potential of Curcumin Preclinical and Clinical Studies. *Anticancer Res* **23**: 363–398.

Aggarwal, B. B., Surh, Y. J., & Shishodia, S. (Eds.). (2007). *The molecular targets and therapeutic uses of curcumin in health and disease* (Vol. 595). Springer Science & Business Media.

Ahmad AM (2007) Recent advances in pharmacokinetic modeling. *Biopharm Drug Dispos* **28**: 135–143.

Anand P, Kunnumakkara AB, Newman R a., Aggarwal BB (2007) Bioavailability of curcumin: Problems and promises. *Mol Pharm* **4**: 807–818.

Araújo RG (2014) Validation and optimization of curcumin production by an engineered *Escherichia coli*. Masters Thesis. Universidade do Minho.

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko K a, Tomita M, Wanner BL, Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**: 2006.0008.

Bagchi A (2012) Extraction of Curcumin. *J Environ Sci Toxicol Food Technol* **1**: 1–16.

Becker JVW, Armstrong GO, Van Der Merwe MJ, Lambrechts MG, Vivier M a., Pretorius IS (2003) Metabolic engineering of *Saccharomyces cerevisiae* for the synthesis of the wine-related antioxidant resveratrol. *FEMS Yeast Res* **4**: 79–85.

Berry A (1996) Improving production of aromatic compounds in *Escherichia coli* by metabolic engineering. *Tibtech* **14**: 250–256.

Bren A, Park JO, Towbin BD, Dekel E, Rabinowitz JD, Alon U (2016) Glucose becomes one of the worst carbon sources for *E. coli* on poor nitrogen sources due to suboptimal levels of cAMP. *Sci Rep* **6**: 1–10.

Burgos-Morón E, Calderón-Montaño JM, Salvador J, Robles A, López-Lázaro M (2010) The dark side of curcumin. *Int J Cancer* **126**: 1771–1775.

Carneiro S, Ferreira EC, Rocha I (2013) Metabolic responses to recombinant bioprocesses in *Escherichia coli*. *J Biotechnol* **164**: 396–408.

Casali, N. (2003). *Escherichia coli* host strains. *E. coli Plasmid Vectors: Methods and Applications*, 27-48.

Chattopadhyay I, Biswas K, Bandyopadhyay U, Banerjee RK (2004) Turmeric and curcumin: Biological actions and medicinal applications. *Curr Sci* **87**: 44–53.

Chemler J a., Koffas MAG (2008) Metabolic engineering for plant natural product biosynthesis in microbes. *Curr Opin Biotechnol* **19**: 597–605.

Choi JH, Lee SY (2004) Secretory and extracellular production of recombinant proteins using *Escherichia coli*. *Appl Microbiol Biotechnol* **64**: 625–635.

Demain AL, Vaishnav P (2009) Production of recombinant proteins by microbes and higher organisms. *Biotechnol Adv* **27**: 297–306.

Fehér T, Karcagi I, Györfy Z, Umenhoffer K, Csorgó B, Pósfai G (2007) Scarless Engineering of the *Escherichia coli* Genome. *Methods in Molecular Biology*, Vol. 416 (Gerdes SY & Osterman AL, eds), pp. 249–261. Totowa, NJ.

Felgner S, Kocijancic D, Frahm M, Weiss S (2016) Bacteria in Cancer Therapy: Renaissance of an Old Concept. *Int J Microbiol* **2016**: 1–14.

Fernández-Castané A, Vine CE, Caminal G, López-Santín J (2012) Evidencing the role of lactose permease in IPTG uptake by *Escherichia coli* in fed-batch high cell density cultures. *J Biotechnol* **157**: 391–398.

Fischer F, Künzler P, Ritz D, Hennecke H, Fischer F, Ku P, Institut M (1995) *Escherichia coli* genes required for cytochrome c maturation. *J Bacteriol* **177**: 4321–4326.

Fox ME, Lemmon MJ, Mauchline ML, Davis TO, Giaccia AJ, Minton NP, Brown JM (1996) Anaerobic bacteria as a delivery system for cancer gene therapy: in vitro activation of 5-fluorocytosine by genetically engineered clostridia. *Gene Ther* **3**: 173–178.

Freter R, Brickner H, Fekete J, Vickerman MM, Carey KE, Carey E (1983) Survival and Implantation of *Escherichia coli* in the Intestinal Tract Survival. *Infect Immun* **39**: 686–703.

Gardlik R, Fruehauf JH (2010) Bacterial vectors and delivery systems in cancer therapy. *IDrugs* **13**: 701–706.

Georgiou G, Segatori L (2005) Preparative expression of secreted proteins in bacteria: Status report and future prospects. *Curr Opin Biotechnol* **16**: 538–545.

Gerdes SY, Scholle MD, Campbell JW, *et al.* (2003) Experimental Determination and System Level Analysis of Essential Genes in *Escherichia coli* MG1655. *J Bacteriol* **185**: 5673–5684.

Goel A, Kunnumakkara AB, Aggarwal BB (2008) Curcumin as "Curecumin": From kitchen to clinic. *Biochem Pharmacol* **75**: 787–809.

Gottesman S (1996) Proteases and their targets in *Escherichia coli*. *Annu Rev Genet* **30**: 465–506.

Grodberg J, Dunn JJ (1988) ompT encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. *J Bacteriol* **170**: 1245–1253.

Gudiña EJ, Rocha V, Teixeira JA, Rodrigues LR (2010a) Antimicrobial and antiadhesive properties of a biosurfactant isolated from *Lactobacillus paracasei* ssp. paracasei A20. *Lett Appl Microbiol* **50**: 419–424.

Gudiña EJ, Teixeira JA, Rodrigues LR (2010b) Isolation and functional characterization of a biosurfactant produced by *Lactobacillus paracasei*. *Colloids Surfaces B Biointerfaces* **76**: 298–304.

Gupta SC, Patchva S, Aggarwal BB (2012) Therapeutic Roles of Curcumin: Lessons Learned from Clinical Trials. *AAPS J* **15**: 195–218.

Halls C, Yu O (2008) Potential for metabolic engineering of resveratrol biosynthesis. *Trends Biotechnol* **26**: 77–81.

Horinouchi S (2008) Combinatorial biosynthesis of non-bacterial and unnatural flavonoids, stilbenoids and curcuminoids by microorganisms. *J Antibiot (Tokyo)* **61**: 709–728.

Horinouchi S (2009) Combinatorial biosynthesis of plant medicinal polyketides by microorganisms. *Curr Opin Chem Biol* **13**: 197–204.

Ignatova Z, Mahsunah a., Georgieva M, Kasche V (2003) Improvement of posttranslational bottlenecks in the production of penicillin amidase in recombinant Escherichia coli strains. *Appl Environ Microbiol* **69**: 1237–1245.

International Agency for Research on Cancer (2012) Population Fact Sheets. *GLOBOCAN 2012* http://globocan.iarc.fr/Pages/fact_sheets_population.aspx (Accessed September 29, 2016).

Jayaprakasha GK, Rao LJM, Sakariah KK (2002) Improved HPLC method for the determination of curcumin, demethoxycurcumin, and bisdemethoxycurcumin. *J Agric Food Chem* **50**: 3668–3672.

Jayaprakasha GK, Rao LJM, Sakariah KK (2005) Chemistry and biological activities of *C. longa*. *Trends Food Sci Technol* **16**: 533–548.

Jeandet P, Delaunois B, Aziz A, Donnez D, Vasserot Y, Cordelier S, Courot E (2012) Metabolic engineering of yeast and plants for the production of the biologically active hydroxystilbene, resveratrol. *J Biomed Biotechnol* **2012**: 1–14.

Jeong KJ, Lee SY (2002) Excretion of human beta-endorphin into culture medium by using outer membrane protein F as a fusion partner in recombinant *Escherichia coli*. *Appl Environ Microbiol* **68**: 4979–4985.

Jiao Z, Baba T, Mori H, Shimizu K (2003) Analysis of metabolic and physiological responses to gnd knockout in *Escherichia coli* by using C-13 tracer experiment and enzyme activity measurement. *FEMS Microbiol Lett* **220**: 295–301.

Jordan BC, Mock CD, Thilagavathi R, Selvam C (2016) Review article: Molecular mechanisms of curcumin and its semisynthetic analogues in prostate cancer prevention and treatment. *Life Sci* **152**: 135–144.

Katsuyama Y, Matsuzawa M, Funa N, Horinouchi S (2007) *In vitro* synthesis of curcuminoids by type III polyketide synthase from *Oryza sativa*. *J Biol Chem* **282**: 37702–37709.

Katsuyama Y, Matsuzawa M, Funa N, Horinouchi S (2008) Production of curcuminoids by *Escherichia coli* carrying an artificial biosynthesis pathway. *Microbiology* **154**: 2620–2628.

Katsuyama Y, Kita T, Funa N, Horinouchi S (2009a) Curcuminoid biosynthesis by two type III polyketide synthases in the herb *Curcuma longa*. *J Biol Chem* **284**: 11160–11170.

Katsuyama Y, Kita T, Horinouchi S (2009b) Identification and characterization of multiple curcumin synthases from the herb *Curcuma longa*. *FEBS Lett* **583**: 2799–2803.

Katsuyama Y, Hirose Y, Funa N, Ohnishi Y, Horinouchi S (2010a) Precursor-directed biosynthesis of curcumin analogs in *Escherichia coli*. *Biosci Biotechnol Biochem* **74**: 641–645.

Katsuyama Y, Ohnishi Y, Horinouchi S (2010b) Production of dehydrogingerdione derivatives in *Escherichia coli* by exploiting a curcuminoid synthase from *Oryza sativa* and a β-oxidation pathway from *Saccharomyces cerevisiae*. *ChemBioChem* **11**: 2034–2041.

Katsuyama Y, Miyazono KI, Tanokura M, Ohnishi Y, Horinouchi S (2011) Structural and biochemical elucidation of mechanism for decarboxylative condensation of β-keto acid by curcumin synthase. *J Biol Chem* **286**: 6659–6668.

Kita T, Imai S, Sawada H, Seto H (2009) Isolation of dihydrocurcuminoids from cell clumps and their distribution in various parts of turmeric (*Curcuma longa*). *Biosci Biotechnol Biochem* **73**: 1113–1117.

Kobayashi, T. (Ed.). (2004). *Recent Progress of Biochemical and Biomedical Engineering in Japan I* (Vol. 90). Springer Science & Business Media.

Kolisnychenko V, Kolisnychenko V, Iii GP, *et al.* (2002) Engineering a Reduced *Escherichia coli* Genome. *Genome Res* **12**: 640–647.

Koopman F, Beekwilder J, Crimi B, van Houwelingen A, Hall RD, Bosch D, van Maris AJ a, Pronk JT, Daran J-M (2012) De novo production of the flavonoid naringenin in engineered *Saccharomyces cerevisiae*. *Microb Cell Fact* **11**: 155.

Kotzsch A, Vernet E, Hammarstrom M, Berthelsen J, Weigelt J, Graslund S, Sundstrom M (2011) A secretory system for bacterial production of high-profile protein targets. *Protein Sci* **20**: 597–609.

Kuhlman TE, Cox EC (2010) Site-specific chromosomal integration of large synthetic constructs. *Nucleic Acids Res* **38** (6), e92-e92.

Kumar G, Mittal S, Sak K, Tuli HS (2016) Molecular mechanisms underlying chemopreventive potential of curcumin: Current challenges and future perspectives. *Life Sci* **148**: 313–328.

Kuo ML, Huang TS, Lin JK (1996) Curcumin, an antioxidant and anti-tumor promoter, induces apoptosis in human leukemia cells. *Biochim Biophys Acta - Mol Basis Dis* **1317**: 95–100.

Kurien BT, Scofield RH (2009) Oral administration of heat-solubilized curcumin for potentially increasing curcumin bioavailability in experimental animals. *Int J Cancer* **125**: 1992–1993.

Lampe V, Milobedzka J (1913) Studien über Curcumin. *Berichte der Dtsch Chem Gesellschaft* **46**: 2235–2240.

Lao CD, Ruffin MT, Normolle D, Heath DD, Murray SI, Bailey JM, Boggs ME, Crowell J, Rock CL, Brenner DE (2006) Dose escalation of a curcuminoid formulation. *BMC Complement Altern Med* **6**: 10.

Le DT, Brockstedt DG, Nir-Paz R, *et al.* (2012) A live-attenuated listeria vaccine (ANZ-100) and a live-attenuated listeria vaccine expressing mesothelin (CRS-207) for advanced cancers: Phase I studies of safety and immune induction. *Clin Cancer Res* **18**: 858–868.

Le DT, Wang-Gillam A, Picozzi V, *et al.* (2015) Safety and survival with GVAX pancreas prime and *Listeria monocytogenes*-expressing mesothelin (CRS-207) boost vaccines for metastatic pancreatic cancer. *J Clin Oncol* **33**: 1325–1333.

Lemmon MJ, van Zijl P, Fox ME, Mauchline ML, Giaccia AJ, Minton NP, Brown JM (1997) Anaerobic bacteria as a gene delivery system that is controlled by the tumor microenvironment. *Gene Ther* **4**: 791–796.

Leonard E, Lim KH, Saw PN, Koffas M a G (2007) Engineering central metabolic pathways for high-level flavonoid production in *Escherichia coli*. *Appl Environ Microbiol* **73**: 3877–3886.

Leonard E, Yan Y, Fowler Z, Li Z (2008) Strain Improvement of Recombinant *Escherichia coli* for Efficient Production of Plant Flavonoids. *Mol Pharmac* **5**: 257–265.

Lussier FX, Colatriano D, Wiltshire Z, Page JE, Martin VJ (2012) Engineering Microbes for Plant Polyketide Biosynthesis. *Comput Struct Biotechnol J* **3**(4): 1–11.

Machado D (2012) *Design and optimization of a synthetic pathway for curcumin production in Escherichia coli*. Project report. Braga.

Majumdar APN, Banerjee S, Nautiyal J, Patel BB, Patel V, Du J, Yu Y, Elliott A a, Levi E, Sarkar FH (2009) Curcumin synergizes with resveratrol to inhibit colon cancer. *Nutr Cancer* **61**: 544–553.

Makrides SC (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev* **60**: 512–538.

Mancarella S, Greco V, Baldassarre F, Vergara D, Maffia M, Leporatti S (2015) Polymer-Coated Magnetic Nanoparticles for Curcumin Delivery to Cancer Cells. *Macromol Biosci* **15**(10): 1365-1374.

Marisch K, Bayer K, Scharl T, Mairhofer J, Krempl PM, Hummel K, Razzazi-Fazeli E, Striedner G (2013a) A Comparative Analysis of Industrial *Escherichia coli* K-12 and B Strains in High-Glucose Batch Cultivations on Process-, Transcriptome- and Proteome Level. *PLoS One* **8**(8): e70516.

Marisch K, Bayer K, Cserjan-Puschmann M, Luchner M, Striedner G (2013b) Evaluation of three industrial *Escherichia coli* strains in fed-batch cultivations during high-level SOD protein production. *Microb Cell Fact* **12**(1): 1-11.

Martínez K, de Anda R, Hernández G, Escalante A, Gosset G, Ramírez OT, Bolívar FG (2008) Coutilization of glucose and glycerol enhances the production of aromatic compounds in an *Escherichia coli* strain lacking the phosphoenolpyruvate: carbohydrate phosphotransferase system. *Microb Cell Fact* **7**(1): 1-12.

Martínez-Gómez K, Flores N, Castañeda HM, Martínez-Batallar G, Hernández-Chávez G, Ramírez OT, Gosset G, Encarnación S, Bolivar F (2012) New insights into *Escherichia coli* metabolism: carbon scavenging, acetate metabolism and carbon recycling responses during growth on glycerol. *Microb Cell Fact* **11**: 46.

Massa PE, Paniccia A, Monegal A, De Marco A, Rescigno M (2013) *Salmonella* engineered to express CD20-targeting antibodies and a drug-converting enzyme can eradicate human lymphomas. *Blood* **122**: 705–714.

Matsuoka Y, Shimizu K (2012) Importance of understanding the main metabolic regulation in response to the specific pathway mutation for metabolic engineering of *Escherichia coli*. *Comput Struct Biotechnol J* **3**: 1–10.

Maxwell L (1995) Therapeutic Ultrasound and the Metastasis of a Solid Tumor. *J Sport Rehabil* **4**: 273–281.

Miyahisa I, Kaneko M, Funa N, Kawasaki H, Kojima H, Ohnishi Y, Horinouchi S (2005) Efficient production of (2S)-flavanones by *Escherichia coli* containing an artificial biosynthetic gene cluster. *Appl Microbiol Biotechnol* **68**: 498–504.

Neidhardt FC, Bloch PL, Smith DF (1974) Culture medium for enterobacteria. *J Bacteriol* **119**: 736–747.

Nemunaitis J, Cunningham C, Senzer N, Kuhn J, Cramm J, Litz C, Cavagnolo R, Cahill A, Clairmont C, Sznol M (2003) Pilot trial of genetically modified, attenuated *Salmonella* expressing the *E. coli* cytosine deaminase gene in refractory cancer patients. *Cancer Gene Ther* **10**: 737–744.

Neubauer P, Lin HY, Mathiszik B (2003) Metabolic load of recombinant protein production: Inhibition of cellular capacities for glucose uptake and respiration after induction of a heterologous gene in *Escherichia coli*. *Biotechnol Bioeng* **83**: 53–64.

Neutzling F, Gazal M, Ribeiro C, Pinto M, Ghisleni G (2016) Curcumin in depressive disorders : An overview of potential mechanisms , preclinical and clinical fi ndings. **784**: 192–198.

Nielsen DR, Yoon SH, Yuan CJ, Prather KLJ (2010) Metabolic engineering of acetoin and meso-2,3-butanediol biosynthesis in *E. coli*. *Biotechnol J* **5**: 274–284.

Nugroho AE, Yuniarti N, Istyastono EP, Supardjan, Maeyama K, Hakim L (2009) Anti-Allergic Effects Of 1,5-Bis(4'-Hydroxy-3'-

Methoxyphenyl)-1,4-Pentadiene-3-One On Mast Cell-Mediated Allergy Model. *Malaysian J Pharm Sci* **7**: 51–71.

Nuyts S, Mellaert L Van, Theys J, Landuyt W, Lambin P, Mellaert L Van (2014). The Use of Radiation-Induced Bacterial Promoters in Anaerobic Conditions: A Means to Control Gene Expression in *Clostridium*-Mediated Therapy for Cancer **155**(5): 716–723.

Odot J, Albert P, Carlier A, Tarpin M, Devy J, Madoulet C (2004) *In vitro* and *in vivo* anti-tumoral effect of curcumin against melanoma cells. *Int J Cancer* **111**: 381–387.

Oppenheimer A (1937) Turmeric (Curcumin) in biliary diseases. *Lancet* **229**: 619–621.

Pagliaro M, Ciriminna R, Kimura H, Rossi M, Della Pina C (2007) From glycerol to value-added products. *Angew Chemie - Int Ed* **46**: 4434–4440.

Paramera EI, Konteles SJ, Karathanos VT (2011) Microencapsulation of curcumin in cells of *Saccharomyces cerevisiae*. *Food Chem* **125**: 892–902.

Park SY, Kim DSHL (2002) Discovery of natural products from *Curcuma longa* that protect cells from beta-amyloid insult: A drug discovery effort against Alzheimer's disease. *J Nat Prod* **65**: 1227–1231.

Paton AW, Morona R, Paton JC (2012) Bioengineered microbes in disease therapy. *Trends Mol Med* **18**: 417–425.

Podolsky T, Fong ST, Lee BT (1996) Direct selection of tetracycline-sensitive *Escherichia coli* cells using nickel salts. *Plasmid* **36**: 112–115.

Potvin G, Ahmad A, Zhang Z (2012) Bioprocess engineering aspects of heterologous protein production in *Pichia pastoris*: A review. *Biochem Eng J* **64**: 91–105.

Prasad S, Tyagi AK, Aggarwal BB (2014) Recent developments in delivery, bioavailability, absorption and metabolism of curcumin: The golden pigment from golden spice. *Cancer Res Treat* **46**: 2–18.

Quan J, Tian J (2014) Circular polymerase extension cloning. *Methods Mol Biol* **1116**: 103–117.

Ramírez BG, Blázquez C, Gómez del Pulgar T, Guzmán M, de Ceballos ML (2005) Prevention of Alzheimer's disease pathology by cannabinoids: neuroprotection mediated by blockade of microglial activation. *J Neurosci* **25**: 1904–1913.

Ramirez-Ahumada MDC, Timmermann BN, Gang DR (2006) Biosynthesis of curcuminoids and gingerols in turmeric (*Curcuma longa*) and ginger (*Zingiber officinale*): Identification of curcuminoid synthase and hydroxycinnamoyl-CoA thioesterases. *Phytochemistry* **67**: 2017–2029.

Ravindran J, Prasad S, Aggarwal BB (2009) Curcumin and cancer cells: how many ways can curry kill tumor cells selectively? *AAPS J* **11**: 495–510.

Reeves AZ, Spears WE, Du J, Tan KY, Wagers AJ, Lesser CF (2015) Engineering *E. coli* into a protein delivery system for mammalian cells. *ACS Synth Biol* **4**(5): 644-654.

Reisch CR, Prather KLJ (2015) The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in *Escherichia coli*. *Sci Rep* **5**: 15096.

Resmi MS, Soniya E V (2012) Molecular cloning and differential expressions of two cDNA encoding Type III polyketide synthase in different tissues of *Curcuma longa* L. *Gene* **491**: 278–283.

Ringman JM, Frautschy SA, Cole GM, Masterman DL, Cummings JL (2005) A potential role of the curry spice curcumin in Alzheimer's disease. *Curr Alzheimer Res* **2**: 131–136.

Roberts NJ, Zhang L, Janku F, *et al.* (2014) Intratumoral injection of Clostridium novyi-NT spores induces antitumor responses. *Sci Transl Med* **6**: 249ra111.

Rodrigues JL, Sousa M, Prather KLJ, Kluskens LD, Rodrigues LR (2014) Selection of Escherichia coli heat shock promoters towards their application as stress probes. *J Biotechnol* **188**: 67–71.

Rodrigues JL, Araújo RG, Prather KLJ, Kluskens LD, Rodrigues LR (2015a) Heterologous production of caffeic acid from tyrosine in *Escherichia coli*. *Enzyme Microb Technol* **71**: 36–44.

Rodrigues JL, Prather KLJ, Kluskens LD, Rodrigues LR (2015b) Heterologous Production of Curcuminoids. *Microbiol Mol Biol Rev* **79**: 39–60.

Rodrigues JL, Araújo RG, Prather KLJ, Kluskens LD, Rodrigues LR (2015c) Production of curcuminoids from tyrosine by a metabolically engineered Escherichia coli using caffeic acid as an intermediate. *Biotechnol J* **10**: 1–11.

Rodrigues JL de LC (2014) Design and construction of a new biosynthetic pathway for the production of curcuminoids in *Escherichia coli*. PhD Thesis. Universidade do Minho.

Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Front Microbiol* **5**: 1–17.

Rubinstein A (1990) Microbially controlled drug delivery to the colon. *Biopharm Drug Dispos* **11**: 465–475.

Santos CNS, Koffas M, Stephanopoulos G (2011) Optimization of a heterologous pathway for the production of flavonoids from glucose. *Metab Eng* **13**: 392–400.

Sharan SK, Thomason LC, Kuznetsov SG, Court DL (2009) Recombineering: A Homologous Recombination-Based Method of Genetic Engineering. *Nat Protoc* **4**: 206–223.

Shiloach J, Kaufman J, Guillard a S, Fass R (1996) Effect of Glucose Supply Strategy on Acetate Accumulation, Growth, and Recombinant Protein Production by *Escherichia coli* BL21 (λDE3) and *Escherichia coli* JM109. *Biotechnology* **49**(4):

421–428.

Shiloach J, Reshamwala S, Noronha SB, Negrete A (2010) Analyzing metabolic variations in different bacterial strains, historical perspectives and current trends - example *E. coli*. *Curr Opin Biotechnol* **21**: 21–26.

Shindikar A, Singh A, Nobre M, Kirolikar S (2016) Curcumin and Resveratrol as Promising Natural Remedies with Nanomedicine Approach for the Effective Treatment of Triple Negative Breast Cancer. *J Oncol* **2016**: 1–13.

Silva GP, Mack M, Contiero J (2009) Glycerol: A promising and abundant carbon source for industrial microbiology. *Biotechnol Adv* **27**: 30–39.

Song CW, Lee SY (2013) Rapid one-step inactivation of single or multiple genes in *Escherichia coli*. *Biotechnol J* **8**: 776–784.

Sørensen HP, Mortensen KK (2005) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J Biotechnol* **115**: 113–128.

Stock PG, Bluestone J a (2004) β-cell replacement for type I diabetes. *Annu Rev Med* **55**: 133–156.

Stritzker J, Weibel S, Hill PJ, Oelschlaeger T a., Goebel W, Szalay A a. (2007) Tumor-specific colonization, tissue distribution, and gene induction by probiotic *Escherichia coli* Nissle 1917 in live mice. *Int J Med Microbiol* **297**: 151–162.

Studier FW, Daegelen P, Lenski RE, Maslov S, Kim JF (2009) Understanding the Differences between Genome Sequences of *Escherichia coli* B Strains REL606 and BL21(DE3) and Comparison of the *E. coli* B and K-12 Genomes. *J Mol Biol* **394**: 653–680.

Sun Y-M, Wang R-X, Yuan S-L, Lin X-J, Liu C-B (2004) Theoretical study on the antioxidant activity of curcumin. *Chinese J Chem* **22**: 827–830.

Tas H, Nguyen CT, Patel R, Kim NH, Kuhlman TE (2015) An integrated system for precise genome modification in *Escherichia coli*. *PLoS One* **10**: 1–19.

Tian JQ and J, Quan J (2009) Circular Polymerase Extension Cloning of Complex Gene Libraries and Pathways. *PLoS One* **4**: e6441.

Tohda C, Nakayama N, Hatanaka F, Komatsu K (2006) Comparison of anti-inflammatory activities of six *Curcuma* rhizomes: A possible curcuminoid-independent pathway mediated by *Curcuma phaeocaulis* extract. *Evidence-based Complement Altern Med* **3**: 255–260.

Torre L a, Bray F, Siegel RL, Ferlay J, Lortet-tieulent J, Jemal A (2015) Global Cancer Statistics, 2012. *CA Cancer J Clin* **65**: 87–108.

Toso JF, Gill VJ, Hwu P, *et al.* (2002) Phase I study of the intravenous administration of attenuated *Salmonella typhimurium* to patients with metastatic melanoma. *J Clin Oncol* **20**: 142–152.

Tseng CP, Yu CC, Lin HH, Chang CY, Kuo JT (2001) Oxygen- and growth rate-dependent regulation of *Escherichia coli* fumarase (FumA, FumB, and FumC) activity. *J Bacteriol* **183**: 461–467.

Vogel A PJ (1815) Examen chimique de la racine de *Curcuma*. *J Pharmacol* **1**: 289–300.

Wang H, Wang F, Wang W, Yao X, Wei D, Cheng H, Deng Z (2014) Improving the expression of recombinant proteins in *E. coli* BL21 (DE3) under acetate stress: An alkaline ph shift approach. *PLoS One* **9**: 1–11.

Wang S, Zhang S, Zhou T, Zeng J, Zhan J (2013) Design and application of an *in vivo* reporter assay for phenylalanine ammonia-lyase. *Appl Microbiol Biotechnol* **97**: 7877–7885.

Watts KT, Lee PC, Schmidt-Dannert C (2006) Biosynthesis of plant-specific stilbene polyketides in metabolically engineered *Escherichia coli*. *BMC Biotechnol* **6**(1): 1-12

Weickert MJ, Doherty DH, Best E a, Olins PO (1996) Optimization of heterologous protein production in *Escherichia coli*. *Curr Opin Biotechnol* **7**: 494–499.

Wilken R, Veena MS, Wang MB, Srivatsan ES (2011) Curcumin: A review of anti-cancer properties and therapeutic activity in head and neck squamous cell carcinoma. *Mol Cancer* **10**(1): 1-19.

Wong WKR, Fu Z, Wang YY, Ng KL, Chan AKN (2012) Engineering of efficient *Escherichia coli* excretion systems for the production of heterologous proteins for commercial applications. *Recent Patents Chem Eng* **5**: 45–55.

Wu J, Du G, Zhou J, Chen J (2013) Metabolic engineering of *Escherichia coli* for (2S)-pinocembrin production from glucose by a modular metabolic strategy. *Metab Eng* **16**: 48–55.

Xie Z, Ma X, Gang DR (2009) Modules of co-regulated metabolites in turmeric (*Curcuma longa*) rhizome suggest the existence of biosynthetic modules in plant specialized metabolism. *J Exp Bot* **60**: 87–97.

Xu P, Ranganathan S, Fowler ZL, Maranas CD, Koffas M a G (2011) Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA. *Metab Eng* **13**: 578–587.

Zhang H, Stephanopoulos G (2013) Engineering *E. coli* for caffeic acid biosynthesis from renewable sugars. *Appl Microbiol Biotechnol* **97**: 3333–3341.

Zhang L, Gao B, Wang X, Zhang Z, Liu X, Wang J, Mo T, Liu Y, Shi S, Tu P (2016) Identification of a new curcumin synthase from ginger and construction of a curcuminoid-producing unnatural fusion protein diketide-CoA synthase::curcumin synthase. *RSC Adv* **6**: 12519–12524.

Zhang X, Chen Q, Wang Y, Peng W, Cai H (2014) Effects of curcumin on ion channels and transporters. *Front Physiol* **5**: 1–6.

Zhang Y, Li SZ, Li J, Pan X, Cahoon RE, Jaworski JG, Wang X, Jez JM, Chen F, Yu O (2006) Using unnatural protein fusions to engineer resveratrol biosynthesis in yeast and mammalian cells. *J Am Chem Soc* **128**: 13030–13031.

Zhao M, Yang M, Li XM, Jiang P, Baranov E, Li S, Xu M, Penman S, Hoffman RM, Trump DL (2005) Tumor-targeting bacterial therapy with amino acid auxotrophs of GFP-expressing *Salmonella typhimurium*. *Urol Oncol Semin Orig Investig* **102**(3): 755-760.

APPENDIX I – GENE DELETION BY LAMBDA RED–MEDIATED, LINEAR DNA–BASED DELETION METHOD

The recombineering method used consisted in an optimized procedure of the homologous recombination-mediated genetic engineering, a Lambda Red–Mediated, Linear DNA–Based Deletion Method, well described by Kolisnychenko *et al.* (2002) and Fehér *et al.* (2007).

In this method, a small deletion cassette with Cm resistance is recombineered into the desired location in the genome by λ-Red mediated homologous recombination, and serves as a subsequent target for cleavage by I-*Sce*I and recombination by λ-Red again. Successful integration of the landing pad is positively selected through growth on a medium containing Cm.

For this method, the following plasmids were used: the suicide pKD46 plasmid that expresses the Red operon under the control of arabinose repressor and confers ampicillin (Amp) resistance; the template pSG76-CS plasmid for the generation of the deletion cassette that confers resistance to chloramphenicol (Cm); the pSTKST plasmid expressing I-*Sce*I Meganuclease under the control of Tet repressor and conferring resistance to kanamycin (Kan); the pKSUC1 which expresses the I-*Sce*I constitutively and has Kan resistance.

The pKD46 and pSTKST plasmids were extracted and digested to verify their correct size. pKD46 was digested with *Bam*HI-HF (NEB) and *Nco*I-HF (NEB) and pSTKST was digested with *Nde*I (NEB) and *Sph*I-HF (NEB).

Primers

To accomplish each deletion, 5 primers were required to construct and confirm the deletion cassette. The primers designed in this work for the desired knockouts are described in Table A1. Primers A contained a homologous region to the zone immediately before the gene and a reverse complement region to the primers B. Primers B included a zone that pairs with the beginning of the pSG76-CS plasmid and a zone reverse complement of the region after the gene. Primers C had a region reverse complement of the end of the gene, as well as a zone that pairs with the end of the pSG76-CS for the generation of the deletion cassette. The primers D and E are sequencing primers required to confirm if the deletion occurs and could also be used for colony PCR (also used for the other deletion approaches). As the *fumC* gene is immediately followed by the gene *fumA* in the *E. coli* genome, the primer D needed to confirm *fumC* deletion was the fumA_D because the *fumA* gene would already be absent from the genome.

Besides, primers homologous to the Cm resistance gene (Cm_Fw and Cm_Rv) were also useful to confirm the introduction of the Cm resistance cassette.

Table A 1. Primers used in the construction and confirmation of the deletion cassette with Cm resistance. Primers D and E were designed to confirm the deletion of the gene

Gene to delete	Primer Name	Primer Sequence[a] 5'-3'
gnd	gnd_D	TCAGCATCCCGGTAGGCTTCA
	gnd_E	CGGTTGTTGATTGGTGCGCAGG
fumA	fumA_A	CACCCGCCCAGAGCATAACCAAACCAGGCAGTAAGTGAGAGAACA**ACAGAGCCGCCCTT**
	fumA_B	CTAAGCTCTCATGGCTAACGCGTGCCATGTAAAAAAACCGCCCCG**AAGGGCGGCTCTGT**
	fumA_C	TTATTTCACACAGCGGGTGCATTGTGTGAGTTGTATCTAGAAACTCAGAAGGTTCGTCC
	fumA_D	GCAAAAAGTCGTACTAGTCTC
	fumA_E	GAATCTTTTTCGCTGCGTAC
fumB	fumB_A	CACGCCATTTTCGAATAACAAATACAGAGTTACAGGCTGGAAGCT**CCTCTTCGGCCCAG**
	fumB_B	CTAAGCTCTCATGGCTAACGTCACCTGGCAGCATGCTGCCAGGCG**CTGGGCCGAAGAGG**
	fumB_C	TTACTTAGTGCAGTTCGCGCACTGTTTGTTGACGATTTAGAAACTCAGAAGGTTCGTCC
	fumB_D	CATCAACCACAGCTTTATTCTG
	fumB_E	GAATTAACGGCTTGAGTGTG
fumC	fumC_A	GAAAGCAAAACAGAAAGAAAAAATTAATCAGGTGAGGAGCAGGTC**TCTGCAACATACAG**
	fumC_B	CTAAGCTCTCATGGCTAACGTCGTTTGATCATTCCACGGCTGCAC**CTGTATGTTGCAGA**
	fumC_C	TTAACGCCCGGCTTTCATACTGCCGACCATCTGTTCTGAGAAACTCAGAAGGTTCGTCC
	fumC_E	GATCAAAAACAAGTCCAACACG
-	Cm_Fw	GCCCTGCCACTCATCGCAGT
	Cm_Rv	CCGTTGATATATCCCAATGGCATCGT

[a]The underlined sequence indicates homology regions with pSG76-CS for the deletion cassette construction. The **bold** sequence represents de homology between A and B primers.

Deletion cassette construction

To make the primer AB, 10 pmol of primer A were mixed with 10 pmol of primer B, and a PCR was performed in a total volume of 50 µL with KAPA HiFi DNA Polymerase (Kapa Biosystems). The PCR programs used are described on Table A2.

Table A 2. PCR conditions for the generation of the AB primer, using A and B primers and KAPA HiFi DNA Polymerase

Cycling step		***fumA* deletion**	***fumB* deletion**	***fumC* deletion**
Initial denaturation		95°C, 5 min		
Denaturation	15 x	98°C, 20 s		
Annealing		64°C, 15 s	62°C, 15 s	50°C, 15 s
Extension		72°C, 15 s		
Final extension		72°C, 5 min		

In the next step, 1 µL of the resultant PCR product was mixed with 10 pmol of each A and C primers and 50 ng of pSG76-CS plasmid, and a second PCR was performed (4x50 µL) using KAPA HiFi DNA Polymerase (Kapa Biosystems). The cycle parameters are described on Table A3.

Table A 3. Second PCR conditions for the deletion cassette construction with primers AB, A and C and KAPA HiFi DNA Polymerase

Cycling step		***fumA* deletion**	***fumB* deletion**	***fumC* deletion**
Initial denaturation		95°C, 5 min		
Denaturation	15 x	98°C, 20 s		
Annealing		64°C, 30 s	62°C, 30 s	50°C, 30 s
Extension		72°C, 15 s		
Denaturation	28 x	98°C, 20 s		
Annealing		63°C, 15 s		
Extension		72°C, 53 s		
Final extension		72°C, 5 min		

Deletion protocol

An electrophoresis of the resulting PCR generated fragment was then run in an agarose gel 0.7 % (w/v) to confirm the presence of the deletion cassette at the correct size. The cassette was subsequently purified with a cleaning kit (NucleoSpin® Gel and PCR Clean-up, Macherey-Nagel) and eluted on 20-30 µL water.

E. coli K-12 MG1655 (DE3) electrocompetent cells harboring pKD46 were grown overnight at 30°C and 200 rpm in 5 mL of LB medium with 100 µg/mL of Amp. Then, 50 mL of LB+Amp medium was inoculated with 1 mL of the overnight culture and grown at 30°C. At the early logarithmic phase, an

OD_{600} ≈ 0.3 read in a microplate spectrophotometer (about 2.5 h from the start), it was added 1 mL of 10% L-arabinose and the culture was harvested at an OD_{600} ≈ 0.35 read in a microplate spectrophotometer (about 3 h from the start). The culture was placed on ice and the protocol to prepare *E. coli* competent cells was followed as described in **Section 3.3**.

About 2 µL (100 to 500 ng) targeting DNA fragment, the Cm resistance cassette, was electroporated into 20 µL of electrocompetent cells. After the recovery step for 1 hour at 37°C cells were plated on LB+Cm agar plates.

The insertion of the fragment was checked by colony PCR using primers D and E or, when the deletion cassette was the same length as the gene that needed to be replaced, the combination of primers D and Cm_Rv or Cm_Fw and E were used. The PCR program used when primers D and E were used is described on Table A4. PCR conditions for the *gnd* deletion were used to confirm several colonies from the previous study (Araújo, 2014).

Table A 4. PCR conditions for colony PCR to confirm the insertion of the cassette Cm using primers D and E and KAPA *Taq* DNA polymerase

Cycling step		*gnd* deletion	*fumA* deletion	*fumB* deletion	*fumC* deletion
Initial denaturation		95°C, 10 min			
Denaturation	35 x	95°C, 30 s			
Annealing		56°C, 30 s	46°C, 30 s	47°C, 30 s	47°C, 30 s
Extension		72°C, 1 min	72°C, 170 s	72°C, 170 s	72°C, 170 s
Final extension		72°C 5 min			

Tables A5 and A6 describe the PCR programs used to confirm the presence of the deletion cassette, using primers Cm_Rv and D or Cm_Fw and E, respectively. The same parameters were used to confirm the gene deletion.

Table A 5. PCR conditions for colony PCR to confirm the insertion of the cassette Cm using primers D and Cm_Rv and KAPA *Taq* DNA polymerase

Cycling step		*fumA* deletion	*fumB* deletion	*fumC* deletion
Initial denaturation		95°C, 10 min		
Denaturation	35 x	95°C, 30 s		
Annealing		46°C, 30 s	47°C, 30 s	46°C, 30 s
Extension		72°C, 170 s		
Final extension		72°C, 5 min		

Table A 6. PCR conditions for colony PCR to confirm the insertion of the cassette Cm using primers Cm_Fw and E and KAPA *Taq* DNA polymerase

Cycling step		*fumA* deletion	*fumB* deletion	*fumC* deletion
Initial denaturation		95°C, 10 min		
Denaturation	35 x	95°C, 30 s		
Annealing		47°C, 30 s	47°C, 30 s	48°C, 30 s
Extension		72°C, 170 s		
Final extension		72°C, 5 min		

Electrocompetent cells harboring the deletion cassette were prepared and transformed with pSTKST using a standard protocol (**Section 3.3**). Afterwards, the transformed cells were grown on LB+Kan+Cm plates at 30°C. Around 2-4 colonies were separately inoculated into 10 mL of LB+Kan medium supplemented with heat-treated CTc (chlortetracycline-hydrochloride) (25 μg/L final concentration) and grown for 72 h at 30°C and 200 rpm. 10^{-3} to 10^{-5} fold dilutions of the culture were prepared and spread on LB agar plates with Kan and CTc and incubated for 12 to 24 h at 30°C. The loss of the cassette in 15-28 colonies was screened by colony PCR using different combinations of the primers mentioned above. However, the cassette was never successfully removed. Positive colonies would be inoculated in LB medium and grown at 37°C to lose the helper plasmid pSTKST. The loss of antibiotic resistance should then be confirmed by inoculating separately these cultures in LB with Kan, Amp and Cm.

APPENDIX II – GENE DELETION BY LANDING PAD RECOMBINATION

The method described as the landing pad system in the literature (Kuhlman & Cox, 2010; Tas *et al.*, 2015) relies in the same principles as the previously described lambda Red–mediated, linear DNA–based deletion method in **Appendix I**. The landing pad is actually a deletion cassette, but different names are described for them in this thesis in order to distinguish the methods from the different authors. The small differences between the methods are: instead of having a deletion cassette with 1.6 kb, this landing pad is slightly smaller (1.3 kb) and has a Tet resistance gene instead of a Cm resistance gene. The landing pad has two 25 bp of homology to the genome, while the Cm deletion cassette has 20 bp. Besides, different constructed plasmids are used for protein expression.

The landing pad system was implemented to delete *fumA* and *fumC* (simultaneously) from the *E. coli* K-12 MG1655 (DE3) genome. Plasmids used include: pTKRED, the helper plasmid containing the spectinomycin resistance marker *aadA*, and harboring genes encoding the constitutively expressed *recA*, three λ-Red genes *gam*, *bet* and *exo* driven by a LacI regulated, IPTG inducible promoter and a I-*Sce*I endonuclease gene inducible with L-arabinose (Kuhlman & Cox, 2010). These genes are necessary for both steps of landing pad integration at the desired integration site of the chromosome and the improvement of the final recombination for gene deletion. pTKRED also contains a temperature-sensitive pSC101 replication origin, which maintains the plasmid at low copy number and allows an easy curing by growth at 42°C and screening against spectinomycin resistance. pTKS/CS plasmid was used as a PCR template to amplify the small “landing pad”, consisting of a *tetA* tetracycline resistance gene flanked by unique I-*Sce*I endonuclease recognition sites and small 25-bp landing pad regions (LP1 and LP2: standardized sites for priming and homologous recombination).

After the desired modification, the efficiency of selection for successfully modified cells is enhanced through negative selection against retention of the landing pad: the bacteria are inoculated into medium containing $NiCl_2$ after the integration step; $NiCl_2$ is selectively lethal to *E. coli* expressing *tetA* (Podolsky *et al.*, 1996), and hence those cells which are unsuccessfully modified and retain the *tetA* landing pad are eliminated from the population. The use of *tetA* resistance gene instead of the one that confers Cm resistance represents an advantage because it allows this negative selection.

Primers

The construction of the landing pad for *fumA* and *fumC* deletions was conducted using the primers listed in Table A7. Their design was done in a similar way to the primers for recombineering (**Appendix I**) since the strategy is based in the same principles.

Table A 7. Primers used for fumA and fumC genes deletion using the landing pad system

Primer Name	Primer Sequence[a] 5'-3'
T_fumAC_A	CACCCGCCCAGAGCATAACCAAACCAGGCAGTAAGTGAGAGAACA**TCTGCAACATACAG**
T_fumAC_B	TCACCGTTTGGACCTTGGGGCCGTATCGTTTGATCATTCCACGGCTGCAC**CTGTATGTTGCAGA**
T_fumAC_C	TTAACGCCCGGCTTTCATACTGCCGACCATCTGTTCTGTTGGCTTCAGGGATGAGGCGCCATC

[a] The underlined sequences indicate homology regions with the landing pad regions LP1 (T_fumAC_B) and LP2 (T_fumAC_C) from pTKS/CS for the landing construction. The **bold** sequences indicate homology between A and B primers.

Landing pad construction

The first step for the landing pad construction was the formation of the primer AB. A 50 µL reaction was prepared with 0.5 µM of primer T_fumAC_A, 0.5 µM of primer T_fumAC_B and the remaining components for a Phusion High-Fidelity DNA Polymerase (Thermo Scientific) mix. The PCR conditions used are in Table A8.

Table A 8. PCR conditions for primer AB formation to delete *fumA* and *fumC*. The polymerase used was Phusion High-Fidelity DNA polymerase.

Cycling step		*fumA* and *fumC* deletion
Initial denaturation		98°C, 30 s
Denaturation	15 x	98°C, 15 s
Annealing		42°C, 15 s
Extension		72°C, 15 s
Final extension		72°C, 5 min

Then, pTKS/CS was used as PCR template to amplify the 1.3 kb landing pad using the landing pad regions as standardized priming sites. The primers include 45/39-bp (start/end) sequence homology for the desired insertion location in the chromosome. Four 50 µL PCR reactions were prepared by adding to each reaction: 50 ng of the template (pTKS/CS), 0.5 µM of primer T_fumAC_A, 0.5 µM of the primer T_fumAC_C and the remaining components for a Phusion High-Fidelity DNA Polymerase (Thermo Scientific) mix. The cycling conditions for this PCR are given in Table A9.

Table A 9. PCR conditions for landing pad construction to delete *fumA* and *fumC*. The polymerase used was Phusion DNA polymerase

Cycling step		*fumA* and *fumC* deletion
Initial denaturation		98°C, 30 s
Denaturation	5 x	98°C, 15 s
Annealing		42°C, 15 s
Extension		72°C, 15 s
Denaturation	35 x	98°C, 15 s
Annealing		55°C, 15 s
Extension		72°C, 45 s
Final extension		72°C, 5 min

The resulting PCR product was digested with 1 µL *DpnI* per each 50 µL reaction for at least 2 h at 37°C. All products were run in a 1% agarose gel. The expected 1485 bp band was isolated from the gel and purified.

Deletion protocol

The desired *E. coli* host was transformed with the helper plasmid pTKRED, containing Spec resistance. After pTKRED transformation, cells were grown at 30°C in LB supplemented with Spec and 2 mM of IPTG (to induce the λ-Red genes) until the OD reached 0.5-0.6. Afterwards, cells were made electrocompetent and transformed (as described in **Section 3.3**) with about 100 ng of the landing pad previously constructed. Cells were then plated in LB with Spec+Tet+0.5% glucose.

Landing pad integration into the *E. coli* genome was confirmed by colony PCR using the primers fumA_D and fumC_E and KAPA *Taq* DNA polymerase (Kapa Biosystems). The PCR program is described in Table A10.

Table A 10. Colony PCR conditions for confirming landing pad integration using KAPA *Taq* Polymerase

Cycling step		***fumA* and *fumC* deletion**
Initial denaturation		95°C, 10 min
Denaturation	35 x	95°C, 30 s
Annealing		46°C, 30 s
Extension		72°C, 3:35 min
Final extension		72°C, 5 min

As the landing pad was never successfully introduced, this protocol was not continued. Individual positive colonies would be inoculated into 5 ml of SOC + 0.5% glycerol, 2 mM IPTG, and 0.2% (w/v) L-arabinose. After growing at 37°C for 1 h, 100 mg/ml of Spec would be added to the culture and the tubes would be transferred to a 30°C shaking incubator for 4 h. In this phase, A P_{araBAD}-driven I-*Sce*I endonuclease gene inducible with L-arabinose from pTKRED would be expressed and the endonuclease would cleave the recognition sites in the chromosome.

If the deletions were successful, the pTKRED could be easily cured by growing the culture overnight at 42°C and screening against Spec resistance, since this plasmid bears a temperature-sensitive pSC101 replication origin, which maintains the plasmid at low copy number and allows an easy curing by growth at 42°C.

APPENDIX III– RESULTS FROM RECOMBINEERING STRATEGIES

AIII.1 *fumA*, *fumB* and *fumC* deletion through Cm cassette recombineering

The deletion of *fumA*, *fumB* and *fumC* genes was attempted using the recombineering protocol. The first step of this method is the construction of the deletion cassette, which codifies for Cm resistance and is flanked by I-*Sce*I sites. Using primers with homology to the plasmid pSG76-CS and the genome region chosen for elimination, the PCR leads to a fragment that is then electroporated into target cells. Figure A1 shows the PCR product obtained after amplification using the primers FumA_AB and FumA_C (A) FumB_AB, FumB_C (B) and FumC_AB, FumC_C (C) and pSG76-CS as template. In all cases the expected band size was 1742 bp.

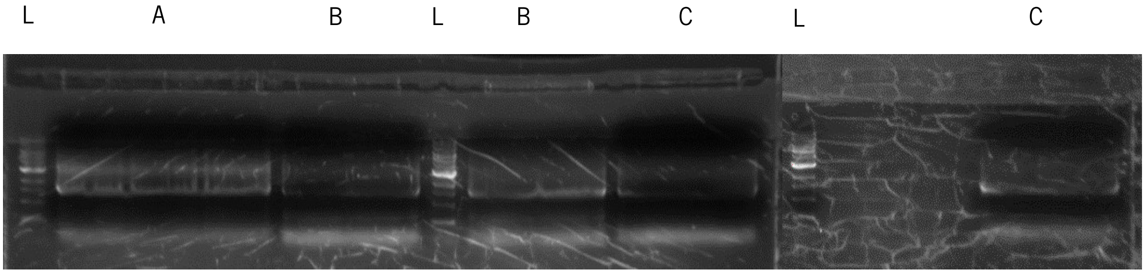


Figure A 1. Agarose gel (0.7% (w/v)) to confirm the presence of the deletion cassette with Cm resistance from pSG76-CS plasmid with primers FumA_AB and FumA_C (A) FumB_AB, FumB_C (B) and FumC_AB, FumC_C (C). The "L" lanes contained 1kb DNA ladder (NEB).

Figure A1 confirms that the deletion cassette for the knockout of *fumA*, *fumB* and *fumC* was successfully constructed. The resultant products were purified and the cassette for *fumA* deletion (A) (333 ng/µL) was electroporated into the *E. coli* K-12 MG1655 (DE3) harboring pKD46. The recombination occurs by replacing the gene to delete from the genome by the cassette that confers Cm resistance. The insertion of the fragment was confirmed by colony PCR by the amplification of a 623 bp fragment (Figure A2).

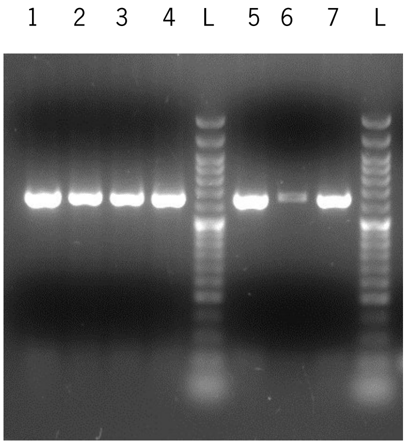


Figure A 2. Agarose gel (0.7% (w/v)) to confirm the insertion of the *fumA* deletion cassette with Cm resistance after electroporation. Colony PCR was performed using the primers Cm_Fw and Cm_Rv. The "L" lanes contained NZYDNA Ladder VI (NZYTech).

All colonies tested had the Cm resistance gene. In order to ensure that the deletion cassette was in the right place of the genome, and the fragment amplified was not from the pSG76-CS plasmid, colony PCR using primers fumA_D and Cm_Rv was performed in three of these positive colonies. The expected PCR product size was 1058 bp in case the cassette was inserted in the correct place. On the contrary, if the *fumA* gene was present, no DNA should be amplified. The results from colony PCR can be observed in Figure A3.

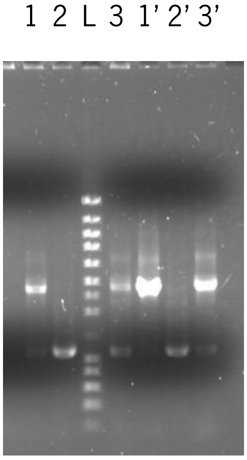


Figure A 3. Agarose gel (0.7% (w/v)) to confirm the correct insertion of the Cm resistance fragment on *E. coli* K-12 MG1655 (DE3) through colony PCR using primers fumA_D and Cm_Rv. NZYDNA ladder III (NZYTech) is present in lane L. Lanes 1-3 show the amplification from three different colonies and lanes 1'-3' correspond to the same colonies but using less cell suspension volume as template.

Besides the expected band (≈ 1000 bp), colonies 1 and 3 exhibited a band between 2500 and 3000 bp, colony 2 exhibited only the expected band. Probably, the deletion cassette was inserted in all three cases in the genome, but only the colony 2 had the cassette in the desired place. For this reason, colony 2 was further used to continue the protocol. pSTKST was transformed into these cells to remove the deletion cassette. One colony was inoculated into LB+Kan also supplemented with CTc to inactivate

Tet repressor, promoting the cleavage of the I-*Sce*I sites present on the inserted cassette, thus eliminating the deletion cassette. Afterwards, the elimination of the deletion cassette that was replacing the *fumA* gene needs to be confirmed. With this purpose, several colonies from the LB+Kan+CTc plates were tested by colony PCR using primers FumA_D and FumA_E. The expected bands were 295 bp if the cassette was successfully removed and 1915 bp in case the Cm resistance fragment was still present. Also, a PCR using the same colonies but using the Cm_Fw and Cm_Rv primers was performed to verify if the Cm gene was present. The results from both PCRs are shown in Figure A4.

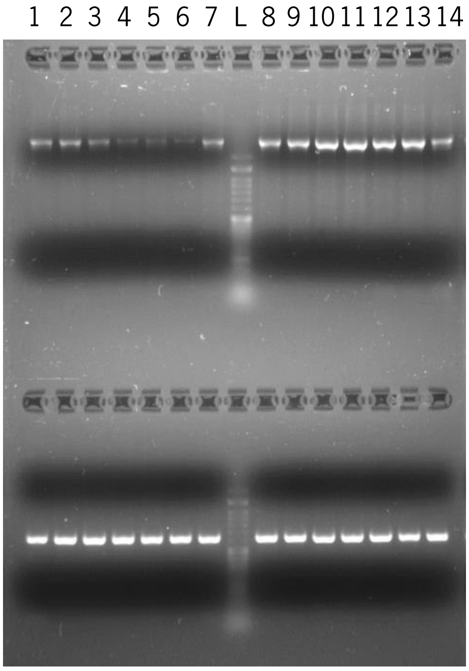


Figure A 4. Agarose gel (1.5% (w/v)) to confirm the elimination of the Cm resistance fragment from the genome of *E. coli* K-12 MG1655 (DE3) using a PCR with primers FumA_D and FumA_E (Up) and using primers Cm_Fw and Cm_Rv (Down). The “L” lanes contained NZYDNALadder VI (NZYTech).

The intended bands were not present in none of the tested colonies and it can be concluded that the Cm resistance gene remained in the genome as the 623 bp band is present when Cm_Fw and Cm_Rv primers were used.

The pKD46 and pSTKST plasmids were then extracted and digested as shown in Figure A5 to confirm their correct size. pKD46 was digested with *Bam*HI-HF and *Nco*I-HF (expected bands were 4094 bp and 2235 bp) and pSTKST was digested with *Nde*I and *Sph*I-HF (expected bands were 1384 bp and 3279 bp).

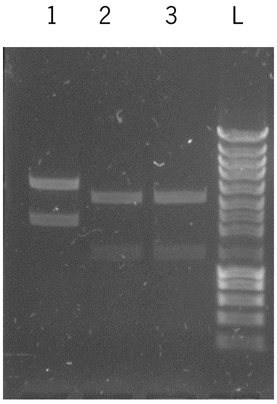
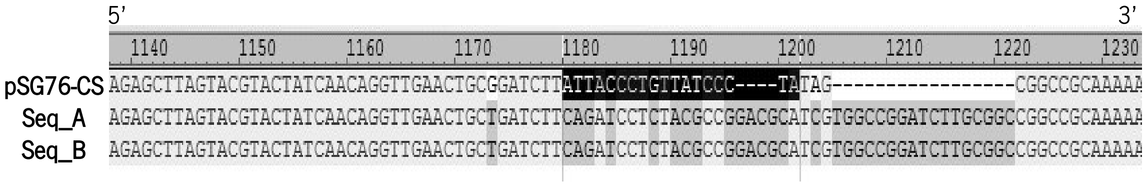


Figure A 5. Agarose gel 0.7% showing pKD46 digestion with *Bam*HI and *Nco*I (1) and pSTKST digestion with *Nde*I and *Sph*I (2 and 3). NZYDNALadder III (NZYTech) can be observed in L lane.

From Figure A5 it was possible to confirm that pKD46 and pSTKST digestions generated the expected bands. Several re-attempts were made to remove the deletion cassette but the same situation kept occurring, the Cm resistance gene continued in the genome, where the *fumA* gene used to be.

The whole protocol was repeated for the deletion of *fumB* but the same situation was observed, it was impossible to remove the deletion cassette from the *E. coli* genome.

As the problem could be related with the efficiency of the restriction enzyme (I-*Sce*I) that is expressed to remove the deletion cassette, a different plasmid that expresses I-*Sce*I constitutively (pKSUC1) was used. However, using this new plasmid it was also impossible to remove the deletion cassette from the genome. Therefore, as the problem of removing the cassette was not probably due to the I-*Sce*I expression (by pSTKST or pKSUC1), the pSG76-CS (used in deletion cassette construction) was sent to sequencing. The sequencing result was compared to the pSG76-CS plasmid sequence (Figure A6).



```
pSG76-CS AGAGCTTAGTACGTACTATCAACAGGTTGAACTGCGGATCTTATTACCCTGTTATCCC----TATAG-----------------CGGCCGCAAAAA
Seq_A    AGAGCTTAGTACGTACTATCAACAGGTTGAACTGCTGATCTTCAGATCCTCTACGCCGGACGCATCGTGGCCGGATCTTGCGGCCGGCCGCAAAAA
Seq_B    AGAGCTTAGTACGTACTATCAACAGGTTGAACTGCTGATCTTCAGATCCTCTACGCCGGACGCATCGTGGCCGGATCTTGCGGCCGGCCGCAAAAA
```

Figure A 6. Alignment between the expected pSG76-CS plasmid sequence (pSG76-CS) and the sequences obtained by sequencing using two different primers. One of the two I-*Sce*I restriction sites (5' A T T A C C C T G T T A T ↑ C C C T A 3') is highlighted in black. The alignment was done with the tool AlignX of Vector NTI Advance (Invitrogen).

In fact, there were several discrepancies between the expected and the obtained sequences. For example, one of the I-*Sce*I restriction sites, that is important to remove the cassette, was not ok, i.e. the sequence was completely different than the expected one. Therefore, the deletion cassette inserted could not be removed from the *fumA* or *fumB* location and this recombineering strategy was no longer used.

AIII.2 *fumA* and *fumC* deletion through recombineering of a landing pad

The protocol for gene deletion using the Landing Pad System starts with the landing pad construction. In order to delete *fumA* and *fumC* with this approach, the primers listed in Table 13 (**Appendix II**) were used since these two genes are near to one another in the *E. coli* genome, the primers were designed to delete them simultaneously. The constructed landing pad was run in an agarose gel (Figure A7).

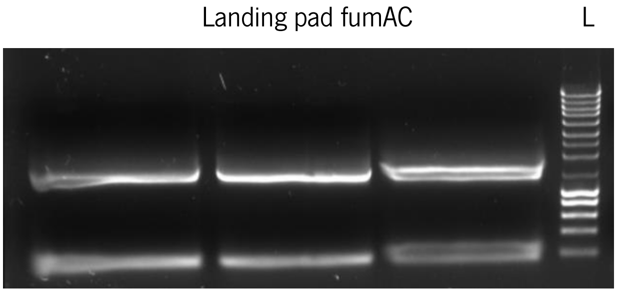


Figure A 7. Agarose gel (1% (w/v)) to visualize the constructed landing pad for *fumA* and *fumC* deletion. In L well it is present NZYDNALadder III (NZYTech).

The construction of the deletion cassette was confirmed by the band of 1485 bp (Figure A7). The other unspecific band (possibly ≈ 124 bp) can be the result of primer dimer formation. After extraction of the band DNA was digested with *DpnI*. *DpnI* cleaves only when its recognition site is methylated, so the plasmid that served as template for landing pad construction will be a substrate and only the linear PCR product should remain in the solution. The landing pad was then electroporated into *E. coli* BL21 (DE3) and *E. coli* K-12 MG1655 (DE3). *E. coli* BL21 (DE3) plate had no colonies in the next day. The insertion in colonies from *E. coli* K-12 MG1655 (DE3) plate was confirmed by colony PCR (Figure A8). The expected positive band was 1591 bp and in case the landing pad was not inserted, a 3505 bp band should be visible.

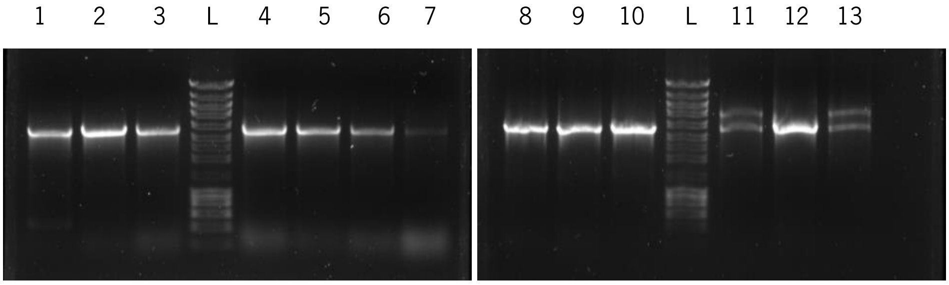


Figure A 8. Agarose gel (1% (w/v)) from colony PCR used to confirm if the deletion cassette was integrated into *E. coli* K-12 MG1655 (DE3) genome. Primers fumA_D and fumC_E were used. L lane represents the NZYDNALadder III (NZYTech).

The transformation of the landing pad was never efficient (especially in BL21 (DE3)). In addition, the *E. coli* K-12 MG1655 (DE3) colonies tested by colony PCR were all negative (Figure A8). Despite the use of *Dpn*I, some pTKS/CS molecules may have remained in solution and generated colonies with Spec resistance due to plasmid incorporation. Hence, the landing pad construction process was repeated and transformed again, but the same situation occurred. This could be due to several facts. For example, the size of the fragment (3193 bp) can be too big, and the bigger the fragment is the more difficult is to delete it given the reduced recombineering efficiency. In addition, the deletion of the two genes at the same time can cause some not predicted biological damage. Therefore, this strategy was no longer used, but the complete protocol is present in **Appendix II**.