Universidade do Minho
Escola de Engenharia

Márcia Regina Linhares Couto

Engineering a biosynthetic pathway for high-value glycosaminoglycans production in *Saccharomyces cerevisiae*

October 2023

Engineering a biosynthetic pathway for high-value glycosaminoglycans production in *Saccharomyces cerevisiae*

Márcia Couto

UMinho | 2023

Universidade do Minho
Escola de Engenharia

Márcia Regina Linhares Couto

Engineering a biosynthetic pathway for high value glycosaminoglycans production in *Saccharomyces cerevisiae*

Tese de Doutoramento
Programa doutoral em Bioengenharia

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

outubro de 2023

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho



Atribuição-NãoComercial-SemDerivações
CC BY-NC-ND

https://creativecommons.org/licenses/by-nc-nd/4.0/

ACKNOWLEDGMENTS

Gostaria de expressar a minha mais profunda gratidão a todos os que de forma direta ou indiretamente colaboraram na realização deste trabalho.

Começo naturalmente por agradecer à minha orientadora, Professora Lígia Rodrigues, pela competência, dedicação e sabedoria com que me orientou. Terei sempre uma grande admiração por si. Agradeço também à minha co-orientadora, Joana Rodrigues, pelo acompanhamento que me deu, pelas discussões científicas e pelos seus sensatos conselhos e pertinentes sugestões. Expresso um agradecimento à Adelaide Braga que tanto me ajudou nos ensaios em reator, obrigada pelos ensinamentos e discussões, pela disponibilidade, mas também pelas boas conversas. Agradeço também ao Óscar Dias por partilhar comigo os conhecimentos bioinformáticos tão relevantes para esta tese.

Às minhas queridas Diana e Rita, com quem desenvolvi uma eterna amizade durante as aulas no programa doutoral do MIT, foram os bens mais preciosos que ganhei com esta experiência! Aos meus amigos que me acolheram e acompanharam dentro e fora do CEB ao longo dos últimos 9 anos, em especial à Débora, Salomé, Ana Silva, Carla, Ana Brandão, Henrique, Eduardo Coelho, Tânia, Daniel e Eduardo Gudiña, muito obrigada pelo vosso apoio. Aos elementos ainda não mencionados da comitiva do sexto dia, João, Zuquim, Joaquim e Carlos, muito obrigada. Não posso também deixar de agradecer aos restantes meus colegas do LBBS, Daniela, João e Beatriz, que foram uma ótima companhia de quem sempre vou recordar-me com muito carinho.

Agradeço à minha tão grande e linda família, que foram o meu maior pilar e sem eles não teria sido possível ultrapassar todos os desafios com que me deparei durante esta jornada. Começo pelos meus pais, José e Maria Amélia, a quem não poderia estar mais grata por todos os seus ensinamentos e os seus esforços para que eu pudesse realizar os meus sonhos! Aos meus irmãos, Marco, Diogo e Joana, que além de muito amor incondicional e apoio trouxeram à minha vida uns maravilhosos cunhados, obrigada a todos vocês. À mais pequena, a minha amorosa sobrinha Mafalda, que durante esta caminhada me trouxe tanta alegria, obrigada pelo teu espírito inocente, a tua energia e as tuas brincadeiras. Ao meu namorado, Manuel Cunha, que me acompanhou mais de perto nesta longa e difícil jornada. Obrigada por me teres ouvido e por me incentivares e desafiares a ser a melhor versão de mim! Sem o teu encorajamento nada disto teria sido possível.

This study was supported by the Portuguese Foundation for Science and Technology (FCT) under the scope of the strategic funding of UIDB/04469/2020 unit. I would also like to acknowledge FCT for funding the individual PhD scholarship SFRH/BD/132998/2017 and consequent extension COVID/BD/152454/2022.

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

REPÚBLICA PORTUGUESA

UNIÃO EUROPEIA
Fundo Europeu de Desenvolvimento Regional

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

University of Minho, October 30st 2023

Márcia Couto

_____

Márcia Couto

ABSTRACT

Engineering a biosynthetic pathway for high value glycosaminoglycans production in *Saccharomyces cerevisiae*

Glycosaminoglycans (GAGs), such as hyaluronic acid, heparosan and chondroitin, have emerged as key substances with diverse applications in both the medical and cosmetic industries. Particularly, chondroitin is a crucial polysaccharide with anti-inflammatory properties, mostly used in the treatment of osteoarthritis. Historically, chondroitin has been extracted from animal sources, including shark fins and animal cartilage. However, rising concerns regarding the risks associated with the use of animal-derived products, coupled with the ecological repercussions of overfishing, have been driving the search for sustainable, alternative methodologies for chondroitin production. Inspired by the inherent biosynthetic capabilities of pathogenic strains of *Pasteurella multocida* and *Escherichia coli* to produce GAG-like polysaccharides, several studies have sought to use non-pathogenic hosts for the biosynthetic production of chondroitin. The aim of this thesis was to design and engineer an artificial pathway to produce chondroitin in *Saccharomyces cerevisiae*, due to its well-established status as industrial host organism, its genetic manipulability, and its robust fermentation capabilities. Alternative key genes were carefully curated and evaluated using literature mining and kinetic assays to streamline and optimize the production of chondroitin. The methodologies for bioinformatic metabolic flux prediction and optimization, as well as for the analysis of chondroitin production and quantification have been firstly evaluated and validated in a simpler organism, *E. coli*. Several targets were identified for under and overexpression that were able to enhance up to 1.9-fold the *in vivo* chondroitin production in *E. coli.* Further bioprocess scale-up with a mutant overexpressing lytic murein transglycosylase (*mltB*) led to 535 mg/L of chondroitin. Subsequently, *S. cerevisiae* harboring chondroitin production pathways was designed and engineered leading to up to 125 mg/L of chondroitin in flask. These titers might be improved in the future with several under and overexpressions enhancing the precursors availability predicted by the computational approach. This thesis represents a step forward toward a sustainable, cost-effective, and scalable platform for the large-scale production of chondroitin, thereby circumventing the challenges posed by conventional extraction methods and alleviating the environmental and availability concerns associated with animal-sourced chondroitin. Also, the methods herein used can be further applied to other GAG production processes.

Keywords: Chondroitin; Glycosaminoglycans; Metabolic engineering; Nutraceuticals; Synthetic biology

RESUMO

Engenharia de uma via biossintética para produção de glicosaminoglicanos de alto valor em *Saccharomyces cerevisiae*

Os glicosaminoglicanos (GAGs), como o ácido hialurónico, a heparosana e a condroitina, têm emergido como substâncias-chave com diversas aplicações tanto nas indústrias médica como cosmética. Em particular, a condroitina é um polissacarídeo crucial com propriedades anti-inflamatórias, utilizado sobretudo no tratamento da osteoartrite. Historicamente, a condroitina tem sido extraída de fontes animais, incluindo barbatana de tubarão e cartilagem animal. No entanto, preocupações crescentes em relação aos riscos associados ao uso de produtos de origem animal, juntamente com as repercussões ecológicas da sobrepesca, têm impulsionado a busca por metodologias alternativas e sustentáveis para a produção de condroitina. Inspirados pelas capacidades biossintéticas inerentes de estirpes patogénicas de *Pasteurella multocida* e *Escherichia coli* na produção de polissacarídeos semelhantes a GAGs, vários estudos têm procurado utilizar hospedeiros não patogénicos para a produção biosintética de condroitina. O objetivo desta tese foi projetar e engenhar uma via artificial para produzir condroitina em *Saccharomyces cerevisiae*, devido à sua consolidação como organismo hospedeiro industrial, à sua manipulabilidade genética e às suas robustas capacidades de fermentação. Genes-chave alternativos foram cuidadosamente selecionados e avaliados através de pesquisa bibliográfica e ensaios cinéticos para otimizar a produção de condroitina. As metodologias bioinformáticas para previsão e otimização do fluxo metabólico, bem como as de análise e quantificação da produção de condroitina, foram inicialmente avaliadas e validadas num organismo mais simples, *E. coli*. Vários alvos foram identificados para sub e sobrexpressão que foram capazes de aumentar a produção *in vivo* de condroitina em até 1.9 vezes. O aumento de escala do bioprocesso com um mutante sobreexpressando murina transglicosilase lítica (*mltB*) resultou em 535 mg/L de condroitina. Posteriormente, *S. cerevisiae* contendo vias de produção de condroitina foi projetada e desenvolvida, levando a até 125 mg/L de condroitina em matraz. Esta produção pode ser aprimorada no futuro com várias sub e sobreexpressões para aumentar a disponibilidade de precursores, como previsto pela ferramenta computacional. Esta tese representa um avanço em direção a uma plataforma sustentável, económica e escalável para a produção em larga escala de condroitina, contornando assim os desafios colocados pelos métodos convencionais de extração e aliviando as preocupações ambientais e de disponibilidade associadas à condroitina de origem animal. Além disso, os métodos aqui utilizados podem ser aplicados a outros processos de produção de GAGs.

Palavras-chave: Biologia sintética; Condroitina; Engenharia metabólica; Glicosaminoglicanos; Nutracêuticos

LIST OF CONTENTS

CHAPTER 1 GENERAL INTRODUCTION 1

1.1. Context and Motivation 2

1.2. Scope and aims 3

1.3. Thesis outline 3

Graphical abstract 5

1.4. References 5

CHAPTER 2 - STATE OF THE ART 7

2.1. Introduction 8

2.2. Occurrence and biological functions of chondroitin and its derivatives 10

2.3. Clinical applications of CS 14

2.4. Enzymatic and chemoenzymatic CS production 16

2.4.1. Enzymatic depolymerization of CS polysaccharides 17

2.4.2. Enzymatic synthesis of CS oligosaccharides and polysaccharides 17

2.5. Microbial production of GAGs 21

2.5.1. Biosynthetic pathways for microbial GAGs production 23

2.5.2. Current state of biotechnological GAGs production using microbes 25

2.5.3. Metabolic engineering to optimize CS production 32

2.5.4. Alternative Hosts 34

2.5.5. Additional optimization strategies 34

2.6. Conclusions 35

2.7. References 36

CHAPTER 3 - CLONING, EXPRESSION, AND CHARACTERIZATION OF UDP-GLUCOSE DEHYDROGENASES 56

3.1. Introduction 57

3.2. Materials and Methods 59

3.2.1. Multiple Sequence Alignment and Evolutionary Relationships 59

3.2.2. Strains and Plasmids 60
3.2.3. Gene Sources and Cloning Strategy 61
3.2.4. Culture Conditions 62
3.2.5. Preparation of *E. coli* and *S. cerevisiae* Protein Extracts 63
3.2.6. Purification of UGD 64
3.2.7. UGD *in vitro* Assays 64
3.2.8. Analysis and Quantification of UDP-sugars 65
3.2.9. Statistical Analysis 65
3.3. Results and Discussion 65
3.3.1. Sequences and Phylogenetic Analysis 65
3.3.2. *E. coli* as UGD Expression Host 69
3.3.3. Three Novel UGDs Expressed in Yeasts 73
3.3.4. Biochemical Characterization of Cloned UGDs 77
3.4. Conclusions 80
3.5. References 82
CHAPTER 4 - OPTIMIZATION OF CHONDROITIN PRODUCTION IN *E. COLI* USING GENOME SCALE MODELS 88
4.1. Introduction 89
4.2. Materials and Methods 90
4.2.1. Model construction and computational approach 90
4.2.2. Heterologous pathway construction 91
4.2.3. Mutant construction 93
4.2.4. Flask fermentations 94
4.2.5. Bioreactor operation 94
4.2.6. Analytic methods 95
4.3. Results and Discussion 96
4.3.1. Bioinformatic results for mutant prediction 96

4.3.2. *In vivo* validation of bioinformatic results 107
4.4. Conclusions 116
4.5. References 117
CHAPTER 5 - *IN SILICO* DESIGN OF *SACCHAROMYCES CEREVISIAE* STRAINS FOR IMPROVED PRODUCTION OF CHONDROITIN 124
5.1. Introduction 125
5.2. Materials and Methods 126
5.2.1. Model construction 126
5.2.2. Conditions for *in silico* simulations and optimization 127
5.2.3. Strains and plasmids 127
5.2.4 Biosynthetic pathway construction 129
5.2.5. Flask fermentation conditions 130
5.2.6. Analytical methods 130
5.3. Results and Discussion 131
5.3.1. Bioinformatic tool for identification of gene targets 131
5.3.2. Flasks fermentations 141
5.4. Conclusions 143
5.5. References 144
CHAPTER 6 - MAIN CONCLUSIONS AND FUTURE PERSPECTIVES 150
APPENDIX - FIGURES 155
APPENDIX - TABLES 163

LIST OF ABBREVIATIONS

APSK	Adenosine-5′-phosphosulfate kinase
ASST IV	Rat aryl sulfotransferase IV
ATPS	Adenosine-5'-triphosphate sulfurylase
BPCY	Biomass-Product Coupled Yield
CAGR	Compound annual growth rate
Cahi	*Capra hircus*
CHSY	Chondroitin synthase/polymerase
CMP	Cytidylate
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPRi	Clustered regularly interspaced short palindromic repeats interference
CS	Chondroitin sulfate
CS-O	Chondroitin oligosaccharides
dCMP	Deoxycytidylate
DMSO	Dimethyl sulfoxide
DNS	Dinitrosalicylic acid
DS	Dermatan sulfate
EA	Evolutionary algorithm
EULAR	European League Against Rheumatism
FBA	Flux balance analysis
fCS	Fucosylated CS
FVA	Flux variability analysis
GAG	Glycosaminoglycan
GalNAc	*N*-acetyl-D-galactosamine
Gal	Galactose
GalN	Galactosamine
GEM	Genome-scale metabolic model
GlcA	Glucuronic acid
GlcA-pNP	*p*-nitrophenyl glucuronide
GlcN	Glucosamine
GlcN-6-P	Glucosamine 6-phosphate
GlcNAc	N-acetylglucosamine
GRAS	Generally regarded as safe
HA	Hyaloronate/ hyaluronic acid
HP	Heparin
HPLC	High performance liquid chromatography
HS	Heparan sulfate
IdoA	Iduronic acid
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KfoA	UDP-GlcNAc 4-epimerase
KfoC	Chondroitin polymerase from *E. coli* K4

LIST OF ABBREVIATIONS (CONT.)

KfoE	Fructosyltransferase from *E. coli* K4
KfoF	UDP-glucose dehydrogenase from *E. coli* K4
KS	Keratan sulfate
Lajo	*Lactobacillus johnsonii*
LB	Lysogeny broth
***Lbj*UGD**	UGD enzyme from *Lactobacillus johnsonii*
LMOMA	Linear MOMA
MOMA	Minimization of Metabolic Adjustment simulation method
MurNAc	*N*-acetylmuramic acid
NAD+	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide reduced
NADPH	Nicotinamide adenine dinucleotide phosphate reduced
NagZ	β-N-acetylglucosaminidase
OD	Optical density
PAPS	3'-Phosphoadenosine-5'-phosphosulfate
PCR	Polymerase Chain Reaction
PEG	Polyethylene glycol
pFBA	Parsimonious FBA
PG	Proteoglycan
PmCS	Chondroitin synthase from *Pasteurella multocida*
RI	Refractive index detector
SA/SEA	Simulated annealing/ evolutionary algorithms
SDS–PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SOC	Super optimal broth
SPEA2	Strength Pareto Evolutionary Algorithm 2
UAE	UDP-glucose/UDP-N-acetylglucosamine 4-epimerase
UDP	Uridine diphosphate
UDP-GalNAc	Uridine diphosphate *N*-acetyl-D-galactosamine
UDP-Glc	UDP-glucose
UDP-GlcA	UDP-glucuronic acid
UDP-GlcNAc	UDP-acetylglucosamine
UGD	UDP-glucose 6-dehydrogenase
UMP	Uridine monophosphate
UTP	Uridine-5′-triphosphate
WYIELD	Weighted Yield
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
YNB	Yeast Nitrogen Base selective minimal medium
YPD	Yeast Extract Peptone Dextrose medium
***Zm*UGD**	UGD enzyme from *Zymomonas mobilis*
Zymo	*Zymomonas mobilis*

LIST OF FIGURES

CHAPTER 2

Figure 2.1. Structures of the main glycosaminoglycans (GAGs) a) hyaluronic acid, b) keratan sulfate, c) chondroitin and chondroitin sulfate, d) dermatan sulfate, and e) heparosan, heparan sulfate and heparin. Monomers of the disaccharide building blocks are abbreviated as GlcA - D-glucuronic acid, GlcNAc – *N*-acetyl-D-glucosamine, Gal – D-galactose, GalNAc – *N*-acetyl-D-galactosamine, IdoA – L-iduronic acid, GlcN, D-glucosamine. Hyaluronic acid (a) does not go under post-polymerization modifications. Keratan sulfate (b) has di-sulfated, mono-sulfated and non-sulfated disaccharide units (each R^6=H or SO_3H) due to *O*-sulfotransferases action. Chondroitin (c) is the simple non-sulfated backbone (R^2, R^3, R^4 and R^6=H) which can be modified by different tissue-specific *O*-sulfotransferases to form chondroitin sulfate (each R^2, R^3, R^4 and R^6=H or SO_3H). Dermatan sulfate (d) is formed from chondroitin through epimerization of GlcA into IdoA by tissue-specific epimerases followed by *O*-sulfotransferases (each R^2, R^4 and R^6=H or SO_3H). Heparosan (e) has non-modified sugar moieties, that can be further modified through actions of tissue-specific *N*-sulfotransferases, C5 epimerases and *O*-sulfotransferases to generate the sulfated forms heparan sulfate and heparin (R^2 from uronic acid=H or SO_3H; when the hexosamine unit is GlcN, R^2 in that unit=SO_3H, while R^2=Ac when the unit is GlcNAc; other groups R^3, R^6=H or SO_3H). Heparin has more sulfate groups and IdoA content than heparan sulfate. Depending on the GAG type and source the molecular size can generally vary between 4 and 200 mer (n = 4-200). Exceptionally, the highest size can be found for hyaluronic acid that can achieve 20,000 repeating units. 9

Figure 2.2. Enzymatic synthesis of chondroitin sulfate (CS). a) Depolymerizing enzymes, such as animal hyaluronidases, can be used to obtain CS oligosaccharides from CS polysaccharides; b) the same type of enzymes is able to, under different conditions, polymerize the CS oligosaccharides through chemoenzymatic approaches. c) Bacterial glycosyltransferases (such as chondroitin synthase from *Escherichia coli* K4, KfoC, or from *Pasteurella multocida* type F, *Pm*CS) act by transferring alternate residues of glucuronic acid (GlcA) and acetylgalactosamine (GalNAc), using uridine diphosphate (UDP)-GlcA and UDP-GalNAc as donors, to the nonreducing end of a chondroitin chain acceptor to elongate the chondroitin oligosaccharide/ polysaccharide backbones. Sulfotransferases such as 4-*O*-sulfotransferase (C4OST), 6-*O*-sulfotransferase (C6OST), *N*-acetylgalactosamine 4-sulfate 6-*O*-sulfotransferase (GalNAc4S-6OST), and 2-*O*-sulfotransferase (2OST) that require the presence of 3′-phosphoadenosine-5′-

phosphosulfate (PAPS) as sulfate donor, convert the unsulfated backbone (CS-O) in CS with different sulfation patterns such as CS-A, CS-C, CS-E and CS-T. Only CSs with a homogenous defined sulfation pattern are shown although a CS chain may have different CS units if a combination of sulfotransferases is used. Dashed arrows represent polymerization steps. 18

Figure 2.3. Production of glycosaminoglycans in microbes and its possible use in the biosynthesis of microbial chondroitin, hyaluronic acid or heparosan. Depending on the microbial host, the heterologous expression of the enzymes shown in orange boxes might be required for glycosaminoglycans production. Enzyme abbreviations: ABC, adenosine triphosphate (ATP)-binding cassette transporters; Aldo, fructose-6-phosphate aldolase; Fbp, fructose-1,6-bisphosphatase; GalU, uridine triphosphate:glucose-1-phosphate uridylyltransferase; Glk, glucokinase; GlmM, phosphoglucosamine mutase; GlmS, glucosamine-6-phosphate synthase; GlmU, glucosamine-1-phosphate N-acetyltransferase/N-acetylglucosamine-1-phosphate uridyltransferase; GlpF, Glycerol uptake facilitator protein; GlpK, glycerol kinase; Gpd, glyceraldehyde-3-phosphate dehydrogenase; HasA, hyaluronan synthase; KfiA, β-1,3-glucuronyltransferase; KfiC, α-1,4-N-acetylglucosaminyltransferase; KfoC, chondroitin synthase; Pfk, 6-phosphofructokinase; Pgi, glucose-6-phosphate isomerase; Pgm, phosphoglucomutase; PTS, phosphotransferase system; Uae, UDP-N-acetylglucosamine 4-epimerase; UGD, uridine diphosphate (UDP)-glucose 6-dehydrogenase..... 24

CHAPTER 3

Figure 3.1. Production of uridine diphosphate (UDP)-glucuronic acid in *Escherichia coli* and its possible use in the biosynthesis of microbial chondroitin, hyaluronic acid, or heparosan. Depending on the microbial host, the heterologous expression of the enzymes shown in orange might be required for glycosaminoglycans production. Enzyme abbreviations: GalU, uridine-triphosphate:glucose-1-phosphate uridylyltransferase; GlmM, phosphoglucosamine mutase; GlmS, glucosamine-6-phosphate synthase; GlmU, glucosamine-1-phosphate N-acetyltransferase/N-acetylglucosamine-1-phosphate uridyltransferase; HasA, hyaluronan synthase; KfiA, β-1,3-glucuronyltransferase; KfiC, α-1,4-*N*-acetylglucosaminyltransferase; KfoC, chondroitin synthase; Pgi, glucose-6-phosphate isomerase; Pgm, phosphoglucomutase; Uae, UDP-*N*-acetylglucosamine 4-epimerase; UGD, UDP-glucose 6-dehydrogenase. 58

Figure 3.2. Evolutionary relationship between 63 uridine diphosphate (UDP)-glucose 6-dehydrogenase (UGD)-glucose/guanosine diphosphate (GDP)-mannose 6-dehydrogenase (GMD) superfamily members. The resulting optimal tree has the sum of branch length = 12.04551231. The tree is drawn to scale, with branch lengths corresponding to the evolutionary distances used to infer the phylogenetic tree. The dataset comprised a total of 991 sites. Enzymes are labelled by the two first letters of the genus followed by two letters of species name. The three UGD expressed in this study are indicated with a black arrow (Lajo, *Lactobacillus johnsonii*; Cahi, *Capra hircus*; and Zymo, *Zymomonas mobilis*). Abbreviation of the protein is also shown after species name when not UGD, or when multiple UGD isoforms exist within same species. Protein abbreviations: gmd, GDP-mannose 6-dehydrogenase; uamad, UDP-*N*-acetyl-D-mannosamine dehydrogenase; uagad, UDP-*N*-acetylglucosamine 6-dehydrogenase; and nagaad, nucleotide diphosphate (NDP)-*N*-acetyl-*D*-galactosaminuronic acid dehydrogenase. 68

Figure 3.3. SDS–PAGE gel showing 6-His-tagged uridine diphosphate-glucose dehydrogenases (UGDs) expression in *Escherichia coli* BL21. Lane L on the left corresponds to the ladder Color Prestained Protein Standard, Broad Range (11–245 kDa). Samples loaded were: (**1**) soluble cell extract of *E. coli* pRSFDuet-1 (15 µg), (**2**) soluble cell extract from *E. coli* pRSFDuet_*Zm*UGD (15 µg), (**3**) 1st eluted fraction from *Zm*UGD purification (5 µg), (**4**) soluble cell extract from *E. coli* pRSFDuet_*Lbj*UGD (15 µg), (**5**) 1st eluted fraction from *Lbj*UGD purification (5 µg), (**6,7**) 2nd eluted fractions from *Zm*UGD and *Lbj*UGD purification, respectively (5 µg). 69

Figure 3.4. *In vivo* uridine diphosphate-glucose (UDP-Glc) and UDP-glucuronic acid (UDP-GlcA) production by *Escherichia coli* BL21 harboring pRSFDuet-1, pRSFDuet_*Zm*UGD or pRSFDuet_*Lbj*UGD, the last two expressing an extra copy of UDP-Glc dehydrogenase (UGD) from *Zymomonas mobilis* (*Zm*UGD) and *Lactobacillus johnsonii* (*Lbj*UGD), respectively. Results are normalized by the cultures' dry weight of cells at the end of the fermentation. The fermentations were performed in triplicate. Four asterisks (****) identify statistically significant differences with adjusted *p* values < 0.0001. 70

Figure 3.5. *In vitro* uridine diphosphate-glucose (UDP-Glc) consumption and UDP-glucuronic acid (UDP-GlcA) production by cell extracts expressed as concentration by UDP-Glc dehydrogenase (UGD) from *Zymomonas mobilis* (*Zm*UGD) and *Lactobacillus johnsonii*

(*Lbj*UGD) expressed in *Escherichia coli* BL21. Values from triplicate reactions. One asterisk (*) identifies adjusted *P* values < 0.05. 72

Figure 3.6. ***In vivo* uridine diphosphate (UDP)-sugars production by *Saccharomyces cerevisiae* CEN.PK2-1C and BY4741 wild types (wt.) and engineered strains expressing UDP-glucose dehydrogenase (UGD) from *Capra hircus* (*Ch*UGD), *Zymomonas mobilis* (*Zm*UGD) and *Lactobacillus johnsonii* (*Lbj*UGD)**. Standard deviations are relative to at least three different transformed colonies tested. Differences in the UDP-glucuronic acid (UDP-GlcA) concentrations between the mutants and the correspondent wild-type are indicated, with statistical significance represented with two asterisks (**) for adjusted *P* values <0.01 and with four asterisks (****) for *p* values <0.0001. Statistically non-significative differences are not shown for simplification..... 74

Figure 3.7. *In vitro* conversion of uridine diphosphate-glucose (UDP-Glc) into UDP-glucuronic acid (UDP-GlcA) by *Saccharomyces cerevisiae* CEN.PK2-1C and BY4741 cell extracts through the action of expressed UDP-Glc dehydrogenases (UGDs) from *Capra hircus* (*Ch*UGD), *Zymomonas mobilis* (*Zm*UGD) and *Lactobacillus johnsonii* (*Lbj*UGD). Standard deviations are relative to at least three independent assays..... 76

Figure 3.8. Characterization of recombinant uridine diphosphate-glucose dehydrogenases (UGDs) at different pH and temperature. *In vitro* reactions were performed using cell extracts containing recombinant UGD expressed in *Escherichia coli* BL21 (DE3) under T7 promoter control, on different pH at 30°C and on different temperatures at pH 8.0, respectively (**A, C**) and cell extracts containing recombinant UGDs expressed in *Saccharomyces cerevisiae* CEN.PK2-1C under PGK1 promoter, on different pH at 30°C and on different temperatures at pH 8.0, respectively (**B, D**). The 100% activity corresponds to the highest activity of (**A**) 31.1 × 10^{-3} U/mL, (**B**) 9.8 × 10^{-3} U/mL, (**C**) 8.4 × 10^{-3} U/mL and (**D**) 3.9 × 10^{-3} U/mL, in all cases achieved when the UGD from *Lactobacillus johnsonni* (*Lbj*UGD) was expressed. One unit of UGD activity was defined as the amount of enzyme that produces 1 μmol of uridine diphosphate-glucuronic acid per minute..... 78

CHAPTER 4

Figure 4.1. Identified targets for gene modification to potentially improve chondroitin heterologous production in *Escherichia coli* and their role in bacterial metabolism. The target

for knock-out (KO) is marked in a red square, underexpressions (UE) are marked in orange squares, and overexpressions (OE) are marked in green squares. MltB, MurJ and NagZ are involved in cell wall biosynthesis and recycling, while GlmU produces the chondroitin precursor uridine-diphosphate (UDP)-*N*-acetylglucosamine and YbiV dephosphorylates the precursor glucose 1-phosphate. SodA is not directly related with the metabolism of chondroitin synthesis intermediates. Enzyme and compounds abbreviations: AdeD: adenine deaminase; AmiABC and AmpD: *N*-acetylmuramoyl-*L*-alanine amidases; AmpG: anhydromuropeptide permease; AnmK: anhydro-N-acetylmuramic acid kinase; AroA: 3-phosphoshikimate 1-carboxyvinyltransferase; BacA: undecaprenyl pyrophosphate phosphatase; C55: *di-trans*,*octa-cis*-undecaprenyl; Cdd: cytidine/deoxycytidine deaminase; CHSY: chondroitin synthase; Cmk: cytidylate kinase; FtsW: peptidoglycan glycosyltransferase; GalU: uridine-triphosphate(UTP):glucose-1-phosphate uridylyltransferase; Glk: glucokinase; GlmU: glucosamine-1-phosphate *N*-acetyltransferase/ UDP-*N*-acetylglucosamine diphosphorylase; GlmM: phosphoglucosamine mutase; GlmS: glucosamine-6-phosphate synthase; KatGE: catalases; LdcA: murein tetrapeptide carboxypeptidase; MltABCDFG: membrane bound lytic transglycosylases; MppA-OppBCDF: oligopeptide permeases complex with muropeptide-binding protein; MraY: phospho-*N*-acetylmuramoyl-pentapeptide-transferase; MurG: *N*-acetylglucosaminyl transferase; MurJ: lipid II flippase; MurQ: *N*-acetylmuramic acid 6-phosphate etherase; NagA: *N*-acetylglucosamine-6-phosphate deacetylase; NagK: *N*-acetyl-D-glucosamine kinase; NagZ: β-*N*-acetylglucosaminidase; Ndk: nucleoside diphosphate kinase; PBPs: penicillin-binding proteins; PDC: pyruvate dehydrogenase complex encoded by genes *pdhA*, *pdhB* and *lpd*; PflB: pyruvate formate-lyase; Pgi: glucose-6-phosphate isomerase; Pgm: phosphoglucomutase; PgpB: phosphatidylglycerophosphatase B; PykAF: pyruvate kinases; PyrG: CTP synthetase; PyrH: UMP kinase; Slt: soluble lytic transglycosylase; SodA: superoxide dismutase; TdcE: 2-ketobutyrate formate-lyase/pyruvate formate-lyase 4; UAE: UDP-*N*-acetylglucosamine 4-epimerase; Udk: uridine/cytidine kinase; UGD: UDP-glucose 6-dehydrogenase; YbjG: undecaprenyl pyrophosphate phosphatase; YbiV: sugar phosphatase. 100

Figure 4.2. Frequency of gene modifications in the 76 solutions from strain optimization using MEWpy tool. The mutant expression (in blue dots) represents the average expression value. Mutant expressions higher than 1 represent overexpressions while values of expression lower than 1 represent underexpressions. Deletions are represented in grey bars. 104

Figure 4.3. Chondroitin production in *Escherichia coli* engineered strains with chondroitin biosynthetic pathway in flask fermentation: A – extracellular chondroitin in mg/L and B – intracellular chondroitin in mg/g CDW, using K12 – *E. coli* K-12 MG1655 (DE3) and BL21 – *E. coli* BL21

(DE3), both carrying pRSFDuet_kfoCA_Zmugd. Assays were performed in triplicate. CDW – cell dry weight. 108

Figure 4.4. Chondroitin production in *E. coli* K-12 MG1655 harboring the chondroitin biosynthetic pathway (K12, genes *kfoC*, *kfoA* and *Zmugd*) with additional modifications: *sodA*, *glmU* and *mltB* overexpressions and/or *ybiV* underexpression. 111

Figure 4.5. Growth and substrate consumptions curves of *E. coli* K-12 MG1655 DE3 mutants: A) carrying pRSFDuet_kfoCA_Zmugd (control - only containing chondroitin biosynthetic pathway); B) carrying pRSFDuet_kfoCA_Zmugd and pETDuet_sodA; C) carrying pRSFDuet_kfoCA_Zmugd and pCDFDuet_mltB. Average of 2 Batch assays starting with 20 g/L glucose. 112

Figure 4.6. Chondroitin production from cultured *E. coli* K-12 MG1655 (DE3) harboring pRSFDuet_kfoCA_Zmugd (K12) and its counterparts with additional *sodA* or *mltB* overexpressions, in bioreactors operated in batch mode..... 113

Figure 4.7. Growth and substrate consumptions curves of: A: *E. coli* K-12 MG1655 (DE3) control (only containing chondroitin biosynthetic pathway - pRSFDuet_kfoCA_Zmugd); B: *E. coli* K-12 MG1655 (DE3) carrying pRSFDuet_kfoCA_Zmugd and pCDFDuet_*mltB*. Average of 2 Fed-Batch assays starting with 10 g/L glucose. The feeding was started after approximately 24 h until 36 h after inoculum. 114

Figure 4.8. Chondroitin production from cultured *E. coli* K-12 MG1655 harboring pRSFDuet_kfoCA_Zmugd (K12) and its counterpart with additional *mltB* overexpression, in bioreactors operated in fed-batch mode..... 115

CHAPTER 5

Figure 5.1. Pathways involved in chondroitin production in engineered *Saccharomyces cerevisiae* cells, and competing pathways that redirect the metabolic flux from chondroitin precursors. Compound abbreviations: CoA – coenzyme A; NAD - nicotinamide adenine dinucleotide; PPi – diphosphate; GPI - glycosylphosphatidylinositol; UDP – uridine diphosphate; UTP - uridine triphosphate; Gene descriptions: *ALG11* - Alpha-1,2-mannosyltransferase; *ALG12* - Alpha-1,6-

mannosyltransferase; *ALG9* – mannosyltransferase; *ALG5* - UDP-glucose:dolichyl-phosphate glucosyltransferase; *ATH1* - acid trehalase; *ATG26* - UDP-glucose:sterol glucosyltransferase; *CDA1*, *CDA2* – chitin deacetylase; *CHS1*, *CHS2*, *CHS3* - chitin synthases; *CHSY* - chondroitin synthase; *DPM1* - dolichol phosphate mannose synthase; *EMI2* - hexokinase; *ERI1* - endoplasmic reticulum-associated Ras Inhibitor; *FBA1* – fructose 1,6-bisphosphate aldolase; *FBP1* - fructose-1,6-bisphosphatase; *FKS1*, *FKS3* - 1,3-β-D-glucan synthase; *GAL10* – UDP-glucose-4-epimerase; *GLG1*, *GLG2* – glycogenin glucosyltransferase; *GLK1* - glucokinase; *GNA1* - glucosamine-6-phosphate *N*-acetyltransferase; *GPD1, GPD2* - glycerol-3-phosphate dehydrogenases; *GPI1*, *GPI10*, *GPI14*, *GPI15*, *GPI18*, *GPI19* – GPI anchor proteins; *GPP1*, *GPP2* - glycerol-3-phosphate phosphatases; *GSC2* - 1,3-β-glucan synthase; *GSY1*, *GSY2* - glycogen synthases; *GUT1* - glycerol kinase; *GUT2* - glycerol-3-phosphate dehydrogenase; *HXK1*, *HXK2* - hexokinases; *INO1* - inositol-3-phosphate synthase; *INM1*, *INM2* - inositol monophosphatases; *NTH1*, *NTH2* - neutral trehalases; *PBN1* - glycosylphosphatidylinositol-mannosyltransferase I; *PCM1* - *N*-acetylglucosamine-phosphate mutase; *PFK1, PFK2* - phosphofructokinase; *PGI1* - phosphoglucose isomerase; *PMT1*, *PMT2*, *PMT3*, *PMT4*, *PMT5*, *PMT6*, *PMT7* - protein *O*-mannosyltransferases; *PSA1* - guanosine diphosphate(GDP)-mannose pyrophosphorylase; *QRI1* - UDP-*N*-acetylglucosamine pyrophosphorylase; *SEC53* - Phosphomannomutase; *SMP3* - alpha 1,2-mannosyltransferase; *SPT14* - UDP-glycosyltransferase; *SUC2* – invertase; *TPS1* - trehalose-6-phosphate synthase; *TPS2* - trehalose-phosphatase; *TPS3* - trehalose-6-phosphatase; *UAE* - UDP-N-acetylglucosamine 4'-epimerase; *UGD* - UDP-glucose 6-dehydrogenase; *ZWF1* - glucose-6-phosphate dehydrogenase. 135

Figure 5.2. Frequency and expression values of genes in the solutions from optimization of *Saccharomyces cerevisiae* model for chondroitin production using MEWpy tool, limiting number of modifications to 10. The mutant expression (in blue dots) represents the average expression value. Mutant expressions higher than 1 represent overexpressions while values of expression lower than 1 represent underexpressions. Deletions are represented in grey bars. 136

Figure 5.3. Frequency and expression values of genes in the solutions from optimization of *Saccharomyces cerevisiae* model using MEWpy tool, limiting number of modifications to 3. The mutant expression (in blue dots) represents the average expression value. Mutant expressions higher than 1 represent overexpressions while values of expression lower than 1 represent underexpressions. Deletions are represented in grey bars. 139

Figure 5.4. Chondroitin production in engineered *Saccharomyces cerevisiae* CEN.PK2-1C and BY4741 strains. 142

FIGURES FROM APPENDIX

CHAPTER 3

Figure A 3.1. Multiple amino acid sequence alignment of the three uridine diphosphate (UDP)-glucose dehydrogenases used in this study generated by Clustal Omega (1.2.4) (EMBL-EBI). Species are abbreviated as the first two letters of the genus followed by the two first letters of the species: Lajo, *Lactobacillus johnsonii*; Cahi, *Capra hircus*; and Zymo, *Zymomonas mobilis*. An asterisk (*) marks the residues that are conserved in the three sequences; a colon (:) indicates conservation between groups of strongly similar properties (scoring > 0.5 in the Gonnet Point Accepted Mutation (PAM) 250 matrix); and a period (.) indicates conservation between groups of weakly similar properties (scoring =< 0.5 in the Gonnet PAM 250 matrix). Conserved amino acids within 63 members of UDP-glucose 6-dehydrogenase/ guanosine diphosphate (GDP)-mannose 6-dehydrogenase (UGD/GMD) superfamily are highlighted in grey while conserved sites within only UGD enzymes are highlighted in yellow. The nicotinamide adenine dinucleotide (NAD)-binding domain contains a characteristic dinucleotide binding Rossman fold, indicated in a blue box, composed by a conserved glycine-rich phosphate-binding loop with the pattern GXGXXG linking the first β-sheet to the following α-helix that, in the case of UGD/GMD dehydrogenase enzymatic family, has a strictly conserved tyrosine so that the region becomes GXGYXG. The central domain is a long α-helical region that serves as the dimerization interface for these enzymes. The C-terminal substrate-binding domain forms an incomplete dinucleotide binding fold, primarily responsible for binding the UDP moiety of the UDP-sugar in a deep pocket. The described active site of UGD contains residues from the three domains that are highly conserved within the superfamily and include: N-terminal threonine (T_{122}), central α-helix lysine, (K_{207}), and asparagine, (N_{211}) and C-terminal contributes with cysteine (C_{263}) and aspartic acid (D_{267}) (reference sites from *Z. mobilis* sequence). 156

Figure A 3.2. *In vitro* reduced nicotinamide adenine dinucleotide (NADH) production over time in engineered *Escherichia coli* at pH 7.2 as obtained by monitoring absorbance at 340 nm. 157

Figure A 3.3. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels from *Saccharomyces cerevisiae* cell extracts (15 µg). In the upper gel, the samples loaded were: cell extract from *S. cerevisiae* CEN.PK2-1C wild-type (C1), cell extract from *S. cerevisiae* CEN.PK2-1C expressing *Ch*UGD (C2) cell extract from *S. cerevisiae* CEN.PK2-1C expressing *Zm*UGD (C3), cell extract from *S. cerevisiae* CEN.PK2-1C expressing *Lbj*UGD (C4), *S. cerevisiae* BY4741 wild-type (B1), *S. cerevisiae* BY4741 expressing *Ch*UGD (B2), *S. cerevisiae* BY4741 expressing *Zm*UGD (B3) and *S. cerevisiae* BY4741 expressing *Lbj*UGD (B4). The lower gel shows the proteins in the corresponding insoluble phases. L: NZYColour Protein Marker II ladder. 158

Figure A 3.4. *In vitro* reduced nicotinamide adenine dinucleotide (NADH) production over time in engineered *S. cerevisiae* yeasts at pH 9.2 obtained by monitoring absorbance at 340 nm. 159

CHAPTER 4

Figure A 4.1. Agarose gel 0.7% showing polymerase chain reaction (PCR) amplification results of genes for chondroitin biosynthetic pathway construction. Ladder: NZYDNA Ladder III, NZYTech..... 159

Figure A 4.2. **Cloning and expression of identified targets for chondroitin production enhancement**: *glmU*, glucosamine-1-phosphate acetyltransferase/*N*-acetylglucosamine-1-phosphate uridyltransferase; *mltB*, membrane-bound lytic murein transglycosylase; *sodA*, superoxide dismutase. **A.** Agarose gel 0.7% showing PCR result of amplification of genes *glmU*, *sodA* and *mltB* from *E. coli* K-12 MG1655 (DE3) genome. **B.** SDS-PAGE gel showing overexpression of genes *glmU*, *sodA* and *mltB* in *E. coli* K-12 MG1655 (DE3). 1 – pETDuet-1, 2 - pETDuet_glmU, 3 – pETDuet_sodA, 4 – pCDFDuet-1, 5 – pCDFDuet_mltB. The predicted sizes were: GlmU 49.2 kDa; SodA 24.91 kDa; MltB 41.2 kDa (Slt35 36 kDa). 160

Figure A 4.3. Attempted CRISPR-Cas9 strategy for *nagZ* deletion. This methodology consists in the cleavage of double stranded DNA by Cas9 and λ-Red recombinase system-facilitated genomic integration of donor DNA. The single-guide RNA (sgRNA) encoding plasmid pKDsgRNA-nagZ to target the *nagZ* has been constructed. The plasmids for the expression of Cas9 and of the sgRNA were transformed. Then, the transformation of the oligonucleotide (donor DNA) that should induce the gene deletion was

performed. During the subsequent screenings, no positive colonies were found in the several attempts of this procedure. 161

Figure A 4.4. Attempted CRISPR interference (CRISPRi) strategy to underexpress *murJ*. In this strategy, a modified version of the caspase 9 protein commonly called as dead Cas9 (dCas9), which does not have the nuclease activity but maintains sequence-specific double stranded DNA-binding capability, is expressed to target the *murJ* gene, ultimately repressing its expression. The cloning of the protospacer in the pCRISPathBrick was performed by Golden Gate Strategy. Sequencing of resulting pCRISPath_murJ showed that constructed plasmids constantly had errors in or next to the protospacer. 162

LIST OF TABLES

CHAPTER 2

Table 2.1. Examples of natural sources of different types of chondroitin, chondroitin sulfate (CS), and dermatan sulfate (DS) and their biological functions in humans. In higher animals, different types of GAGs can occur in different proportions and sizes depending for example on the animal, tissue, age, or diet. 11

Table 2.2. Main applications of chondroitin sulfate and dermatan sulfate under clinical trials, either as the single glycosaminoglycan (GAG) of the formulation or in combination with other GAGs. 15

Table 2.3. Advantages and disadvantages of different chondroitin and chondroitin derivatives production methods 22

Table 2.4. Last decade studies on microbial production of chondroitin and chondroitin sulfate by natural producer microbial strains and engineered hosts. 28

CHAPTER 3

Table 3.1. Strains and plasmids used in this study. 60

Table 3.2. Kinetic parameters V_{max} and K_m for each enzyme expressed in this study. *Zm*UGD: Uridine diphosphate-glucose dehydrogenase (UGD) from *Zymomonas mobilis*; *Lbj*UGD: UGD from *Lactobacillus johnsonii*; *Ch*UGD: UGD from *Capra hircus*. 80

CHAPTER 4

Table 4.1. Strains and plasmids used in this study. 92

Table 4.2. Optimization of chondroitin production using OptFlux. The optimization algorithm was run at least four times for each model. The predicted phenotype for the unmodified and modified strains (from the resulting solutions with highest biomass-product coupled yield (BPCY)) are shown. BPCY

is calculated by OptFlux by multiplying biomass by product and then dividing by substrate consumed (in all cases being 10 mol/gDW/h), as predicted by pFBA simulation. Flux variability analysis (FVA) results are shown as minimum and maximum chondroitin obtained through pFBA for fixed biomass. Predicted biomass and chondroitin values are in units of mol/gDW/h. 97

Table 4.3. Optimization results obtained for iB21_1397_c model using the MEWpy tool in Python, and the corresponding relevant fluxes according to phenotype simulations using parsimonious flux balance analysis (FBA) in OptFlux. BPCY was calculated by multiplying biomass by product and then dividing by substrate consumed, as predicted by pFBA. WYIELD is the weighed sum of the minimum and maximum product fluxes. Flux variability analysis (FVA) results are shown as minimum and maximum chondroitin obtained for fixed biomass. Predicted biomass and chondroitin values are in units of mol/gDW/h. 105

CHAPTER 5

Table 5.1. Strains and plasmids used in this study. 128

Table 5.2. Intermediates and their stoichiometric values in growth equation in the model yeast-GEM 8.4.2 before and after including 1% of chitin. 132

Table 5.3. Optimization of chondroitin production in yeast-GEM_c model using OptFlux. The optimization algorithm for under and overexpression identification was run three times. The predicted phenotype for the unmodified and modified strains (from the resulting solutions with highest biomass-product coupled yield (BPCY)) are shown. BPCY is calculated by OptFlux by multiplying biomass by product and then dividing by substrate consumed (in all cases being 10 mol/gDW/h), as predicted by pFBA simulation. Flux variability analysis (FVA) results are shown as minimum and maximum chondroitin obtained through pFBA for fixed biomass. Predicted biomass and chondroitin values are in units of mol/gDW/h..... 132

Table 5.4. Optimization results obtained for *Saccharomyces cerevisiae* model using MEWpy tool, allowing a maximum of 10 modifications. Growth rate and chondroitin production rate were predicted by phenotype simulations using parsimonious flux balance analysis (pFBA) and are presented in units of mol/gDW/h. BPCY was calculated by multiplying biomass by product and then

dividing by substrate consumed, as predicted by pFBA. WYIELD is the weighed sum of the minimum and maximum product fluxes. Flux variability analysis (FVA) results are shown as minimum and maximum chondroitin obtained for fixed biomass. 138

Table 5.5. Optimization results obtained for *Saccharomyces cerevisiae* model using MEWpy tool, allowing a maximum of 3 modifications. Growth rate and chondroitin production rate were predicted by phenotype simulations using parsimonious flux balance analysis (pFBA) and are presented in units of mol/gDW/h. BPCY was calculated by multiplying biomass by product and then dividing by substrate consumed, as predicted by pFBA. WYIELD is the weighed sum of the minimum and maximum product fluxes. Flux variability analysis (FVA) results are shown as minimum and maximum chondroitin obtained for fixed biomass. 139

Table 5.6. Attempts to obtain higher chondroitin production *in silico* by combining promising gene overexpressions. Results from simulations in OptFlux using parsimonious flux balance analysis (pFBA) as simulation method for growth rate (biomass) and chondroitin production rate, presented in units of mol/gDW/h. BPCY was calculated by multiplying biomass by product and then dividing by substrate consumed, as predicted by pFBA. 141

TABLES FROM APPENDIX

CHAPTER 2

Table A 2.1. Reported applications of genome-scale metabolic modeling in identification of targets for metabolic engineering. Abbreviations of computational methods: CMA, cofactor modification analysis; cMCS, constrained minimal cut set; CiED, cipher for evolutionary design; EMA, elementary mode analysis; FBA, flux balance analysis; FDCA, flux distribution comparison analysis; FSEOF, flux scanning based on enforced objective flux; FVSEOF, flux variability scanning based on enforced objective flux; GDLS, genetic design by local search, LMOMA, linear minimization of metabolic adjustment; MFA, metabolic flux analysis; pFBA, parsimonious enzyme usage FBA; tSOT, transcriptomics-based strain optimization tool. Enzymes and genes: *acr1*, fatty acyl-CoA reductase; *adhE*, acetaldehyde/alcohol dehydrogenase; *aspC*, aspartate aminotransferase; *dgkA*, diacylglycerol kinase; *dptI*,α-ketoglutarate methyltransferase; *dptJ*, tryptophan oxygenase; FumC, fumarase; GapA, glyceraldehyde-3-phosphate dehydrogenase; GDH1, Glutamate dehydrogenase; *ldh*, lactate

dehydrogenase; *mdh*, malate dehydrogenase; PDC, pyruvate decarboxylase; PDH, pyruvate dehydrogenase multi-enzyme complex; PGK, phosphoglycerate kinase; pfl, pyruvate formate lyase; poxB, pyruvate dehydrogenase (cytochrome); Sdh3, cytochrome b subunit of the succinate dehydrogenase complex; Ser3p/Ser33p (isoenzymes), 3-phosphoglycerate dehydrogenase; tyrB, tyrosine aminotransferase; zwf2, glucose-6-phosphate hydrogenase. 164

CHAPTER 3

Table A 3.1. Primers used in this study for gene amplification, colony PCR and sequencing. Histidine Tag sequence is underlined. 171

Table A 3.2. Gene sequence of UGD from *Capra hircus* codon optimized (NZYTech) for *Saccharomyces cerevisiae.* 172

Table A 3.3. Results of BLASTp sequences search against UniProtKB/Swiss-Prot database ordered by the identity percentages obtained. Results with score <50 were removed to exclude enzymes that showed similarity with only one of the proteins domains. 172

CHAPTER 4

Table A 4.1. Primers used in this study. 177

Table A 4.2. Primers used in the attempts of *nagZ* deletion and *murJ* underexpression. 178

CHAPTER 5

Table A 5.1. Primers used in this study. 179

Table A 5.2. Codon-optimized gene sequences (5′ → 3′) for *Saccharomyces cerevisiae.* 180

SCIENTIFC OUTPUTS

Part of the work described in this thesis has been published in international peer-reviewed journals and presented in national and international scientific conferences.

Articles published in peer-reviewed journals

M.R. Couto, J.L. Rodrigues, L.R. Rodrigues, Heterologous production of chondroitin, Biotechnol. Reports. 33 (2022). https://doi.org/10.1016/j.btre.2022.e00710.

M.R. Couto, J.L. Rodrigues, L.R. Rodrigues, Cloning, expression and characterization of UDP-glucose dehydrogenases, Life. 11 (2021) 1201. https://doi.org/10.3390/life11111201.

Communications at international scientific conferences

M.R. Couto, Ó. Dias, J.L. Rodrigues, L.R. Rodrigues, Metabolic engineering of *Escherichia coli* for biotechnological chondroitin production. Abstract and Poster. The 2nd International BioDesign Research Conference. Online, December 6th – 17th, 2021.

M.R. Couto, J.L. Rodrigues, L.R. Rodrigues, Alternative enzymes for glycosaminoglycans biosynthetic pathway design. Flash presentation. Workshop Biotechnology to Build a Brighter Future: Food, Health and Environmental Applications. Online, May 10th, 2021.

M.R. Couto, Ó. Dias, J.L. Rodrigues, L.R. Rodrigues, Stoichiometric genome scale models for the chondroitin production in *Escherichia coli*. Abstract and Poster. Bioinformatics Open Days - 10th edition. Online, May 5th -7th, 2021.

M.R. Couto, J.L. Rodrigues, L.R. Rodrigues, UDP-glucuronic acid heterologous production in *Escherichia coli* and *Saccharomyces cerevisiae* chassis. Abstract, Poster and flash presentation. ASBE 2020 - 5th Applied Synthetic Biology in Europe. Delft, Netherlands (Online), November 2nd – 4th, 2020.

M.R. Couto, J.L. Rodrigues, L.R. Rodrigues, Engineering a biosynthetic pathway for high-value glycosaminoglycans production. Abstract and Poster. MIT Portugal 2018 Annual Conference - An International Partnership Towards the Challenges of Complex Systems. Lisbon, Portugal, October 1st, 2018.

Papers in preparation for submission to peer reviewed journals

M.R. Couto, J.L. Rodrigues, A. Braga, Ó. Dias, L.R. Rodrigues, Optimization of chondroitin production in *E. coli* using genome scale models.

M.R. Couto, J.L. Rodrigues, Ó. Dias, L.R. Rodrigues, *In silico* design of *Saccharomyces cerevisiae* strains for improved production of chondroitin.

CHAPTER 1

GENERAL INTRODUCTION

Abstract

Glycosaminoglycans are important molecules with several applications. In particular, chondroitin sulfate has numerous reported and potential applications, due to its structural diversity, including in human and veterinarian joint health, in pharmaceutical solutions for dry-eyes, and in the cosmetic industry. This chapter presents an overview of the thesis, showing the motivation behind the research, the research questions aimed to be answered and the main scope of this work. Moreover, the thesis outline is herein presented.

1.1. Context and Motivation

Nutraceuticals are natural bioactive compounds that promote general well-being and some medical benefits to humans [1]. The global interest and demand for nutraceuticals to prevent and treat human diseases is reflected in its multi-billion-dollar market [2]. Nevertheless, their extraction from natural sources such as plants, animals or fungi is very difficult, limiting the large-scale production of nutraceuticals [3]. An eco-friendly and safer alternative consists in engineering microbes to become cellular factories that produce nutraceuticals from simple carbon sources. *Escherichia coli* and *Saccharomyces cerevisiae* are the most widely engineered heterologous organisms to produce diverse and complex value-added chemicals such as phytochemicals, prebiotics, polysaccharides and poly amino acids [2], due to the great knowledge on their genome and existing molecular techniques for their manipulation.

Glycosaminoglycans (GAG), such as hyaluronic acid, heparosan and chondroitin are natural polysaccharides with applications in the medical, veterinary, pharmaceutical and cosmetic fields. Particularly, chondroitin sulfate (CS) is a nutraceutical with known benefits for osteoarthritis treatment [4]. However, the biotechnological production of chondroitin at industrial scale is currently limited due to low yields and safety issues related to the cultivation of some pathogenic microorganisms [5]. Several pathogenic microorganisms are reported as GAG-like polysaccharides producers, as reviewed in [3], namely: *Pasteurella multocida* type A that produces hyaluronic acid identical to the one of vertebrates; *E. coli* K5 that can synthesize a capsular polysaccharide structurally analogous to desulfated heparin; *P. multocida* type D that produces *N*-acetylheparosan; *E. coli* O5:K4:H4 that synthesizes a capsular polysaccharide with a backbone similar to that of the unsulfated chondroitin, except for the presence of fructose residues; and *P. multocida* type F that produces unsulfated chondroitin. Therefore, inspired by these endogenous microbial pathways, in the last decade, research efforts have been focused on the improvement of the production process and on the design of better heterologous hosts to produce chondroitin. Currently, the main heterologous chondroitin production pathway includes the genes from the pathogenic strain *E. coli* K4, namely the *kfoF* (UDP-glucose dehydrogenase (UGD) EC 1.1.1.22), *kfoA* (UDP-glucose 4-epimerase (UAE) EC 5.1.3.7), and *kfoC* (chondroitin polymerase/synthase (CHSY) EC 2.4.1.-). This pathway has been used in the non-pathogenic *E. coli* BL21 (DE3) [6]. Also, chondroitin production by an engineered *Bacillus subtilis* was accomplished through the incorporation of *kfoC* and *kfoA* and up-regulation of *tuaD*, an endogenous gene which encodes UGD, similarly to *kfoF* [7]. Despite these advances towards the biotechnological production of chondroitin, there is only one recent study exploring a eukaryotic microorganism as host (*Pichia Pastoris*) [8].

As an alternative host, *S. cerevisiae* is a possible eukaryotic host for the design and construction of an artificial chondroitin production pathway. This requires the introduction of three genes to encode for UGD, UAE and CHSY. UGD from *Arabidopsis thaliana* was already successfully cloned and expressed in *S. cerevisiae* with a different goal [9]. Furthermore, alternative genes can be used to optimize the chondroitin production, such as the ones from *P. multocida* or even from eukaryotic sources. The construction of the complete heterologous pathway for chondroitin production *in S. cerevisiae* is expected to be highly advantageous for the industrial implementation of this bioprocess since it is a suitable host for expression of recombinant enzymes that need post-translational modifications making it more attractive for expressing some animal enzymes [10]. Also, this yeast has a generally regarded as safe (GRAS) status, which is essential for nutraceutical applications, and its use in several industries is well implemented [2].

1.2. Scope and aims

The main goal of this thesis was to construct a chondroitin production pathway in a novel host, *S. cerevisiae.* To achieve this goal, this work was initially focused on identifying and exploring enzymes from different sources with potential improved activities towards chondroitin production, different from the ones used so far in the biotechnological production of chondroitin, based on kinetic information and literature mining. For constructing the biosynthetic pathway, *E. coli* was initially used as a proof-of-concept because it was an established microorganism for chondroitin production. The identification of metabolic engineering targets through the use of metabolic models to simulate phenotype and aid in strain optimization was also an important goal of this work. Both *E. coli* and *S. cerevisiae* models have been used for the bioinformatic approach.

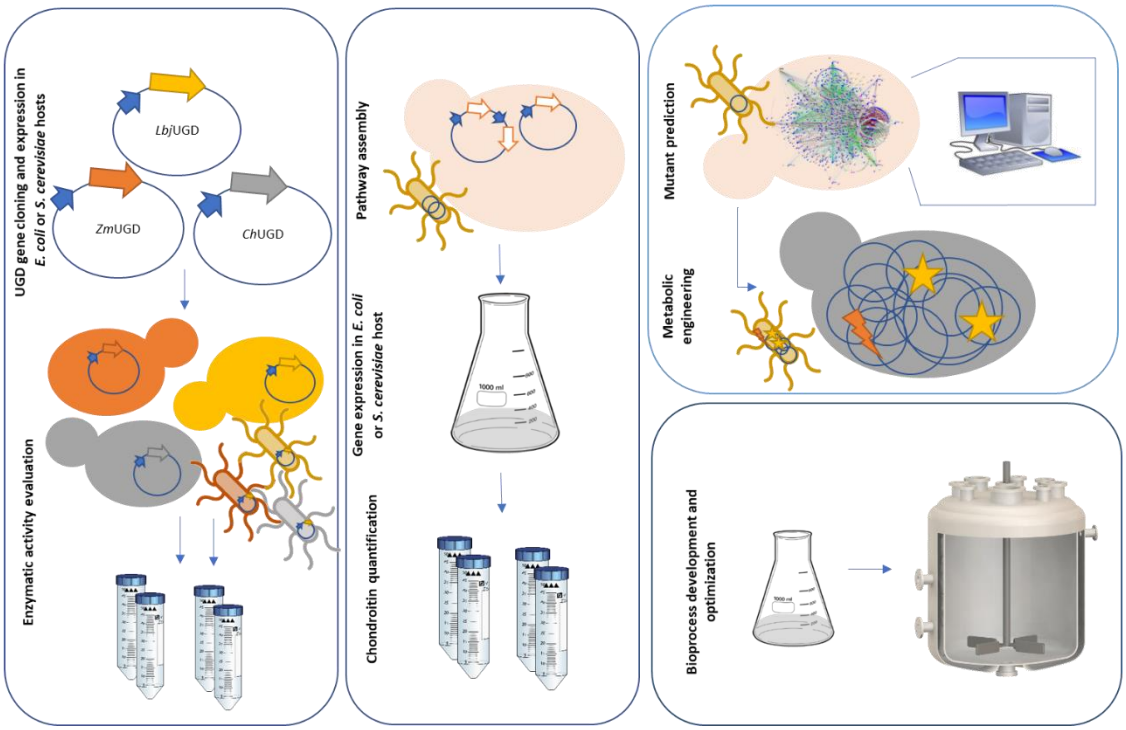
1.3. Thesis outline

This thesis describes the main achievements from the research developed at the Centre of Biological Engineering (CEB), University of Minho (Braga, Portugal) under the supervision of Professor Ligia Rodrigues and Doctor Joana Rodrigues. It is organized in 6 chapters that cover the research aims as follows:

- In the current chapter, **Chapter 1**, it is presented the general introduction, including the motivation, scope and goals of the work.

- **Chapter 2** describes the state-of-the-art on chondroitin applications, production methods, and the heterologous pathways and genetic engineering strategies used so far in the literature to produce chondroitin and related compounds.
- In **Chapter 3** *E. coli* and *S. cerevisiae* were used as hosts for the expression of genes from the first and limiting step on chondroitin production (UGD). Different sources were evaluated and the enzymes kinetics were determined.
- **Chapter 4** describes the construction of a complete heterologous pathway for chondroitin production in *E. coli*. A metabolic model representing this microorganism was used to perform gene target prediction to potential improve chondroitin production. This microorganism was initially used for chondroitin production to evaluate if the different genes could be properly expressed and to anticipate any issue in the computational and analytical techniques due to its simpler nature.
- In **Chapter 5** *S. cerevisiae* was the ultimate selected host for this work as its genome and metabolome is well known and due to its genetic engineering versatility. This model yeast was used for constructing novel chondroitin biosynthetic pathways. Also, a metabolic model of *S. cerevisiae* was used in a computational approach to identify metabolic engineering targets for further improving chondroitin production.
- In **Chapter 6** the main contributions of this thesis for the progression of research on chondroitin production (and potentially useful for other glycosaminoglycans) are summarized and some future perspectives from the current work are addressed.

Graphical abstract



1.4. References

[1] M. Ganapathy, S. Bhunia, Nutraceuticals: The new generation therapeutics, Adv. Tech. Biol. Med. 04 (2015). https://doi.org/10.4172/2379-1764.1000179.

[2] J. Wang, S. Guleria, M.A.G. Koffas, Y. Yan, Microbial production of value-added nutraceuticals, Curr. Opin. Biotechnol. 37 (2016) 97–104. https://doi.org/10.1016/j.copbio.2015.11.003.

[3] M.R. Couto, J.L. Rodrigues, L.R. Rodrigues, Heterologous production of chondroitin, Biotechnol. Reports. 33 (2022). https://doi.org/10.1016/j.btre.2022.e00710.

[4] C. Bottegoni, R.A.A. Muzzarelli, F. Giovannini, A. Busilacchi, A. Gigante, Oral chondroprotection with nutraceuticals made of chondroitin sulphate plus glucosamine sulphate in osteoarthritis, Carbohydr. Polym. 109 (2014) 126–138. https://doi.org/10.1016/j.carbpol.2014.03.033.

[5] C. Schiraldi, D. Cimini, M. De Rosa, Production of chondroitin sulfate and chondroitin, Appl. Microbiol. Biotechnol. 87 (2010) 1209–1220. https://doi.org/10.1007/s00253-010-2677-1.

[6] W. He, L. Fu, G. Li, J. Andrew Jones, R.J. Linhardt, M. Koffas, J.A. Jones, R.J. Linhardt, M. Koffas, J. Andrew Jones, R.J. Linhardt, M. Koffas, Production of chondroitin in metabolically engineered *E. coli*, Metab. Eng. 27 (2015) 92–100. https://doi.org/10.1016/j.ymben.2014.11.003.

[7] P. Jin, L. Zhang, P. Yuan, Z. Kang, G. Du, J. Chen, Efficient biosynthesis of polysaccharides chondroitin and heparosan by metabolically engineered *Bacillus subtilis*, Carbohydr. Polym. 140

(2016) 424–432. https://doi.org/10.1016/j.carbpol.2015.12.065.

[8] X. Jin, W. Zhang, Y. Wang, J. Sheng, R. Xu, J. Li, G. Du, Z. Kang, Biosynthesis of non-animal chondroitin sulfate from methanol using genetically engineered *Pichia pastoris*, Green Chem. 23 (2021) 4365–4374. https://doi.org/10.1039/d1gc00260k.

[9] T. Oka, Y. Jigami, Reconstruction of *de novo* pathway for synthesis of UDP-glucuronic acid and UDP-xylose from intrinsic UDP-glucose in *Saccharomyces cerevisiae*, FEBS J. 273 (2006) 2645–2657. https://doi.org/10.1111/j.1742-4658.2006.05281.x.

[10] J. Nielsen, Yeast systems biology: Model organism and cell factory, Biotechnol. J. 14 (2019) 1–9. https://doi.org/10.1002/biot.201800421

CHAPTER 2

STATE OF THE ART

Abstract

Chondroitin sulfate (CS) is a glycosaminoglycan with a broad range of applications being a popular dietary supplement for osteoarthritis. Usually, CS is extracted from animal sources. However, the known risks of animal products use have been driving the search for alternative methods and sources to obtain this compound. Several pathogenic bacteria naturally produce chondroitin-like polysaccharides through well-known pathways and, therefore, have been the basis for numerous studies that aim to produce chondroitin using non-pathogenic hosts. However, the yields obtained are not enough to meet the high demand for this glycosaminoglycan. Metabolic engineering strategies have been used to construct improved heterologous hosts. The identification of metabolic bottlenecks and regulation points, and the screening for efficient enzymes are key points for constructing microbial cell factories with improved chondroitin yields to achieve industrial CS production. The recent advances on enzymatic and microbial strategies to produce non-animal chondroitin are herein reviewed. Challenges and prospects for future research are also discussed.

Keywords: chondroitin, glycosaminoglycans, biosynthetic pathway, metabolic engineering, microbial fermentation, heterologous production

This chapter was partially published on the following publication:

M.R. Couto, J.L. Rodrigues, L.R. Rodrigues, Heterologous production of chondroitin, Biotechnol. Reports. 33 (2022). https://doi.org/10.1016/j.btre.2022.e00710

2.1. Introduction

Glycosaminoglycans (GAGs) are naturally occurring heteropolysaccharides with polyanionic character. They consist of repeating disaccharide units composed of an *N*-acetylated or *N*-sulfated hexosamine (glucosamine, GlcN or galactosamine, GalN) and either a uronic acid (glucuronic acid, GlcA or iduronic acid, IdoA) or galactose (Gal). There are four main types of GAGs that vary in their monomeric composition, glycosidic bonds and sulfation pattern, hyaluronic acid (hyaluronate, hyaluronan, HA); keratan sulfate (KS); chondroitin sulfate/dermatan sulfate (CS/DS, stereoisomers); and heparin/heparan sulfate (HP/HS) [1]. The chemical structures of the repeating disaccharide units of different GAGs are represented in Figure 2.1. While HA lacks sulfate groups, the remaining GAGs can contain sulfates at various positions. DS is distinguished from CS by the presence of IdoA. HS is the only glycosaminoglycan that contains a *N*-sulfated hexosamine and differs from HP by possessing less sulfated units and less content of epimerized entities into IdoA. Heparosan is the non-sulfated precursor to HS and HP. Keratan sulfate lacks uronic acids and instead contains sulfated Gal residues.

In animals, GAGs usually exist as long chains covalently bound to a protein core, as part of proteoglycans (PGs) [2]. Depending on their molecular structure, GAGs have a wide distribution through tissues and perform different physiologic functions including in structural support, cell recognition and signaling, matrix organization, inflammation, cell division, and tissue repair and development [3,4], being particularly common in connective tissues such as the skin, bone, cartilage, and the intervertebral discs [5]. Some bacterial pathogens produce GAG-like polysaccharides as part of their capsule composition, which contribute to their pathogenicity [6–8].

The wide applications of GAGs in medicine, veterinary, cosmetics, and pharmaceutics have been extensively reviewed [5,9,10]. The most famous GAG clinical application is the use of HP as an anticoagulant drug. Furthermore, GAGs also exist as nutraceuticals in human and veterinary supplements, serve as cosmetic or pharmaceutical ingredients, or act as materials for several biomedical applications such as tissue engineering.

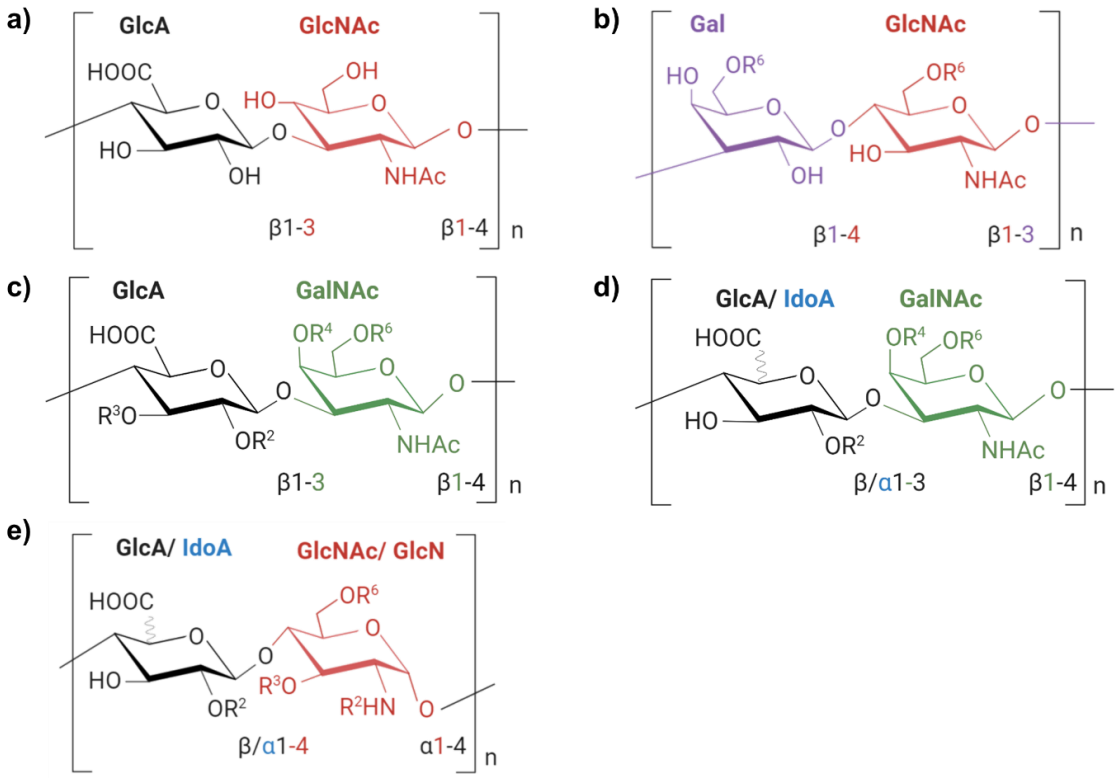


Figure 2.1. Structures of the main glycosaminoglycans (GAGs) a) hyaluronic acid, b) keratan sulfate, c) chondroitin and chondroitin sulfate, d) dermatan sulfate, and e) heparosan, heparan sulfate and heparin. Monomers of the disaccharide building blocks are abbreviated as GlcA - D-glucuronic acid, GlcNAc – *N*-acetyl-D-glucosamine, Gal – D-galactose, GalNAc – *N*-acetyl-D-galactosamine, IdoA – L-iduronic acid, GlcN, D-glucosamine. Hyaluronic acid (a) does not go under post-polymerization modifications. Keratan sulfate (b) has di-sulfated, mono-sulfated and non-sulfated disaccharide units (each R^6=H or SO_3H) due to *O*-sulfotransferases action. Chondroitin (c) is the simple non-sulfated backbone (R^2, R^3, R^4 and R^6=H) which can be modified by different tissue-specific *O*-sulfotransferases to form chondroitin sulfate (each R^2, R^3, R^4 and R^6=H or SO_3H). Dermatan sulfate (d) is formed from chondroitin through epimerization of GlcA into IdoA by tissue-specific epimerases followed by *O*-sulfotransferases (each R^2, R^4 and R^6=H or SO_3H). Heparosan (e) has non-modified sugar moieties, that can be further modified through actions of tissue-specific *N*-sulfotransferases, C5 epimerases and *O*-sulfotransferases to generate the sulfated forms heparan sulfate and heparin (R^2 from uronic acid=H or SO_3H; when the hexosamine unit is GlcN, R^2 in that unit=SO_3H, while R^2=Ac when the unit is GlcNAc; other groups R^3, R^6=H or SO_3H). Heparin has more sulfate groups and IdoA content than heparan sulfate. Depending on the GAG type and source the molecular size can generally vary between 4 and 200 mer (n = 4-200). Exceptionally, the highest size can be found for hyaluronic acid that can achieve 20,000 repeating units.

The industrial production of GAGs can be classified in 4 types: (i) extraction from animal sources; (ii) chemical synthesis; (iii) enzymatic or chemoenzymatic production or (iv) microbial or animal cells

fermentation. Regarding microbial fermentation, the industrial production of HA is currently the most well established due to its simple structure that does not undergo sulfation or epimerization contrarily to the other GAGs. However, sulfated GAGs such as CS, HS, HP and KS are usually chemically extracted from animal sources, through a laborious process that requires a large amount of environmentally hazardous compounds [11,12]. Also, the use of animal sources brings concerns on contamination with other animal products, viruses, and prions. In addition, religious motivations and vegetarianism trends have also led to the search for non-animal sources [13].

This review is focused on the recent advances of production processes for animal-free chondroitin and CS production, including enzymatic and microbial strategies, which are areas with great advances in the last few years. With the increasing knowledge on biosynthesis, regulation, transportation, bottlenecks, metabolic engineering strategies, bioprocess optimization, and with the discovery of new enzymes, a sustainable process for the microbial production of value-added natural and artificial chondroitin derivatives is closer to becoming a reality.

2.2. Occurrence and biological functions of chondroitin and its derivatives

Chondroitin is the unsulfated precursor of CS, the most abundant GAG in the human body. It comprises a repeated disaccharide structurally composed by a glucuronic acid (GlcA) residue and a *N*-acetyl-D-galactosamine (GalNAc) residue linked by β(1-3) and β(1-4) bonds, respectively (Figure 2.1, Table 2.1). According to its sulfation pattern, CS is mainly classified into CS-A, CS-C, CS-D, CS-E and CS-O (unsulfated chondroitin) [2,14]. However, novel CS disaccharide units have been discovered in natural sources (CS-F, CS-G, CS-K, CS-L, CS-M, CS-S and CS-T) [15–17] or chemically synthesized (CS-R, CS-U and CS-V) [18]. Considered by some authors as a type of CS, DS (CS-B) is distinguished by the presence of at least one IdoA residue, resulting from the epimerization of GlcA. In higher animals, CS exists anchored to proteins as part of proteoglycans, being CS-A and CS-C the predominant types, and is present in connective tissues, for example, cartilage, cornea, bone, skin, arterial walls [19]. Examples of natural sources of CS and DS are presented in Table 2.1.

Table 2.1. Examples of natural sources of different types of chondroitin, chondroitin sulfate (CS), and dermatan sulfate (DS) and their biological functions in humans. In higher animals, different types of GAGs can occur in different proportions and sizes depending for example on the animal, tissue, age, or diet.

GAG type	Disaccharide repeat	Natural sources	Biological functions	References
CS-A	GlcA(β1-3)GalNAc(4S)(β1-4)	Dogfish, shark, and whale cartilage; human, bovine, porcine and chicken cartilaginous tissues	mediates malaria-infected erythrocytes adhesion; negatively regulates axonal guidance and growth; activates metastatic cascate	[13,22,38–40]
DS (CS-B)	GlcA/IdoA(2S)(β/α1-3)GalNAc(4S)(β1-4)	Animal skin/hide, cornea, cartilage, heart valve, tendons, blood vessels, and bone	regulates growth factors activity; has anticoagulant activity; promotes proliferation of serveral cell lines; mediates homeostasis, tumorigenesis, infection, wound repair, collagen organization, fibrosis and stabilization of the basement membrane	[25–32]
CS-C	GlcA(β1-3)GalNAc(6S)(β1-4)	Dogfish and shark cartilage; human, bovine, porcine and chicken cartilaginous tissues	may promote progression of epilepsy; neuroprotective properties in Alzheimer's disease	[13,38,41–43]
CS-D	GlcA(2S)(β1-3)GalNAc(6S)(β1-4)	Shark cartilage; animal brain	promotes neuron growth; interacts with humoral factors	[22,44–46]
CS-E	GlcA(β1-3)GalNAc(4,6diS)(β1-4)	Squid cartilage; animal lung	mediates angiogenesis; acts as cell surface receptor for herpes virus; modulates humoral factors; stimulates neurite outgrowth; promotes neural stem cells proliferation; mediates osteogenic differentiation	[22,24,32,46–48]
Fucosylated CS (fCS, CS-F)	different types of CS with L-fucosyl branches usually attached to the O-3 of GlcA unit	Sea cucumbers	-	[49–54]

Table 2.1. Examples of natural sources of different types of chondroitin, chondroitin sulfate (CS), and dermatan sulfate (DS) and their biological functions in humans. In higher animals, different types of GAGs can occur in different proportions and sizes depending for example on the animal, tissue, age, or diet (*continuation*).

GAG type	Disaccharide repeat	Natural sources	Biological functions	References
CS-G	different types of CS with glucose attached to the O-6 of GalNAc unit	Squid cartilage	-	[55]
CS-H (highly sulfated DS)	IdoA(α1-3)GalNAc(4S,6S)(β1-4)	Hagfish notochord	-	[56]
CS-K	GlcA(3S)(β1−3)GalNAc(4S)(β1-4)	Squid, king crab and octopus	-	[57–60]
CS-L	GlcA(3S)(β1−3)GalNAc(6S)(β1-4)	Squid	-	[60]
CS-M	GlcA(3S)(β1−3)GalNAc(4,6diS)(β1-4)	Squid	-	[60]
Unsulfated chondrotin (CS-O)	GlcA(β1-3)GalNAc(β1-4)	*Pasteurella multocida* Type F; *Avibacterium paragallinarium* genotype I	-	[35,36]
Fructosylated chondroitin	GlcA(3Fru)(β1-3) GalNAc(β1-4)	*Escherichia coli* K4	-	[37,61–65]
***N*-glycolyl chondroitin (Gc-CS)**	GlcA(β1-3)GalNGc(β1-4)	Serum of humans who eat red meat (beef, lamb, and pork)	-	[66]

CS has biological roles in inflammation prevention, immune modulation, maintenance of the structure, elasticity and shock-absorbing properties of cartilage, regulation of cell adhesion to the extracellular matrix, facilitation of nutrient and oxygen diffusion, mediation of tumor growth and metastasis, pathogen adhesion, angiogenesis, osteogenic differentiation and in brain development, plasticity, and regeneration [20–24]. Interestingly, CS biological functions depend on the chain length, sulfation pattern and percentage of the structural units, which vary upon animal age and tissue. DS is predominant in skin, heart valve, tendons, and blood vessels [25–27]. The physiologic roles of DS include regulation of transforming growth factor-β activity, cell proliferation, cell development, cell adhesion, homeostasis, collagen organization, anticoagulant activity, tumorigenesis, infection, wound repair, fibrosis, and stabilization of the basement membrane [25,28–32]. Particularly, the well-studied anticoagulant activity of DS occurs by its binding to HP cofactor II. The resulting complex can inhibit the procoagulant effect of thrombin, while not affecting the clotting cascade (factor X) or the platelet function. DS selective inhibition of thrombin makes it an interesting alternative to HP [33,34].

In bacteria, unsulfated chondroitin backbone is produced as capsular polysaccharide by pathogens *Pasteurella multocida* Type F and *Avibacterium paragallinarium* genotype I [35,36], while a chondroitin decorated with fructose residues is produced by *Escherichia coli* K4 [37]. Bacteria use these capsular polysaccharides to mask their infection and thus increase their pathogenicity [6].

Commercial CS is usually provided as a mixture of CS-A and CS-C, the most common CS in animal cartilage [67], which is obtained through chemical extraction and purification from animal cartilaginous tissues, mainly from bovine, porcine, shark and chicken. The structural composition, molecular weight and yields of CS obtained from extraction vary not only with the extraction method but also with the animal and tissue used. Some of the reported CS yields include extraction from shark fin resulted in 150.5 mg/g dry cartilage [68]; from blue shark head wastes 120.8 mg/g dry cartilage was obtained; crocodile hyoid, rib, sternum, trachea cartilages generated 91 to 274 mg/g of dry cartilage [68]; 75 mg/g of dry weight has been achieved by extraction from ray [68]; buffalo nasal, joint and tracheal cartilages generated 60 to 62 mg/g of dry cartilage [69]; using chicken keel, 24.8 mg/g of wet cartilage were obtained [70]; from different sea cucumbers, fucosylated CS (fCS) yields varied from 63 to 110 mg/g weight [54]; and from fish by-products, yields from 19 to 137 mg/g dry cartilage have been achieved [71]. In addition to the relatively low yields, the shortage of such materials and the concerns on intraspecies contamination led to studies on artificial synthesis of the chondroitin backbone, either using chemical synthesis, enzymatic, chemo-enzymatic, and microbial fermentation strategies.

2.3. Clinical applications of CS

Given the growing potential applications of CS, and the rising prevalence of osteoarthritis due to the increase in average life expectancy among the population, a boost in its demand and market volume has been registered. The global CS market size was valued as USD 1.17 billion in 2020 and is expected to grow at a compound annual growth rate (CAGR) of 3.0% from 2020 to 2028 [72]. In 2018, China was the biggest producer accounting for 83.7% of CS produced globally [73].

Depending on their application, different systems have been implemented for the administration of chondroitin and its derivatives, including oral administration as food additives, in pharmaceutical preparations, nutraceuticals or veterinary supplements [74–77]; incorporation in intra-articular or intravenous injectables [78]; integration in skin dermatology/cosmeceutical and ophthalmic products [79–81]; as part of medical devices [82–85]; or as biomaterials for regenerative medicine [9,86].

The current main application of CS is as nutraceutical to treat osteoarthritis symptoms and retard cartilage degradation [77,87]. These supplements have been widely prescribed for humans and animals and are usually formulated with glucosamine. Although the results on the efficacy of chondroitin or its combination with glucosamine in joint repair and pain relief compared with placebo have not always been concurrent [88–91], CS and glucosamine have been recommended in guidelines, prescribed by general practitioners and rheumatologists as over-the-counter drugs, and suggested by the European League Against Rheumatism (EULAR) to patients with knee and hand osteoarthritis [92]. The variable efficacy of CS is in part attributed to its highly variable composition. Since they are over-the-counter supplements, they are not tightly controlled and can differ in the biological source, purification method and ingredient amount and type, thus resulting in heterogenous CS mixtures with different sulfation patterns or even containing other GAGs as contaminants. Condrosulf (fish CS, IBSA Institut Biochimique SA, Lugano, Switzerland), Condrosan (bovine CS, Bioiberica S.A.U, Barcelona, Spain), and Structum (avian CS, Laboratoires Pierre Fabre, Paris, France) are examples of commercialized pharmaceutical-grade CS products that have shown benefits for osteoarthritis in clinical trials [93–100]. In these studies, CS has shown to be safe and, generally, exhibited slight to moderate efficacy over placebo in pain relief and function improvement in osteoarthritis patients. Also, CS is included in eye drop solutions for dry eyes [79,80]. Preservative-free ophthalmic solutions combining xanthan gum and CS or combining sodium HA and CS showed similar efficacy compared to polyethylene glycol and propylene glycol solutions in the treatment of dry eye disease. The CS oral supplementation in combination with glucosamine hydrochloride has also shown improved therapeutic benefits in reducing systemic inflammation in overweighted individuals, over placebo [101]. For interstitial cystitis, the intravesical treatment with CS achieved improved efficiency in

pain reduction and nocturia and had superior tolerability over DMSO, a standard approved therapy [84]. CS is also found in solutions for preserving corneas [102], or in injectable devices to protect the eye during cataract surgery (phacoemulsification) [82,83]. The ophthalmic viscosurgical devices containing HA and CS (DisCoVisc, Viscoat and Duovisc) showed greater efficiency during phacoemulsification, and improved protection of corneal endothelium, compared to hydroxypropylmethylcellulose [82] and to HA and lidocaine hydrochloride (Visthesia) injectables [83].

In addition, the potential applications of exogenous CS are wide-ranging since CS exhibits anti-inflammatory activity [50,103–106], anticoagulant properties [53,54,107], promotes the regeneration of different tissues [108–110], has antiviral activity [111–116], and can be used in cancer treatment [47,105]. Specifically, CS-E has been considered a powerful antiviral agent against flavivirus dengue, herpes viruses and T-cell leukemia virus type I [111,112,114], tobacco mosaic virus is inhibited by CS-A and CS-C [113], fCS showed inhibitory activity against human immunodeficiency virus type-1 [115] and against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [116]. CS has also potential applications in engineering scaffolds for regenerative medicine and tissue engineering, by combination with other biopolymers such as HA and collagen [117,118].

DS exhibits several reported therapeutic applications. Drug formulations of DS have been marketed for the prevention of venous thromboembolism under the commercial names Mistral (Mediolanum Farmaceutici, Milan, Italy) [119] and Venorix (Laboratório Medinfar, Amadora, Portugal). It is also a component of sulodexide, a more widely used antithrombotic agent [120].

Table 2.2 lists clinical trials with recently published results where CS or DS are used as treatment for different conditions, either as single GAG in the formulation or when formulated with other GAGs.

Table 2.2. Main applications of chondroitin sulfate and dermatan sulfate under clinical trials, either as the single glycosaminoglycan (GAG) of the formulation or in combination with other GAGs.

Glycosaminoglycan	Condition / Potential application	Clinical trials ID
Chondroitin sulfate	Osteoarthritis (knee, hand)	NCT00291499 [93] NCT00604539 [95] NCT01354145 [96,97] NCT00955552 [121] NCT01425853 [122] NCT01893905 [123] NCT02830919 [124] NCT01271218 [125] NCT00513422 [126,127]

Table 2.2. Main applications of chondroitin sulfate and dermatan sulfate under clinical trials, either as the single glycosaminoglycan (GAG) of the formulation or in combination with other GAGs (*continuation*).

Glycosaminoglycan	Condition / Potential application	Clinical trials ID
Chondroitin sulfate (*continuation*)	Inflammation reduction and prevention	NCT01682694 [101]
	Interstitial cystitis	NCT04268810 [84]
	Dry eye	NCT01657253 [79]
	Hemostasis in surgeries	NCT03725098 [128]
	Corneal storage medium	NCT01657500 [102]
Chondroitin sulfate and **hyaluronic acid**	Knee osteoarthritis	NCT04352322 [129]
	Recurrent urinary tract infections	NCT02016118 [85]
	Adjuvant during phacoemulsification	NCT01387620 [82] NCT02304861 [83]
	Post-surgical rehabilitation	NCT03355651 [130]
Dermatan sulfate and **heparan sulfate** (mesoglycan)	Post-operative thrombosis	NCT04481698 [131]
Dermatan sulfate and **heparin** (sulodexide)	Anticoagulant	NCT04257487 [132]
	Diabetic retinopathy	NCT01295775 [133]
	Diabetic kidney disease (albuminuria, nephropathy)	NCT00130312 [134] NCT00130208 [135]

2.4. Enzymatic and chemoenzymatic CS production

Enzymes can be used to produce CS oligosaccharides by degrading CS polymers or by synthesizing CS oligosaccharides/polymers through polymerization. The main advantages of using enzymes over chemical approaches is that enzymes catalyze stereoselective and regioselective reactions, potentially resulting in homogeneous CS in an eco-friendlier and faster manner [136,137]. However, these enzymatic steps are often combined with chemical synthesis or modifications (chemoenzymatic strategies) to either provide synthetic precursors or to perform the sulfation step. The depolymerizing/ degrading enzymes and the polysaccharide substrates used are relatively inexpensive [136]. However, the commercial polysaccharides are usually from animal sources which raises the concerns already mentioned. The degrading enzymes will be briefly discussed since they can give insights on structural analysis of CS, and because they can perform the contrary polymerization reaction under certain conditions. Enzymatic synthesis of CS from natural and unnatural precursors is the alternative for the preparation of homogeneous CS that most resembles the biological process. However, the high cost of such precursors generally limits the industrial implementation of these enzymatic methods.

2.4.1. Enzymatic depolymerization of CS polysaccharides

CS depolymerizing enzymes able to produce CS oligosaccharides are CS lyases [138–140] and CS hydrolases [141]. While polysaccharide hydrolases are found in almost every organism, polysaccharide lyases do not occur in vertebrates. According to their substrate specificity, the polysaccharide lyases acting on CS can be classified into chondroitinases ABC, AC, B, and C. They perform the depolymerization of CS/DS polymers via endolytic or exolytic β-elimination resulting mainly in unsaturated oligosaccharides [138,140]. *Proteus vulgaris, Flavobacterium heparinum* (*Pedobacter heparinus*), *Sphingomonas paucimobilis, Bacteroides stercoris, Pseudopedobacter saltans* (*Pedobacter saltans*), and *Bacteroides thetaiotaomicron* are some examples of different CS lyase producers [138,140,142,143]. Differently from CS lyases, CS hydrolases usually depolymerize CS with no unsaturated bonds formed, which make them more attractive for the enzymatic production of CS [1]. Some animal HA hydrolases, hyaluronidases, show depolymerizing activity against CS and some isoforms can only accept CS as substrate [144,145]. Therefore, these hyaluronidases have been used to depolymerize CS polysaccharides to obtain CS oligosaccharides (Figure 2.2a) up to tetrasaccharides. The substrate specificity of CS lyases and CS hydrolases can retrieve functional characterization of GAGs.

2.4.2. Enzymatic synthesis of CS oligosaccharides and polysaccharides

For the synthesis of CS oligosaccharides and polysaccharides, degrading enzymes that in the reverse direction synthesize polymers (such as hyaluronidases, Figure 2.2b), or synthases (such as glycosyltransferases, Figure 2.2c) can be used. Further post-polymerization modifications of the chondroitin backbone are required to achieve the CS oligosaccharides or polysaccharides. Specialized CS sulfotransferases, including 4-*O*-sulfotransferase (C4OST), 6-*O*-sulfotransferase (C6OST), 2-*O*-sulfotransferase (2OST) and *N*-acetylgalactosamine 4-sulfate 6-*O*-sulfotransferase (GalNAc4S-6OST), sulfate the chondroitin composing units in the presence of the co-factor 3′-phosphoadenosine-5′-phosphosulfate (PAPS) to generate CS with different sulfation patterns and consequently different biological activities.

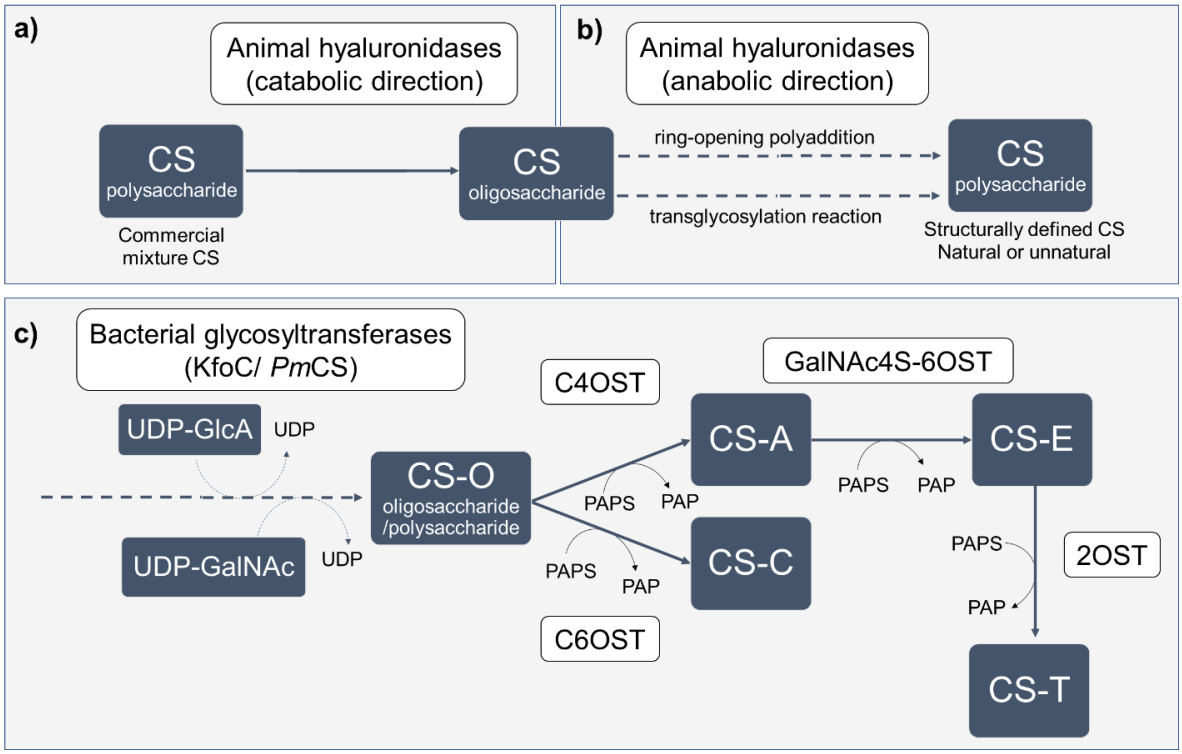


Figure 2.2. Enzymatic synthesis of chondroitin sulfate (CS). a) Depolymerizing enzymes, such as animal hyaluronidases, can be used to obtain CS oligosaccharides from CS polysaccharides; b) the same type of enzymes is able to, under different conditions, polymerize the CS oligosaccharides through chemoenzymatic approaches. c) Bacterial glycosyltransferases (such as chondroitin synthase from *Escherichia coli* K4, KfoC, or from *Pasteurella multocida* type F, *Pm*CS) act by transferring alternate residues of glucuronic acid (GlcA) and acetylgalactosamine (GalNAc), using uridine diphosphate (UDP)-GlcA and UDP-GalNAc as donors, to the nonreducing end of a chondroitin chain acceptor to elongate the chondroitin oligosaccharide/ polysaccharide backbones. Sulfotransferases such as 4-*O*-sulfotransferase (C4OST), 6-*O*-sulfotransferase (C6OST), *N*-acetylgalactosamine 4-sulfate 6-*O*-sulfotransferase (GalNAc4S-6OST), and 2-*O*-sulfotransferase (2OST) that require the presence of 3′-phosphoadenosine-5′-phosphosulfate (PAPS) as sulfate donor, convert the unsulfated backbone (CS-O) in CS with different sulfation patterns such as CS-A, CS-C, CS-E and CS-T. Only CSs with a homogenous defined sulfation pattern are shown although a CS chain may have different CS units if a combination of sulfotransferases is used. Dashed arrows represent polymerization steps.

Chemoenzymatic methods using HA or CS degrading enzymes in the anabolic direction to produce CS and its derivatives have been established, by either ring-opening polyaddition [146,147] or by transglycosylation reaction [148]. Kobayashi *et al.* [146] and Fujikawa *et al.* [147] used testicular hyaluronidases from ovine or bovine to catalyze the synthesis of structurally defined CS (natural and unnatural) from chemically synthesized disaccharide oxazoline precursors which acted as transition state analogues. Kakizaki *et al.* [148] used immobilized bovine testicular hyaluronidase to produce hybrid

polysaccharides composed by both HA and CS units. First, the hyaluronidase was used to prepare oligosaccharides by digestion of commercial CS and HA polysaccharides. The oligosaccharides were fluorolabeled (to facilitate analytical identification) generating pyridylaminated acceptors which underwent transglycosylation catalyzed by immobilized hyaluronidase using HA and CS donor polysaccharides. These methods enabled to control the molecular weight of the resulting CS and the construction of unnatural CS derivatives. However, the use of expensive sugar precursors and dedicated reactors, as well as the low product yields are drawbacks that limited large-scale application [149].

The discovery of microbial counterparts of glycosyltransferases in pathogenic bacteria further boosted the development of enzymatic strategies to produce chondroitin and CS. Bacterial glycosyltransferases used for chondroitin synthesis, also named chondroitin synthases or chondroitin polymerases, include *Pm*CS from *Pasteurella multocida* and KfoC from *E. coli* K4 strains. These enzymes act by transferring alternate residues of GlcA and GalNAc, using uridine diphosphate (UDP)-GlcA and UDP-GalNAc as donors, to the nonreducing end of a chondroitin chain acceptor to form chondroitin oligosaccharide backbones (Figure 2.2c). DeAngelis *et al.* [150] identified for the first time a microbial gene encoding the chondroitin synthase *Pm*CS and cloned it in *E. coli.* The same authors further used this recombinant glycosyltransferase for *in vitro* production of chondroitin from the substrates UDP-GlcA and UDP-GalNAc [151]. Other microbial glycosyltransferase KfoC was used by Sugiura *et al.* [152] to produce chondroitin oligosaccharides (CS-O) with controlled size from 7-mer to 16-mer. The authors first digested a commercial unsulfated chondroitin polymer with hyaluronidase to achieve even-numbered oligosaccharides. Further digestion with β-glucuronidase allowed to obtain odd-numbered oligosaccharides. The oligosaccharides were subjected to pyridylamination to facilitate their analysis. Then, two mutated recombinant KfoC enzymes were expressed each being selective for one of the substrates (UDP-GlcA or UDP-GalNAc) and the engineered enzymes, immobilized in beads, were used to elongate the chondroitin chain. The same group used chemoenzymatic strategies for the preparation of CS oligosaccharides and polysaccharides with different sulfation patterns using the CS polymerase KfoC and recombinant sulfotransferases C4OST, C6OST, GalNAc4S-6OST and 2OST [153–155]. The chondroitin backbone was obtained by desulfating commercial CS from different sources and HEK293T cells were used to express the human sulfotransferases. However, the resulting products were heterogeneous, low amounts were achieved, and the process required laborious techniques for the separation of oligosaccharides with different sizes. Afterwards, the enzymatic production of 15 structurally diversified and homogeneous CS-A and CS-C oligosaccharides was achieved using KfoC expressed in *E. coli* BL21 [156]. Amounts of 4 to 30 mg of CS oligosaccharides were obtained. The authors also

synthesized the expensive substrate UDP-GalNAc using *Bifidobacterium longum N*-acetylhexosamine kinase and human GalNAc pyrophosphorylase both expressed in *E. coli* BL21, which significantly reduced the process cost. Human C4OST and mouse C6OST were afterwards used for sulfation of chondroitin backbones. Sf9 insect cells were used to express C4OST and C6OST. The authors were unsuccessful expressing C6OST in *E. coli* which they rationalized to be due to the glycosylation being required for the sulfotransferase activity as shown by Yusa *et. al.* [157]. Accordingly, attempts to express soluble mouse C4OST in *E. coli* were also unsuccessful, and C4OST and C6OST were effectively expressed and secreted by the *Pichia pastoris* [158]. However, active human C4OST has been expressed in *E. coli* BL21 and *P. pastoris* [159]. Those authors showed that the glycosylated and non-glycosylated forms of the enzyme exhibited similar activities. More recently, the soluble expression of human C4OST enzyme was also achieved in *E. coli* K4 and in *E. coli* K-12 MG1655 [160].

In addition to the most common CS types, CS-A and CS-C, efforts to enzymatically produce alternative CS structures have resulted in the synthesis of CS-E. Despite the enzymatic synthesis of structurally heterogeneous CS-E polysaccharides has been primarily reported [161], the enzymatic production of homogenous CS-E oligosaccharides was more recently achieved [106]. The process included a serial elongation of *p*-nitrophenyl glucuronide (GlcA-*p*NP) with UDP-GalNAc using KfoC followed by the sulfotransferase modifications on the unsulfated backbone by C4OST and GalNAc4S-6OST. Mouse GalNAc4S-6OST was expressed in Sf9 cells, the glycotransferase KfoC and C4OST were expressed as reported in Li *et al.* [156] and the structural units UDP-GlcA, UDP-GalNAc, and the co-factor PAPS were synthesized by enzymatic approaches.

Another recently reported chemoenzymatic strategy for the synthesis of homogenous chondroitin polymers combined stepwise oligosaccharides synthesis from GalNAc and GlcA units with further enzymatic polymerization by *Pm*CS [162]. The authors showed that *Pm*CS needs a suitable oligosaccharide (at least a chondroitin trisaccharide) as the acceptor to trigger the *in vitro* chondroitin chain polymerization. The polymers were further chemically sulfated to form CS.

Even when the CS backbone is not obtained through enzymatic polymerization, enzymatic *in vitro* reactions can be used to participate in the post-polymerization modification of the chondroitin backbone. Rat aryl sulfotransferase IV (ASST IV) expressed in *E. coli* and mouse C4OST and C6OST sulfotransferases expressed in *P. pastoris* were used to modify a chondroitin backbone produced *in vivo* [158]. Instead of using sulfotransferases to add sulfate groups to the chondroitin, an alternative way to produce structurally defined CS was performed by employing a chemoenzymatic strategy for regioselective desulfation using a recombinant 4-*O*-endo-sulfatase from *B. thetaiotaomicron* [163].

Regarding the sulfate donor for sulfotransferases, PAPS, a chemoenzymatic synthesis [164], enzymatic [165] or *in vivo* production using engineered *E. coli* [166] are recent attractable strategies to supply this co-factor for the synthesis of CS. These strategies have been applied for CS-A synthesis *in vitro* [166] and *in vivo* [160].

The main drawbacks of the enzymatic CS synthesis include difficulty in controlling glycosyltransferase and sulfotransferase activities for a homogenous compound with a defined sulfation pattern [136]. The chemoenzymatic CS synthesis can provide the flexibility of chemical synthesis along with the efficiency and selectivity offered by enzymes to produce more defined CS oligosaccharides [167]. Nevertheless, the poor availability of substrates and co-factors, the poor enzymes activity and consequent low CS yields are still major limitations for the implementation of enzymatic and chemoenzymatic methods at industrial scale [168].

2.5. Microbial production of GAGs

Biotechnological processes using microorganisms have been developed to produce GAGs and its oligosaccharides, based on the ability of certain pathogenic microbes to produce them. This strategy offers several advantages over animal extraction or chemical synthesis, such as avoiding using expensive substrates and environment-hazard chemicals, avoiding interspecies infection or contaminations, improving homogeneity, and precisely controlling the degree of sulfation and molecular weight [1]. The main benefit of producing GAGs *in vivo* over the enzymatic methods is taking advantage of biological machinery to avoid the use of expensive substrates and co-factors (and instead use cheap simple carbon sources as substrates). However, the industrial biotechnological production of GAGs, and specially of sulfated GAGs such as CS, is limited due to safety concerns for culturing pathogenic microorganisms and low yields [169]. Therefore, research efforts have been focused on the improvement of the production process, as well as on the design of better microbial cell factories for production of GAGs. An overview on the metabolic pathways from native microbial GAG producers will be provided, which can give insights on the design of biosynthetic pathways for chondroitin and CS production. Then it is discussed the current state of the biotechnological production of chondroitin and CS. In Table 2.3, advantages and disadvantages of the available chondroitin production methods are summarized.

Table 2.3. Advantages and disadvantages of different chondroitin and chondroitin derivatives production methods

	Animal extraction	**Chemical synthesis**	**Enzymatic production**	**Microbial production**
Substrate cost	Cheap	Expensive	Expensive	Cheap
Shortage of materials	Yes	No	No	No
Presence of contaminants (prions, viruses, growth factors)	Yes	No	No	No
Vegan/ vegetarian	No	Yes	Yes	Yes
Requires feeding co-factors	No	No	Yes	No
Chemical steps to obtain substrates or sulfation	No	Yes	Most times	Sometimes
Stereoselective and regioselective reactions	*not applicable*	No	Yes	Yes
Time-consuming process	Yes	Yes	Yes, for protein purification	No
Harsh conditions (pH, temperature, pressure)	Yes	Yes	No	No
Scale-up	Limited and expensive	Difficult	Difficult	Easy
Control of degree of sulfation and size	No	Yes	Yes	Yes
Yields	Highest	Low	Lowest	Low
Environmental-friendly process	No	No	Yes	Yes
Possibility to obtain unnatural compounds	No	Yes	Yes	Yes
Final product purification complexity	Very heterogeneous, polydisperse and usually contaminated with other glycosaminoglycans	Easy	Easy	May require cell lysis and purification. For pathogenic hosts, endotoxins need to be removed

2.5.1. Biosynthetic pathways for microbial GAGs production

As already mentioned, native microbial GAG and GAG-like producers are pathogens that use these compounds as constituents of their capsule to camouflage their infection and improve pathogenicity [6]. In these microorganisms, the genes involved in the capsular polysaccharide production from UDP-sugars and its transport are usually expressed in the form of an operon. The pathways for the biosynthesis of microbial GAGs are represented in Figure 2.3. Representative native GAG-producers are *E. coli* K4 for chondroitin, *E. coli* K5 for heparosan, and *P. multocida* type A for HA. Capsular polysaccharide synthesis gene cluster includes three functional regions. Region 1 and region 3 are mainly responsible for the modification, transport, and localization of newly synthesized polysaccharide chain. Region 2 is mostly responsible for encoding enzymes related to the synthesis of polysaccharides and their precursors, so it is the most variable region across strains, according to the capsular GAG produced. The polymerization of these capsular polysaccharides occurs entirely on the cytoplasm through the action of glycosyltransferases followed by transportation, known to be adenosine triphosphate (ATP)-binding cassette (ABC) transporter-dependent [4,170–172].

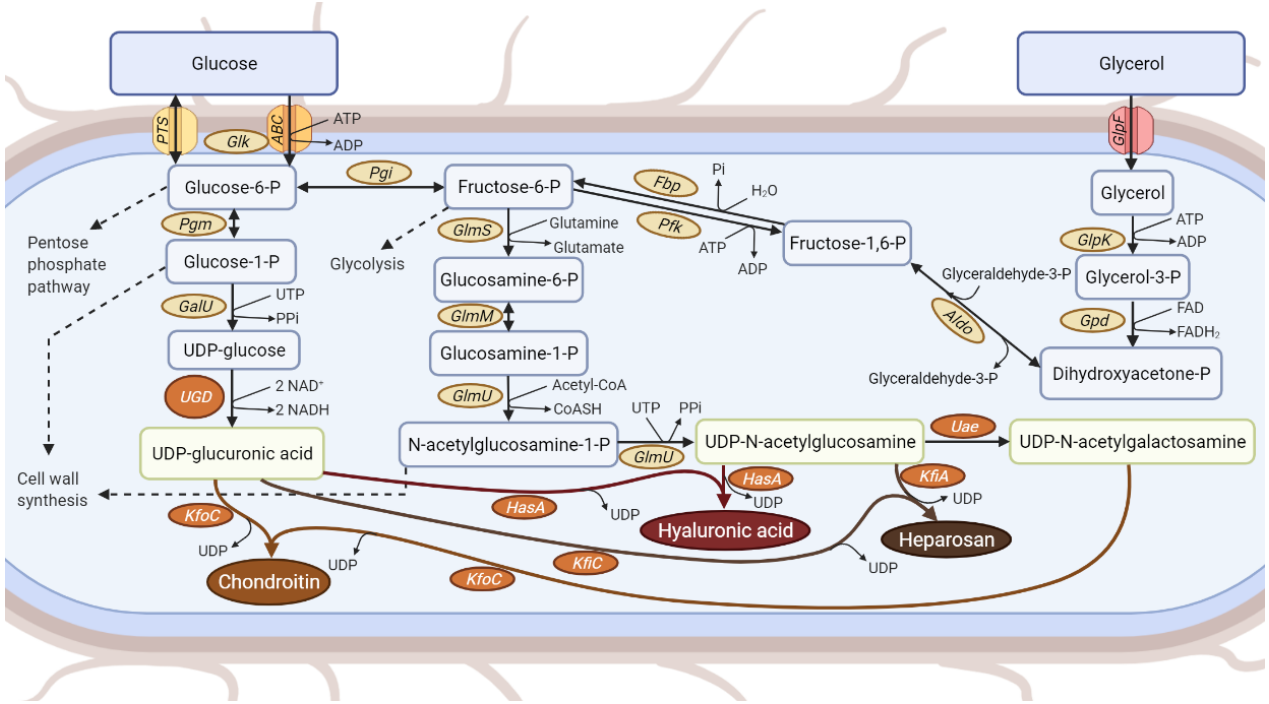


Figure 2.3. Production of glycosaminoglycans in microbes and its possible use in the biosynthesis of microbial chondroitin, hyaluronic acid or heparosan. Depending on the microbial host, the heterologous expression of the enzymes shown in orange boxes might be required for glycosaminoglycans production. Enzyme abbreviations: ABC, adenosine triphosphate (ATP)-binding cassette transporters; Aldo, fructose-6-phosphate aldolase; Fbp, fructose-1,6-bisphosphatase; GalU, uridine triphosphate:glucose-1-phosphate uridylyltransferase; Glk, glucokinase; GlmM, phosphoglucosamine mutase; GlmS, glucosamine-6-phosphate synthase; GlmU, glucosamine-1-phosphate N-acetyltransferase/N-acetylglucosamine-1-phosphate uridyltransferase; GlpF, Glycerol uptake facilitator protein; GlpK, glycerol kinase; Gpd, glyceraldehyde-3-phosphate dehydrogenase; HasA, hyaluronan synthase; KfiA, β-1,3-glucuronyltransferase; KfiC, α-1,4-N-acetylglucosaminyltransferase; KfoC, chondroitin synthase; Pfk, 6-phosphofructokinase; Pgi, glucose-6-phosphate isomerase; Pgm, phosphoglucomutase; PTS, phosphotransferase system; Uae, UDP-*N*-acetylglucosamine 4-epimerase; UGD, uridine diphosphate (UDP)-glucose 6-dehydrogenase.

2.5.2. Current state of biotechnological GAGs production using microbes

Organisms producing other GAGs may comprise an important source of genes with relevance for the chondroitin heterologous pathway given that most of the steps involved in such production are the same. HA is naturally produced by *Streptococcus* group A or C and by *P. multocida* type A [6,172–176]. The genes required for HA biosynthesis are encoded by the HA synthesis (HAS) operon (*hasA-E*) in Streptococci. Since HA does not require modifications after polymerization, the product from the fermentation can be directly isolated and purified. Commercial HA has been obtained mainly by microbial fermentation of *Streptococcus equi* and *S. zooepidermidis* [176,177].

Industrial production of sulfated GAGs is more challenging because it requires post-polymerization modifications of the microbially-produced polysaccharide backbone.

Biosynthesis of HP from microbial sources requires the microbial production of the precursor heparosan [178–180] with further chemical and/or enzymatic modifications [19,181–184]. *E. coli* K5, *P. multocida* type D and *Avibacterium paragallinarum* genotype II can synthesize heparosan [35,36,185,186]. High titers of heparosan have been achieved using *E. coli* K5 (up to 15 g/L in bioreactor) [178,179] which has the heparosan biosynthetic genes encoding glycosyltransferases and dehydrogenases organized in an operon (*kfiA-D,* Figure 2.3). Regarding DS and KS, there have been no reports on their production using microorganisms.

Unsulfated chondroitin, a starting material for CS, is produced by *P. multocida* type F and *A. paragallinarium* genotype I [35,36]. The pathogenic strain *E. coli* K4 natively synthesizes a capsular polysaccharide that shares a similar structure with the unsulfated chondroitin but contains a residue of fructose [187]. Table 2.4 presents recent reported studies on *in vivo* production of chondroitin and CS, including in native producers and in heterologous hosts.

E. coli K4 became the first and most studied microorganism to be used for the biotechnological production of chondroitin [37]. In fact, food grade CS has been industrially produced by microbial fermentation. The native producer *E. coli* K4 is used to produce microbial K4 polysaccharide followed by its chemical defructosylation and regioselective sulfation into CS sodium. The product is sold under the name of Mythocondro™ (marketed since 2017 by Gnosis S.p.A., which was acquired by Lesaffre in 2018).

The genes required for fructosylated chondroitin production and transportation in *E. coli* K4 are organized in an operon containing *kfoA-G* genes. From the seven constituent genes, the most relevant genes in the biosynthesis of fructosylated chondroitin are: *kfoA* encoding UDP-GlcNAc 4-epimerase that provides UDP-GalNAc; *kfoC* encoding chondroitin synthase (homolog to *Pm*CS); and *kfoF* encoding UDP-glucose dehydrogenase that provides UDP-GlcA. Fructosylated chondroitin can undergo a subsequent step of

hydrolysis of fructose monomer and chemical sulfation [188–190]. The identification of the enzyme responsible for inserting the fructose residue (KfoE, fructosyltransferase) and its consequent deletion led to the production of unsulfated chondroitin [191]. Genetic engineering has been applied to *E. coli* K4 to further improve the polysaccharide yields [61,63,187,190,192–195]. In fact, the highest reported yields so far for biotechnological production of chondroitin have been achieved with genetically engineered *E. coli* K4 through fed-batch fermentations, reaching 8.43 to 9.2 g/L of fructosylated chondroitin [187,192]. Feeding alternative synthetic precursors to the fermentation medium is a possible strategy to synthesize rare or unnatural chondroitin derivatives *in vivo*. In the work of Awofiranye *et al.* [196], *N*-glycolyl chondroitin has been produced from chemically synthesized precursors (*N*-glycolylglucosamine replacing GlcNAc in the pathway) in *E. coli* K4 lacking *kfoE* gene.

The first study aiming to produce chondroitin in a heterologous host used the pathogenic bacteria *E. coli* K5 expressing *kfoC* and *kfoA* from *E. coli* K4 and 52.6 mg/L of chondroitin was produced without the need for defructosylation [197]. Other pathogen, HA producer *S. zooepidemicus*, was also attempted as an heterologous host for chondroitin using the same genes [176] but it also resulted in low yields (300 mg/L).

To avoid the contamination of the final product with virulence factors and toxins and to eliminate the defructosylation step, alternative microorganisms have been engineered with K4 polysaccharide production pathway. The chondroitin production in non-pathogenic heterologous hosts has been achieved in *E. coli* BL21 (DE3) [198], *B. subtilis* [158,180,199], *Corynebacterium glutamicum* [200] and *E. coli* K-12 [160,201]. Using safer heterologous hosts, the highest chondroitin yield reported (2.4 g/L) was achieved by expressing expressing *kfoC*, *kfoA*, *kfoF* in *E. coli* BL21 using the high copy number vector pETM6, with a pseudo-operon gene configuration [198].

The chondroitin backbone can be further modified through the action of various sulfotransferases to produce CS, or it can undergo epimerization of GlcA into IdoA and subsequent sulfation to generate DS. Several studies combined the microbial *in vivo* production of chondroitin with enzymatic or chemical strategies to produce CS [158,166,202]. By combining the *in vivo* production of chondroitin backbone and *in vitro* enzymatic sulfation, CS-A and CS-C production has been achieved [158]. *Bacillus subtilis* harboring the *kfoA* and *kfoC* genes from *E. coli* K4 was engineered. *E. coli* was used to express rat aryl sulfotransferase IV (ASST IV) and *P. pastoris* was the expression host for mouse C4OST and C6OST sulfotransferases.

In the work of Erenler [202], a non-pathogenic *E. coli* strain was engineered to express a previously described biosynthetic pathway for chondroitin production [198] along with a *Vitreoscilla* hemoglobin gene

(*vgb*) whose expression has previously been reported to provide benefits to bacteria growth and recombinant protein expression and improved accumulation of different biopolymers [203]. The authors produced and purified the microbial chondroitin and performed chemical sulfation. Recently, the construction of the complete pathway for *in vivo* production of sulfated chondroitin in *E. coli* has been reported [160]. The authors evaluated both *E. coli* K4 and K-12 (containing K4 *kfoCAF* chondroitin genes) with expression of C4OST and PAPS reductase (*cysH*) deletion (to improve the pool of sulfate donor PAPS) for CS production. Only *E. coli* K4 was able to produce extracellular CS (10.76 μg/L in shaken flasks), and the intracellular yield was also higher than the one found for the engineered *E. coli* K-12. However, the percentage of sulfation of intracellular CS was 36% higher in the K-12 strain. After improving the activity of C4OST through mutation, optimizing growth and C4OST induction conditions, and inhibiting transport system through clustered regularly interspaced short palindromic repeats interference (CRISPRi), the levels of sulfation (~55%) reached the same ones achieved in K-12. The authors also optimized the fermentation process in bioreactor using K-12 which resulted in 27 μg/g dry cell weight of intracellular CS-A with a sulfation degree of 96%, a level of 4-*O*-sulfation similar to the ones found in animals. More recently, *P. pastoris* has been successfully engineered towards the production of CS-A using methanol as substrate [204]. The expression of *kfoC* and *kfoA* from *E. coli* K4, *tuaD* from *B. subtilis*, engineered mouse C4OST and overexpression of endogenous genes coding adenosine-5′-triphosphate sulfurylase (ATPS) and adenosine-5′-phosphosulfate kinase (APSK) to improve PAPS supply and sulfation, resulted in a production of 2.1 g/L CS-A with 4.0% sulfation using a fed-batch fermentation. The integrated approach from these two recent studies [160,204] provides significant cost reduction to the process. However, the CS yields in *E. coli* are still low while the CS produced by engineered *P. pastoris* has a low sulfation degree. Therefore, more efforts are required to genetically improve the host and to optimize the process towards interesting yields and sulfation degrees for biotechnological applications.

Table 2.4. Last decade studies on microbial production of chondroitin and chondroitin sulfate by natural producer microbial strains and engineered hosts.

GAG	Host	Substrate(s)	Genetic modification(s)	Process (working volume)	Maximal Yield (mg/L)	Reference
Fructosylated chondroitin	*Escherichia coli* O5:K4:H4	Glycerol/ glucose	Insertion of multiple copies of the autologous *rfaH* gene	Shake flask	280/ 300	[192]
				Batch (2.5 L)	475/ 525	
				Fed-batch (2.5 L)	4000/ 5100	
				Three-phase fermentation (2.5 L): batch-fed batch-in microfiltration regimen	8400/ 9200	
				Three-phase fermentation (1000 L)	9000/(*not tested with glucose)*	
	E. coli O5:K4:H4	Glycerol/ glucose	Overexpression of *rfaH*	Shake-flask (200 mL)	212/ 283	[63]
				Batch (2 L)	466	
				Fed- batch (2 L)	5300	
	E. coli O5:K4:H4	Glycerol	Overexpression of the transcriptional regulator *slyA*	Shake-flask (70 mL)	1000	[194]
				Batch (4 L)	2600	
	E. coli O5:K4:H4	Glycerol	IS2 transposon-mediated *kfoC* overexpression	Shake flasks	302 (plasmid)	[205]
				Batch (2 L)	425 (integrative)	
				Fed-batch (2.5–22 L)	3470 (integrative)	
	E. coli O5:K4:H4	Glycerol	Overexpression of *rfaH, pgm* and *galU*	Shake-flask (200 mL)	391 (with glutamine supplementation)	[65]
				Batch (4 mL)	592	

Table 2.4. Last decade studies on microbial production of chondroitin and chondroitin sulfate by natural producer microbial strains and engineered hosts (*continuation*).

GAG	Host	Substrate(s)	Genetic modification(s)	Process (working volume)	Maximal Yield (mg/L)	Reference
Fructosylated chondroitin	*E. coli* O5:K4:H4	Glycerol	*pfkA* deletion, overexpression of *pgm*, *galU*, *glmS*, *glmM* and mutated *kfoC*, and RBS optimization	Fed-batch (30 L)	8430	[187]
	E. coli O5:K4:H4	Glucose	Overexpression of *kfoA* and *kfoF*	Shake-flask (0.2 L)	~1739 (61 mg/OD)	[193]
	E. coli O5:K4:H4	Glycerol	*glmM*, *glmS*	Fed-batch (2.5 L)	3990	[206]
	E. coli O5:K4:H4	Glucose	Overexpression of *kfoF* Overexpression of *pgm*, *galU* and *kfoF*	Fed-batch (2 L)	2000 2090	[190]
			Overexpression of *pgm*, *galU* and *kfoF*	Fed-batch (22 L)	2140	
	E. coli O5:K4:H4	Glucose/ glycerol	*Wild-type*	Batch in microfermenter (4 mL) and in bioreactor (1.8 L)	315 (microfermenter); 300 (bioreactor)	[195]
				Fed-batch microfermenter (3 mL) and in bioreactor (1.6 L)	1410 (microfermenter); 1570 (bioreactor)	
Chondroitin	*E. coli* O10:K5:H4	Glucose	*kfoC and kfoA* expression	Shake flask (20 mL)	52.6	[197]
	Streptococcus equi subsp. zooepidemicus	Sucrose	*kfoA* and *kfoC* expression	Shake-flask (0.6 L)	90	[176]
				Batch (1.6 L)	300	
	E. coli BL21 Star (DE3)	Glucose	*kfoC, kfoA and kfoF* expression in pseudo-operon gene configuration	Shake flask (25 mL)	213	[198]
				Fed-batch (1 L)	2400	
	Bacillus subtilis 168	Sucrose	*kfoC, kfoA* expression and *tuaD* up-regulation	Shake flask (50 mL)	2540	[180]
				Fed-batch (1.35 L)	5220	

Table 2.4. Last decade studies on microbial production of chondroitin and chondroitin sulfate by natural producer microbial strains and engineered hosts (*continuation*).

GAG	Host	Substrate(s)	Genetic modification(s)	Process (working volume)	Maximal Yield (mg/L)	Reference
Chondroitin	*Corynebacterium glutamicum*	Glucose	*kfoC, kfoA* expression (codon-optimized), *ldh* deletion, and *ugd* overexpression	Fed-batch (2 L)	1910	[200]
Furyl-terminated chondroitin	*E. coli* DH1 (K-12 derivative, lacking *kps* genes)	Glucose-glycerol-lactosides	Expression of *kfoC, kfoG*, *wbpP*, *kfiD* and mouse *glcAT*	Fed-batch (0.2 L)	~2500	[201]
Chondroitin sulfate A	*E. coli* C2987	Catabolizable amino acids from LB	Expression of *kfoA*, *kfoC*, *kfoF* (plasmid from [198]) and *vgb*	Shake flask (25 mL); Chemical sulfation	*Not reported*	[202]
Chondroitin sulfate A and C	*B. subtilis* (chondroitin production) *E. coli* BL21 and *Pichia pastoris* (STs expression)	Sucrose	Expression of *tuaD, glmU, gtaB, glmM, glmS*, and *kfoA* from *B. subtilis* E168C [180]; expression of ASST IV, C4OST and C6OST for sulfation	Fed-batch (1.35 L); Enzymatic sulfation	7150	[158]
Chondroitin sulfate A	*E. coli* O5:K4:H4 (chondroitin production) *E. coli* K-12 MG1655 (cell lysate with PAPS) *E. coli* BL21 Star (DE3) (ST expression)	Glucose (supplementation of sodium sulfate)	*kfoE* deletion in *E. coli* K4 for chondroitin backbone production; deletion of *cysH*, overexpression of *cysDNCQ* in *E. coli* MG1655 – provided cell lysate with PAPS for *in vitro* sulfation reaction; expression of C4OST in *E. coli* BL21 for sulfation	Shake flask (1 L) Enzymatic sulfation (50 µL)	Not reported for *in vivo* production *In vitro:* 8.3 ng chondroitin sulfate A with 0.035% sulfation	[166]

Table 2.4. Last decade studies on microbial production of chondroitin and chondroitin sulfate by natural producer microbial strains and engineered hosts (*continuation*).

GAG	Host	Substrate(s)	Genetic modification(s)	Process (working volume)	Maximal Yield (mg/L)	Reference
Chondroitin sulfate A	*E. coli* O5:K4:H4	Glucose	*kfoE* deletion, C4OST expression and engineering, *cysH* deletion	Shake-flask (25 mL)	0.01076 (extracellular) (126.64 μg/g dry cell weight intracellular)	[160]
	E. coli K-12 MG1655	Glucose	*kfoC, kfoA*, *kfoF*, engineered C4OST expression, *cysH* deletion	Shake-flask (25 mL) Fed-batch (1 L)	0 (extracellular; chondroitin sulfate was only produced intracellularly- 13.14 μg/g dry cell weight intracellular)	
	Pichia pastoris	Methanol	*kfoC*, *kfoA*, *tuaD*, C4OST expression, overexpression of endogenous genes coding ATPS and APSK	Fed-batch (0.9 L)	2100 with 4.0% sulfation	[204]
N-glycolyl chondroitin	*E. coli* O5:K4:H4	Glucose and *N*-glycolylglucosamine	*kfoE* deletion	Shake-flask (100 mL)	~300	[196]

Genes: APSK, adenosine-5′-phosphosulfate kinase; ASST IV, aryl sulfotransferase IV; ATPS, adenosine-5′-triphosphate sulfurylase; C4OST, chondroitin 4-sulfotransferase; C6OST, chondroitin 6-sulfotransferase; *cysC*, adenylyl-sulfate kinase; *cysDN*, adenosine triphosphate sulfurylase; *cysH*, 3′-phosphoadenosine-5′-phosphosulfate (PAPS) reductase; *cysQ*, adenosine-3',5'-bisphosphate nucleotidase; *galU*, uridine triphosphate-glucose-1-phosphate uridylyltransferase; *glcAT*, β-1,3-glucuronyl transferase; *glmM*, phosphoglucosamine mutase; *glmS*, glucosamine-6-phosphate synthase; *glmU*, glucosamine-1-phosphate N-acetyltransferase/N-acetylglucosamine-1-phosphate uridyltransferase; *kfiD*, uridine diphosphate-glucose 6-dehydrogenase from *Escherichia coli* K5; *kfoA*, uridine diphosphate-acetylglucosamine 4-epimerase from *E. coli* K4; *kfoC*, chondroitin synthase from *E. coli* K4; *kfoE*, fructosyltransferase; *kfoF*, uridine diphosphate-glucose 6-dehydrogenase from *E. coli* K4; *kfoG*, chondroitin synthase protein helper from *E. coli* K4; *kps*, surface polysaccharide synthesis genes; *ldh*, lactate dehydrogenase; *pfkA*, adenosine triphosphate-dependent 6-phosphofructokinase; *pgm*, phosphoglucomutase; *rfaH*, transcription antitermination protein; *slyA*, transcriptional regulator; *tuaD*, uridine diphosphate-glucose 6-dehydrogenase from *Bacillus subtilis*; *ugd*, uridine diphosphate-glucose 6-dehydrogenase; *vgb*, *Vitreoscilla* hemoglobin; *wbpP*, uridine diphosphate-acetylglucosamine 4-epimerase from *Pseudomonas aeruginosa*. Other abbreviations: PAPS, 3′-phosphoadenosine-5′-phosphosulfate; RBS, ribosome binding site; ST, sulfotransferase.

2.5.3. Metabolic engineering to optimize CS production

Despite the recent advances on biotechnological production of CS, the yields are still not sufficient to meet the increasing demand of this widely used nutraceutical. Metabolic engineering of the host microorganisms for improving CS precursors (UDP-GlcA, UDP-GlcNAc and UDP-GalNAc) production pathways should be performed to improve intermediate pools and their subsequent availability.

To redirect the metabolic flux towards chondroitin or CS production in microbes, the most obvious modification is the overexpression of chondroitin polymerase/synthase which in *E. coli* K4 is encoded by *kfoC*. When *kfoC* was overexpressed in *E. coli* K4 using plasmids, the fructosylated chondroitin yield improved 2-fold compared to the wild-type [61]. When the *kfoC* overexpression was IS2 transposon-mediated, the K4 production was 2.5 times higher than the wild-type [205]. Strategies to improve KfoC enzymatic activity have also been addressed. By using random mutagenesis, it was possible to improve KfoC activity (R313Q) that led to 82% improvement on fructosylated chondroitin over the wild-type [207]. Regulatory elements of expression of the bacterial GAG-like polysaccharides have also been targets for genetic engineering. The overexpression of the transcriptional activators *rfaH* [63] and s*lyA* [194] have enhanced the expression of capsular genes and consequently improved *E. coli* K4 polysaccharide production by 58% and by a 1.5-fold, respectively, over the wild-type. RfaH acts on capsular gene cluster by preventing the transcript termination of genes related to the modification, transport, and localization of newly synthesized polysaccharide chain (region 3 genes), and consequently improve the expression of genes related to the synthesis of polysaccharides and their precursors (region 2 genes) [208]. SlyA activates the transcription of the whole capsular gene operon [194].

Regarding the precursor availability, balancing UDP-sugars is an established strategy for metabolic engineering of HA [209] and fructosylated chondroitin [187]. The genes *pgm*, *galU* and *ugd/kfoF* (encoding phosphoglucomutase, uridine triphosphate-glucose-1-phosphate uridylyltransferase and UDP-glucose dehydrogenase, respectively) for UDP-GlcA synthesis and *glmS, glmM and glmU* (encoding glucosamine-6-phosphate synthase, phosphoglucosamine mutase and glucosamine-1-phosphate *N*-acetyltransferase/*N*-acetylglucosamine-1-phosphate uridyltransferase, respectively) for UDP-GalNAc production might be interesting targets for overexpression (Figure 2.3). Overexpression of these genes have resulted in improved production of K4 and other polysaccharides. Levander and coworkers [210] found that the overexpression of the endogenous *galU* gene in *S. thermophilus* LY03 led to a 10-fold increase in *galU* activity, however exopolysaccharide yield was not affected. Nevertheless, when *galU* was overexpressed in combination with *pgmA* (*pgm* homolog), the exopolysaccharide yield increased. Overexpression of *pgm* and *galU* in *E. coli* AD202 also resulted in increased UDP-galactose derived

disaccharides from 2.5 mM to 20 mM by improvement of carbon flux through UDP-glucose synthesis pathway [211]. Engineering *E. coli* K4 with one copy of endogenous genes *pgm* and *galU* also improved capsular polysaccharide production [65]. The authors further increased capsular polysaccharide yields in 45% using glutamine supplementation to boost UDP-GalNAc precursor production [63,65]. Engineering *Saccharomyces cerevisiae* with the endogenous genes *pgm2* and *ugp1* (equivalent to *galU*) resulted in a 17% improvement of scutellarein 7-O-glucoside production rate by improving UDP-glucose pool [212]. Overexpressing *glmM* and *glmU* genes resulted in higher capsular polysaccharide production in *E. coli* K4 [206]. Despite the step catalyzed by *ugd*-codified enzyme was considered in previous works as the limiting factor of GAG biosynthesis in homologous and heterologous organisms [61,180,213,214], few examples evaluating alternative genes for this step have been reported, possibly missing interesting catalysts.

On the other hand, the repression of genes from competing pathways can also be beneficial for redirecting the metabolic flux towards the production of GAG precursors. Down-regulating the expression of three genes (glucose-6-phosphate 1-dehydrogenase *zwf,* adenosine triphosphate-dependent 6-phosphofructokinase *pfkA, and glmM*) that control the major competing reactions (cell wall recycling pathways) of GlcNAc synthesis by CRISPRi improved GlcNAc titers in *B. subtilis* [215]. Silencing *zwf* and *pfkA* also resulted in improved HA production in *B. subtilis* [216]. By favoring the production of fructose 6-phosphate, *pfkA* knock-out resulted in increased chondroitin production in *E. coli* K4 [187].

Combination of both repression and activation can provide optimized microbial cell flux towards GAGs production. Knock-out of *pfkA* and *zwf* genes coupled with overexpression of the genes *galU-ugd* and *glmS-glmM-glmU* improved HA production [217] in *E. coli* K-12 W3110 harboring HA synthase gene (*hasA*) from *S. zooepidemicus*, and with transcriptional repressors genes *galR* and *galS* deleted.

When the intended biosynthetic pathway has intermediates that are critical for biomass production, as it is the case for chondroitin, redirecting the carbon flux might result in lower yield and growth rate. To assess this question, there have been interesting advances on dynamical control to balance the metabolic flux according to the intermediate concentration. This strategy has been applied for overproduction of GlcNAc in *B. subtilis* which consisted in engineering the native glucosamine-6-phosphate responsive *glmS* ribozyme switch to act as a sensor and dynamically control the metabolic flux [218]. This strategy increased about 2-fold the native GlcNAc titer, being a potential strategy to improve chondroitin producing strains.

Finally, computational methods can be useful to aid on the prediction of potential targets for improving CS production in heterologous hosts. The use of stoichiometric or kinetic models of the heterologous host

to perform *in silico* predictions has been shown to be useful to guide engineering strategies towards the improved production of valuable compounds. In fact, a computational approach was already performed to engineer an optimal HA production in *C. glutamicum* [219]. As this field of knowledge is complex and growing, a summary on reported applications of computational approaches for aiding metabolic engineering is presented on Table A 2.1.

Sophisticated genetic engineering tools for genome editing, such as CRISPR [220], have been reported to be rapid and cost-effective and, therefore, can be implemented on the intended system to generate engineered hosts with improved chondroitin production.

2.5.4. Alternative Hosts

Alternative hosts can provide more suitable platforms for CS production. Besides replacing the native pathogenic bacteria and, consequently, avoiding the presence of virulence factors in the final product, the application of industrially used *E. coli*, *B. subtilis*, and *Corynebacterium* for CS production has the advantages of being fast-growing organisms, with well-characterized genetic and physiological backgrounds, and many available tools for gene expression and genome editing [221].

However, biosynthetic enzymes expressed in bacterial systems usually present low solubility, stability, and activity, which may limit CS synthesis. On the other hand, enzymes expressed in eukaryotes typically show enhanced activity but are expressed at lower amounts [167]. Specially for the expression of eukaryotic genes, in particular sulfotransferases, eukaryotic microbial cells such as the broadly used *S. cerevisiae* and *P. pastoris* can provide beneficial conditions regarding codon preference, enzyme folding and glycosylation patterns more similar to the original enzymes. Also, since mammalian CHO cells have shown to be able to produce HP [222], animal cell fermentation can also be a possible, however expensive, alternative strategy to produce CS.

2.5.5. Additional optimization strategies

Besides the metabolic engineering strategies to improve intermediate pools, or the use of alternative hosts, other approaches to make biotechnological CS production more cost-effective include screening for more efficient enzymes, modification of the enzymes through protein engineering and immobilization to improve activity and productivity. Also, salvage pathways can be added to improve precursors – for example, UDP-GlcA can also be produced in two different ways directly from GlcA using glucuronokinase, UDP-sugar pyrophosphorylase, and inorganic phosphatase instead of depending on UDP-glucose

conversion through *ugd* step. Increasing energy through ATP supply and implementing ATP regeneration systems can also benefit the CS *in vivo* production. The control of CS molecular weight in fermentation is also desirable to achieve a more monodisperse product [190,223]. In addition, optimization of the fermentation process, media composition, temperature and pH conditions have been shown to be effective for improved GAG yields [178,224]. Using low-cost media can provide a more competitive process, namely agro-industrial by-products such as sugarcane molasses and corn steep liquor which have already been used for HA production [225–227], therefore they present potential applicability for CS production. Additionally, engineering of CS transportation can aid to improve CS titers, by even repressing transporters for increased CS intracellular *in vivo* sulfation [160], or by contrarily strengthening cell CS exportation, which would be more valuable from an industrial point of view but would require the extracellular expression of the sulfotransferases. Finally, optimizing environmental conditions for maximal enzyme activity to perform the post-polymerization modifications of biotechnological CS polysaccharides *in vivo* or *in vitro* would aid to obtain a final commercial sulfated product at competitive yields.

2.6. Conclusions

Compared to animal-derived products, heterologous microbes can provide pathogen-free chondroitin without the use of hazardous chemicals. With the increasing applications of CS, the studies on its microbial production have been growing with much focus on the native GAG producers' pathway to polymerize the backbone starting material, and on its post-polymerization modifications through enzymatic *in vitro* reactions. However, there are bottlenecks such as insufficient supply of precursors, low activity and/or stability of the enzymes and difficult and costly sulfotransferase enzymes purification. Strategies that can help solving these issues include genetic engineering for improving the precursors pool and transport; evaluation of alternative hosts for efficient expression of enzymes (since most genes required for sulfated GAGs production are from animals, expression in eukaryotic hosts could provide more soluble and active enzymes); improvement of enzyme activity and solubility through enzyme engineering and environment optimization; and the use of integrated approaches of microbial production and post-polymerization. The use of computational approaches to aid the identification of metabolic engineering targets is also a powerful tool in the construction of optimized microbial cell factories. The knowledge improvement on GAGs enzymes for polymerization, transport, and post-polymerization, together with the development of efficient technologies for gene editing, are expected to lead to the establishment of efficient microbial cell factories for the CS production.

2.7. References

[1] Z. Kang, Z. Zhou, Y. Wang, H. Huang, G. Du, J. Chen, Bio-based strategies for producing glycosaminoglycans and their oligosaccharides, Trends Biotechnol. 36 (2018) 806–818. https://doi.org/10.1016/j.tibtech.2018.03.010.

[2] M. Mende, C. Bednarek, M. Wawryszyn, P. Sauter, M.B. Biskup, U. Schepers, S. Bräse, Chemical synthesis of glycosaminoglycans, Chem. Rev. 116 (2016) 8193–8255. https://doi.org/10.1021/acs.chemrev.6b00010.

[3] J. Casale, J.S. Crane, Biochemistry, Glycosaminoglycans, Treasure Island (FL): StatPearls Publishing, 2021. http://www.ncbi.nlm.nih.gov/pubmed/31335015 (accessed October, 2023).

[4] A. Varki, R.D. Cummings, J.D. Esko, P. Stanley, G.W. Hart, M. Aebi, A.G. Darvill, T. Kinoshita, N.H. Packer, J.H. Prestegard, R.L. Schnaar, P.H. Seeberger, Essentials of glycobiology, Cold Spring Harb. (2017) 823. https://www.ncbi.nlm.nih.gov/books/NBK310274/ (accessed October, 2023).

[5] J. Kovensky, E. Grand, M.L. Uhrig, Applications of glycosaminoglycans in the medical, veterinary, pharmaceutical, and cosmetic fields, in: S.N. Goyanes, N.B. D'Accorso (Eds.), Ind. Appl. Renew. Biomass Prod., Springer International Publishing, Cham, 2017: pp. 135–164. https://doi.org/10.1007/978-3-319-61288-1_5.

[6] B.F. Cress, J.A. Englaender, W. He, D. Kasper, R.J. Linhardt, M.A.G. Koffas, Masquerading microbial pathogens: capsular polysaccharides mimic host-tissue molecules, FEMS Microbiol. Rev. 38 (2014) 660–697. https://doi.org/10.1111/1574-6976.12056.

[7] K.W. Chang, S.F. Weng, Y.H. Tseng, UDP-glucose dehydrogenase gene of *Xanthomonas campestris* is required for virulence, Biochem. Biophys. Res. Commun. 287 (2001) 550–555. https://doi.org/10.1006/bbrc.2001.5591.

[8] L. Guan, L. Zhang, Y. Xue, J. Yang, Z. Zhao, Molecular pathogenesis of the hyaluronic acid capsule of *Pasteurella multocida*, Microb. Pathog. 149 (2020) 104380. https://doi.org/10.1016/j.micpath.2020.104380.

[9] A. Köwitsch, G. Zhou, T. Groth, Medical application of glycosaminoglycans: a review, J. Tissue Eng. Regen. Med. 12 (2018) e23–e41. https://doi.org/10.1002/term.2398.

[10] N. Volpi, Therapeutic Applications of glycosaminoglycans, Curr. Med. Chem. 13 (2006) 1799–1810. https://doi.org/10.2174/092986706777452470.

[11] F. Zhang, Z. Zhang, R.J. Linhardt, Glycosaminoglycans, in: Handb. Glycomics, 2009: pp. 59–80. https://doi.org/10.1016/B978-0-12-373600-0.00003-2.

[12] M.M. Abdallah, N. Fernández, A.A. Matias, M. do R. Bronze, Hyaluronic acid and chondroitin sulfate from marine and terrestrial sources: extraction and purification methods, Carbohydr. Polym. 243 (2020) 116441. https://doi.org/10.1016/j.carbpol.2020.116441.

[13] D. Cimini, O.F. Restaino, C. Schiraldi, Microbial production and metabolic engineering of chondroitin and chondroitin sulfate, Emerg. Top. Life Sci. 2 (2018) 349–361.

https://doi.org/10.1042/ETLS20180006.

[14] R. Tian, Y. Liu, L. Liu, Microbial production of oligosaccharides and polysaccharides, in: L. Liu, J. Chen (Eds.), Syst. Synth. Biotechnol. Prod. Nutraceuticals, Springer Singapore, Singapore, 2019: pp. 75–91. https://doi.org/10.1007/978-981-15-0446-4_4.

[15] V.H. Pomin, W.P. Vignovich, A. V. Gonzales, A.A. Vasconcelos, B. Mulloy, Galactosaminoglycans: medical applications and drawbacks, Molecules. 24 (2019) 1–33. https://doi.org/10.3390/molecules24152803.

[16] G. Vessella, J.A. Vázquez, J. Valcárcel, L. Lagartera, D.T. Monterrey, A. Bastida, E. García-junceda, E. Bedini, A. Fernández-mayoralas, J. Revuelta, Deciphering structural determinants in chondroitin sulfate binding to FGF-2: paving the way to enhanced predictability of their biological functions, Polymers (Basel). 13 (2021) 1–15. https://doi.org/10.3390/polym13020313.

[17] C. Malavaki, S. Mizumoto, N. Karamanos, K. Sugahara, Recent advances in the structural study of functional chondroitin sulfate and dermatan sulfate in health and disease, Connect. Tissue Res. 49 (2008) 133–139. https://doi.org/10.1080/03008200802148546.

[18] G. Vessella, S. Traboni, D. Cimini, A. Iadonisi, C. Schiraldi, E. Bedini, Development of semisynthetic, regioselective pathways for accessing the missing sulfation patterns of chondroitin sulfate, Biomacromolecules. 20 (2019) 3021–3030. https://doi.org/10.1021/acs.biomac.9b00590.

[19] P. Datta, R.J. Linhardt, S.T. Sharfstein, Industrial production of glycosaminoglycans, in: Encycl. Microbiol., Elsevier, 2019: pp. 681–690. https://doi.org/10.1016/B978-0-12-809633-8.12224-1.

[20] N. Maeda, M. Ishii, K. Nishimura, K. Kamimura, Functions of chondroitin sulfate and heparan sulfate in the developing brain, Neurochem. Res. 36 (2011) 1228–1240. https://doi.org/10.1007/s11064-010-0324-y.

[21] J.C.F. Kwok, P. Warren, J.W. Fawcett, Chondroitin sulfate: a key molecule in the brain matrix, Int. J. Biochem. Cell Biol. 44 (2012) 582–586. https://doi.org/10.1016/j.biocel.2012.01.004.

[22] T. Mikami, H. Kitagawa, Biosynthesis and function of chondroitin sulfate, Biochim. Biophys. Acta - Gen. Subj. 1830 (2013) 4719–4733. https://doi.org/10.1016/j.bbagen.2013.06.006.

[23] R.R. Dinglasan, A. Alaganan, A.K. Ghosh, A. Saito, T.H. Van Kuppevelt, M. Jacobs-Lorena, *Plasmodium falciparum* ookinetes require mosquito midgut chondroitin sulfate proteoglycans for cell invasion, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 15882–15887. https://doi.org/10.1073/pnas.0706340104.

[24] P. Kastana, E. Choleva, E. Poimenidi, N. Karamanos, K. Sugahara, E. Papadimitriou, Insight into the role of chondroitin sulfate E in angiogenesis, FEBS J. 286 (2019) 2921–2936. https://doi.org/10.1111/febs.14830.

[25] J.M. Trowbridge, R.L. Gallo, Dermatan sulfate: new functions from an old glycosaminoglycan, Glycobiology. 12 (2002) 117R–125R. https://doi.org/10.1093/glycob/cwf066.

[26] L.C. Rosenberg, H.U. Choi, L.H. Tang, T.L. Johnson, S. Pal, C. Webber, A. Reiner, A.R. Poole, Isolation of dermatan sulfate proteoglycans from mature bovine articular cartilages, J. Biol. Chem. 260 (1985) 6304–6313. https://doi.org/10.1016/s0021-9258(18)88971-2.

[27] K. Sugahara, T. Mikami, T. Uyama, S. Mizuguchi, K. Nomura, H. Kitagawa, Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate, Curr. Opin. Struct. Biol. 13 (2003) 612–620. https://doi.org/10.1016/j.sbi.2003.09.011.

[28] E.H. Song, J. Shang, D.M. Ratner, Polysaccharides, in: Polym. Sci. A Compr. Ref. 10 Vol. Set, Elsevier, 2012: pp. 137–155. https://doi.org/10.1016/B978-0-444-53349-4.00246-6.

[29] M. Vallières, P. du Souich, Modulation of inflammation by chondroitin sulfate, Osteoarthr. Cartil. 18 (2010) S1–S6. https://doi.org/10.1016/j.joca.2010.02.017.

[30] E. Tykesson, M. Maccarana, H. Thorsson, J. Liu, A. Malmström, U. Ellervik, G. Westergren-Thorsson, Recombinant dermatan sulfate is a potent activator of heparin cofactor II-dependent inhibition of thrombin, Glycobiology. 29 (2019) 446–451. https://doi.org/10.1093/glycob/cwz019.

[31] A.M.F. Tovar, D.A. De Mattos, M.P. Stelling, B.S.L. Sarcinelli-Luz, R.A. Nazareth, P.A.S. Mourão, Dermatan sulfate is the predominant antithrombotic glycosaminoglycan in vessel walls: Implications for a possible physiological function of heparin cofactor II, Biochim. Biophys. Acta - Mol. Basis Dis. 1740 (2005) 45–53. https://doi.org/10.1016/j.bbadis.2005.02.008.

[32] M. Ida, T. Shuo, K. Hirano, Y. Tokita, K. Nakanishi, F. Matsui, S. Aono, H. Fujita, Y. Fujiwara, T. Kaji, A. Oohira, Identification and functions of chondroitin sulfate in the milieu of neural stem cells, J. Biol. Chem. 281 (2006) 5982–5991. https://doi.org/10.1074/jbc.M507130200.

[33] M.K. Sarangi, M.E.B. Rao, V. Parcha, D.K. Yi, S.S. Nanda, Marine polysaccharides for drug delivery in tissue engineering, in: Nat. Polysaccharides Drug Deliv. Biomed. Appl., Elsevier, 2019: pp. 513–530. https://doi.org/10.1016/B978-0-12-817055-7.00022-4.

[34] P.C.Y. Liaw, D.L. Becker, A.R. Stafford, J.C. Fredenburgh, J.I. Weitz, Molecular basis for the susceptibility of fibrin-bound thrombin to inactivation by heparin cofactor II in the presence of dermatan sulfate but not heparin, J. Biol. Chem. 276 (2001) 20959–20965. https://doi.org/10.1074/jbc.M010584200.

[35] P.L. DeAngelis, N.S. Gunay, T. Toida, W. Mao, R.J. Linhardt, Identification of the capsular polysaccharides of Type D and F *Pasteurella multocida* as unmodified heparin and chondroitin, respectively, Carbohydr. Res. 337 (2002) 1547–1552. https://doi.org/10.1016/S0008-6215(02)00219-7.

[36] J.-R. Wu, P.-Y. Chen, J.-H. Shien, C.-L. Shyu, H.K. Shieh, F. Chang, P.-C. Chang, Analysis of the biosynthesis genes and chemical components of the capsule of *Avibacterium paragallinarum*, Vet. Microbiol. 145 (2010) 90–99. https://doi.org/10.1016/j.vetmic.2010.03.002.

[37] D. Cimini, O.F. Restaino, A. Catapano, M. De Rosa, C. Schiraldi, Production of capsular polysaccharide from *Escherichia coli* K4 for biotechnological applications, Appl. Microbiol. Biotechnol. 85 (2010) 1779–1787. https://doi.org/10.1007/s00253-009-2261-8.

[38] V. Gargiulo, R. Lanzetta, M. Parrilli, C. De Castro, Structural analysis of chondroitin sulfate from *Scyliorhinus canicula*: a useful source of this polysaccharide, Glycobiology. 19 (2009) 1485–1491. https://doi.org/10.1093/glycob/cwp123.

[39] K. Sugahara, M. Masuda, T. Harada, I. Yamashina, P. de Waard, J.F.G. Vliegenthart, Structural studies on sulfated oligosaccharides derived from the carbohydrate-protein linkage region of chondroitin sulfate proteoglycans of whale cartilage, J. Biol. Chem. 202 (1991) 805–811. https://doi.org/10.1111/j.1432-1033.1991.tb16436.x.

[40] E. Bedini, M.M. Corsaro, A. Fernández-Mayoralas, A. Iadonisi, Chondroitin, dermatan, heparan, and keratan sulfate: structure and functions, Springer, Cham, 2019. https://doi.org/10.1007/978-3-030-12919-4_5.

[41] K. Sugahara, Y. Ohi, T. Harada, P. de Waard, J.F.G. Vliegenthart, Structural studies on sulfated oligosaccharides derived from the carbohydrate-protein linkage region of chondroitin 6-sulfate proteoglycans of shark cartilage. I. Six compounds containing 0 or 1 sulfate and/or phosphate residue, J. Biol. Chem. 267 (1992) 6027–6035. https://doi.org/10.1016/s0021-9258(18)42658-0.

[42] N. Yutsudo, H. Kitagawa, Involvement of chondroitin 6-sulfation in temporal lobe epilepsy, Exp. Neurol. 274 (2015) 126–133. https://doi.org/10.1016/j.expneurol.2015.07.009.

[43] Q. Zhang, J. Li, C. Liu, C. Song, P. Li, F. Yin, Y. Xiao, J. Li, W. Jiang, A. Zong, X. Zhang, F. Wang, Protective effects of low molecular weight chondroitin sulfate on amyloid beta (Aβ)-induced damage *in vitro* and *in vivo*, Neuroscience. 305 (2015) 169–182. https://doi.org/10.1016/j.neuroscience.2015.08.002.

[44] S. Nadanaka, A. Clement, K. Masayama, A. Faissner, K. Sugahara, Characteristic hexasaccharide sequences in octasaccharides derived from shark cartilage chondroitin sulfate D with a neurite outgrowth promoting activity, J. Biol. Chem. 273 (1998) 3296–3307. https://doi.org/10.1074/jbc.273.6.3296.

[45] U. Rauch, P. Gao, A. Janetzko, A. Flaccus, L. Hilgenberg, H. Tekotte, R.K. Margolis, R.U. Margolis, Isolation and characterization of developmentally regulated chondroitin sulfate and chondroitin/keratin sulfate proteoglycans of brain identified with monoclonal antibodies, J. Biol. Chem. 266 (1991) 14785–14801. https://doi.org/10.1016/s0021-9258(18)98755-7.

[46] B. Zhang, L. Chi, Chondroitin sulfate/dermatan sulfate-protein interactions and their biological functions in human diseases: implications and analytical tools, Front. Cell Dev. Biol. 9 (2021) 1–13. https://doi.org/10.3389/fcell.2021.693563.

[47] C. Peng, Q. Wang, R. Jiao, Y. Xu, N. Han, W. Wang, C. Zhu, F. Li, A novel chondroitin sulfate E from *Dosidicus gigas* cartilage and its antitumor metastatic activity, Carbohydr. Polym. 262 (2021) 117971. https://doi.org/10.1016/j.carbpol.2021.117971.

[48] R.L. Stevens, C.C. Fox, L.M. Lichtenstein, K.F. Austen, Identification of chondroitin sulfate E proteoglycans and heparin proteoglycans in the secretory granules of human lung mast cells, Proc. Natl. Acad. Sci. U. S. A. 85 (1988) 2284–2287. https://doi.org/10.1073/pnas.85.7.2284.

[49] N.E. Ustyuzhanina, M.I. Bilan, A.S. Dmitrenok, N.E. Nifantiev, A.I. Usov, Two fucosylated chondroitin sulfates from the sea cucumber *Eupentacta fraudatrix*, Carbohydr. Polym. 164 (2017) 8–12. https://doi.org/10.1016/j.carbpol.2017.01.034.

[50] N.E. Ustyuzhanina, M.I. Bilan, E.G. Panina, N.P. Sanamyan, A.S. Dmitrenok, E.A. Tsvetkova, N.A. Ushakova, A.S. Shashkov, N.E. Nifantiev, A.I. Usov, Structure and anti-inflammatory activity of a new unusual fucosylated chondroitin sulfate from *Cucumaria djakonovi*, Mar. Drugs. 16 (2018) 389. https://doi.org/10.3390/md16100389.

[51] N.E. Ustyuzhanina, M.I. Bilan, A.S. Dmitrenok, E.Y. Borodina, V.A. Stonik, N.E. Nifantiev, A.I. Usov, A highly regular fucosylated chondroitin sulfate from the sea cucumber *Massinium magnum*: Structure and effects on coagulation, Carbohydr. Polym. 167 (2017) 20–26. https://doi.org/10.1016/j.carbpol.2017.02.101.

[52] N.E. Ustyuzhanina, M.I. Bilan, A.S. Dmitrenok, A.S. Shashkov, N.E. Nifantiev, A.I. Usov, The structure of a fucosylated chondroitin sulfate from the sea cucumber *Cucumaria frondosa*, Carbohydr. Polym. 165 (2017) 7–12. https://doi.org/10.1016/j.carbpol.2017.02.003.

[53] J. Mou, C. Wang, W. Li, J. Yang, Purification, structural characterization and anticoagulant properties of fucosylated chondroitin sulfate isolated from *Holothuria mexicana*, Int. J. Biol. Macromol. 98 (2017) 208–215. https://doi.org/10.1016/j.ijbiomac.2017.01.123.

[54] S. Chen, C. Xue, L. Yin, Q. Tang, G. Yu, W. Chai, Comparison of structures and anticoagulant activities of fucosylated chondroitin sulfates from different sea cucumbers, Carbohydr. Polym. 83 (2011) 688–696. https://doi.org/10.1016/j.carbpol.2010.08.040.

[55] A. Kinoshita-Toyoda, S. Yamada, S.M. Haslam, K.H. Khoo, M. Sugiura, H.R. Morris, A. Dell, K. Sugahara, Structural determination of five novel tetrasaccharides containing 3-*O*-sulfated *D*-glucuronic acid and two rare oligosaccharides containing a β-*D*-glucose branch isolated from squid cartilage chondroitin sulfate E, Biochemistry. 43 (2004) 11063–11074. https://doi.org/10.1021/bi049622d.

[56] C. Ueoka, S. Nadanaka, N. Seno, K.H. Khoo, K. Sugahara, Structural determination of novel tetra- and hexasaccharide sequences isolated from chondroitin sulfate H (oversulfated dermatan sulfate) of hagfish notochord, Glycoconj. J. 16 (1999) 291–305. https://doi.org/10.1023/A:1007022229813.

[57] K. Higashi, Y. Okamoto, A. Mukuno, J. Wakai, S. Hosoyama, R.J. Linhardt, T. Toida, Functional chondroitin sulfate from *Enteroctopus dofleini* containing a 3-*O*-sulfo glucuronic acid residue, Carbohydr. Polym. 134 (2015) 557–565. https://doi.org/10.1016/j.carbpol.2015.07.082.

[58] K. Sugahara, Y. Tanaka, S. Yamada, N. Seno, H. Kitagawa, S.M. Haslam, H.R. Morris, A. Dell, Novel sulfated oligosaccharides containing 3-*O*-sulfated glucuronic acid from king crab cartilage chondroitin sulfate K: Unexpected degradation by chondroitinase ABC, J. Biol. Chem. 271 (1996) 26745–26754. https://doi.org/10.1074/jbc.271.43.26745.

[59] H. Kitagawa, Y. Tanaka, S. Yamada, N. Seno, S.M. Haslam, H.R. Morris, A. Dell, K. Sugahara, A novel pentasaccharide sequence GlcA(3-sulfate)(β1-3)GalNAc(4-sulfate)(β1-4)(Fucα1-3)GlcA(β1-

3)GalNAc(4-sulfate) in the oligosaccharides isolated from king crab cartilage chondroitin sulfate K and its differential susceptibility to chondroitinases and hyalurinidase, Biochemistry. 36 (1997) 3998–4008. https://doi.org/10.1021/bi962740j.

[60] A. Kinoshita, S. Yamada, S.M. Haslam, H.R. Morris, A. Dell, K. Sugahara, Isolation and structural determination of novel sulfated hexasaccharides from squid cartilage chondroitin sulfate E that exhibits neuroregulatory activities, Biochemistry. 40 (2001) 12654–12665. https://doi.org/10.1021/bi015577n.

[61] D. Cimini, M. De Rosa, A. Viggiani, O.F. Restaino, E. Carlino, C. Schiraldi, Improved fructosylated chondroitin production by *kfoC* overexpression in *E. coli* K4, J. Biotechnol. 150 (2010) 324–331. https://doi.org/10.1016/j.jbiotec.2010.09.954.

[62] C. Schiraldi, A. Alfano, D. Cimini, M. De Rosa, A. Panariello, O.F. Restaino, M. De Rosa, Application of a 22L scale membrane bioreactor and cross-flow ultrafiltration to obtain purified chondroitin, Biotechnol. Prog. 28 (2012) 1012–1018. https://doi.org/10.1002/btpr.1566.

[63] D. Cimini, De Rosa, E. Carlino, A. Ruggiero, C. Schiraldi, Homologous overexpression of *rfaH* in *E. coli* K4 improves the production of chondroitin-like capsular polysaccharide, Microb. Cell Fact. 12 (2013) 1. https://doi.org/10.1186/1475-2859-12-46.

[64] J. Liu, A. Yang, J. Liu, X. Ding, L. Liu, Z. Shi, *KfoE* encodes a fructosyltransferase involved in capsular polysaccharide biosynthesis in *Escherichia coli* K4, Biotechnol. Lett. 36 (2014) 1469–1477. https://doi.org/10.1007/s10529-014-1502-9.

[65] D. Cimini, E. Carlino, A. Giovane, O. Argenzio, I. Dello Iacono, M. De Rosa, C. Schiraldi, Engineering a branch of the UDP-precursor biosynthesis pathway enhances the production of capsular polysaccharide in *Escherichia coli* O5:K4:H4, Biotechnol. J. 10 (2015) 1307–1315. https://doi.org/10.1002/biot.201400602.

[66] A.N. Samraj, O.M.T. Pearce, H. Läubli, A.N. Crittenden, A.K. Bergfeld, K. Band, C.J. Gregg, A.E. Bingman, P. Secrest, S.L. Diaz, N.M. Varki, A. Varki, S.A. Kornfeld, A red meat-derived glycan promotes inflammation and cancer progression, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 542–547. https://doi.org/10.1073/pnas.1417508112.

[67] T. Zhao, J. Zhang, X. Liu, F. Wang, Analysis of chondroitin sulfate from different sources of cartilage by electrophoretically mediated microanalysis, RSC Adv. 5 (2015) 52314–52319. https://doi.org/10.1039/c5ra05576h.

[68] W. Garnjanagoonchorn, L. Wongekalak, A. Engkagul, Determination of chondroitin sulfate from different sources of cartilage, Chem. Eng. Process. Process Intensif. 46 (2007) 465–471. https://doi.org/10.1016/j.cep.2006.05.019.

[69] G. Sundaresan, R.J.J. Abraham, V. Appa Rao, R. Narendra Babu, V. Govind, M.F. Meti, Established method of chondroitin sulphate extraction from buffalo (*Bubalus bubalis*) cartilages and its identification by FTIR, J. Food Sci. Technol. 55 (2018) 3439–3445. https://doi.org/10.1007/s13197-018-3253-4.

[70] X.M. Luo, G.J. Fosmire, R.M. Leach, Chicken keel cartilage as a source of chondroitin sulfate,

Poult. Sci. 81 (2002) 1086–1089. https://doi.org/10.1093/ps/81.7.1086.

[71] R. Novoa-Carballal, R. Pérez-Martín, M. Blanco, C.G. Sotelo, D. Fassini, C. Nunes, M.A. Coimbra, T.H. Silva, R.L. Reis, J.A. Vázquez, By-products of *Scyliorhinus canicula*, *Prionace glauca* and *Raja clavata*: a valuable source of predominantly 6S sulfated chondroitin sulfate, Carbohydr. Polym. 157 (2017) 31–37. https://doi.org/10.1016/j.carbpol.2016.09.050.

[72] Grand View Research, Chondroitin sulfate market size, share & trends analysis report by source (bovine, swine, poultry), by application (nutraceuticals, pharmaceuticals, animal feed, personal care & cosmetics), and segment forecascts, 2021-2028, 2021. https://www.grandviewresearch.com/industry-analysis/chondroitin-sulfate-market#.

[73] Beroe Inc, Chondroitin Sulfate Market Intelligence. Price, Forecast, Cost Models, Market Analysis - Market Research, Beroe. (2018). https://www.beroeinc.com/category-intelligence/chondroitin-sulfate-market/.

[74] A. Bhathal, M. Spryszak, C. Louizos, G. Frankel, Glucosamine and chondroitin use in canines for osteoarthritis: A review, Open Vet. J. 7 (2017) 36–49. https://doi.org/10.4314/ovj.v7i1.6.

[75] F. Oliviero, R. Ramonda, A. Hoxha, A. Scanu, P. Galozzi, M. Favero, P. Frallonardo, L. Punzi, Effect of an oral preparation containing hyaluronic acid, chondroitin sulfate, hydrolyzed collagen type II and hydrolyzed keratin on synovial fluid features and clinical indices in knee osteoarthritis. A pilot study, Reumatismo. 72 (2020) 125–130.

[76] G.R.C. Santos, A.A. Piquet, B.F. Glauser, A.M.F. Tovar, M.S. Pereira, E. Vilanova, P.A.S. Mourão, Systematic analysis of pharmaceutical preparations of chondroitin sulfate combined with glucosamine, Pharmaceuticals. 10 (2017). https://doi.org/10.3390/ph10020038.

[77] C. Bottegoni, R.A.A. Muzzarelli, F. Giovannini, A. Busilacchi, A. Gigante, Oral chondroprotection with nutraceuticals made of chondroitin sulphate plus glucosamine sulphate in osteoarthritis, Carbohydr. Polym. 109 (2014) 126–138. https://doi.org/10.1016/j.carbpol.2014.03.033.

[78] D. Uebelhart, Clinical review of chondroitin sulfate in osteoarthritis, Osteoarthr. Cartil. 16 (2008) 19–21. https://doi.org/10.1016/j.joca.2008.06.006.

[79] A.L. Pérez-Balbuena, J.C. Ochoa-Tabares, S. Belalcazar-Rey, C. Urzúa-Salinas, L.R. Saucedo-Rodríguez, R. Velasco-Ramos, R.G. Suárez-Sánchez, A.D. Rodríguez-Carrizalez, A.A. Oregón-Miranda, Efficacy of a fixed combination of 0.09 % xanthan gum/0.1 % chondroitin sulfate preservative free vs polyethylene glycol/propylene glycol in subjects with dry eye disease: A multicenter randomized controlled trial, BMC Ophthalmol. 16 (2016) 1–6. https://doi.org/10.1186/s12886-016-0343-9.

[80] S. Belalcázar-Rey, V. Sánchez Huerta, J.C. Ochoa-Tabares, S. Altamirano Vallejo, A. Soto-Gómez, R. Suárez-Velasco, F. García-Félix, L. Baiza-Durán, O. Olvera-Montaño, P. Muñoz-Villegas, Efficacy and safety of sodium hyaluronate/chondroitin sulfate preservative-free ophthalmic solution in the treatment of dry eye: A clinical trial, Curr. Eye Res. 46 (2021) 919–929. https://doi.org/10.1080/02713683.2020.1849733.

[81] J. Theodosakis, M.L. Grove, M. Cohen, A randomized, double blind, placebo controlled trial of a

topical cream containing glucosamine sulfate, chondroitin sulfate, and camphor for osteoarthritis of the knee, J. Rheumatol. 31 (2004) 826–827.

[82] R.F. Espíndola, E.F.S. Castro, M.R. Santhiago, N. Kara, A clinical comparison between DisCoVisc and 2% hydroxypropylmethylcellulose in phacoemulsification: A fellow eye study, Clinics. 67 (2012) 1059–1062. https://doi.org/10.6061/clinics/2012(09)13.

[83] G. Labiris, H. Sideroudi, K. Rousopoulos, V.P. Kozobolis, Cohesive versus dispersive-cohesive ophthalmic viscosurgical device in torsional intelligent phaco, J. Cataract Refract. Surg. 41 (2015) 681–682. https://doi.org/10.1016/j.jcrs.2015.01.010.

[84] M. Tutolo, E. Ammirati, G. Castagna, K. Klockaerts, H. Plancke, D. Ost, F. van der Aa, D. Ridder, A prospective randomized controlled multicentre trial comparing intravesical DMSO and chondroïtin sulphate 2% for painful bladder syndrome/interstitial cystitis, Int. Braz J Urol. 43 (2017) 134–141. https://doi.org/10.1590/S1677-5538.IBJU.2016.0302.

[85] O. Ciani, E. Arendsen, M. Romancik, R. Lunik, E. Costantini, M. Di Biase, G. Morgia, E. Fragalà, T. Roman, M. Bernat, G. Guazzoni, R. Tarricone, M. Lazzeri, Intravesical administration of combined hyaluronic acid (HA) and chondroitin sulfate (CS) for the treatment of female recurrent urinary tract infections: A European multicentre nested case-control study, BMJ Open. 6 (2016) 1–8. https://doi.org/10.1136/bmjopen-2015-009669.

[86] J. Yang, M. Shen, H. Wen, Y. Luo, R. Huang, L. Rong, J. Xie, Recent advance in delivery system and tissue engineering applications of chondroitin sulfate, Carbohydr. Polym. 230 (2020) 115650. https://doi.org/10.1016/j.carbpol.2019.115650.

[87] N. Volpi, Chondroitin sulphate for the treatment of osteoarthritis, Curr. Med. Chem. Anti-Inflamm. Anti-Allergy Agents. 4 (2005) 221–234. https://doi.org/10.2174/1568014054065276.

[88] X. Zhu, L. Sang, D. Wu, J. Rong, L. Jiang, Effectiveness and safety of glucosamine and chondroitin for the treatment of osteoarthritis: a meta-analysis of randomized controlled trials, J. Orthop. Surg. Res. 13 (2018) 1–9. https://doi.org/10.1186/s13018-018-0871-5.

[89] H.S. Vasiliadis, K. Tsikopoulos, Glucosamine and chondroitin for the treatment of osteoarthritis, World J. Orthop. 8 (2017) 1–11. https://doi.org/10.5312/wjo.v8.i1.1.

[90] S. Wandel, P. Jüni, B. Tendal, E. Nüesch, P.M. Villiger, N.J. Welton, S. Reichenbach, S. Trelle, Effects of glucosamine, chondroitin, or placebo in patients with osteoarthritis of hip or knee: network meta-analysis, BMJ. 341 (2010) 711. https://doi.org/10.1136/bmj.c4675.

[91] J.A. Singh, S. Noorbaloochi, R. Macdonald, L.J. Maxwell, Chondroitin for osteoarthritis, Cochrane Database Syst. Rev. 2017 (2015) 1–297. https://doi.org/10.1002/14651858.CD005614.pub2.

[92] M. Kloppenburg, F.P.B. Kroon, F.J. Blanco, M. Doherty, K.S. Dziedzic, E. Greibrokk, I.K. Haugen, G. Herrero-Beaumont, H. Jonsson, I. Kjeken, E. Maheu, R. Ramonda, M.J.P.F. Ritt, W. Smeets, J.S. Smolen, T.A. Stamm, Z. Szekanecz, R. Wittoek, L. Carmona, 2018 update of the EULAR recommendations for the management of hand osteoarthritis, Ann. Rheum. Dis. 78 (2019) 16–24. https://doi.org/10.1136/annrheumdis-2018-213826.

[93] C. Gabay, C. Medinger-Sadowski, D. Gascon, F. Kolo, A. Finckh, Symptomatic effects of chondroitin 4 and chondroitin 6 sulfate on hand osteoarthritis: a randomized, double-blind, placebo-controlled clinical trial at a single center, Arthritis Rheum. 63 (2011) 3383–3391. https://doi.org/10.1002/art.30574.

[94] I. Möller, M. Pérez, J. Monfort, P. Benito, J. Cuevas, C. Perna, G. Doménech, M. Herrero, E. Montell, J. Vergés, Effectiveness of chondroitin sulphate in patients with concomitant knee osteoarthritis and psoriasis: a randomized, double-blind, placebo-controlled study, Osteoarthr. Cartil. 18 (2010) S32–S40. https://doi.org/10.1016/j.joca.2010.01.018.

[95] L.M. Wildi, J.P. Raynauld, J. Martel-Pelletier, A. Beaulieu, L. Bessette, F. Morin, F. Abram, M. Dorais, J.P. Pelletier, Chondroitin sulphate reduces both cartilage volume loss and bone marrow lesions in knee osteoarthritis patients starting as early as 6 months after initiation of therapy: a randomised, double-blind, placebo-controlled pilot study using MRI, Ann. Rheum. Dis. 70 (2011) 982–989. https://doi.org/10.1136/ard.2010.140848.

[96] J. Martel-Pelletier, J.P. Raynauld, F. Mineau, F. Abram, P. Paiement, P. Delorme, J.P. Pelletier, Levels of serum biomarkers from a two-year multicentre trial are associated with treatment response on knee osteoarthritis cartilage loss as assessed by magnetic resonance imaging: an exploratory study, Arthritis Res. Ther. 19 (2017) 1–11. https://doi.org/10.1186/s13075-017-1377-y.

[97] J.P. Pelletier, J.P. Raynauld, A.D. Beaulieu, L. Bessette, F. Morin, A.J. de Brum-Fernandes, P. Delorme, M. Dorais, P. Paiement, F. Abram, J. Martel-Pelletier, Chondroitin sulfate efficacy versus celecoxib on knee osteoarthritis structural changes using magnetic resonance imaging: a 2-year multicentre exploratory study, Arthritis Res. Ther. 18 (2016) 1–12. https://doi.org/10.1186/s13075-016-1149-0.

[98] B. Mazières, M. Hucher, M. Zaïm, P. Garnero, Effect of chondroitin sulphate in symptomatic knee osteoarthritis: a multicentre, randomised, double-blind, placebo-controlled study, Ann. Rheum. Dis. 66 (2007) 639–645. https://doi.org/10.1136/ard.2006.059899.

[99] J.-J. Railhac, M. Zaim, A.-S. Saurel, J. Vial, B. Fournie, Effect of 12 months treatment with chondroitin sulfate on cartilage volume in knee osteoarthritis patients: a randomized, double-blind, placebo-controlled pilot study using MRI, Clin. Rheumatol. 31 (2012) 1347–1357. https://doi.org/10.1007/s10067-012-2022-4.

[100] P. Fardellone, M. Zaim, A.-S. Saurel, E. Maheu, Comparative efficacy and safety study of two chondroitin sulfate preparations from different origin (avian and bovine) in symptomatic osteoarthritis of the knee, Open Rheumatol. J. 7 (2013) 1–12. https://doi.org/10.2174/1874312901307010001.

[101] S.L. Navarro, E. White, E.D. Kantor, Y. Zhang, J. Rho, X. Song, G.L. Milne, P.D. Lampe, J.W. Lampe, Randomized trial of glucosamine and chondroitin supplementation on inflammation and oxidative stress biomarkers and plasma proteomics profiles in healthy humans, PLoS One. 10 (2015) 1–20. https://doi.org/10.1371/journal.pone.0117534.

[102] M.O. Price, O.J. Knight, B.A. Benetz, S.M. Debanne, D.D. Verdier, G.O. Rosenwasser, M.

Rosenwasser, F.W. Price, J.H. Lass, Randomized, prospective, single-masked clinical trial of endothelial keratoplasty performance with 2donor cornea 4°C storage solutions and associated chambers, Cornea. 34 (2015) 253–256. https://doi.org/10.1097/ICO.0000000000000354.

[103] N. Volpi, Anti-inflammatory activity of chondroitin sulphate: new functions from an old natural macromolecule, Inflammopharmacology. 19 (2011) 299–306. https://doi.org/10.1007/s10787-011-0098-0.

[104] M. Iovu, G. Dumais, P. du Souich, Anti-inflammatory activity of chondroitin sulfate, Osteoarthr. Cartil. 16 (2008) S14–S18. https://doi.org/10.1016/j.joca.2008.06.008.

[105] L. Borsig, L. Wang, M.C.M. Cavalcante, L. Cardilo-Reis, P.L. Ferreira, P.A.S. Mourão, J.D. Esko, M.S.G. Pavão, Selectin blocking activity of a fucosylated chondroitin sulfate glycosaminoglycan from sea cucumber: effect on tumor metastasis and neutrophil recruitment, J. Biol. Chem. 282 (2007) 14984–14991. https://doi.org/10.1074/jbc.M610560200.

[106] J. Li, E.M. Sparkenbaugh, G. Su, F. Zhang, Y. Xu, K. Xia, P. He, S. Baytas, S. Pechauer, A. Padmanabhan, R.J. Linhardt, R. Pawlinski, J. Liu, Enzymatic synthesis of chondroitin sulfate E to attenuate bacteria lipopolysaccharide-induced organ damage, ACS Cent. Sci. 6 (2020) 1199–1207. https://doi.org/10.1021/acscentsci.0c00712.

[107] S. Chen, G. Li, N. Wu, X. Guo, N. Liao, X. Ye, D. Liu, C. Xue, W. Chai, Sulfation pattern of the fucose branch is important for the anticoagulant and antithrombotic activities of fucosylated chondroitin sulfates, Biochim. Biophys. Acta - Gen. Subj. 1830 (2013) 3054–3066. https://doi.org/10.1016/j.bbagen.2013.01.001.

[108] R.A.A. Muzzarelli, F. Greco, A. Busilacchi, V. Sollazzo, A. Gigante, Chitosan, hyaluronan and chondroitin sulfate in tissue engineering for cartilage regeneration: a review, Carbohydr. Polym. 89 (2012) 723–739. https://doi.org/10.1016/j.carbpol.2012.04.057.

[109] S. Yamada, K. Sugahara, Potential therapeutic application of chondroitin sulfate/dermatan sulfate, Curr. Drug Discov. Technol. 5 (2008) 289–301. https://doi.org/10.2174/157016308786733564.

[110] T. Laabs, D. Carulli, H.M. Geller, J.W. Fawcett, Chondroitin sulfate proteoglycans in neural development and regeneration, Curr. Opin. Neurobiol. 15 (2005) 116–120. https://doi.org/10.1016/j.conb.2005.01.014.

[111] K. Bergefall, E. Trybala, M. Johansson, T. Uyama, S. Naito, S. Yamada, H. Kitagawa, K. Sugahara, T. Bergström, Chondroitin sulfate characterized by the E-disaccharide unit is a potent inhibitor of herpes simplex virus infectivity and provides the virus binding sites on gro2C cells, J. Biol. Chem. 280 (2005) 32193–32199. https://doi.org/10.1074/jbc.M503645200.

[112] D. Kato, S. Era, I. Watanabe, M. Arihara, N. Sugiura, K. Kimata, Y. Suzuki, K. Morita, K.I.P.J. Hidari, T. Suzuki, Antiviral activity of chondroitin sulphate E targeting dengue virus envelope protein, Antiviral Res. 88 (2010) 236–243. https://doi.org/10.1016/j.antiviral.2010.09.002.

[113] Y. Sano, Antiviral activity of chondroitin sulfate against infection, Carbohydr. Polym. 33 (1997) 125–129. https://doi.org/10.1016/S0144-8617(98)00119-2.

[114] A. Jinno-Oue, A. Tanaka, N. Shimizu, T. Mori, N. Sugiura, K. Kimata, H. Isomura, H. Hoshino, Inhibitory effect of chondroitin sulfate type E on the binding step of human T-cell leukemia virus type 1, AIDS Res. Hum. Retroviruses. 29 (2013) 621–629. https://doi.org/10.1089/aid.2012.0156.

[115] N. Huang, M.Y. Wu, C.B. Zheng, L. Zhu, J.H. Zhao, Y.T. Zheng, The depolymerized fucosylated chondroitin sulfate from sea cucumber potently inhibits HIV replication via interfering with virus entry, Carbohydr. Res. 380 (2013) 64–69. https://doi.org/10.1016/j.carres.2013.07.010.

[116] S. Song, H. Peng, Q. Wang, Z. Liu, X. Dong, C. Wen, C. Ai, Y. Zhang, Z. Wang, B. Zhu, Inhibitory activities of marine sulfated polysaccharides against SARS-CoV-2, Food Funct. 11 (2020) 7415–7420. https://doi.org/10.1039/d0fo02017f.

[117] L. Zhang, K. Li, W. Xiao, L. Zheng, Y. Xiao, H. Fan, X. Zhang, Preparation of collagen-chondroitin sulfate-hyaluronic acid hybrid hydrogel scaffolds and cell compatibility *in vitro*, Carbohydr. Polym. 84 (2011) 118–125. https://doi.org/10.1016/j.carbpol.2010.11.009.

[118] Y. Liu, H. Lv, L. Ren, G. Xue, Y. Wang, Improving the moisturizing properties of collagen film by surface grafting of chondroitin sulfate for corneal tissue engineering, J. Biomater. Sci. Polym. Ed. 27 (2016) 758–772. https://doi.org/10.1080/09205063.2016.1160561.

[119] S. Saivin, J.-P. Cambus, C. Thalamas, G. Lau, B. Boneu, G. Houin, F. Gianese, Pharmacokinetics and pharmacodynamics of intramuscular dermatan sulfate revisited, Clin. Drug Investig. 23 (2003) 533–543. https://doi.org/10.2165/00044011-200323080-00006.

[120] S. Coccheri, F. Mannello, Development and use of sulodexide in vascular diseases: implications for treatment, Drug Des. Devel. Ther. 8 (2013) 49–65. https://doi.org/10.2147/DDDT.S6762.

[121] A.B.V. Lomonte, J.A. Mendonça, G. de Castro Brandão, M.L. Castro, Multicenter, randomized, double-blind clinical trial to evaluate efficacy and safety of combined glucosamine sulfate and chondroitin sulfate capsules for treating knee osteoarthritis, Adv. Rheumatol. (London, England). 58 (2018). https://doi.org/10.1186/s42358-018-0041-9.

[122] M.C. Hochberg, J. Martel-Pelletier, J. Monfort, I. Möller, J.R. Castillo, N. Arden, F. Berenbaum, F.J. Blanco, P.G. Conaghan, G. Doménech, Y. Henrotin, T. Pap, P. Richette, A. Sawitzke, P. Du Souich, J.P. Pelletier, Combined chondroitin sulfate and glucosamine for painful knee osteoarthritis: a multicentre, randomised, double-blind, non-inferiority trial versus celecoxib, Ann. Rheum. Dis. 75 (2016) 37–44. https://doi.org/10.1136/annrheumdis-2014-206792.

[123] J.A. Roman-Blas, S. Castañeda, O. Sánchez-Pernaute, R. Largo, G. Herrero-Beaumont, Combined treatment with chondroitin sulfate and glucosamine sulfate shows no superiority over placebo for reduction of joint pain and functional impairment in patients with knee osteoarthritis: a six-month multicenter, randomized, double-blind, placebo-controlled clinical trial, Arthritis Rheumatol. 69 (2017) 77–85. https://doi.org/10.1002/art.39819.

[124] A.B.V. Lomonte, E. Gimenez, A.C. da Silva, S.C. Radominski, M.A. Scheinberg, A.C. Ximenes, C.A. de Freitas Zerbini, Treatment of knee osteoarthritis with a new formulation of a fixed-dose combination of glucosamine sulfate and bovine chondroitin: a multicenter, randomized, single-

blind, non-inferiority clinical trial, Adv. Rheumatol. 61 (2021). https://doi.org/10.1186/s42358-021-00165-9.

[125] T. Magrans-Courtney, C. Wilborn, C. Rasmussen, M. Ferreira, L. Greenwood, B. Campbell, C.M. Kerksick, E. Nassar, R. Li, M. Iosia, M. Cooke, K. Dugan, D. Willoughby, L.A. Soliah, R.B. Kreider, Effects of diet type and supplementation of glucosamine, chondroitin, and MSM on body composition, functional status, and markers of health in women with knee osteoarthritis initiating a resistance-based exercise and weight loss program, J. Int. Soc. Sports Nutr. 8 (2011) 1–17. https://doi.org/10.1186/1550-2783-8-8.

[126] M. Fransen, M. Agaliotis, L. Nairn, M. Votrubec, L. Bridgett, S. Su, S. Jan, L. March, J. Edmonds, R. Norton, M. Woodward, R. Day, C. Buckland-Wright, C. Bridges-Webb, S. Heritier, G. Jones, P. Sambrook, C. Speechly, E. Vignon, G. Jovanovic, S. Jeram, M. Dowd, M. Champion, S. Grealy, D. Watson, A. Vrataric, C. Johnson, R. White, P. O'Connell, F. Milwraith, D. Kennedy, C. Faulkes, O. Bulton, K. Johnson, Glucosamine and chondroitin for knee osteoarthritis: a double-blind randomised placebo-controlled clinical trial evaluating single and combination regimens, Ann. Rheum. Dis. 74 (2015) 851–858. https://doi.org/10.1136/annrheumdis-2013-203954.

[127] M. Simic, A.R. Harmer, M. Agaliotis, L. Nairn, L. Bridgett, L. March, M. Votrubec, J. Edmonds, M. Woodward, R. Day, M. Fransen, Clinical risk factors associated with radiographic osteoarthritis progression among people with knee pain: a longitudinal study, Arthritis Res. Ther. 23 (2021) 1–10. https://doi.org/10.1186/s13075-021-02540-9.

[128] N.C. Dang, A. Ardehali, B.A. Bruckner, P.E. Parrino, D.L. Gillen, R.W. Hoffman, R. Spotnitz, S. Cavoores, I.J. Shorn, R.J. Manson, W.D. Spotnitz, Prospective, multicenter, randomized, controlled trial evaluating the performance of a novel combination powder vs hemostatic matrix in cardiothoracic operations, J. Card. Surg. 35 (2020) 313–319. https://doi.org/10.1111/jocs.14376.

[129] S.J. Wang, Y.H. Wang, L.C. Huang, The effect of oral low molecular weight liquid hyaluronic acid combination with glucosamine and chondroitin on knee osteoarthritis patients with mild knee pain: An 8-week randomized double-blind placebo-controlled trial, Medicine (Baltimore). 100 (2021) e24252. https://doi.org/10.1097/md.0000000000024252.

[130] E. López-Vidriero, R. Olivé-Vilas, D. López-Capapé, L. Varela-Sende, R. López-Vidriero, L. Til-Pérez, Efficacy and tolerability of progen, a nutritional supplement based on innovative plasma proteins, in ACL reconstruction: a multicenter randomized controlled trial, Orthop. J. Sport. Med. 7 (2019) 232596711982723. https://doi.org/10.1177/2325967119827237.

[131] G. Gallo, M. Mistrangelo, R. Passera, V. Testa, M. Pozzo, R. Perinotti, I. Lanati, I. Lazzari, P. Tonello, E. Ugliono, E. De Luca, A.R. Luc, G. Clerico, M. Trompetto, Efficacy of mesoglycan in pain control after excisional hemorrhoidectomy: a pilot comparative prospective multicenter study, Gastroenterol. Res. Pract. 2018 (2018) 1–8. https://doi.org/10.1155/2018/6423895.

[132] G. Palareti, C. Legnani, E. Antonucci, S. Zorzi, A.A. Bignamini, C. Lodigiani, A. Tosetto, L. Bertù, V. Pengo, S. Testa, W. Ageno, D. Prisco, P. Prandoni, D. Poli, Design and rationale of a randomized, placebo-controlled trial on the efficacy and safety of sulodexide for extended

treatment in elderly patients after a first venous thromboembolism, Intern. Emerg. Med. 16 (2021) 359–368. https://doi.org/10.1007/s11739-020-02381-5.

[133] J.H. Song, H.S. Chin, O.W. Kwon, S.J. Lim, H.K. Kim, Effect of sulodexide in patients with non-proliferative diabetic retinopathy: diabetic retinopathy sulodexide study (DRESS), Graefe's Arch. Clin. Exp. Ophthalmol. 253 (2015) 829–837. https://doi.org/10.1007/s00417-014-2746-8.

[134] A. Bidadkosh, S.P.H. Lambooy, H.J. Heerspink, M.J. Pena, R.H. Henning, H. Buikema, L.E. Deelman, Predictive properties of biomarkers GDF-15, NTproBNP, and hs-TnT for morbidity and mortality in patients with type 2 diabetes with nephropathy, Diabetes Care. 40 (2017) 784–792. https://doi.org/10.2337/dc16-2175.

[135] E.J. Lewis, J.B. Lewis, T. Greene, L.G. Hunsicker, T. Berl, M.A. Pohl, D. de Zeeuw, H.L. Heerspink, R.D. Rohde, R.C. Atkins, A.T. Reutens, D.K. Packham, I. Raz, Sulodexide for kidney protection in type 2 diabetes patients with microalbuminuria: a randomized controlled trial, Am. J. Kidney Dis. 58 (2011) 729–736. https://doi.org/10.1053/j.ajkd.2011.06.020.

[136] Y. Ji, S. Zhang, M. Qiao, R. Jiao, J. Li, P. Song, X. Zhang, H. Huang, Synthesis of structurally defined chondroitin sulfate: paving the way to the structure-activity relationship studies, Carbohydr. Polym. 248 (2020). https://doi.org/10.1016/j.carbpol.2020.116796.

[137] P.L. DeAngelis, J. Liu, R.J. Linhardt, Chemoenzymatic synthesis of glycosaminoglycans: Re-creating, re-modeling and re-designing nature's longest or most complex carbohydrate chains, Glycobiology. 23 (2013) 764–777. https://doi.org/10.1093/glycob/cwt016.

[138] E. Petit, C. Delattre, D. Papy-Garcia, P. Michaud, Chondroitin sulfate lyases: applications in analysis and glycobiology, in: Adv. Pharmacol., Elsevier, 2006: pp. 167–186. https://doi.org/10.1016/S1054-3589(05)53008-4.

[139] Y. Li, Z. Zhou, Z. Chen, High-level production of ChSase ABC I by co-expressing molecular chaperones in *Escherichia coli*, Int. J. Biol. Macromol. 119 (2018) 779–784. https://doi.org/10.1016/j.ijbiomac.2018.08.009.

[140] R.J. Linhardt, F.Y. Avci, T. Toida, Y.S. Kim, M. Cygler, CS lyases: structure, activity, and applications in analysis and the treatment of diseases, Adv. Pharmacol. 53 (2006) 187–215. https://doi.org/10.1016/S1054-3589(05)53009-6.

[141] S. Yamada, Chondroitin sulfate-specific novel hydrolase in human, in: P.R. Sudhakaran, A. Surolia (Eds.), Adv. Exp. Med. Biol., Springer New York, 2012: pp. 47–56. https://doi.org/10.1007/978-1-4614-3381-1_4.

[142] J. Fu, Z. Jiang, J. Chang, B. Han, W. Liu, Y. Peng, Purification, characterization of chondroitinase ABC from *Sphingomonas paucimobilis* and *in vitro* cardiocytoprotection of the enzymatically degraded CS-A, Int. J. Biol. Macromol. 115 (2018) 737–745. https://doi.org/10.1016/j.ijbiomac.2018.04.117.

[143] A. Rani, A. Goyal, A new member of family 8 polysaccharide lyase chondroitin AC lyase (*Ps*PL8A) from *Pedobacter saltans* displays endo- and exo-lytic catalysis, J. Mol. Catal. B Enzym. 134 (2016) 215–224. https://doi.org/10.1016/j.molcatb.2016.11.001.

[144] L. Gushulak, R. Hemming, D. Martin, V. Seyrantepe, A. Pshezhetsky, B. Triggs-Raine, Hyaluronidase 1 and β-hexosaminidase have redundant functions in hyaluronan and chondroitin sulfate degradation, J. Biol. Chem. 287 (2012) 16689–16697. https://doi.org/10.1074/jbc.M112.350447.

[145] T. Kaneiwa, A. Miyazaki, R. Kogawa, S. Mizumoto, K. Sugahara, S. Yamada, Identification of amino acid residues required for the substrate specificity of human and mouse chondroitin sulfate hydrolase (Conventional hyaluronidase-4), J. Biol. Chem. 287 (2012) 42119–42128. https://doi.org/10.1074/jbc.M112.360693.

[146] S. Kobayashi, S.I. Fujikawa, M. Ohmae, Enzymatic synthesis of chondroitin and its derivatives catalyzed by hyaluronidase, J. Am. Chem. Soc. 125 (2003) 14357–14369. https://doi.org/10.1021/ja036584x.

[147] S.I. Fujikawa, M. Ohmae, S. Kobayashi, Enzymatic synthesis of chondroitin 4-sulfate with well-defined structure, Biomacromolecules. 6 (2005) 2935–2942. https://doi.org/10.1021/bm050364p.

[148] I. Kakizaki, S. Suto, Y. Tatara, T. Nakamura, M. Endo, Hyaluronan-chondroitin hybrid oligosaccharides as new life science research tools, Biochem. Biophys. Res. Commun. 423 (2012) 344–349. https://doi.org/10.1016/j.bbrc.2012.05.127.

[149] Y. Shi, Y. Meng, J. Li, J. Chen, Y. Liu, X. Bai, Chondroitin sulfate: extraction, purification, microbial and chemical synthesis, J. Chem. Technol. Biotechnol. 89 (2014) 1445–1465. https://doi.org/10.1002/jctb.4454.

[150] P.L. DeAngelis, A.J. Padgett-McCue, Identification and molecular cloning of a chondroitin synthase from *Pasteurella multocida* type F, J. Biol. Chem. 275 (2000) 24124–24129. https://doi.org/10.1074/jbc.M003385200.

[151] P.L. DeAngelis, Chondroitin synthase gene and methods of making and using same, US Patent 0164984 A1, 2005.

[152] N. Sugiura, S. Shimokata, T. Minamisawa, J. Hirabayashi, K. Kimata, H. Watanabe, Sequential synthesis of chondroitin oligosaccharides by immobilized chondroitin polymerase mutants, Glycoconj. J. 25 (2008) 521–530. https://doi.org/10.1007/s10719-008-9105-0.

[153] N. Sugiura, T. Shioiri, M. Chiba, T. Sato, H. Narimatsu, K. Kimata, H. Watanabe, Construction of a chondroitin sulfate library with defined structures and analysis of molecular interactions, J. Biol. Chem. 287 (2012) 43390–43400. https://doi.org/10.1074/jbc.M112.412676.

[154] N. Sugiura, T.M. Clausen, T. Shioiri, T. Gustavsson, H. Watanabe, A. Salanti, Molecular dissection of placental malaria protein VAR2CSA interaction with a chemo-enzymatically synthesized chondroitin sulfate library, Glycoconj. J. 33 (2016) 985–994. https://doi.org/10.1007/s10719-016-9685-z.

[155] T. Shioiri, J. Tsuchimoto, H. Watanabe, N. Sugiura, Sequence determination of synthesized chondroitin sulfate dodecasaccharides, Glycobiology. 26 (2016) 592–606. https://doi.org/10.1093/glycob/cww008.

[156] J. Li, G. Su, J. Liu, Enzymatic synthesis of homogeneous chondroitin sulfate oligosaccharides, Angew. Chemie - Int. Ed. 56 (2017) 11784–11787. https://doi.org/10.1002/anie.201705638.

[157] A. Yusa, K. Kitajima, O. Habuchi, *N*-linked oligosaccharides on chondroitin 6-sulfotransferase-1 are required for production of the active enzyme, Golgi localization, and sulfotransferase activity toward keratan sulfate, J. Biol. Chem. 281 (2006) 20393–20403. https://doi.org/10.1074/jbc.M600140200.

[158] Z. Zhou, Q. Li, H. Huang, H. Wang, Y. Wang, G. Du, J. Chen, Z. Kang, A microbial–enzymatic strategy for producing chondroitin sulfate glycosaminoglycans, Biotechnol. Bioeng. 115 (2018) 1561–1570. https://doi.org/10.1002/bit.26577.

[159] W. He, Y. Zhu, A. Shirke, X. Sun, J. Liu, R.A. Gross, M.A.G. Koffas, R.J. Linhardt, M. Li, Expression of chondroitin-4-*O*-sulfotransferase in *Escherichia coli* and *Pichia pastoris*, Appl. Microbiol. Biotechnol. 101 (2017) 6919–6928. https://doi.org/10.1007/s00253-017-8411-5.

[160] A. Badri, A. Williams, A. Awofiranye, P. Datta, K. Xia, W. He, K. Fraser, J.S. Dordick, R.J. Linhardt, M.A.G. Koffas, Complete biosynthesis of a sulfated chondroitin in *Escherichia coli*, Nat. Commun. 12 (2021). https://doi.org/10.1038/s41467-021-21692-5.

[161] O. Habuchi, R. Moroi, S. Ohtake, Enzymatic synthesis of chondroitin sulfate E by *N*-acetylgalactosamine 4-sulfate 6-*O*-sulfotransferase purified from squid cartilage, Anal. Biochem. 310 (2002) 129–136. https://doi.org/10.1016/S0003-2697(02)00277-4.

[162] Y. Wang, S. Li, X. Xu, Y. Tan, X. Liu, J. Fang, Chemoenzymatic synthesis of homogeneous chondroitin polymers and its derivatives, Carbohydr. Polym. 232 (2020) 115822. https://doi.org/10.1016/j.carbpol.2019.115822.

[163] R. Benito-Arenas, E. Doncel-Pérez, M. Fernández-Gutiérrez, L. Garrido, E. García-Junceda, J. Revuelta, A. Bastida, A. Fernández-Mayoralas, A holistic approach to unravelling chondroitin sulfation: correlations between surface charge, structure and binding to growth factors, Carbohydr. Polym. 202 (2018) 211–218. https://doi.org/10.1016/j.carbpol.2018.08.120.

[164] C. An, L. Zhao, Z. Wei, X. Zhou, Chemoenzymatic synthesis of 3′-phosphoadenosine-5′-phosphosulfate coupling with an ATP regeneration system, Appl. Microbiol. Biotechnol. 101 (2017) 7535–7544. https://doi.org/10.1007/s00253-017-8511-2.

[165] P. Datta, L. Fu, W. He, M.A.G. Koffas, J.S. Dordick, R.J. Linhardt, Expression of enzymes for 3′-phosphoadenosine-5′-phosphosulfate (PAPS) biosynthesis and their preparation for PAPS synthesis and regeneration, Appl. Microbiol. Biotechnol. 104 (2020) 7067–7078. https://doi.org/10.1007/s00253-020-10709-6.

[166] A. Badri, A. Williams, K. Xia, R.J. Linhardt, M.A.G. Koffas, Increased 3′-phosphoadenosine-5′-phosphosulfate levels in engineered *Escherichia coli* cell lysate facilitate the *in vitro* synthesis of chondroitin sulfate A, Biotechnol. J. 14 (2019) 1800436. https://doi.org/10.1002/biot.201800436.

[167] X. Zhang, L. Lin, H. Huang, R.J. Linhardt, Chemoenzymatic synthesis of glycosaminoglycans, Acc. Chem. Res. 53 (2020) 335–346. https://doi.org/10.1021/acs.accounts.9b00420.

[168] L. Lin, M. Qiao, X. Zhang, R.J. Linhardt, Site-selective reactions for the synthesis of glycoconjugates in polysaccharide vaccine development, Carbohydr. Polym. 230 (2020) 115643. https://doi.org/10.1016/j.carbpol.2019.115643.

[169] C. Schiraldi, D. Cimini, M. De Rosa, Production of chondroitin sulfate and chondroitin, Appl. Microbiol. Biotechnol. 87 (2010) 1209–1220. https://doi.org/10.1007/s00253-010-2677-1.

[170] L.M. Willis, C. Whitfield, Structure, biosynthesis, and function of bacterial capsular polysaccharides synthesized by ABC transporter-dependent pathways, Carbohydr. Res. 378 (2013) 35–44. https://doi.org/10.1016/j.carres.2013.05.007.

[171] C. Whitfield, S.S. Wear, C. Sande, Assembly of bacterial capsular polysaccharides and exopolysaccharides, Annu. Rev. Microbiol. 74 (2020) 521–543. https://doi.org/10.1146/annurev-micro-011420-075607.

[172] L. Guan, Y. Xue, W. Ding, Z. Zhao, Biosynthesis and regulation mechanisms of the *Pasteurella multocida* capsule, Res. Vet. Sci. 127 (2019) 82–90. https://doi.org/10.1016/j.rvsc.2019.10.011.

[173] J.D. de Oliveira, L.S. Carvalho, A.M.V. Gomes, L.R. Queiroz, B.S. Magalhães, N.S. Parachin, Genetic basis for hyper production of hyaluronic acid in natural and engineered microorganisms, Microb. Cell Fact. 15 (2016) 1–19. https://doi.org/10.1186/s12934-016-0517-4.

[174] A. Zakeri, M.J. Rasaee, N. Pourzardosht, Enhanced hyluronic acid production in *Streptococcus zooepidemicus* by over expressing HasA and molecular weight control with Niscin and glucose, Biotechnol. Reports. 16 (2017) 65–70. https://doi.org/10.1016/j.btre.2017.02.007.

[175] X. Chu, J. Han, D. Guo, Z. Fu, W. Liu, Y. Tao, Characterization of UDP-glucose dehydrogenase from *Pasteurella multocida* CVCC 408 and its application in hyaluronic acid biosynthesis, Enzyme Microb. Technol. 85 (2016) 64–70. https://doi.org/10.1016/j.enzmictec.2015.12.009.

[176] D. Cimini, I. Dello Iacono, E. Carlino, R. Finamore, O.F. Restaino, P. Diana, E. Bedini, C. Schiraldi, Engineering *S. equi* subsp. *zooepidemicus* towards concurrent production of hyaluronic acid and chondroitin biopolymers of biomedical interest, AMB Express. 7 (2017). https://doi.org/10.1186/s13568-017-0364-7.

[177] C. Schiraldi, A. La, M. De, Biotechnological production and application of hyaluronan, in: Biopolymers, Sciyo, 2010. https://doi.org/10.5772/10271.

[178] Z. Wang, M. Ly, F. Zhang, W. Zhong, A. Suen, A.M. Hickey, J.S. Dordick, R.J. Linhardt, *E. coli* K5 fermentation and the preparation of heparosan, a bioengineered heparin precursor, Biotechnol. Bioeng. 107 (2010) 964–973. https://doi.org/10.1002/bit.22898.

[179] Z. Wang, J.S. Dordick, R.J. Linhardt, *Escherichia coli* K5 heparosan fermentation and improvement by genetic engineering, Bioeng. Bugs. 2 (2011) 63–67. https://doi.org/10.4161/bbug.2.1.14201.

[180] P. Jin, L. Zhang, P. Yuan, Z. Kang, G. Du, J. Chen, Efficient biosynthesis of polysaccharides chondroitin and heparosan by metabolically engineered *Bacillus subtilis*, Carbohydr. Polym. 140 (2016) 424–432. https://doi.org/10.1016/j.carbpol.2015.12.065.

[181] Z. Wang, J. Li, S. Cheong, U. Bhaskar, O. Akihiro, F. Zhang, J.S. Dordick, R.J. Linhardt, Response surface optimization of the heparosan *N*-deacetylation in producing bioengineered heparin, J. Biotechnol. 156 (2011) 188–196. https://doi.org/10.1016/j.jbiotec.2011.08.013.

[182] A.A.E. Chavaroche, L.A.M. van den Broek, G. Eggink, Production methods for heparosan, a precursor of heparin and heparan sulfate, Carbohydr. Polym. 93 (2013) 38–47. https://doi.org/10.1016/j.carbpol.2012.04.046.

[183] S. Peterson, A. Frick, J. Liu, Design of biologically active heparan sulfate and heparin using an enzyme-based approach, Nat. Prod. Rep. 26 (2009) 610–627. https://doi.org/10.1039/b803795g.

[184] U. Bhaskar, G. Li, L. Fu, A. Onishi, M. Suflita, J.S. Dordick, R.J. Linhardt, Combinatorial one-pot chemoenzymatic synthesis of heparin, Carbohydr. Polym. 122 (2015) 399–407. https://doi.org/10.1016/j.carbpol.2014.10.054.

[185] O.F. Restaino, I. di Lauro, R. Di Nuzzo, M. De Rosa, C. Schiraldi, New insight into chondroitin and heparosan-like capsular polysaccharide synthesis by profiling of the nucleotide sugar precursors, Biosci. Rep. 37 (2017) 1–12. https://doi.org/10.1042/BSR20160548.

[186] T.-T. Wang, C.-Y. Zhu, S. Zheng, C.-C. Meng, T.-T. Wang, D.-H. Meng, Y.-J. Li, H.-M. Zhu, F.-S. Wang, J.-Z. Sheng, Identification and characterization of a chondroitin synthase from *Avibacterium paragallinarum*, Appl. Microbiol. Biotechnol. 102 (2018) 4785–4797. https://doi.org/10.1007/s00253-018-8926-4.

[187] Q. Zhang, R. Yao, X. Chen, L. Liu, S. Xu, J. Chen, J. Wu, Enhancing fructosylated chondroitin production in *Escherichia coli* K4 by balancing the UDP-precursors, Metab. Eng. 47 (2018) 314–322. https://doi.org/10.1016/j.ymben.2018.04.006.

[188] E. Bedini, C. Decastro, M. Derosa, A. Dinola, A. Iadonisi, O.F. Restaino, C. Schiraldi, M. Parrilli, A microbiological-chemical strategy to produce chondroitin sulfate A,C, Angew. Chemie - Int. Ed. 50 (2011) 6160–6163. https://doi.org/10.1002/anie.201101142.

[189] J.A. Vázquez, I. Rodríguez-Amado, M.I. Montemayor, J. Fraguas, M. del P. González, M.A. Murado, Chondroitin sulfate, hyaluronic acid and chitin/chitosan production using marine waste sources: characteristics, applications and eco-friendly processes: a review, Mar. Drugs. 11 (2013) 747–774. https://doi.org/10.3390/md11030747.

[190] S. D'ambrosio, A. Alfano, E. Cassese, O.F. Restaino, S. Barbuto Ferraiuolo, R. Finamore, M. Cammarota, C. Schiraldi, D. Cimini, Production and purification of higher molecular weight chondroitin by metabolically engineered *Escherichia coli* K4 strains, Sci. Rep. 10 (2020) 1–10. https://doi.org/10.1038/s41598-020-70027-9.

[191] A. Trilli, I. Busiello, S. Daly, F. Bagatin, Biotechnological production of chondroitin, WO Patent 2012/004063 A1, 2012.

[192] M. De Rosa, C. Schiraldi, D. Cimini, Biotechnological production of chondroitin, US Patent 8,592,186 B2, 2013.

[193] D. Cimini, R. Russo, S. D'Ambrosio, I. Dello Iacono, C. Rega, E. Carlino, O. Argenzio, L. Russo, B.

D'Abrosca, A. Chambery, C. Schiraldi, Physiological characterization and quantitative proteomic analyses of metabolically engineered *E. coli* K4 strains with improved pathways for capsular polysaccharide biosynthesis, Biotechnol. Bioeng. 115 (2018) 1801–1814. https://doi.org/10.1002/bit.26597.

[194] Q. Wu, A. Yang, W. Zou, Z. Duan, J. Liu, J. Chen, L. Liu, Transcriptional engineering of *Escherichia coli* K4 for fructosylated chondroitin production, Biotechnol. Prog. 29 (2013) 1140–1149. https://doi.org/10.1002/btpr.1777.

[195] S. D'ambrosio, M. Ventrone, A. Alfano, C. Schiraldi, D. Cimini, Microbioreactor (micro-Matrix) potential in aerobic and anaerobic conditions with different industrially relevant microbial strains, Biotechnol. Prog. (2021) 1–7. https://doi.org/10.1002/btpr.3184.

[196] A.E. Awofiranye, S.N. Baytas, K. Xia, A. Badri, W. He, A. Varki, M. Koffas, R.J. Linhardt, *N*-glycolyl chondroitin synthesis using metabolically engineered *E. coli*, AMB Express. 10 (2020) 1–17. https://doi.org/10.1186/s13568-020-01084-6.

[197] K. Suzuki, K. Miyamoto, H. Kaseyama, Chondroitin-producing bacterium and method of producing chondroitin, US Patent 8,283,145 B2, 2012.

[198] W. He, L. Fu, G. Li, J. Andrew Jones, R.J. Linhardt, M. Koffas, Production of chondroitin in metabolically engineered *E. coli*, Metab. Eng. 27 (2015) 92–100. https://doi.org/10.1016/j.ymben.2014.11.003.

[199] P.L. DeAngelis, Methods of expressing gram-negative glycosaminoglycan synthase genes in gram-positive hosts, US Patent 7,642,071 B2, 2010.

[200] F. Cheng, S. Luozhong, H. Yu, Z. Guo, Biosynthesis of chondroitin in engineered *Corynebacterium glutamicum*, J. Microbiol. Biotechnol. 29 (2019) 392–400. https://doi.org/10.4014/jmb.1810.10062.

[201] B. Priem, J. Peroux, P. Colin-Morel, S. Drouillard, S. Fort, Chemo-bacterial synthesis of conjugatable glycosaminoglycans, Carbohydr. Polym. 167 (2017) 123–128. https://doi.org/10.1016/j.carbpol.2017.03.026.

[202] A.S. Erenler, Capsular polysaccharide biosynthesis from recombinant *E. coli* and chondroitin sülfate production, Cell. Mol. Biol. 65 (2019) 17–21. https://doi.org/10.14715/cmb/2019.65.6.4.

[203] X.-X. Wei, G.-Q. Chen, Applications of the VHb gene *vgb* for improved microbial fermentation processes, in: Methods Enzymol., 2008: pp. 273–287. https://doi.org/10.1016/S0076-6879(08)36015-7.

[204] X. Jin, W. Zhang, Y. Wang, J. Sheng, R. Xu, J. Li, G. Du, Z. Kang, Biosynthesis of non-animal chondroitin sulfate from methanol using genetically engineered *Pichia pastoris*, Green Chem. 23 (2021) 4365–4374. https://doi.org/10.1039/d1gc00260k.

[205] D. Cimini, S. Fantaccione, F. Volpe, M. De Rosa, O.F. Restaino, G. Aquino, C. Schiraldi, IS2-mediated overexpression of *kfoC* in *E. coli* K4 increases chondroitin-like capsular polysaccharide production, Appl. Microbiol. Biotechnol. 98 (2014) 3955–3964.

https://doi.org/10.1007/s00253-014-5506-0.

[206] J. Wu, Q. Zhang, L. Liu, X. Chen, J. Liu, Q. Luo, Recombinant *Escherichia coli* for high efficiency production of fructosylated chondroitin and method for making thereof, US Patent 10,508,279 B2, 2019.

[207] A. Zanfardino, O.F. Restaino, E. Notomista, D. Cimini, C. Schiraldi, M. De Rosa, M. De Felice, M. Varcamonti, Isolation of an *Escherichia coli* K4 *kfoC* mutant over-producing capsular chondroitin, Microb. Cell Fact. 9 (2010) 1–8. https://doi.org/10.1186/1475-2859-9-34.

[208] M.P. Stevens, B.R. Clarke, I.S. Roberts, Regulation of the *Escherichia coli* K5 capsule gene cluster by transcription antitermination, Mol. Microbiol. 24 (1997) 1001–1012. https://doi.org/10.1046/j.1365-2958.1997.4241780.x.

[209] S.S. Badle, G. Jayaraman, K.B. Ramachandran, Ratio of intracellular precursors concentration and their flux influences hyaluronic acid molecular weight in *Streptococcus zooepidemicus* and recombinant *Lactococcus lactis*, Bioresour. Technol. 163 (2014) 222–227. https://doi.org/10.1016/j.biortech.2014.04.027.

[210] F. Levander, M. Svensson, P. Rådström, Enhanced exopolysaccharide production by metabolic engineering of *Streptococcus thermophilus*, Appl. Environ. Microbiol. 68 (2002) 784–790. https://doi.org/10.1128/AEM.68.2.784-790.2002.

[211] Z. Mao, H.D. Shin, R.R. Chen, Engineering the *E. coli* UDP-glucose synthesis pathway for oligosaccharide synthesis, Biotechnol. Prog. 22 (2006) 369–374. https://doi.org/10.1021/bp0503181.

[212] H. Wang, Y. Yang, L. Lin, W. Zhou, M. Liu, K. Cheng, W. Wang, Engineering *Saccharomyces cerevisiae* with the deletion of endogenous glucosidases for the production of flavonoid glucosides, Microb. Cell Fact. 15 (2016) 1–12. https://doi.org/10.1186/s12934-016-0535-2.

[213] B. Widner, R. Behr, S. Von Dollen, M. Tang, T. Heu, A. Sloma, D. Sternberg, P.L. DeAngelis, P.H. Weigel, S. Brown, Hyaluronic acid production in *Bacillus subtilis*, Appl. Environ. Microbiol. 71 (2005) 3747–3752. https://doi.org/10.1128/AEM.71.7.3747-3752.2005.

[214] H. Yu, G. Stephanopoulos, Metabolic engineering of *Escherichia coli* for biosynthesis of hyaluronic acid, Metab. Eng. 10 (2008) 24–32. https://doi.org/10.1016/j.ymben.2007.09.001.

[215] Y. Wu, T. Chen, Y. Liu, X. Lv, J. Li, G. Du, R. Ledesma-Amaro, L. Liu, CRISPRi allows optimal temporal control of *N*-acetylglucosamine bioproduction by a dynamic coordination of glucose and xylose metabolism in *Bacillus subtilis*, Metab. Eng. 49 (2018) 232–241. https://doi.org/10.1016/j.ymben.2018.08.012.

[216] A.W. Westbrook, X. Ren, J. Oh, M. Moo-Young, C.P. Chou, Metabolic engineering to enhance heterologous production of hyaluronic acid in *Bacillus subtilis*, Metab. Eng. 47 (2018) 401–413. https://doi.org/10.1016/j.ymben.2018.04.016.

[217] J.E. Woo, H.J. Seong, S.Y. Lee, Y.S. Jang, Metabolic engineering of *Escherichia coli* for the production of hyaluronic acid from glucose and galactose, Front. Bioeng. Biotechnol. 7 (2019) 1–9. https://doi.org/10.3389/fbioe.2019.00351.

[218] T. Niu, Y. Liu, J. Li, M. Koffas, G. Du, H.S. Alper, L. Liu, Engineering a glucosamine-6-phosphate responsive *glmS* ribozyme switch enables dynamic control of metabolic flux in *Bacillus subtilis* for overproduction of *N*-acetylglucosamine, ACS Synth. Biol. 7 (2018) 2423–2435. https://doi.org/10.1021/acssynbio.8b00196.

[219] F. Cheng, H. Yu, G. Stephanopoulos, Engineering *Corynebacterium glutamicum* for high-titer biosynthesis of hyaluronic acid, Metab. Eng. 55 (2019) 276–289. https://doi.org/10.1016/j.ymben.2019.07.003.

[220] J. Rainha, J.L. Rodrigues, L.R. Rodrigues, CRISPR-Cas9: a powerful tool to efficiently engineer *Saccharomyces cerevisiae*, Life. 11 (2021) 1–16. https://doi.org/10.3390/life11010013.

[221] A. Williams, R.J. Linhardt, M.A.G. Koffas, Metabolic engineering of capsular polysaccharides, Emerg. Top. Life Sci. 2 (2018) 337–348. https://doi.org/10.1042/ETLS20180003.

[222] J.Y. Baik, L. Gasimli, B. Yang, P. Datta, F. Zhang, C.A. Glass, J.D. Esko, R.J. Linhardt, S.T. Sharfstein, Metabolic engineering of Chinese hamster ovary cells: towards a bioengineered heparin, Metab. Eng. 14 (2012) 81–90. https://doi.org/10.1016/j.ymben.2012.01.008.

[223] O.F. Restaino, S. D'ambrosio, E. Cassese, S.B. Ferraiuolo, A. Alfano, R. Ventriglia, A. Marrazzo, C. Schiraldi, D. Cimini, Molecular weight determination of heparosan- and chondroitin-like capsular polysaccharides: figuring out differences between wild -type and engineered *Escherichia coli* strains, Appl. Microbiol. Biotechnol. 103 (2019) 6771–6782. https://doi.org/10.1007/s00253-019-09969-8.

[224] O.F. Restaino, D. Cimini, M. De Rosa, A. Catapano, M. De Rosa, C. Schiraldi, High cell density cultivation of *Escherichia coli* K4 in a microfiltration bioreactor: a step towards improvement of chondroitin precursor production, Microb. Cell Fact. 10 (2011) 1–10. https://doi.org/10.1186/1475-2859-10-10.

[225] I.R. Amado, J.A. Vázquez, L. Pastrana, J.A. Teixeira, Cheese whey: a cost-effective alternative for hyaluronic acid production by *Streptococcus zooepidemicus*, Food Chem. 198 (2016) 54–61. https://doi.org/10.1016/j.foodchem.2015.11.062.

[226] I.R. Amado, J.A. Vázquez, L. Pastrana, J.A. Teixeira, Microbial production of hyaluronic acid from agro-industrial by-products: molasses and corn steep liquor, Biochem. Eng. J. 117 (2017) 181–187. https://doi.org/10.1016/j.bej.2016.09.017.

[227] N.C. Pan, H.C.B. Pereira, M.L.C. da Silva, A.F.D. Vasconcelos, M.A.P.C. Celligoi, Improvement production of hyaluronic acid by *Streptococcus zooepidemicus* in sugarcane molasses, Appl. Biochem. Biotechnol. 182 (2017) 276–293. https://doi.org/10.1007/s12010-016-2326-y.

CHAPTER 3

CLONING, EXPRESSION, AND CHARACTERIZATION OF UDP-GLUCOSE DEHYDROGENASES

Abstract

Uridine diphosphate-glucose dehydrogenase (UGD) is an enzyme that produces uridine diphosphate-glucuronic acid (UDP-GlcA), which is an intermediate in glycosaminoglycans (GAGs) production pathways. GAGs are generally extracted from animal tissues. Efforts to produce GAGs in a safer way have been conducted by constructing artificial biosynthetic pathways in heterologous microbial hosts. This work characterizes novel enzymes with potential for UDP-GlcA biotechnological production. The UGD enzymes from *Zymomonas mobilis* (*Zm*UGD) and from *Lactobacillus johnsonii* (*Lbj*UGD) were expressed in *Escherichia coli*. These two enzymes and an additional eukaryotic one from *Capra hircus* (*Ch*UGD) were also expressed in *Saccharomyces cerevisiae* strains. The three enzymes herein studied represent different UGD phylogenetic groups. The UGD activity was evaluated through UDP-GlcA quantification *in vivo* and after *in vitro* reactions. Engineered *E. coli* strains expressing *Zm*UGD and *Lbj*UGD were able to produce *in vivo* 28.4 μM and 14.9 μM UDP-GlcA, respectively. Using *S. cerevisiae* as the expression host, the highest *in vivo* UDP-GlcA production was obtained for the strain CEN.PK2-1C expressing *Zm*UGD (17.9 μM) or *Ch*UGD (14.6 μM). Regarding the *in vitro* assays, under the optimal conditions, *E. coli* cell extract containing *Lbj*UGD was able to produce about 1800 μM, while *Zm*UGD produced 407 μM UDP-GlcA, after 1 h of reaction. Using engineered yeasts, the *in vitro* production of UDP-GlcA reached a maximum of 533 μM using *S. cerevisiae* CEN.PK2-1C_pSP-GM_*Lbj*UGD cell extract. The UGD enzymes were active in both prokaryotic and eukaryotic hosts, therefore the genes and expression chassis herein used can be valuable alternatives for further industrial applications.

Keywords: UDP-glucose dehydrogenase; UDP-glucuronic acid; glycosaminoglycans biosynthesis; heterologous production; *Escherichia coli*; *Saccharomyces cerevisiae*

This chapter was partially published on the following publication:

M.R. Couto, J.L. Rodrigues, L.R. Rodrigues, Cloning, expression and characterization of UDP-glucose dehydrogenases, Life. 11 (2021) 1201. https://doi.org/10.3390/life11111201.

3.1. Introduction

Uridine diphosphate-glucose dehydrogenase (UDP-glucose dehydrogenase, UGD, EC 1.1.1.22) is an oxidoreductase that catalyzes the production of UDP-glucuronic acid (UDP-GlcA) from UDP-glucose (UDP-Glc). UDP-GlcA is a compound with diverse applications whose major biotechnological use is as a precursor for glycosaminoglycans (GAGs) production [1].

Glycosaminoglycans (GAGs) are a class of natural long linear polysaccharides that include hyaluronic acid, heparan sulfate, and chondroitin sulfate which have medical, veterinary, pharmaceutical, and cosmetic applications. Nevertheless, except for hyaluronic acid that is produced by microbial fermentation, GAGs are mainly extracted from animal sources using laborious methods and hazardous reagents, generating products with heterogeneous compositions, raising concerns of adulteration and risks of prion contamination, which ultimately limits their large-scale production [2]. Additionally, religious, and ethical issues alongside with vegetarianism trends are key factors driving the search for non-animal sources of GAGs. An eco-friendly and controllable alternative consists of the metabolic engineering of microbial cell factories to produce GAGs from simple carbon sources [3–5].

Several pathogenic microorganisms have been reported as GAG-like polysaccharides producers through well-known pathways (Figure 3.1). For example, group A and C streptococci and *Pasteurella multocida* type A produce hyaluronic acid [6,7]; *Escherichia coli* K5, *P. multocida* type D and *Avibacterium paragallinarum* genotype II can synthesize a capsular polysaccharide, heparosan, structurally analogous to unsulfated heparin [8–10]; *E. coli* O5:K4:H4 synthesizes a capsular polysaccharide composed by unsulfated fructosylated chondroitin [8]; and *P. multocida* type F and *A. paragallinarium* genotype I produce unsulfated chondroitin [9,10]. However, the biotechnological production of GAGs at an industrial-scale, and especially of sulfated GAGs, is limited, due to low yields and safety issues related to the cultivation of pathogenic microorganisms [11]. Therefore, in the past few years research efforts have been focused on the improvement of the production process, as well as on the design of better heterologous hosts to produce GAGs. Recent advances in the microbial production of GAGs using non-pathogenic hosts (mostly prokaryotic organisms) have been reported [12–19].

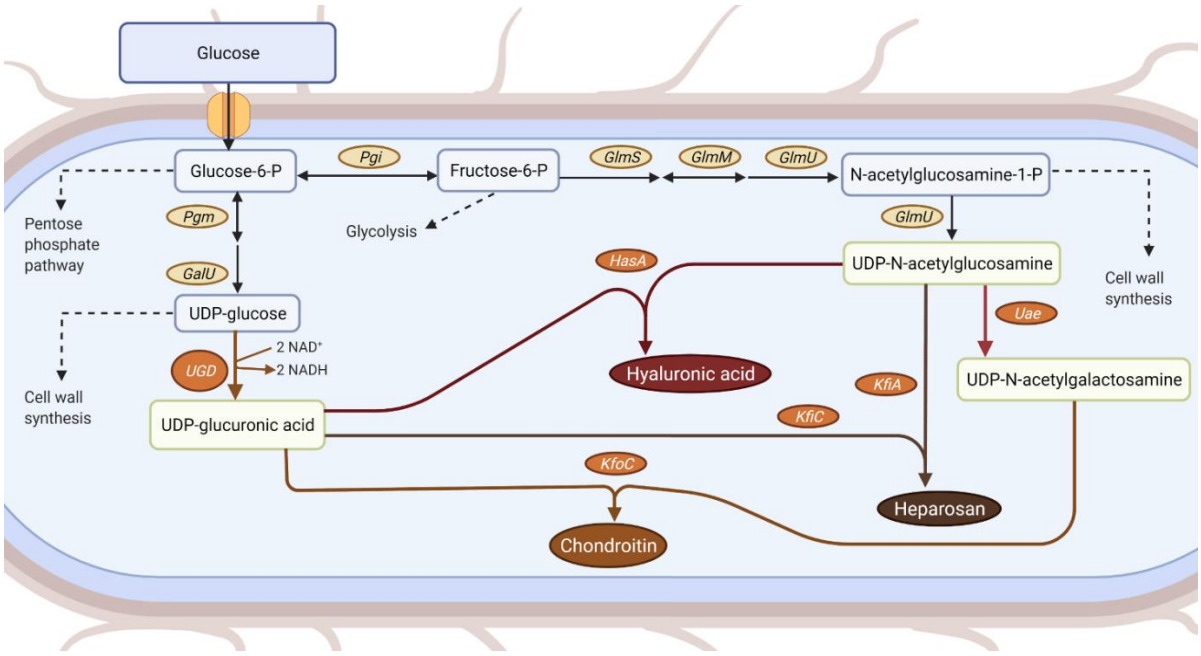


Figure 3.1. Production of uridine diphosphate (UDP)-glucuronic acid in *Escherichia coli* and its possible use in the biosynthesis of microbial chondroitin, hyaluronic acid, or heparosan. Depending on the microbial host, the heterologous expression of the enzymes shown in orange might be required for glycosaminoglycans production. Enzyme abbreviations: GalU, uridine-triphosphate:glucose-1-phosphate uridylyltransferase; GlmM, phosphoglucosamine mutase; GlmS, glucosamine-6-phosphate synthase; GlmU, glucosamine-1-phosphate N-acetyltransferase/N-acetylglucosamine-1-phosphate uridyltransferase; HasA, hyaluronan synthase; KfiA, β-1,3-glucuronyltransferase; KfiC, α-1,4-*N*-acetylglucosaminyltransferase; KfoC, chondroitin synthase; Pgi, glucose-6-phosphate isomerase; Pgm, phosphoglucomutase; Uae, UDP-*N*-acetylglucosamine 4-epimerase; UGD, UDP-glucose 6-dehydrogenase.

The heterologous pathway design should consider every enzymatic step to maximize GAGs yields. The step catalyzed by UGD was previously considered to be the limiting factor of GAG biosynthesis in homologous and heterologous organisms [13,16,20]. Since the first discovered UGD [21], genes encoding UGD have been identified in almost all organisms, from which alternative genes can be extracted for further evaluation as biocatalytic producers of UDP-GlcA. However, although UGD genes are widespread across nature, few UGD have been biochemically characterized or evaluated in heterologous hosts. The overexpression of UGD enzyme in a heterologous host is needed even in hosts containing the gene as it is often over-regulated limiting its expression [22]. Additionally, no eukaryotic heterologous host has yet been evaluated to produce some GAGs. Eukaryotic hosts can be a favorable alternative for eukaryotic gene expression offering an interesting platform for complex integrated production and sulfation systems.

In this study, the expression of three novel genes encoding UGD was evaluated to expand the knowledge of potential candidates for the construction of GAGs biosynthetic pathways. Two prokaryotic UGDs from *Zymomonas mobilis* (*Zm*UGD) and *Lactobacillus johnsonii* (*Lbj*UGD) were expressed in *E. coli* BL21 (DE3). In addition, a eukaryotic system was also used for heterologous expression of these UGD enzymes to assess if this chassis could be a proper platform for GAGs production. *Saccharomyces cerevisiae* CEN.PK2-1C and *S. cerevisiae* BY4741 were used to express *Zm*UGD and *Lbj*UGD and an additional UGD from *Capra hircus* (*Ch*UGD) that has never been evaluated. These enzymes are representatives of each of the three evolutionary groups identified by an evolutionary analysis, namely from group I and II of prokaryotic UGD and from the eukaryotic group. The expressed UGD enzymes were characterized for the first time in a prokaryotic and eukaryotic host through *in vivo* and *in vitro* reactions by quantification of UDP-GlcA production. These enzymes were found to have a high catalytic efficiency and to be suitable for use in the construction of biosynthetic pathways for GAGs production. In addition, *S. cerevisiae* was demonstrated to be a suitable alternative platform for UDP-GlcA production.

3.2. Materials and Methods

3.2.1. Multiple Sequence Alignment and Evolutionary Relationships

The evolutionary relationship of the three UGD enzymes explored in this work, *Ch*UGD, *Zm*UGD, and *Lbj*UGD (GenBank sequences XP_013820045.1 WP_012817132.1, ABM21411.1, respectively), was accessed. Clustal Omega (1.2.4) tool of European Molecular Biology Laboratory—European Bioinformatics Institute (EMBL-EBI, https://www.ebi.ac.uk/Tools/msa/clustalo/) was used to perform a multiple amino acid sequences alignment of the three enzymes and their domains were predicted using the HMMER web service (EMBL-EBI, https://www.ebi.ac.uk/Tools/hmmer/search/phmmer). Using *Zm*UGD amino acid sequence as query, a sequence similarity search was performed using BLASTp tool at National Center for Biotechnology Information (NCBI, https://blast.ncbi.nlm.nih.gov/), against the reviewed proteins in the SwissProt/UniProt database. The identified homologous protein sequences and 5 additional relevant protein sequences from known GAG producers (a total of 63 representative enzymes) were analyzed through multiple sequence alignment (ClustalW algorithm) and evolutionary analysis in MEGA-X [23]. The phylogenetic tree was constructed in MEGA-X using the Neighbor-Joining method [24]. The Poisson correction method was used to compute evolutionary distances [25] which are in the units of the number of amino acid substitutions per site. All ambiguous positions were removed for each sequence pair (pairwise deletion option).

3.2.2. Strains and Plasmids

The strains and plasmids used in this study are listed in Table 3.1. *Z. mobilis* ATCC 29191 and *L. johnsonii* ATCC 11506 were obtained from the American Type Culture Collection (ATCC, Manassas, USA) and used as genomic DNA (gDNA) source for amplification of UGD genes. *E. coli* NZY5α (NZYTech, Lisbon, Portugal) was used for cloning procedures and vector propagation. *E. coli* BL21 (DE3) (NZYTech) was used as expression host of *Zm*UGD or *Lbj*UGD genes using pRSFDuet-1 (Novagen, Madison, USA) as the expression vector. *S. cerevisiae* CEN.PK2-1C and *S. cerevisiae* BY4741 strains were obtained from Euroscarf (Oberursel, Germany) and used for UGD expression. In this case, two different shuttle vectors were also tested, namely pSP-GM1 (PGK1 promoter, Addgene, Watertown, USA) and p426GPD (ATCC).

Table 3.1. Strains and plasmids used in this study.

Strains	**Relevant Genotype**	**Source**
Zymomonas mobilis subsp. *mobilis* ATCC 29191	Wild type	ATCC 29191
Lactobacillus johnsonii ATCC 11506	Wild type	ATCC 11506
Escherichia coli NZY5α	*fhuA2* Δ(*argF lacZ*)U169 *phoA glnV44*Φ80 Δ(*lacZ*)M15 *gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	NZYTech (MB00401)
E. coli BL21 (DE3)	F- *ompT gal dcm lon hsd*SB(rB- mB-) λ(DE3 * *lacI lacUV5-T7 gene 1 ind1 sam7 nin5*)	NZYTech (MB006)
Saccharomyces cerevisiae CEN.PK2-1C	*MATa ura3-52 his3Δ 1 leu2-3,112 trp1-289 MAL2-8c SUC2*	Euroscarf 30000A [26]
S. cerevisiae BY4741	*MATa his3Δ 1 leu2Δ 0 met15Δ 0 ura3Δ 0*	Euroscarf Y00000 [27]
Plasmids	**Description**	**Source**
pGEM®-T Easy	ColE1/pMB1/pBR322/pUC *ori*, *lacZ*, Amp^R	Promega
pRSFDuet-1	RSF1030 *ori*, *lacI*, double P_{T7lac}, Kan^R	Novagen
pGEM_*Lbj*UGD	pGEM®-T Easy harboring *L. johnsonii* UGD (*Lbj*UGD) gene	This study
pRSFDuet_*Zm*UGD	pRSFDuet-1 carrying *Z. mobilis* UGD (*Zm*UGD) gene	This study
pRSFDuet_*Lbj*UGD	pRSFDuet-1 carrying *Lbj*UGD gene	This study
pUC57_*Ch*UGD	pMB1 *ori*, Amp^R; pUC57 carrying *Capra hircus* UGD (*Ch*UGD) gene codon-optimized for *S. cerevisiae*	NZYTech
pSP-GM1	pUC *ori*, Amp^R, 2 μ *ori*, *URA3* $P_{TEF1}P_{PGK1}$	Addgene #64739 [28]
p426GPD	pUC *ori*, Amp^R, 2 μ *ori*, *URA3* P_{GPD}	ATCC 87361
pSP-GM1_*Zm*UGD	pSP-GM1 carrying *Zm*UGD gene	This study
p426GPD_*Zm*UGD	p426GPD carrying *Zm*UGD gene	This study
pSP-GM1_*Lbj*UGD	pSP-GM1 carrying *Lbj*UGD gene	This study
p426GPD_*Lbj*UGD	p426GPD carrying *Lbj*UGD gene	This study
pSP-GM1_*Ch*UGD	pSP-GM1 carrying *Ch*UGD gene	This study
p426GPD_*Ch*UGD	p426GPD carrying *Ch*UGD gene	This study

3.2.3. Gene Sources and Cloning Strategy

Table A 3.1 compiles the primers used for gene cloning. UGD genes from *Z. mobilis* and *L. johnsonii* (GenBank accession numbers CP003704.1:c550193-548880 and EF138834.1:15898-17145, respectively) were amplified from gDNA. Phusion High Fidelity DNA Polymerase (Thermo Fisher Scientific, Wilmington, United States) was used for all gene amplifications unless otherwise stated. Amplification of the *Zm*UGD gene from the *Z. mobilis* genome was achieved directly through colony Polymerase Chain Reaction (PCR) using the primers Zm_Fw and Zm_Rv. The gDNA from the Gram-positive *L. johnsonii* was extracted from a 2 mL culture using Wizard® Genomic DNA Purification Kit (Promega, Madison, USA) according to the manufacturer instructions with slight modifications as follows: the pellet was treated only with 10 mg/mL of lysozyme prior nuclei lysis; in the end of the procedure, the DNA was rehydrated with 50 µL of DNA Rehydration Buffer and stored at 4 °C. This DNA was used as template for *Lbj*UGD gene amplification with primers Lbj_Fw and Lbj_Rv. Since the amplification yield was low, the cleaned PCR product was further A-tailored to be cloned in Promega's pGEM®-T Easy Vector and the final construction was used as amplification template. A-tailored product was achieved by adding 5 units of Dream *Taq* DNA Polymerase (Thermo Scientific) to the PCR product and 0.2 mM dNTPs in a final volume of 10 µL and incubating the mixture at 70 °C for 30 min. Then, 1 µL of this reaction was added to a ligation reaction with Promega's pGEM®-T Easy Vector according to the kit instructions originating pGEM_*Lbj*UGD. UGD gene from *C. hircus* (GenBank accession number XM_018049379.1) was codon-optimized for *S. cerevisiae* and synthesized by NZYTech (Table A 3.2). Afterwards, it was amplified from pUC57_*Ch*UGD using the Ch_Fw and Ch_Rv primers. After amplification, *Zm*UGD*, Lbj*UGD*, and Ch*UGD genes were cloned in p426GPD and pSP-GM1. To clone *Zm*UGD and *Lbj*UGD genes in pRSFDuet-1, the genes were amplified from p426GPD_*Zm*UGD and pGEM_*Lbj*UGD using the Zm_prsf_FW/Zm_prsf_Rv and Lbj_prsf_Fw/Lbj_prsf_Rv primer pairs, respectively.

All genes were cloned in frame with a N-terminal 6 x Histidine tag originally present in pRSFDuet-1 or in the primers used to clone the genes in pSP-GM1 and p426GPD.

Plasmid DNA was extracted using NucleoSpin® Plasmid Miniprep Kit (Macherey-Nagel, Düren, Germany). Amplified DNA fragments were purified from agarose using NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel). Plasmid DNA and PCR products were quantified using a NanoDrop instrument (ND-1000, Thermo Fisher Scientific) and were digested with the proper restriction endonucleases (Thermo Fisher Scientific) for 1 h at 37 °C and purified using NucleoSpin® Gel and PCR Clean-up Kit. Ligations were performed for 1 h at room temperature using a T4 DNA ligase (Thermo Fisher Scientific). The constructions were transformed by heat shock into *E. coli* NZY5α competent cells (NZYTech). Super

optimal broth with catabolite repression (SOC; NZYTech) was used for transformants recovery. In the case of pGEM_*Lbj*UGD construction, positive transformants were selected using the blue-white screening method. White colonies were picked from lysogeny broth (LB) (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl; NZYTech) agar plates (20 g/L agar, JMGS, Odivelas, Portugal) containing ampicillin (0.1 mg/mL, NZYTech), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal, 0.1 mg/mL; NZYTech) and isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.5 mM; NZYTech). All plasmids herein constructed were verified by colony PCR using Dream *Taq* polymerase and digestion and their sequences were further confirmed by sequencing (GATC Biotech, Konstanz, Germany). After sequence confirmation, *E. coli* BL21 (DE3) competent cells (NZYTech) were transformed with the constructed plasmids pRSFDuet_*Zm*UGD, pRSFDuet_*Lbj*UGD and empty pRSFDuet-1 (used as negative control). Transformations of *S. cerevisiae* were performed by lithium acetate/single-stranded carrier DNA/polyethylene glycol method [29]. Lithium acetate, salmon sperm DNA and polyethylene glycol (PEG-3350) were obtained from Sigma-Aldrich. Selection of transformants was performed in selective minimal medium Yeast Nitrogen Base (YNB) with ammonium sulfate without amino acids (Sigma-Aldrich, St. Louis, USA) with required amino acids in the absence of uracil.

3.2.4. Culture Conditions

Z. mobilis was grown at 30 °C in solid medium with glucose 20 g/L (Acros Organics, New Jersey, USA), yeast extract 5 g/L (Panreac AppliChem, Darmstadt, Germany) and agar 15 g/L. *L. johnsonii* was grown in de Man, Rogosa and Sharpe (MRS) medium (Himedia, Mumbai, India) at 37 °C and 200 rpm.

E. coli cultures were incubated at 37 °C and 200 rpm in LB, or on LB agar plates. Ampicillin at a final concentration of 100 μg/mL (NZYTech) or kanamycin at 50 μg/mL (Fisher Scientific, New Jersey, USA) were supplemented when necessary. For UGD enzyme expression and consequent UDP-GlcA production, engineered *E. coli* BL21 (DE3) were grown by inoculating 1% (v/v) of an overnight pre-culture in 50 mL LB medium supplemented with kanamycin in 250 mL Erlenmeyer flasks (initial optical density at 600 nm (OD_{600nm}) = 0.1), which were then shaken at 37 °C and 200 rpm. *E. coli* transformed with the empty vector, pRSFDuet-1, was also grown and used as negative control. When an OD_{600nm} of 0.6 was reached, IPTG was added at a final concentration of 1 mM. Cells were then grown at 30 °C and 200 rpm for additional 3 h. The absorbance was measured in a 96-well plate spectrophotometric reader Synergy HT (BioTek, Winooski, USA). Dry weight of cells was determined through a calibration curve with $OD_{600\ nm}$ and expressed as g/L.

Wild-type *S. cerevisiae* growth was performed at 30 °C, 200 rpm, in Yeast Extract Peptone Dextrose (YPD, 20 g/L bacteriological peptone (HiMedia), 10 g/L yeast extract, 20 g/L glucose) or in plates with same composition supplemented with agar. After transformation, cells were grown in YNB plates supplemented with 20 g/L glucose and the required amino acids depending on the strain (tryptophan, histidine, methionine (Panreac AppliChem) and leucine (Fisher Scientific) and, in the case of wild-type strains, the pyrimidine uracil (Sigma)), at final concentrations of 100 mg/L. A single colony was picked from the plate and grown 24 h in 8 mL YNB for pre-culture. Afterwards, 50 mL of the selective minimal medium YNB supplemented with the required amino acids in 250 mL flasks were inoculated to an initial $OD_{600\ nm}$ of 0.1. Cells were grown at 30 °C and 200 rpm for 24–30 h.

3.2.5. Preparation of *E. coli* and *S. cerevisiae* Protein Extracts

In the end of the fermentation, the final culture of *E. coli* cells (~50 mL) was centrifuged (5,000 × *g* 15 min) and the pellet resuspended in 10 mL of Tris–HCl buffer (10 mM, pH 7.8). Cells were then lysed by sonication with a microtip probe linked to Vibra-cell processor (Sonics, Newtown, USA). Keeping the solution on ice during the procedure, short pulses of 3 s ON and 9 s OFF at 30% amplitude were performed until 5 min of active sonication was reached. The resulting lysate was centrifuged (12000 × *g* 15 min). The soluble fraction was collected and filtered with 0.2 µm NY filter and the insoluble phase was saved separately after resuspending the pellet in 1 mL of 10 mM Tris–HCl buffer. The filtered soluble samples were analyzed to quantify *in vivo* UDP-GlcA production through ultra-high-performance liquid chromatography (UHPLC) and were further used to perform *in vitro* reactions.

Cells from *S. cerevisiae* grown cultures (~50 mL) were harvested by centrifugation (5000 × *g* 15 min). For each 0.1 g of wet cells, 0.2 g of glass beads (425–600 µm, Sigma-Aldrich) were added to the cells pellet, as well as 2 mL of Tris–HCl (10 mM, pH 7.8) and 1x protease inhibitor cocktail (NZYTech). FastPrep-24 (MP Biomedicals, Salon, United States) was then used during 5 cycles of 1 min at 6-6.5 m/s interspersed with 1 min cooling on ice. Lysed samples were centrifuged (16000 × g 15 min) and the supernatant was filtered with 0.2 µm filter. Soluble and insoluble fractions were kept separately. The soluble fraction samples were analyzed using UHPLC to quantify *in vivo* UDP-GlcA production and were used for *in vitro* reactions.

Protein concentration in the soluble and insoluble fractions was measured using Pierce™ Coomassie (Bradford) Protein Assay Kit according to the manufacturers' instructions.

The UGD expression of *E. coli* and *S. cerevisiae* was evaluated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE) [30] (4% stacking gel and 12% running gel). Samples were mixed with

2x Laemmli Sample Buffer (65.8 mM Tris–HCl pH 6.8, 2.1% SDS, 26.3% glycerol, 0.01% bromophenol blue, from Fisher Scientific, JMGS, and Sigma-Aldrich, respectively) with β-mercaptoethanol (AppliChem) and were denatured at 95 °C for 5 min. The protein marker used was Color Protein Standard—Broad Range (NEB, #77125) or NZYColour Protein Marker II (NZYTech). After electrophoresis, the gel was stained using Coomassie Blue R-250 (AppliChem) for 15 min and de-stained using distilled water.

3.2.6. Purification of UGD

Purification of engineered 6 x His-tagged UGD from *E. coli* cell extracts was performed through affinity chromatography using nickel NTA agarose resin (ABT, Doral, USA) in 2 mL Pierce™ centrifuge columns (Thermo Scientific) according to ABT's instructions. The first elution was performed with 250 mM imidazole (Merck, Kenilworth, USA) followed by a second and third elution with 500 mM imidazole. Purified protein was quantified again with the Bradford Protein Assay Kit and analyzed by SDS–PAGE.

3.2.7. UGD *in vitro* Assays

UGD enzymatic activity was evaluated through *in vitro* reactions using the different cell extracts as enzyme source. The *in vitro* reaction was set by mixing 150 µL of reaction mixture with 50 µL of cell extracts from *E. coli* or *S. cerevisiae* up to a final concentration of 0.25 g/L protein. The reaction mixture was composed by 2.5 mM UDP-Glc (Carbosynth, Compton, Berkshire, United Kingdom), 0.5 mM nicotinamide adenine dinucleotide (NAD^+) (Panreac AppliChem), 1x protease inhibitor and 50 mM Tris–HCl pH 8.6, final concentrations in the 200 µL total volume. Immediately after adding the reaction mixture to the cell extract, the microplate was incubated at 30 °C for 60 min during which NADH (reduced form of NAD^+) formation was monitored in a microplate reader. A calibration curve was performed measuring the absorbance at 340 nm of different NADH (Panreac AppliChem) concentrations (0–750 µM) in a UV-transparent 96-well polystyrene microplate (Sigma-Aldrich). To evaluate the optimal pH of recombinant UGDs, cell extracts from UGD-producing recombinants were incubated in reaction solutions with different pHs (from 3.5 to 9.1) for 1 h at 30 °C. To evaluate the optimal temperature, cell extracts were reacted at different temperatures (from 15 to 40 °C) with a pre-heated reaction mixture of pH 8.0 during 1 h. To determine the kinetic parameters of the UGD enzymes, the cell extract at a final protein concentration of 0.5 g/L was incubated with a pre-heated reaction mixture (0.5 mM NAD^+,1x protease inhibitor and 50 mM Tris–HCl pH 8.0, at 30 °C) with variable concentrations of UDP-Glc (1–5 mM). Samples were taken at several timepoints during the incubation (0–15 min). The kinetic parameters K_m and V_{max} were calculated

by fitting the curve through nonlinear regression to the Michaelis-Menten model using GraphPad Prism 8.

3.2.8. Analysis and Quantification of UDP-sugars

UHPLC analysis was performed to quantify UDP-sugars in bacteria and yeasts' cell extracts (*in vivo* production) or reacted supernatants from *in vitro* assays. Analysis was performed using the Shimadzu Nexera-X2 system (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 Luna® Omega 3 µm Polar C18 (150 mm × 4.6 mm, Phenomenex, Torrance, USA). The column was equilibrated at 25 °C in 20 mM tetraethylammonium acetate (pH 6.0) during 8 min followed by a 22 min gradient from 0 to 5% of acetonitrile. The injection volume was 5 µL and the working flow rate was 0.5 mL/min. UDP-sugars were detected at UV_{260nm} absorbance. Retention times of UDP-Glc and UDP-GlcA were 23.8 min and 27.9 min, respectively.

Calibration curves based on UDP-Glc and UDP-GlcA standards (Carbosynth), in concentrations ranging from 50 to 2000 µM, were performed to quantify UDP-sugars in the samples.

3.2.9. Statistical Analysis

GraphPad Prism was used for all statistical analysis and the statistical significance of the differences between results was based on 2-way ANOVA with multiple comparisons (Tukey's or Durnett's multiple comparisons test).

3.3. Results and Discussion

3.3.1. Sequences and Phylogenetic Analysis

Despite the wide range of biological sources known to express UGD, there are few studies exploring their cloning or biochemical, structural, and functional characterization. UniProt entries for this enzyme include 45,705 unreviewed and only 44 reviewed sequences (data retrieved on 4th October 2021). The enzymes used in this study, *Zm*UGD, *Lbj*UGD and *Ch*UGD, are part of the unreviewed group. None of them has been expressed in heterologous hosts or studied for biotechnological applications. Expressing these UGD representatives in both prokaryotic and eukaryotic hosts can give new insights e.g., into the gene selection

for biotechnological production of GAGs. Moreover, the selection of *Ch*UGD as a eukaryotic representative enzyme was because the highest reported specific activity of UGD was obtained using an isolated UGD from caprine liver [31].

UGD belongs to the UDP-glucose/guanosine diphosphate (GDP)-mannose 6-dehydrogenase (UGD/GMD) superfamily. These enzymes have a wide range of functions, often involved in the biosynthesis of polysaccharides and are frequently part of operons dedicated to that purpose. The sequences of the three enzymes herein studied were aligned (Figure A 3.1) and their identity was calculated. Percent identity matrix showed that *Lbj*UGD shares 25% identity with *Ch*UGD and 30% identity with *Zm*UGD, while *Ch*UGD and *Zm*UGD exhibit 37% identify. Afterwards, *Zm*UGD sequence was screened against the proteins in the SwissProt database using BLASTp which demonstrated that this enzyme shares high levels of identity with other enzymes in the UGD/GMD superfamily. The sequences of the proteins identified as similar to *Zm*UGD, putative UGDs used in this study and UGDs from well-known GAG producers were analyzed through multiple sequence alignment. The resulting 63 representative amino acid sequences included 29 bacterial UGD, 1 archaeal UGD, 19 eukaryotic UGD, 5 GMD, 7 UAMAD, 1 UAGAD, and 1 NAGAAD. The multiple sequence analysis allowed to identify highly conserved amino acids in UGD/GMD superfamily and specifically within UGDs (highlighted residues in Figure A 3.1).

Although the amino acid sequences can be very variable within species, all the secondary structural features were found to be conserved across the UGD/GMD superfamily sequences, with the largest sequence diversity occurring close to the C-terminus. The three UGD enzymes cloned in this work presented, as predicted, the described conserved domains, namely the NAD(+)-binding domain, the central domain, and UDP-sugar binding domain for the N-terminal, central, and C-terminal regions, respectively (Figure A 3.1). Regarding the specific active site of UGD, the multiple sequences alignment confirmed that all sequences of the enzymes evaluated have the conserved residues required for UGD activity.

Afterwards, the sequences of each of the three studied enzymes were compared with NCBI sequences using BLASTp. These BLASTp results are described in Table A 3.3. The BLASTp analysis showed that *Zm*UGD is significantly similar to the UGD from *Sinorhizobium meliloti*, followed by the UGDs from *Bacillus subtilis* subsp. *subtilis* (YwqF and TuaD), and UGD from *Pseudomonas aeruginosa*. Interestingly, these UGDs are from known polysaccharide producers. Although *Z. mobilis* itself is a recognized levan producer [32], *S. melitoti* is known to produce succinoglycan [33], *B. subtilis* is also able to synthesize levan, teichuronic acid and other exopolymeric substances [34,35], and *P. aeruginosa* produces alginate [36]. Remarkably, *Zm*UGD exhibited a higher identity with archaeal UGD from *Haloferax volcanii*, eukaryotic

UGDs (such as from *Drosophila melanogaster, Homo sapiens, Pongo abelii, Bos taurus and Mus musculus*), and even with GMDs, than with other prokaryotic UGDs such as the heparosan producer *E. coli* K5. The probiotic *L. johnsonii* produces the polysaccharide inulin [37,38], and its UGD showed the highest identity with UGD from *E. coli* O111:H-. This strain, as well as the ones with the following greater identity are also known producers of capsular polysaccharides, including GAGs, namely *E. coli* O8:K40, *E. coli* CFT073/ O6:K2:H1, *E. coli* O157:H7, *E. coli* K-12, *Salmonella enterica* and *Streptococcus pyogenes* [39–42]. Expectably, *Ch*UGD was highly identical to UGDs from other mammals (identity from 99.8 to 96% with UGDs from *B. taurus*, *H. sapiens*, *P. abelli*, *M. muscullus and R. novergicus*). The following enzymes with similar sequences are UGDs from plants (59.8 to 61.8% identity with *Glycine max*, *Arabidopsis thaliana* and *Oryza sativa*) followed by the group of enzymes most similar to *Zm*UGD.

Finally, a phylogenetic tree was built to show the evolutionary distances between the 63 aligned amino acid sequences (Figure 3.2). Five main groups of enzymes emerged, namely eukaryotic UGD, prokaryotic UGD—Group I, prokaryotic UGD—Group II, prokaryotic GMD and other prokaryotic members of UGD/GMD superfamily. The enzymes used in this study, indicated in Figure 3.2 with black arrows, are representative of three different phylogenetic groups of UGDs. *Lbj*UGD belongs to the group II prokaryotic UGDs, *Zm*UGD belongs to group I prokaryotic UGDs, while *Ch*UGD belongs to eukaryotic UGDs group. Granja et al. [43] previously reported the divergence between the two groups of prokaryotic UGDs. Enzymes belonging to the so-called group I are more related to eukaryotic UGDs than to the group II prokaryotic UGDs. This analysis gives insights on which enzymes exhibit functional similarities, demonstrating that *Lbj*UGD is the most similar to microbial GAG producers, *Zm*UGD is included in groups of other types of polysaccharides producing organisms, while *Ch*UGD is included in the animal GAGs producers.

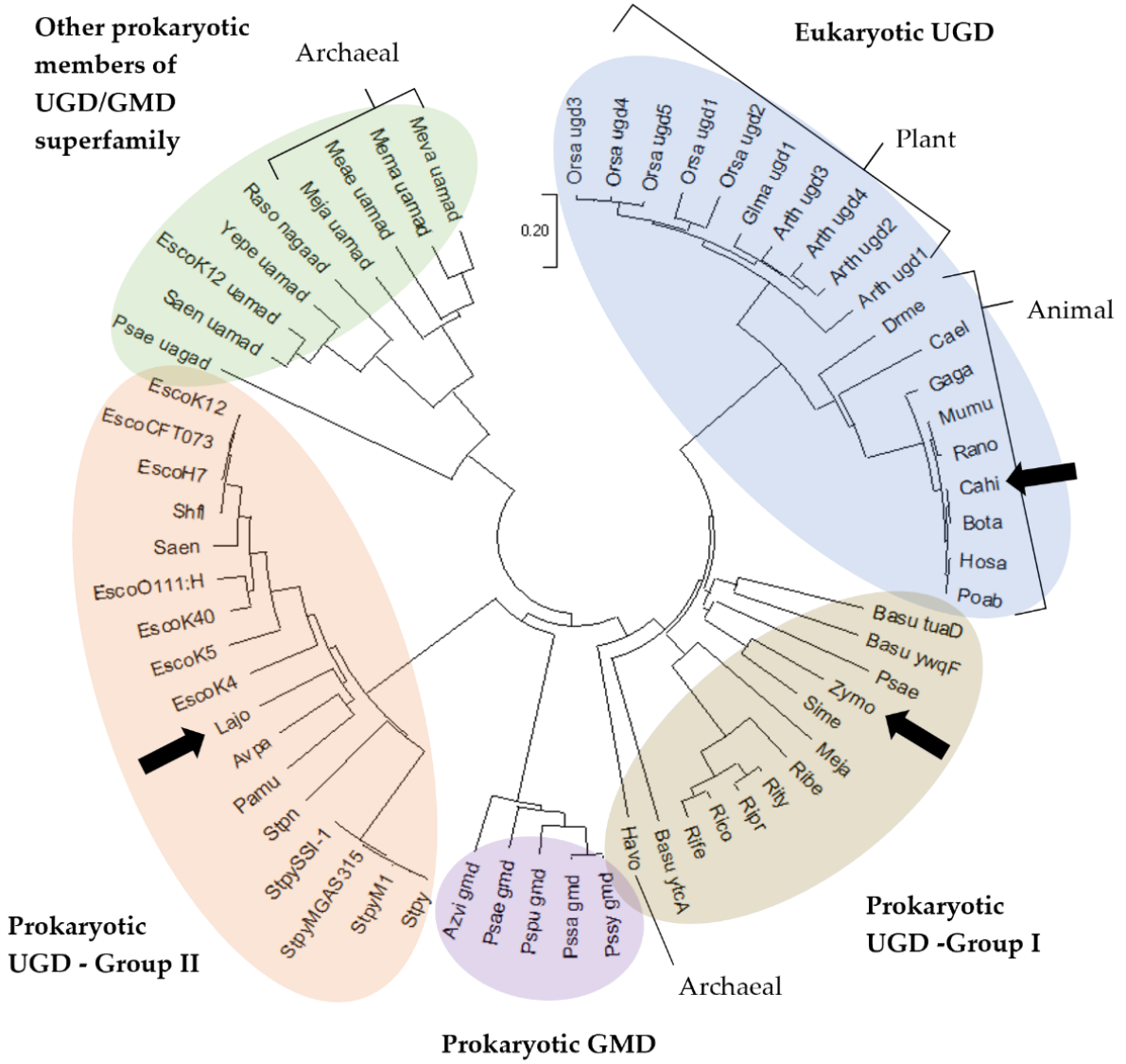


Figure 3.2. Evolutionary relationship between 63 uridine diphosphate (UDP)-glucose 6-dehydrogenase (UGD)-glucose/guanosine diphosphate (GDP)-mannose 6-dehydrogenase (GMD) superfamily members. The resulting optimal tree has the sum of branch length = 12.04551231. The tree is drawn to scale, with branch lengths corresponding to the evolutionary distances used to infer the phylogenetic tree. The dataset comprised a total of 991 sites. Enzymes are labelled by the two first letters of the genus followed by two letters of species name. The three UGD expressed in this study are indicated with a black arrow (Lajo, *Lactobacillus johnsonii*; Cahi, *Capra hircus*; and Zymo, *Zymomonas mobilis*). Abbreviation of the protein is also shown after species name when not UGD, or when multiple UGD isoforms exist within same species. Protein abbreviations: gmd, GDP-mannose 6-dehydrogenase; uamad, UDP-*N*-acetyl-D-mannosamine dehydrogenase; uagad, UDP-*N*-acetylglucosamine 6-dehydrogenase; and nagaad, nucleotide diphosphate (NDP)-*N*-acetyl-*D*-galactosaminuronic acid dehydrogenase.

3.3.2. *E. coli* as UGD Expression Host

Since there are several reports on eukaryotic UGDs expressed in *E. coli* that were not active [44–47], only the two prokaryotic genes were here expressed in *E. coli* BL21 (DE3). This host was evaluated for the expression of *Zm*UGD and *Lbj*UGD by SDS–PAGE (Figure 3.3). Based on their amino acid sequence, the expected sizes for the recombinant 6 × His *Zm*UGD and 6 × His *Lbj*UGD were calculated to be 49.2 kDa and 48.0 kDa, respectively.

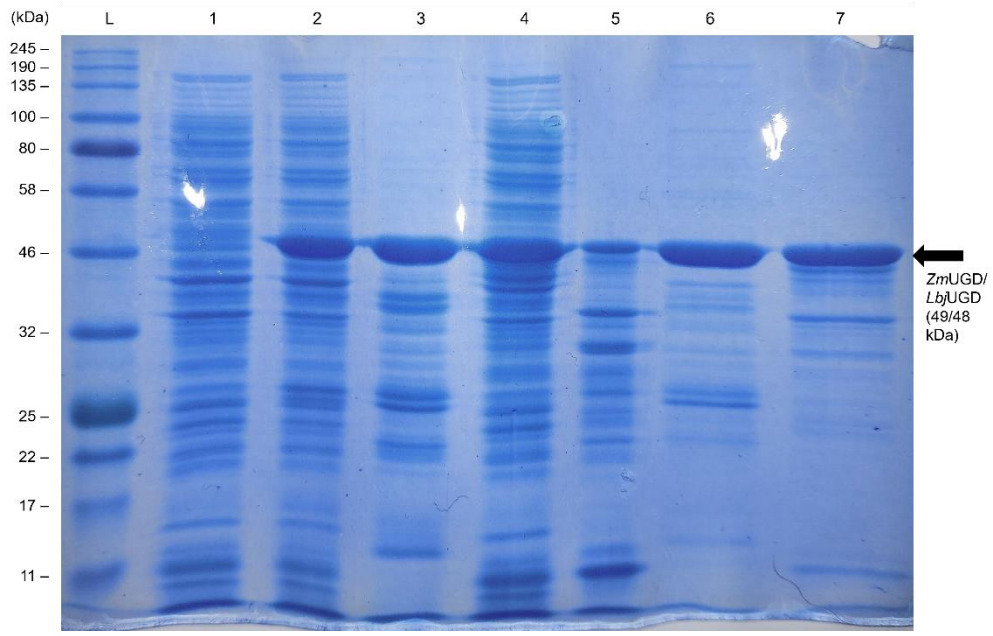


Figure 3.3. SDS–PAGE gel showing 6-His-tagged uridine diphosphate-glucose dehydrogenases (UGDs) expression in *Escherichia coli* BL21. Lane L on the left corresponds to the ladder Color Prestained Protein Standard, Broad Range (11–245 kDa). Samples loaded were: (**1**) soluble cell extract of *E. coli* pRSFDuet-1 (15 µg), (**2**) soluble cell extract from *E. coli* pRSFDuet_*Zm*UGD (15 µg), (**3**) 1st eluted fraction from *Zm*UGD purification (5 µg), (**4**) soluble cell extract from *E. coli* pRSFDuet_*Lbj*UGD (15 µg), (**5**) 1st eluted fraction from *Lbj*UGD purification (5 µg), (**6,7**) 2nd eluted fractions from *Zm*UGD and *Lbj*UGD purification, respectively (5 µg).

As it can be observed, *Zm*UGD and *Lbj*UGD were efficiently overexpressed in *E. coli* and exhibited the correct size. After UGD purification, as expected, the fraction from the 2nd elution showed a cleaner band. The UDP-GlcA production *in vivo* by the engineered *E. coli* was quantified by UHPLC. Figure 3.4 shows the UDP-Glc and UDP-GlcA quantification in bacterial cell extracts, normalized by the dry cell weight obtained at the end of the fermentation.

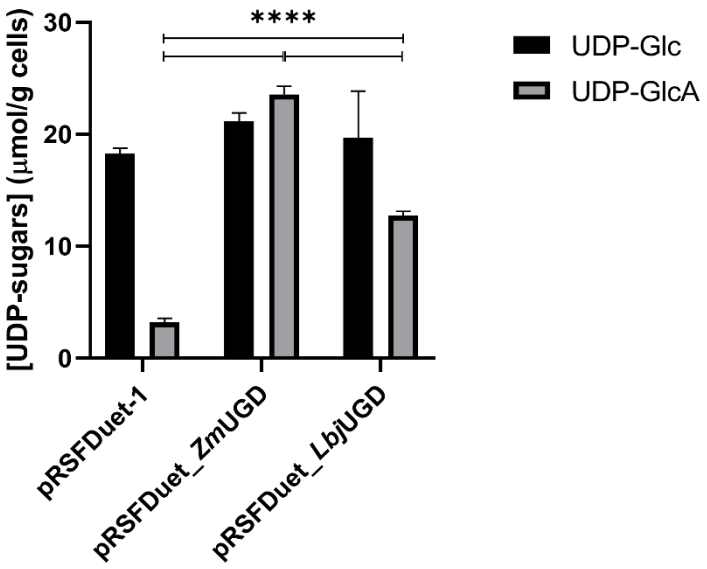


Figure 3.4. *In vivo* uridine diphosphate-glucose (UDP-Glc) and UDP-glucuronic acid (UDP-GlcA) production by *Escherichia coli* BL21 harboring pRSFDuet-1, pRSFDuet_*Zm*UGD or pRSFDuet_*Lbj*UGD, the last two expressing an extra copy of UDP-Glc dehydrogenase (UGD) from *Zymomonas mobilis* (*Zm*UGD) and *Lactobacillus johnsonii* (*Lbj*UGD), respectively. Results are normalized by the cultures' dry weight of cells at the end of the fermentation. The fermentations were performed in triplicate. Four asterisks (****) identify statistically significant differences with adjusted *p* values < 0.0001.

UDP-Glc is naturally produced from glucose in *E. coli* BL21. The UDP-Glc in this strain can be used for the biosynthesis of trehalose, lipid A-core, cellulose, or for galactose degradation via the Leloir pathway. Although there is still no wet lab confirmation, the gene prediction computational methods by protein homology suggest that this strain contains a UGD gene (NCBI sequences NC_012892 region: 1994244-1995410; NC_012971 region: 1994250-1995416), which justifies the presence of small amounts of UDP-GlcA in the negative controls. Additionally, it is known that *E. coli* BL21 has a polymyxin resistance pathway [48] that requires the use of UDP-GlcA. Additionally, there is evidence that UDP-Glc regulates the activity of enzymes in *E. coli*. UDP-Glc enhances adenosine diphosphate-sugar pyrophosphatase activity, which can lead to reduced glycogen production by the cell [49].

When recombinant UGDs were expressed in *E. coli* BL21 cells, UDP-GlcA formation became evidently higher in their cell extracts compared to the one carrying the empty vector. The highest UDP-GlcA *in vivo* production at the end of fermentation was achieved by *E. coli* expressing *Zm*UGD, 28.4 µM (or 23.6 $\mu mol \cdot g^{-1}$ cell dry weight, CDW), while the *E. coli* expressing *Lbj*UGD produced 14.9 µM (or 12.7 $\mu mol \cdot g^{-1}$ CDW). Interestingly, the UDP-GlcA production did not result in a reduction of the UDP-Glc pool in the

engineered *E. coli*. In the work of Cimini et al. [50] the UDP-sugars levels were accessed in wild-type and in an *E. coli* K4 engineered strain. UDP-Glc varied approximately from 0.1 to 0.4 μmol·g^{-1} CDW while UDP-GlcA varied from 0.1 to 1.2 μmol·g^{-1} CDW, depending on the strain and on the time of fermentation. Restaino et al. [8] reported that UDP-sugars concentrations during microbial fermentations of the pathogenic *E. coli* K4 or *E. coli* K5 in different media varied between 0.25 and 11.0 μM·g^{-1} CDW (from 0.2 to 12.0 μM), depending on the strain, type of precursor, growth phase, and culture conditions. Specifically, using *E. coli* K4 grown in a glycerol-based medium, UDP-Glc concentrations varied from 8.1 to less than 0.5 μM·g^{-1} CDW and UDP-GlcA from 0.25 to approximately 0.48 μM·g^{-1} CDW. When using *E. coli* K5 in the same medium, UDP-Glc and UDP-GlcA concentrations varied from around 11 to 1 μM·g^{-1} CDW and from almost undetectable to approximately 1.5 μM·g^{-1} CDW, respectively. The authors showed that the UDP-precursors availability considerably influenced the total chondroitin or heparosan produced. Fermentation using recombinant strains of *S. zooepidemicus* [51] resulted in the same range of UDP-sugar precursors in the exponential phase. UDP-Glc varied from approximately 1.1 to 3.9 μmol·g^{-1} CDW and UDP-GlcA varied from almost 0 to 17.2 μmol·g^{-1} CDW, according to the strain. It is important to notice that in the referred studies there is not much accumulation of UDP-GlcA as in the current work because it is being redirected for GAGs production. Nevertheless, the results from the present work show that the *E. coli* BL21 was able to produce UDP-GlcA in the same range as the reported for GAGs-producing pathogenic strains.

After evaluating UDP-GlcA productions *in vivo*, cell extracts were used to perform *in vitro* reactions that were analyzed by UHPLC (Figure 3.5). Cell extracts were used because after His-Tag purification, recombinant UGDs showed no activity *in vitro* (*data not shown*).

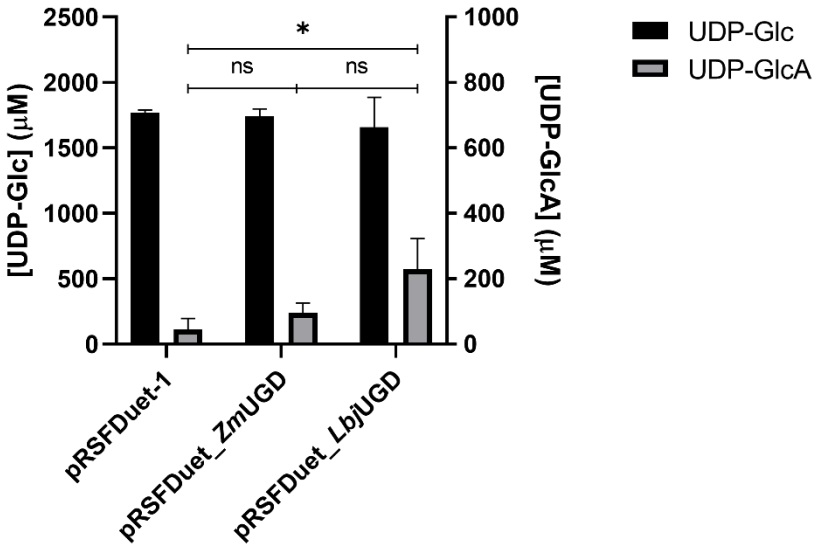


Figure 3.5. *In vitro* uridine diphosphate-glucose (UDP-Glc) consumption and UDP-glucuronic acid (UDP-GlcA) production by cell extracts expressed as concentration by UDP-Glc dehydrogenase (UGD) from *Zymomonas mobilis* (*Zm*UGD) and *Lactobacillus johnsonii* (*Lbj*UGD) expressed in *Escherichia coli* BL21. Values from triplicate reactions. One asterisk (*) identifies adjusted *P* values < 0.05.

Differently from the *in vivo* results, in the *in vitro* assay using cell extracts, the enzyme *Lbj*UGD resulted in the highest UDP-GlcA production, despite not being statistically different than *Zm*UGD. The variations from *in vitro* and *in vivo* reactions may be explained by the different pH and temperature conditions at which the assays are conducted. After 1 h of reaction, the cell extract containing *Lbj*UGD led to a 229 μM production *in vitro*, while the one containing *Zm*UGD led to a 95 μM production under the same conditions. Differences on substrate and product concentrations between reactions of *Lbj*UGD recombinant enzyme and negative control were statistically significant at a 95% confidence level. The initial UDP-Glc concentration was 2.5 mM and from Figure 3.5 it is possible to conclude that more UDP-Glc was consumed than the one required to produce the detected amounts of UDP-GlcA. This was observed also in the negative control which again shows the ability of the enzymes from *E. coli* BL21 to use UDP-Glc as substrate for competitive reactions, such as for the biosynthesis of trehalose or lipid A-core, and the ability of endogenous UGD to convert UDP-Glc to UDP-GlcA that can be afterwards metabolized in the polymyxin pathway or for colanic acid biosynthesis.

The NADH production was also evaluated through the monitoring of the absorbance at 340 nm (Figure A 3.2) during the *in vitro* reactions. Absorbance increases with the NADH production. For each mol of UDP-Glc that is converted into UDP-GlcA by UGD, 2 mol of NADH are produced. However, NADH is not a suitable indicator of UGD enzymatic activity in crude cell extracts since it can also be produced by other

enzymes. Nevertheless, the highest NADH productions were obtained in the reactions using cell extracts from engineered *E. coli* expressing *Lbj*UGD, as observed by the quantification of UDP-GlcA by UHPLC. *S. cerevisiae* was further used as expression host for UGD to achieve higher enzymatic activities and consequently higher UDP-GlcA titers.

3.3.3. Three Novel UGDs Expressed in Yeasts

In this work, the budding yeast was evaluated as an expression system for UGD enzymes since the use of eukaryotic hosts could improve the solubility and activity of the expressed enzymes. Eukaryotic enzymes that are not easily expressed in *E. coli* can become viable alternatives for the construction of biosynthetic pathways.

Besides the same two prokaryotic genes previously expressed in *E. coli*, one eukaryotic UGD, codon-optimized for *S. cerevisiae*, was evaluated for UDP-GlcA production. The expression of the UGDs, fused to a His-Tag, was evaluated through SDS–PAGE (Figure A 3.3). The transformed *S. cerevisiae* strains did not show an increased UGD expression compared to the other proteins in their cell extract and to the controls. The expected sizes of *Ch*UGD, *Zm*UGD, and *Lbj*UGD enzymes were 55.9 kDa, 48.2 kDa and 47.4 kDa, respectively. The purification of the his-tagged proteins was performed but the proteins could not be observed in the gels and the amounts produced were not enough to further use in *in vitro* assays. Although the constitutive promoters used (phosphoglycerate kinase, PGK1, or glyceraldehyde-3-phosphate dehydrogenase, GPD, promoters) are widely applied for high constitutively expression of recombinant proteins in yeasts [52–55], the expression is not comparable to the prokaryotic *E. coli* system where the exceptionally strong T7 promoter and a high copy number plasmid were used.

The cell extracts were analyzed through UHPLC and UDP-GlcA, although in very small amounts, was detected for all engineered *S. cerevisiae* strains. Contrarily to *E. coli*, *S. cerevisiae* does not harbor the UGD gene in its genome. Therefore, UDP-GlcA was not detected in cell extracts of wild-type yeasts. The *in vivo* production of UDP-GlcA by *S. cerevisiae* expressing recombinant ChUGD, *Lbj*UGD and *Zm*UGD is shown in Figure 3.6.

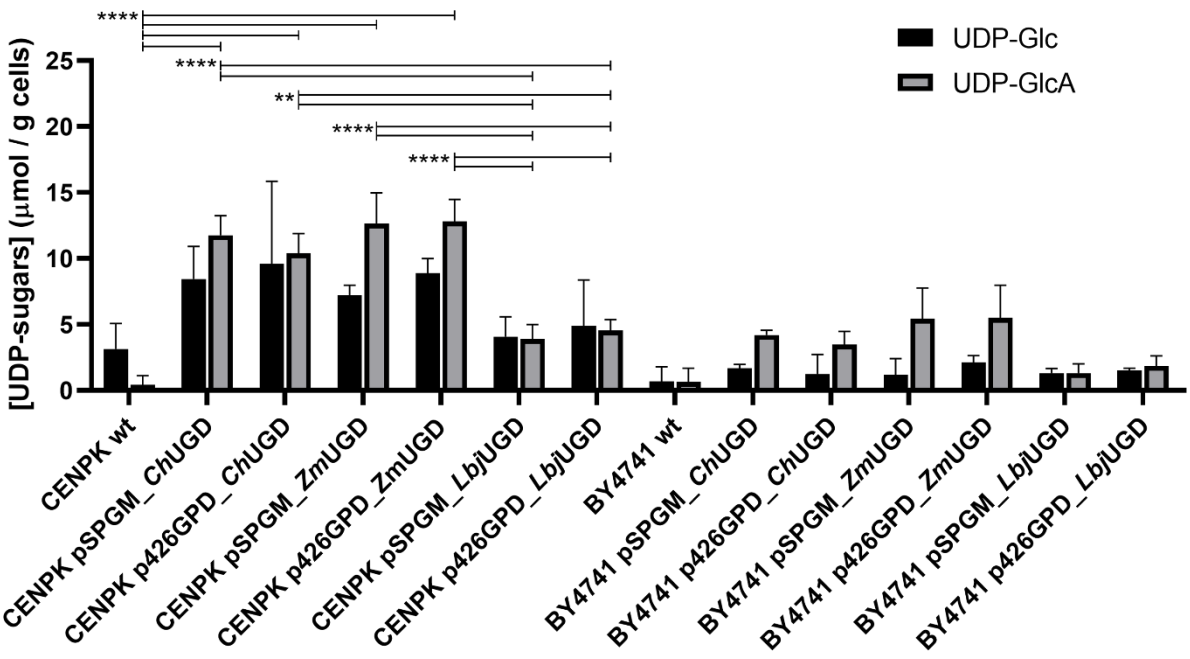


Figure 3.6. *In vivo* uridine diphosphate (UDP)-sugars production by *Saccharomyces cerevisiae* CEN.PK2-1C and BY4741 wild types (wt.) and engineered strains expressing UDP-glucose dehydrogenase (UGD) from *Capra hircus* (*Ch*UGD), *Zymomonas mobilis* (*Zm*UGD) and *Lactobacillus johnsonii* (*Lbj*UGD). Standard deviations are relative to at least three different transformed colonies tested. Differences in the UDP-glucuronic acid (UDP-GlcA) concentrations between the mutants and the correspondent wild-type are indicated, with statistical significance represented with two asterisks (**) for adjusted *P* values <0.01 and with four asterisks (****) for *p* values <0.0001. Statistically non-significative differences are not shown for simplification.

It has been reported that in *S. cerevisiae*, UDP-Glc is a naturally abundant substrate in the cytoplasm, required for the synthesis of β-glucan polysaccharides [56], but also for trehalose and glycogen biosynthesis. The UDP-Glc concentration herein obtained ranged from 1.3 to 9.9 µM (0.7 to 8.4 $\mu mol \cdot g^{-1}$ CDW), depending on the strain and plasmid used. In general, within each strain, differences between UDP-Glc concentration of all the different combinations were not statistically significant. In terms of UDP-GlcA production, all mutants showed increased UDP-GlcA production, due to UGD overexpression, as compared to the negative controls. However, differences were not statistically significative in *S. cerevisiae* BY4741 engineered strains compared to the wild-type strain. It is important to mention that as episomal plasmids were used as expression vectors, different yeast transformed colonies can have huge differences in terms of plasmid copy number and, consequently, in product concentration. Hence, future integration in the genome, as single or multiple gene copies, could result in more reproducible effects [57].

S. cerevisiae CEN.PK2-1C harboring pSP-GM_*Ch*UGD, p426GPD_*Ch*UGD, pSP-GM_*Zm*UGD and p426GPD_*Zm*UGD showed the highest statistically significant production compared to the wild-type, producing on average 14.6, 10.5, 17.9 and 15.9 μM, respectively (corresponding to 13.1, 9.7, 13.5 and 13.6 μmol·g^{-1} CDW). As it can be observed, the engineered strains expressing the eukaryotic enzyme *Ch*UGD and the ones expressing the prokaryotic enzyme *Zm*UGD exhibited greater UDP-GlcA amounts, while the systems expressing *Lbj*UGD generated lower UDP-GlcA levels *in vivo*. This observation suggests that UGD enzymes from the eukaryotic and group I prokaryotic phylogenetic groups (Figure 3.2) have improved *in vivo* activity compared to group II prokaryotic UGDs. The work of Oka & Jigami [56] was the only one that so far reported the use of *S. cerevisiae*, specifically the strain W303a, to express UGD. The expression of one of the UGD isoforms from *A. thaliana*, allowed to obtain an *in vivo* UDP-GlcA production of 5.71 ± 1.06 μmol·g^{-1} CDW, and the substrate UDP-Glc was reduced 54% (from 2.21 μmol·g^{-1} CDW to 1.20 μmol·g^{-1} CDW) compared to the wild-type. This decrease in UDP-Glc pool was not observed in the current work. Nevertheless, the values of UDP-GlcA herein achieved are 2 to 3 times higher (depending on the strain) than the ones obtained by those authors. In a different work, Zhou et al. [47] expressed a fungal UGD using a different yeast as host, namely *Pichia pastoris*. However, no values of UDP-GlcA production are reported in this study.

The slightly reduced *in vivo* UDP-GlcA production obtained with *S. cerevisiae* compared to *E. coli* might be due to the absence of a UGD homolog gene in *S. cerevisiae* as it occurs in bacteria.

Afterwards, the *in vitro* UDP-GlcA production was evaluated. Figure 3.7 shows the UHPLC results obtained from the *in vitro* reactions using the yeast cell extracts.

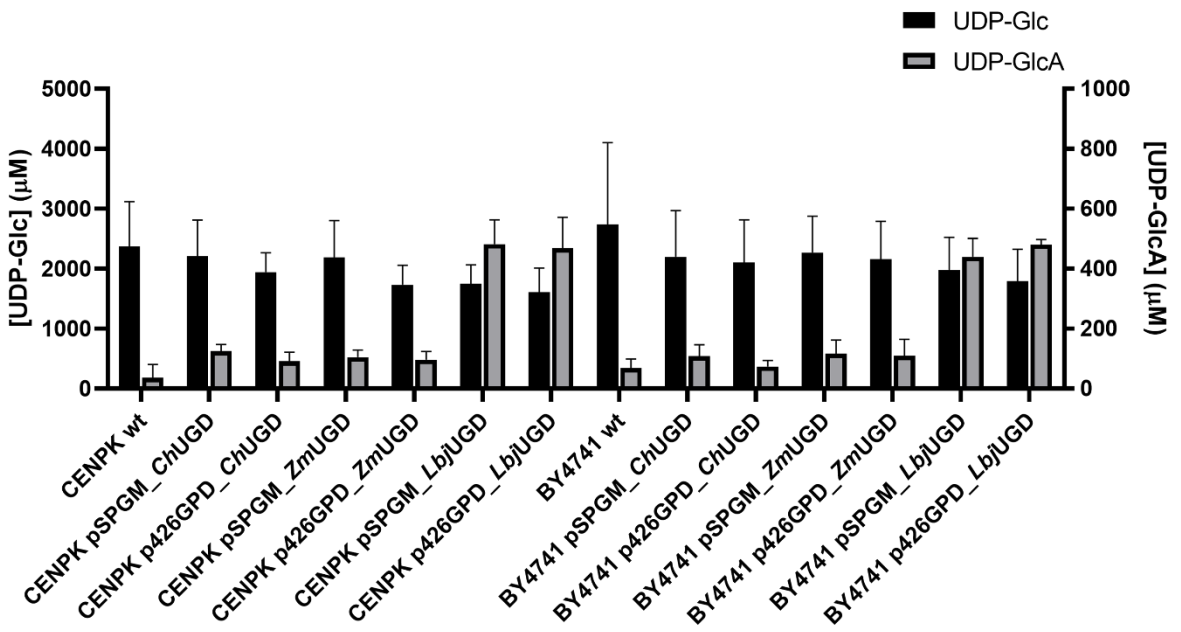


Figure 3.7. *In vitro* conversion of uridine diphosphate-glucose (UDP-Glc) into UDP-glucuronic acid (UDP-GlcA) by *Saccharomyces cerevisiae* CEN.PK2-1C and BY4741 cell extracts through the action of expressed UDP-Glc dehydrogenases (UGDs) from *Capra hircus* (*Ch*UGD), *Zymomonas mobilis* (*Zm*UGD) and *Lactobacillus johnsonii* (*Lbj*UGD). Standard deviations are relative to at least three independent assays.

Although the differences in UDP-GlcA concentrations were not statistically significant, *Lbj*UGD in yeast cell extracts exhibited the highest UDP-GlcA concentrations *in vitro*, resembling what was observed in the *in vitro* reactions using *E. coli* cell extracts. The higher final UDP-GlcA concentrations obtained were on average: 482 µM using *S. cerevisiae* CEN.PK2-1C pSP-GM_*Lbj*UGD; 480 µM using BY4741 p426GPD_*Lbj*UGD; 469 µM using CEN.PK2-1C p426GPD_*Lbj*UGD; and 439 µM with BY4741 pSP-GM_*Lbj*UGD. The high standard deviation values are related to the use of cell extracts from different transformed colonies for the *in vitro* reactions. As it can be observed, the UDP-GlcA production *in vitro* was not significantly affected using different vectors or promoters for UGD expression. Using plasmids pSP-GM1 and p426GPD, containing PGK1 and GPD promoters, respectively, for UGD expression did not seem to result in statistically significative differences in UDP-GlcA production *in vitro.*

The NADH production was also evaluated through the monitoring of the absorbance at 340 nm (Figure A 3.4). The highest NADH productions were obtained in the reactions using cell extracts from engineered yeasts expressing *Lbj*UGD, which agrees with the obtained results of UDP-GlcA production.

The obtained results for the *in vitro* reactions using yeast cell extracts resembled the ones obtained with *E. coli* cell extracts. The enzyme that generally led to the highest UDP-GlcA concentration *in vitro* in

engineered yeast, *Lbj*UGD, was not the one that allowed the highest production of UDP-GlcA *in vivo.* Again, *in vivo* and *in vitro* conditions, such as the pH, were different which may justify the different relative activities found to the enzymes.

Comparing the prokaryotic and eukaryotic hosts for UGD expression, while the highest *in vivo* UDP-GlcA production was achieved with engineered *E. coli*, the enzymes expressed in *S. cerevisiae* exhibited higher UDP-GlcA production *in vitro*. This suggests that the conditions used for *in vivo* UDP-GlcA production in yeasts are probably not appropriate for optimal UGD activity. For example, the pH of yeast cell extracts at the end of fermentation was slightly lower (around 6.0) than the pH of cells extracts obtained after culturing *E. coli* (around 7.0), and UGDs usually have an optimal activity at basic pH as further discussed below. Therefore, if an *in vivo* strategy is envisaged, *S. cerevisiae* fermentation needs to be further optimized.

It was demonstrated that *Zm*UGD was the enzyme that more actively converted UDP-Glc to UDP-GlcA *in vivo* in both prokaryotic and eukaryotic hosts. In the yeast, *Ch*UGD allowed to produce the same UDP-GlcA levels as *Zm*UGD. However, cell extracts containing *Lbj*UGD produced the highest UDP-GlcA, in both prokaryotic and eukaryotic hosts. Contrarily to what might be expected, the activities *in vivo* and *in vitro* of *Zm*UGD and *Lbj*UGD were very distinct, regardless of the host, despite being both from prokaryotic sources; being *Zm*UGD and *Ch*UGD more similar in terms of enzymatic activities as shown by UDP-GlcA production. This can be related to the fact that *Zm*UGD sequence is more similar to *Ch*UGD than *Lbj*UGD as observed from the phylogenetic analysis (Figure 3.2).

3.3.4. Biochemical Characterization of Cloned UGDs

The cell extracts from *E. coli* and *S. cerevisiae* were used to determine the optimal pH and temperature for UGD recombinant enzymes under *in vitro* conditions. Figure 3.8 shows the *in vitro* relative enzymatic activity between different cell extracts, when varying the pH (Figure 3.8A,B) and the temperature (Figure 3.8C,D), according to the final UDP-GlcA concentration determined by UHPLC. The *S. cerevisiae* strain used for UGD expression was CEN.PK2-1C as it was the strain that previously led to higher productions *in vivo* and the results *in vitro* of both strains were similar. The genes were expressed under control of the PGK1 promoter as no significant differences among the two promoters tested were previously observed.

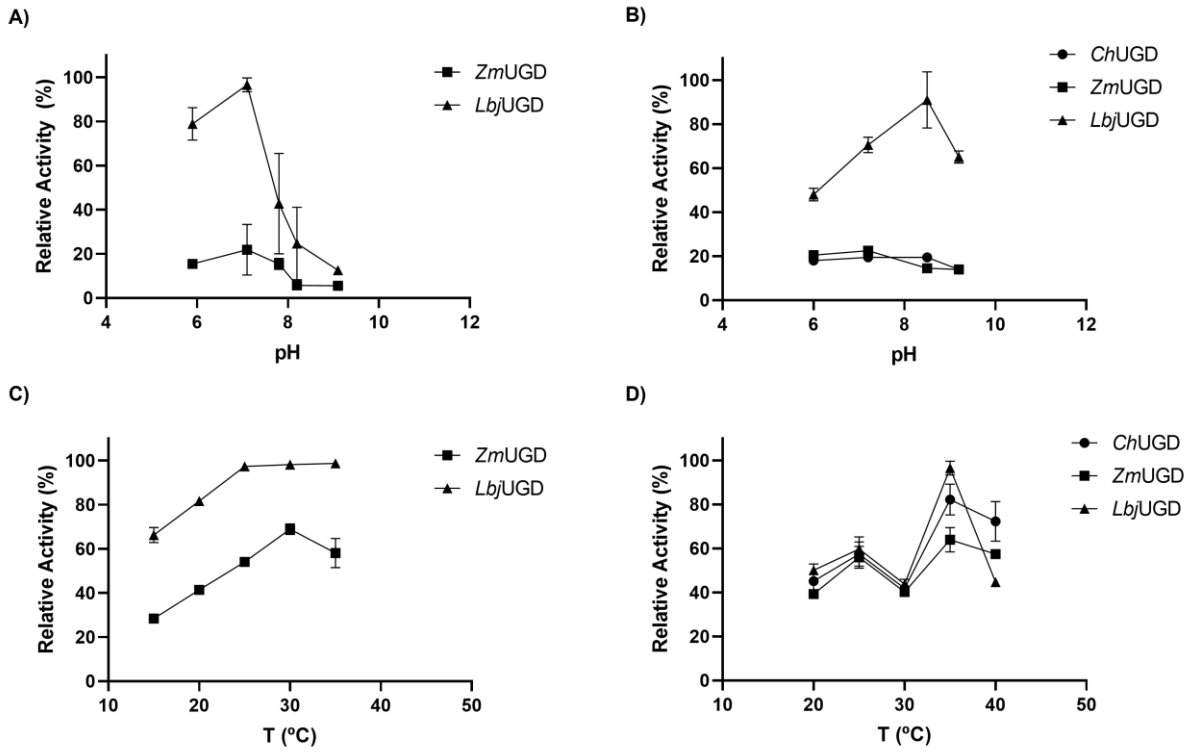


Figure 3.8. Characterization of recombinant uridine diphosphate-glucose dehydrogenases (UGDs) at different pH and temperature. *In vitro* reactions were performed using cell extracts containing recombinant UGD expressed in *Escherichia coli* BL21 (DE3) under T7 promoter control, on different pH at 30°C and on different temperatures at pH 8.0, respectively (**A, C**) and cell extracts containing recombinant UGDs expressed in *Saccharomyces cerevisiae* CEN.PK2-1C under PGK1 promoter, on different pH at 30°C and on different temperatures at pH 8.0, respectively (**B, D**). The 100% activity corresponds to the highest activity of (**A**) 31.1 × 10^{-3} U/mL, (**B**) 9.8 × 10^{-3} U/mL, (**C**) 8.4 × 10^{-3} U/mL and (**D**) 3.9 × 10^{-3} U/mL, in all cases achieved when the UGD from *Lactobacillus johnsonni* (*Lbj*UGD) was expressed. One unit of UGD activity was defined as the amount of enzyme that produces 1 µmol of uridine diphosphate-glucuronic acid per minute.

Cell extracts from engineered *E. coli* were used and the two expressed UGD enzymes, *Zm*UGD and *Lbj*UGD, showed slightly different properties (Figure 3.8A,C). Again, *Lbj*UGD allowed to produce more UDP-GlcA *in vitro* than *Zm*UGD. Regarding the different pH, higher UDP-GlcA content was obtained at the pH 7.1. The optimal temperature for *E. coli* expressed *Zm*UGD was 30 °C while the *Lbj*UGD conserved its maximal activity between 25 °Cand 35°C. From the UGDs expressed in *S. cerevisiae*, *Lbj*UGD generally exhibited the highest activity *in vitro* under different pH and temperatures (Figure 3.8B,D). Its activity was maximal at the temperature of 35°C and a pH of 8.5. The other UGDs showed lower activities under all pH tested and exhibited the maximal activity at 35°C.

Nevertheless, these results are in accordance with the reported studies on UGD activities from other organisms. For example, the pH optimum for *Cryptococcus laurentii* UGD activity was in the range of 7.3–7.8 [58]; the optimum pH and temperature of soybean UGD were reported to be 8.7 and 30°C [59]; optimal pH for UGD activity of sugarcane was 8.4 [60]; UGD from *Sphingomonas elodea* exhibited optimum pH and temperature of 8.7 and 37°C [43]; the UGD from the fungus *Phoma herbarum* showed the highest activity at a pH of 8.6 and at 35°C [47]; *P. multocida's* UGD activity was optimal at a pH of 10 and 37°C [7]; and UGDs from *Granulibacter bethesdensis* and *Akkermansia muciniphila* showed optimal activity at 37°C and pH 9.0 [61,62].

Expectably, these results show a general increase in UDP-GlcA concentration compared to the *in vitro* assays previously performed under non-optimal conditions. In these experiments, UDP-GlcA concentrations using cell extracts from *E. coli* expressing *Lbj*UGD and *Zm*UGD reached on average 1800 µM and 407 µM, respectively (pH 7.1 and 30 °C, Figure 3.8A). When using cell extracts from engineered *S. cerevisiae* strains, the maximum UDP-GlcA concentration achieved with *Ch*UGD and *Zm*UGD was 116 µM and 134 µM (pH 7.2 and 30°C, Figure 3.8B), while *Lbj*UGD allowed to obtain 533 µM (pH 8.5 and 30 °C, Figure 3.8B).

This biochemical characterization of UGDs could explain the differences previously obtained in *in vivo* and *in vitro* UDP-GlcA production using the different expression systems. It was observed that UGD enzymes produced more UDP-GlcA *in vivo* using *E. coli* than *S. cerevisiae*. That result could be justified by the acidic environment of yeast fermentation being far from the optimal pH of UGDs. Additionally, it is demonstrated that the *Lbj*UGD activity is highly affected by the pH, so in the buffered *in vitro* reactions it has shown a much better performance than *in vivo* where pH is changing with the fermentation time. Additionally, in *E. coli* the expression temperature was 30°C which was optimal for the enzymes expressed in that system (Figure 3.8C). However, the temperature of 30°C, used for *S. cerevisiae* fermentation, was not the optimal for UGDs expressed in that chassis (Figure 3.8D).

Moreover, in the case of the *in vitro* reactions performed, presented in Section 3.3.2 and 3.3.3, it is important to notice that they were performed at 30°C and at a pH of 8.0 which was here demonstrated to be close to, but not the optimal conditions for UGDs expressed in *E. coli* and in *S. cerevisiae*, thus suggesting that the UDP-GlcA production could be further improved. The kinetic parameters (K_m and V_{max}) (Table 3.2). were also determined under these conditions.

Table 3.2. Kinetic parameters V_{max} and K_m for each enzyme expressed in this study. *Zm*UGD: Uridine diphosphate-glucose dehydrogenase (UGD) from *Zymomonas mobilis*; *Lbj*UGD: UGD from *Lactobacillus johnsonii*; *Ch*UGD: UGD from *Capra hircus*.

Expression system	Enzyme Expressed	V_{max} ($\mu mol.min^{-1}.g_{protein}^{-1}$)	K_m (μM)	V_{max}/K_m ($L.min^{-1}.g_{protein}^{-1}$)
Escherichia coli	*Zm*UGD	81	2586	3.1×10^{-2}
	*Lbj*UGD	66	1939	3.4×10^{-2}
Saccharomyces cerevisiae	*Ch*UGD	2.2	841	2.6×10^{-3}
	*Zm*UGD	3.7	1234	3.0×10^{-3}
	*Lbj*UGD	1.4	314	4.5×10^{-3}

As it can be seen through the V_{max}/K_m ratio, the enzyme that showed the highest potential was *Lbj*UGD both in *E. coli* and *S. cerevisiae* systems, which agrees with the previous observed results from *in vitro* reactions that showed that *Lbj*UGD was able to produce more UDP-GlcA in the assay conditions. Since the kinetic parameters of characterized UGD in the literature are generally calculated with purified enzyme, it was expected that the parameters herein estimated would be significantly different from previous reports. Since the protein concentration in the cell extracts accounts for much more proteins than UGD, the reaction maximum velocities of the expressed UGDs were lower and required higher concentrations of UDP-Glc to obtain half of the maximum velocity (higher K_m). This effect is obviously more pronounced in the enzymes expressed in yeasts since, as mentioned before, the expression of enzymes was not as evident as the ones expressed in *E. coli*, where the overexpression was easily observed in the SDS-PAGE gels. Examples of reported UGD kinetic parameters are: K_m of 230 μM and a V_{max} of 8.0×10^3 $\mu mol.min^{-1}.g_{protein}^{-1}$ for UGD from *Burkholderia cepacia* [63]; K_m and V_{max} of 110 μM and 9.8×10^9 $\mu mol.min^{-1}.g_{protein}^{-1}$, respectively, for UGD from *P. multocida* [7]; an apparent K_m of 137.2 μM [64] for a recombinant UGD from *Bacillus cereus* expressed in *E. coli*; and a K_m of 34 μM and a V_{max} of 313 $\mu mol.min^{-1}.g_{protein}^{-1}$ for the human UGD recombinantly expressed in *E. coli* [65].

3.4. Conclusions

Although UGD is considered a bottleneck step for GAGs biotechnological production, few alternative UGDs have been characterized and evaluated for such purpose. In addition, the use of eukaryotic hosts to express enzymes from the GAG production pathway has been absent in the literature although they are candidate systems that could possibly address the low solubility and activity of recombinant enzymes.

Herein, three UGD enzymes (*Zm*UGD, *Lbj*UGD and *Ch*UGD) were expressed in *S. cerevisiae* and two prokaryotic enzymes (*Zm*UGD and *Lbj*UGD) were expressed in *E. coli*.

The sequences of the UGDs in this study were compared with other reviewed enzymes using bioinformatic tools and a phylogenetic analysis was performed, being the three genes cloned in this study representative of three UGD phylogenetic groups. *Zm*UGD was found to be functionally closer to *Ch*UGD than to the other prokaryotic *Lbj*UGD.

The activity of the UGDs was evaluated through quantification of UDP-sugars production *in vivo* in both *E. coli* and *S. cerevisiae* cell extracts, with and without expressing the recombinant UGD. *E. coli* was found to be a useful system for protein expression since the recombinant proteins *Zm*UGD and *Lbj*UGD were clearly overexpressed in a soluble form and a significant increase in the UDP-GlcA pool in cell extracts of engineered *E. coli* was observed as compared to the negative control. However, *in vitro*, the UDP-GlcA concentration did not increase very significantly. Moreover, the same two prokaryotic enzymes and one eukaryotic UGD, codon-optimized for *S. cerevisiae* (*Ch*UGD), were evaluated for UDP-GlcA production, using two different yeast strains and expression vectors. When using *S. cerevisiae* as the heterologous host, the expression of recombinant enzymes was not demonstrated in SDS–PAGE. However, the quantification by UHPLC demonstrated that all engineered yeasts produced more UDP-GlcA *in vivo* than the control strains. The strains with highest *in vivo* UDP-GlcA production were *S. cerevisiae* CEN.PK2-1C expressing *Zm*UGD or expressing *Ch*UGD. Although the highest *in vivo* UDP-GlcA production was achieved with engineered *E. coli*, the enzymes expressed by *S. cerevisiae* exhibited higher performance *in vitro*. The cell extracts were further used to determine the optimum pH and temperature for the UGD recombinant enzymes under *in vitro* conditions. Generally, the optimum conditions that maximize the UGD enzymes activity were a pH between 7 and 9 and a temperature between 30 and 35 °C. Remarkably, the enzyme that generally showed highest activity *in vitro* in both *E. coli* and yeast mutants, *Lbj*UGD, was not the one that allowed the highest production of UDP-GlcA *in vivo*, thus suggesting that the *in vitro* and *in vivo* conditions were sufficiently different leading to different relative activities. Depending on the strategy intended for GAGs production, i.e., *in vivo* or enzymatic production, different enzymes should be used, according to their kinetics, towards higher product yields.

This work evaluates alternative genes for biotechnological applications, including GAGs microbial production, recombinant protein production for enzymatic production of GAGs or to produce expensive UDP-precursors for further chemical and/or enzymatic production of GAGs. These novel UGD genes showed to efficiently convert UDP-Glc into UDP-GlcA (up to 74% conversion using *Lbj*UGD in the optimal conditions) and should be explored for the construction of a complete heterologous pathway for *in vivo*

GAG production. Additionally, metabolic engineering of host microorganisms for improving UDP-Glc availability should be considered, either by overexpressing the pathway genes and regulators and/or by performing the knockout of competitive reactions (essentially cell wall recycling pathways) to improve intermediate pools and subsequent availability.

3.5. References

[1] Q. Zhang, R. Yao, X. Chen, L. Liu, S. Xu, J. Chen, J. Wu, Enhancing fructosylated chondroitin production in *Escherichia coli* K4 by balancing the UDP-precursors, Metab. Eng. 47 (2018) 314–322. https://doi.org/10.1016/j.ymben.2018.04.006.

[2] X. Zhang, L. Lin, H. Huang, R.J. Linhardt, Chemoenzymatic Synthesis of Glycosaminoglycans, Acc. Chem. Res. 53 (2020) 335–346. https://doi.org/10.1021/acs.accounts.9b00420.

[3] A. Badri, A. Williams, R.J. Linhardt, M.A. Koffas, The road to animal-free glycosaminoglycan production: current efforts and bottlenecks, Curr. Opin. Biotechnol. 53 (2018) 85–92. https://doi.org/10.1016/j.copbio.2017.12.018.

[4] Z. Kang, Z. Zhou, Y. Wang, H. Huang, G. Du, J. Chen, Bio-Based Strategies for Producing Glycosaminoglycans and Their Oligosaccharides, Trends Biotechnol. 36 (2018) 806–818. https://doi.org/10.1016/j.tibtech.2018.03.010.

[5] D. Cimini, O.F. Restaino, C. Schiraldi, Microbial production and metabolic engineering of chondroitin and chondroitin sulfate, Emerg. Top. Life Sci. 2 (2018) 349–361. https://doi.org/10.1042/ETLS20180006.

[6] A. Zakeri, M.J. Rasaee, N. Pourzardosht, Enhanced hyluronic acid production in *Streptococcus zooepidemicus* by over expressing HasA and molecular weight control with Niscin and glucose, Biotechnol. Reports. 16 (2017) 65–70. https://doi.org/10.1016/j.btre.2017.02.007.

[7] X. Chu, J. Han, D. Guo, Z. Fu, W. Liu, Y. Tao, Characterization of UDP-glucose dehydrogenase from *Pasteurella multocida* CVCC 408 and its application in hyaluronic acid biosynthesis, Enzyme Microb. Technol. 85 (2016) 64–70. https://doi.org/10.1016/j.enzmictec.2015.12.009.

[8] O.F. Restaino, I. di Lauro, R. Di Nuzzo, M. De Rosa, C. Schiraldi, R. Di Nuzzo, M. De Rosa, C. Schiraldi, R. Di Nuzzo, M. De Rosa, C. Schiraldi, New insight into chondroitin and heparosan-like capsular polysaccharide synthesis by profiling of the nucleotide sugar precursors, Biosci. Rep. 37 (2017) 1–12. https://doi.org/10.1042/BSR20160548.

[9] P.L. DeAngelis, N.S. Gunay, T. Toida, W. Mao, R.J. Linhardt, Identification of the capsular polysaccharides of Type D and F *Pasteurella multocida* as unmodified heparin and chondroitin, respectively, Carbohydr. Res. 337 (2002) 1547–1552. https://doi.org/10.1016/S0008-6215(02)00219-7.

[10] J.-R. Wu, P.-Y. Chen, J.-H. Shien, C.-L. Shyu, H.K. Shieh, F. Chang, P.-C. Chang, Analysis of the biosynthesis genes and chemical components of the capsule of *Avibacterium paragallinarum*, Vet.

Microbiol. 145 (2010) 90–99. https://doi.org/10.1016/j.vetmic.2010.03.002.

[11] C. Schiraldi, D. Cimini, M. De Rosa, Production of chondroitin sulfate and chondroitin, Appl. Microbiol. Biotechnol. 87 (2010) 1209–1220. https://doi.org/10.1007/s00253-010-2677-1.

[12] W. He, L. Fu, G. Li, J. Andrew Jones, R.J. Linhardt, M. Koffas, J.A. Jones, R.J. Linhardt, M. Koffas, J. Andrew Jones, R.J. Linhardt, M. Koffas, Production of chondroitin in metabolically engineered *E. coli*, Metab. Eng. 27 (2015) 92–100. https://doi.org/10.1016/j.ymben.2014.11.003.

[13] P. Jin, L. Zhang, P. Yuan, Z. Kang, G. Du, J. Chen, Efficient biosynthesis of polysaccharides chondroitin and heparosan by metabolically engineered *Bacillus subtilis*, Carbohydr. Polym. 140 (2016) 424–432. https://doi.org/10.1016/j.carbpol.2015.12.065.

[14] C. Zhang, L. Liu, L. Teng, J. Chen, J. Liu, J. Li, G. Du, J. Chen, Metabolic engineering of *Escherichia coli* BL21 for biosynthesis of heparosan, a bioengineered heparin precursor, Metab. Eng. 14 (2012) 521–527. https://doi.org/10.1016/j.ymben.2012.06.005.

[15] F. Cheng, Q. Gong, H. Yu, G. Stephanopoulos, High-titer biosynthesis of hyaluronic acid by recombinant *Corynebacterium glutamicum*, Biotechnol. J. 11 (2016) 574–584. https://doi.org/10.1002/biot.201500404.

[16] H. Yu, G. Stephanopoulos, Metabolic engineering of *Escherichia coli* for biosynthesis of hyaluronic acid, Metab. Eng. 10 (2008) 24–32. https://doi.org/10.1016/j.ymben.2007.09.001.

[17] J.Z. Sheng, P.X. Ling, X.Q. Zhu, X.P. Guo, T.M. Zhang, Y.L. He, F.S. Wang, Use of induction promoters to regulate hyaluronan synthase and UDP-glucose-6-dehydrogenase of *Streptococcus zooepidemicus* expression in *Lactococcus lactis*: a case study of the regulation mechanism of hyaluronic acid polymer, J. Appl. Microbiol. 107 (2009) 136–144. https://doi.org/10.1111/j.1365-2672.2009.04185.x.

[18] P. Jin, Z. Kang, P. Yuan, G. Du, J. Chen, Production of specific-molecular-weight hyaluronan by metabolically engineered *Bacillus subtilis* 168, Metab. Eng. 35 (2016) 21–30. https://doi.org/10.1016/j.ymben.2016.01.008.

[19] A. Badri, A. Williams, A. Awofiranye, P. Datta, K. Xia, W. He, K. Fraser, J.S. Dordick, R.J. Linhardt, M.A.G.G. Koffas, Complete biosynthesis of a sulfated chondroitin in *Escherichia coli*, Nat. Commun. 12 (2021) 1–10. https://doi.org/10.1038/s41467-021-21692-5.

[20] D. Cimini, M. De Rosa, A. Viggiani, O.F. Restaino, E. Carlino, C. Schiraldi, Improved fructosylated chondroitin production by *kfoC* overexpression in *E. coli* K4, J. Biotechnol. 150 (2010) 324–331. https://doi.org/10.1016/j.jbiotec.2010.09.954.

[21] J.L. Strominger, H.M. Kalckar, J. Axelrod, E.S. Maxwell, Enzymatic oxidation of uridine diphosphate glucose to uridine diphosphate glucuronic acid, J. Am. Chem. Soc. 76 (1954) 6411–6412. https://doi.org/10.1021/ja01653a051.

[22] I.A. Rodionova, Z. Zhang, M. Aboulwafa, J.M.H. Saier, M.H.S. Jr., J.M.H. Saier, UDP-glucose dehydrogenase Ugd in *E. coli* is activated by Gmd and RffD, is inhibited by CheY, and regulates swarming, BioRxiv. (2020) 1–14. https://doi.org/10.1101/2020.01.08.899336.

[23] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms, Mol. Biol. Evol. 35 (2018) 1547–1549. https://doi.org/10.1093/molbev/msy096.

[24] N. Saitou, M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees., Mol. Biol. Evol. 4 (1987) 406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454.

[25] M. Nei, J. Zhang, Evolutionary Distance: Estimation, in: ELS, American Cancer Society, 2006: pp. 1–4. https://doi.org/https://doi.org/10.1038/npg.els.0005108.

[26] K.-D. Entian, P. Kötter, 25 Yeast Genetic Strain and Plasmid Collections, in: I. Stansfield, M.J.R. Stark (Eds.), Methods Microbiol., 2nd ed., Elsevier, 2007: pp. 629–666. https://doi.org/10.1016/S0580-9517(06)36025-4.

[27] C.B. Brachmann, A. Davies, G.J. Cost, E. Caputo, J. Li, P. Hieter, J.D. Boeke, Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications, Yeast. 14 (1998) 115–132. https://doi.org/10.1002/(SICI)1097-0061(19980130)14:2<115::AID-YEA204>3.0.CO;2-2.

[28] Y. Chen, S. Partow, G. Scalcinati, V. Siewers, J. Nielsen, Enhancing the copy number of episomal plasmids in *Saccharomyces cerevisiae* for improved protein production, FEMS Yeast Res. 12 (2012) 598–607. https://doi.org/10.1111/j.1567-1364.2012.00809.x.

[29] R.D. Gietz, R.A. Woods, Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method, in: Methods Enzymol., 2002: pp. 87–96. https://doi.org/10.1016/S0076-6879(02)50957-5.

[30] A. Waśko, M. Kieliszek, Z. Targoński, Purification and characterization of a proteinase from the probiotic *Lactobacillus rhamnosus* OXY, Prep. Biochem. Biotechnol. 42 (2012) 476–488. https://doi.org/10.1080/10826068.2012.656869.

[31] N. Banerjee, D. Bhattacharyya, UDP-glucose dehydrogenase from *Capra hircus* liver: Purification, partial characterization and evaluation as a coupling enzyme in UDP-galactose 4-epimerase assay, J. Mol. Catal. B Enzym. 68 (2011) 37–43. https://doi.org/10.1016/j.molcatb.2010.09.007.

[32] S. Silbir, S. Dagbagli, S. Yegin, T. Baysal, Y. Goksungur, Levan production by *Zymomonas mobilis* in batch and continuous fermentation systems, Carbohydr. Polym. 99 (2014) 454–461. https://doi.org/10.1016/j.carbpol.2013.08.031.

[33] E. Jofré, A. Becker, Production of Succinoglycan Polymer in *Sinorhizobium meliloti* Is Affected by SMb21506 and Requires the N-terminal Domain of ExoP, Mol. Plant-Microbe Interact. 22 (2009) 1656–1668. https://doi.org/10.1094/MPMI-22-12-1656.

[34] M. Marvasi, P.T. Visscher, L. Casillas Martinez, Exopolymeric substances (EPS) from *Bacillus subtilis*: polymers and genes encoding their synthesis, FEMS Microbiol. Lett. 313 (2010) 1–9. https://doi.org/10.1111/j.1574-6968.2010.02085.x.

[35] A.P. Bhavsar, L.K. Erdman, J.W. Schertzer, E.D. Brown, Teichoic Acid Is an Essential Polymer in *Bacillus subtilis*, J. Bacteriol. 186 (2004) 7865–7873.

https://doi.org/10.1128/JB.186.23.7865–7873.2004.

[36] M.E. Valentine, B.D. Kirby, T.R. Withers, S.L. Johnson, T.E. Long, Y. Hao, J.S. Lam, R.M. Niles, H.D. Yu, Generation of a highly attenuated strain of *Pseudomonas aeruginosa* for commercial production of alginate, Microb. Biotechnol. 13 (2020) 162–175. https://doi.org/10.1111/1751-7915.13411.

[37] S. Górska-Frączek, C. Sandström, L. Kenne, M. Paściak, E. Brzozowska, M. Strus, P. Heczko, A. Gamian, The structure and immunoreactivity of exopolysaccharide isolated from *Lactobacillus johnsonii* strain 151, Carbohydr. Res. 378 (2013) 148–153. https://doi.org/10.1016/j.carres.2013.05.012.

[38] M.A. Anwar, S. Kralj, M.J.E.C. van der Maarel, L. Dijkhuizen, The Probiotic *Lactobacillus johnsonii* NCC 533 Produces High-Molecular-Mass Inulin from Sucrose by Using an Inulosucrase Enzyme, Appl. Environ. Microbiol. 74 (2008) 3426–3433. https://doi.org/10.1128/AEM.00377-08.

[39] M.R. Wessels, Capsular Polysaccharide of Group A Streptococcus, Microbiol. Spectr. 7 (2019) 1–12. https://doi.org/10.1128/microbiolspec.GPP3-0050-2018.

[40] P. Li, Q. Liu, H. Luo, K. Liang, Y. Han, K.L. Roland, R. Curtiss, Q. Kong, Bi-valent polysaccharides of Vi capsular and O9 O-antigen in attenuated *Salmonella Typhimurium* induce strong immune responses against these two antigens, Npj Vaccines. 3 (2018) 1. https://doi.org/10.1038/s41541-017-0041-5.

[41] C. Whitfield, Biosynthesis and Assembly of Capsular Polysaccharides in *Escherichia coli*, Annu. Rev. Biochem. 75 (2006) 39–68. https://doi.org/10.1146/annurev.biochem.75.103004.142545.

[42] R. Stenutz, A. Weintraub, G. Widmalm, The structures of *Escherichia coli* O-polysaccharide antigens, FEMS Microbiol. Rev. 30 (2006) 382–403. https://doi.org/10.1111/j.1574-6976.2006.00016.x.

[43] A.T. Granja, A. Popescu, A.R. Marques, I. Sá-Correia, A.M. Fialho, Biochemical characterization and phylogenetic analysis of UDP-glucose dehydrogenase from the gellan gum producer *Sphingomonas elodea* ATCC 31461, Appl. Microbiol. Biotechnol. 76 (2007) 1319–1327. https://doi.org/10.1007/s00253-007-1112-8.

[44] M. Klinghammer, R. Tenhaken, Genome-wide analysis of the UDP-glucose dehydrogenase gene family in *Arabidopsis*, a key enzyme for matrix polysaccharides in cell walls, J. Exp. Bot. 58 (2007) 3609–3621. https://doi.org/10.1093/jxb/erm209.

[45] R. Tenhaken, O. Thulke, Cloning of an Enzyme That Synthesizes a Key Nucleotide-Sugar Precursor of Hemicellulose Biosynthesis from Soybean: UDP-Glucose Dehydrogenase, Plant Physiol. 112 (1996) 1127–1134. https://doi.org/10.1104/pp.112.3.1127.

[46] L. V. Bindschedler, E. Wheatley, E. Gay, J. Cole, A. Cottage, G.P. Bolwell, Characterisation and expression of the pathway from UDP-glucose to UDP-xylose in differentiating tobacco tissue, Plant Mol. Biol. 57 (2005) 285–301. https://doi.org/10.1007/s11103-004-7795-7.

[47] T. Zhou, Y.C. Song, M.Q. Feng, R.X. Tan, Cloning, expression and biochemical characterization of

UDP-glucose 6-dehydrogenase, a key enzyme in the biosynthesis of an anti-tumor polysaccharide from the marine fungus *Phoma herbarum* YS4108, Process Biochem. 46 (2011) 2263–2268. https://doi.org/10.1016/j.procbio.2011.09.006.

[48] C. Robichon, J. Luo, T.B. Causey, J.S. Benner, J.C. Samuelson, Engineering Escherichia coli BL21(DE3) Derivative Strains To Minimize E. coli Protein Contamination after Purification by Immobilized Metal Affinity Chromatography, Appl. Environ. Microbiol. 77 (2011) 4634–4646. https://doi.org/10.1128/AEM.00119-11.

[49] M.T. Morán-Zorzano, A.M. Viale, F.J. Muñoz, N. Alonso-Casajús, G.G. Eydallín, B. Zugasti, E. Baroja-Fernández, J. Pozueta-Romero, *Escherichia coli* AspP activity is enhanced by macromolecular crowding and by both glucose-1,6-bisphosphate and nucleotide-sugars, FEBS Lett. 581 (2007) 1035–1040. https://doi.org/10.1016/j.febslet.2007.02.004.

[50] D. Cimini, E. Carlino, A. Giovane, O. Argenzio, I. Dello Iacono, M. De Rosa, C. Schiraldi, Engineering a branch of the UDP-precursor biosynthesis pathway enhances the production of capsular polysaccharide in *Escherichia coli* O5:K4:H4, Biotechnol. J. 10 (2015) 1307–1315. https://doi.org/10.1002/biot.201400602.

[51] W.Y. Chen, E. Marcellin, J. Hung, L.K. Nielsen, Hyaluronan Molecular Weight Is Controlled by UDP-N-acetylglucosamine Concentration in *Streptococcus zooepidemicus*, J. Biol. Chem. 284 (2009) 18007–18014. https://doi.org/10.1074/jbc.M109.011999.

[52] S. Gao, H. Zhou, J. Zhou, J. Chen, Promoter-Library-Based Pathway Optimization for Efficient (2S)-Naringenin Production from p-Coumaric Acid in *Saccharomyces cerevisiae*, J. Agric. Food Chem. 68 (2020) 6884–6891. https://doi.org/10.1021/acs.jafc.0c01130.

[53] S. Partow, V. Siewers, S. Bjørn, J. Nielsen, J. Maury, Characterization of different promoters for designing a new expression vector in *Saccharomyces cerevisiae*, Yeast. 27 (2010) 955–964. https://doi.org/10.1002/yea.1806.

[54] J. Sun, Z. Shao, H. Zhao, N. Nair, F. Wen, J.-H. Xu, H. Zhao, Cloning and characterization of a panel of constitutive promoters for applications in pathway engineering in *Saccharomyces cerevisiae*, Biotechnol. Bioeng. 109 (2012) 2082–2092. https://doi.org/10.1002/bit.24481.

[55] L. Xiong, Y. Zeng, R.-Q. Tang, H.S. Alper, F.-W. Bai, X.-Q. Zhao, Condition-specific promoter activities in *Saccharomyces cerevisiae*, Microb. Cell Fact. 17 (2018) 58. https://doi.org/10.1186/s12934-018-0899-6.

[56] T. Oka, Y. Jigami, Reconstruction of de novo pathway for synthesis of UDP-glucuronic acid and UDP-xylose from intrinsic UDP-glucose in *Saccharomyces cerevisiae*, FEBS J. 273 (2006) 2645–2657. https://doi.org/10.1111/j.1742-4658.2006.05281.x.

[57] S. Isogai, N. Okahashi, R. Asama, T. Nakamura, T. Hasunuma, F. Matsuda, J. Ishii, A. Kondo, Synthetic production of prenylated naringenins in yeast using promiscuous microbial prenyltransferases, Metab. Eng. Commun. 12 (2021) e00169. https://doi.org/10.1016/j.mec.2021.e00169.

[58] H. Ankel, E. Ankel, D.S. Feingold, Biosynthesis of Uridine Diphosphate D-Xylose. III. Uridine

Diphosphate D-Glucose Dehydrogenase of *Cryptococcus laurentii**, Biochemistry. 5 (1966) 1864–1869. https://doi.org/10.1021/bi00870a012.

[59] B. Hinterberg, C. Klos, R. Tenhaken, Recombinant UDP-glucose dehydrogenase from soybean, Plant Physiol. Biochem. 40 (2002) 1011–1017. https://doi.org/10.1016/S0981-9428(02)01465-1.

[60] W. Turner, F.C. Botha, Purification and kinetic properties of UDP-glucose dehydrogenase from sugarcane, Arch. Biochem. Biophys. 407 (2002) 209–216. https://doi.org/10.1016/S0003-9861(02)00500-3.

[61] S. Wei, A. Kulinich, X. Duan, L. Liu, J. Voglmeir, Discovery and Biochemical Characterization of UDP-Glucose Dehydrogenase from *Granulibacter bethesdensis*, Protein Pept. Lett. 22 (2015) 628–634. https://doi.org/10.2174/0929866522666150526092818.

[62] S. Wei, X.-Y. Zhang, Y. Sun, L.P. Conway, L. Liu, Discovery and Biochemical Characterization of UDP-Glucose Dehydrogenase from *Akkermansia muciniphila*, Protein Pept. Lett. 24 (2017) 735–741. https://doi.org/10.2174/0929866524666170724111147.

[63] J. Rocha, A.O. Popescu, P. Borges, D. Mil-Homens, L.M. Moreira, I. Sá-Correia, A.M. Fialho, C. Frazão, Structure of *Burkholderia cepacia* UDP-Glucose Dehydrogenase (UGD) BceC and Role of Tyr10 in Final Hydrolysis of UGD Thioester Intermediate, J. Bacteriol. 193 (2011) 3978–3987. https://doi.org/10.1128/JB.01076-10.

[64] B. Broach, X. Gu, M. Bar-Peled, Biosynthesis of UDP-glucuronic acid and UDP-galacturonic acid in *Bacillus cereus* subsp. *cytotoxis* NVH 391-98, FEBS J. 279 (2012) 100–112. https://doi.org/10.1111/j.1742-4658.2011.08402.x.

[65] A.S. Hyde, A.M. Thelen, J.J. Barycki, M.A. Simpson, UDP-glucose dehydrogenase activity and optimal downstream cellular function require dynamic reorganization at the dimer-dimer subunit interfaces, J. Biol. Chem. 288 (2013) 35049–35057. https://doi.org/10.1074/jbc.M113.519090.

CHAPTER 4

OPTIMIZATION OF CHONDROITIN PRODUCTION IN *E. COLI* USING GENOME SCALE MODELS

Abstract

Chondroitin is a natural occurring glycosaminoglycan with applications as a nutraceutical and pharmaceutical ingredient. It can be extracted from animal tissues, though microbial chondroitin-like polysaccharides emerged as a safer and more sustainable alternative source. However, chondroitin titers using either natural or recombinant microorganisms are still far from meeting the increasing demand. The use of genome-scale models and computational predictions can help to design microbial cell factories with possible improved titers of these value-added compounds. Herein, genome-scale models have been used to predict gene modifications in *Escherichia coli* engineered strains that would potentially lead to improved chondroitin production. Additionally, using synthetic biology approaches, a pathway for producing chondroitin has been designed and engineered in *E. coli.* Afterwards, the most promising mutants according to bioinformatic predictions have been constructed and evaluated for chondroitin production in flask fermentation leading to 118 mg/L extracellular chondroitin by overexpressing both superoxide dismutase (*sodA*) and a lytic murein transglycosylase (*mltB*) Then, fermentations in bioreactor operating in batch and fed-batch were also evaluated in which the mutant overexpressing *mltB* led to a production of extracellular chondroitin of 427 mg/L and 535 mg/L, respectively. The computational approach herein described identified several potential novel targets for improved chondroitin biosynthesis, which might lead to a more efficient production of this glycosaminoglycan.

Keywords: *Escherichia coli*; chondroitin; heterologous production; metabolic engineering; genome-scale models

This chapter is in preparation to be submitted for publication:

M.R. Couto, J.L. Rodrigues, A. Braga, Ó. Dias, L.R. Rodrigues, Optimization of chondroitin production in *E. coli* using genome scale models.

4.1. Introduction

Genome-scale metabolic models (GEMs) are mathematical representations of the entire metabolic network of an organism (or consortium), that include a description of genes, enzymes, reactions, metabolites, as well as their associations and compartments, ultimately allowing to predict biological capabilities [1]. Along with the increasing knowledge and data provided by genome sequencing technologies, the number, quality, and applications of available GEMs have been growing. These models provide valuable information for metabolic engineering strategies as they allow to predict phenotypic behavior of either wild-type or mutated strains under different environmental conditions and to identify targets to improve the metabolic flux towards a target product [2,3]. Applications of GEMs on drug target identification/ drug discovery/ drug design and human disease studying have also been described [2]. The BiGG Models repository [4] currently provides 108 open-source manually curated GEMs, that exhibit robustness of growth predictions.

GEMs representation follows a matrix configuration where the columns represent reactions and the rows metabolites (stoichiometric matrix). For each metabolite, its stoichiometric coefficient in each reaction is presented. Negative coefficients correspond to consumed substrates (reactants), and positive coefficients represent produced metabolites (products). The bounds for each reaction are defined to constrain the range of possible fluxes. The phenotype can then be estimated by different tools. Flux balance analysis (FBA) is a common mathematical method for analyzing the flux distributions through the metabolic network. This approach relies on mass balance equations assuming steady-state growth (all mass that enters the system must leave, so no metabolite is accumulated) [1]. Considering the constraints imposed by mass balance equations, environmental conditions and model metabolic bounds, FBA computes a possible flux distribution by maximizing an objective function by linear programming. Commonly, the biomass production reaction/equation is used as the objective function for simulating microbial phenotypes. FBA accurately predicts wild-type behavior, although for mutant phenotype prediction the simulation method Minimization of Metabolic Adjustment (MOMA) [5] and linear MOMA (LMOMA) [6,7] have been more commonly used. These methods assume that engineered strains will no longer grow to optimize biomass, but instead they grow to maintain the flux distribution as close to the wild-type as possible.

While simulation methods are fundamental for phenotype prediction, finding combinations of gene modifications to reach a desired phenotype require more complex computational tools to iteratively generate and evaluate candidate solutions until a desired phenotype or other termination criteria is achieved [8]. Examples of such computational strain optimization methods include OptKnock [9],

OptStrain [10], OptGene [11], simulated annealing/ evolutionary algorithms (SA/SEA) [12], FluxDesign [13], OptORF [14] and OptForce [15]. These tools differ in the defined objective function, optimization algorithms, simulation methods, the information they support, and/or in the mathematical formulation [8].

Biological behavior prediction through computational modulation has been widely and successfully used to increase the biotechnological production of several high-value compounds such as fatty acids [16], organic acids [17–20], lipids [21–23], polymers [3,24–26], amino acids [27], butanol [28], naringenin [29] and the glycosaminoglycan hyaluronic acid [30].

As reviewed in Chapter 2, chondroitin is a glycosaminoglycan (GAG) with several applications mainly used as a chondroprotective ingredient in human and veterinary medical prescriptions [31]. The usual chondroitin source in current chondroitin-based market products is cartilages such as shark or cows, and efforts are being implemented to shift the production process to a more sustainable and safer one such as the biological production using microorganisms [31]. The biotechnological production of chondroitin is still challenging mainly due to the low titers obtained [32]. Strategies to improve microbial strains for producing chondroitin are therefore essential. *E. coli* has been the most used and well-known host for biotechnological applications, participating as a living catalyst in well-established industrial processes. Therefore, this microorganism was herein used as the host for heterologous chondroitin production and *in silico* strategies have been developed to find potential gene manipulation targets.

4.2. Materials and Methods

4.2.1. Model construction and computational approach

E. coli BL21 stoichiometric models iEC1356_Bl21DE3, iB21_1397, iECBD_1354 and the parental model iJO1366, from *E. coli* K-12, were obtained from the BiGG Models database (http://bigg.ucsd.edu/). These models were modified to include the heterologous pathway for chondroitin production *in silico*. The included reactions were uridine-diphosphate (UDP)-*N*-acetylglucosamine 4-epimerase (UAE), chondroitin synthase/polymerase (CHSY) and their corresponding genes, as well as a chondroitin exchange reaction. Chondroitin was added as new species in all models, and UDP-*N*-acetyl-D-galactosamine was also included, except for iEC1356_Bl21DE3 model that already contained this intermediate. The resulting models harboring the chondroitin pathway were named iEC1356_Bl21DE3_c, iB21_1397_c, iECBD_1354_c and iJO1366_c.

The constructed models were uploaded/imported in the workbench OptFlux Version 3.3.5 [33]. CPLEX Optimization Studio Version 12.9.0 (IBM) was used as Linear Programming solver. Evolutionary optimization was performed for gene deletion and for gene under and overexpression predictions, to search for mutants with enhanced flux for the chondroitin production reaction. Strength Pareto Evolutionary Algorithm 2 (SPEA2) was used as optimization method. Simulations were performed with parsimonious FBA (pFBA) [34] and using the Biomass-Product Coupled Yield (BPCY) as the objective function. The maximum for evaluation functions was set to 50000. A maximum of 10 modifications was allowed. The environmental conditions were set to aerobic (without any restriction of oxygen, lower bound set to -1000) and the substrate glucose to 10 mol/gDW/h (lower bound set to -10).

To find more robust solutions, a recently described alternative computational approach was further performed using the MEWpy workbench developed for Python [35]. *E. coli* BL21 model iB21_1397 was used for this approach. The same previously described environmental conditions were set, and the evolutionary optimization was run. Using this tool, the objective functions, to be maximized, were BPCY and the Weighted Yield (WYIELD, weighed sum of the minimum and maximum product fluxes). pFBA was used as the phenotype prediction method. The evolutionary algorithm (EA) Non-dominated Sorting Genetic Algorithm II [36] was used for the optimization approach.

To evaluate the robustness of the obtained solutions from strain optimization, flux variability analysis (FVA) was performed by fixing biomass as the obtained in the mutant solution and alternating the pFBA objective function between maximizing and minimizing chondroitin production [37].

4.2.2. Heterologous pathway construction

The strains and plasmids used to construct a chondroitin production pathway *in vivo* in *E. coli* are described in Table 4.1. Two reactions were included in *E. coli* enzymatic machinery, namely UAE and CHSY. The selected genes to perform these enzymatic reactions were the ones from the pathogenic strain *E. coli* O5:K4:H4 which naturally produces a compound analogue to chondroitin as part of its capsule [38]. The genes *kfoA* and *kfoC* (encoding UAE and CHSY, respectively) present in pETM6 plasmid, have been kindly provided by Dr. Matheos Koffas (Rensselaer Polytechnic Institute, Troy, NY) [32] and were cloned in pRSFDuet-1 plasmid (Novagen, Madison, USA) in multiple cloning site 1 (MCS1). Also, since the UDP-glucose dehydrogenase (UGD) overexpression has been determined to be crucial for glycosaminoglycan production [39–42], *Zymomonas mobilis* UGD gene (*Zmugd*) [42] was also cloned in the same plasmid in MCS2. The three genes were cloned in pseudo-operon configuration, i.e., the DNA

sequence of each gene follows its own lac operator, T7 promoter and RBS, and a single T7 terminator exists in the end of all genes.

Table 4.1. Strains and plasmids used in this study.

Strains	Relevant Genotype	Source
Escherichia coli NZY5α	*fhuA*2 Δ(*argF-lacZ*)U169 *phoA glnV*44Φ80 Δ(*lacZ*)M15 *gyrA*96 *recA*1 *relA*1 *endA*1 *thi*-1 *hsdR*17	NZYTech (MB00401)
E. coli K-12 MG1655 (DE3)	*F-* λ- *ilvG- rfb*-50 *rph*-1 λ(DE3)	[43]
E. coli BL21 (DE3)	*F- ompT gal dcm lon hsdSB*(*rB- mB*-) λ(DE3 * *lacI lacUV5-T7* gene 1 *ind*1 *sam7 nin*5)	NZYTech (MB006)
Plasmids	**Description**	**Source**
pRSFDuet-1	RSF1030 *ori*, *lacI*, double P_{T7lac}, Kan^R	Novagen
pRSFDuet_Zmugd	pRSFDuet-1 carrying *Z. mobilis* UGD (*Zm*UGD) gene in the MCS1	[42]
pETM6_PCA	pETM6 harboring the genes *kfoC*, *kfoA* from *E. coli* K4 in pseudo-operon configuration	[32]
pRSFDuet_kfoCA_Zmugd	pRSFDuet-1 harboring *kfoA* and *kfoC* in MCS1 and *Zmugd* in MCS2	This study
pETDuet-1	ColE1(pBR322) *ori*, *lacI*, double P_{T7lac}, Amp^R	Novagen
pCDFDuet-1	CloDF13 *ori*, *lacI*, double P_{T7lac}, $Strep^R$	Novagen
pCRISPathBrick	P15A ori, Cm^R, expression of *S. pyogenes* dCas9, tracrRNA, and nontargeting CRISPR array	Addgene #65006
pETDuet_sodA	pETDuet-1 harboring *sodA* from *E. coli* K-12 MG1655 (DE3)	This study
pETDuet_glmU	pETDuet-1 harboring *glmU* from *E. coli* K-12 MG1655 (DE3)	This study
pETDuet_sodA_glmU	pETDuet-1 harboring *sodA* and *glmU*	This study
pCDFDuet_mltB	pCDFDuet-1 harboring *mltB* from *E. coli* K-12 MG1655 (DE3)	This study
pCRISPath_ybiV	pCRISPathBrick harboring protospacer to target *ybiV* for underexpression	This study

Plasmid DNA was extracted using NucleoSpin® Plasmid Miniprep Kit (Macherey-Nagel, Düren, Germany). Genes have been amplified using Phusion High Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, USA) and the primers (Metabion, Steinkirchen, Germany; Eurofins, Ebersberg, Germany) are described in Table A 4.1. Amplified DNA fragments were purified from agarose using NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel). Plasmid DNA and PCR products were quantified using a NanoDrop One instrument (Thermo Fisher Scientific) and were digested with the proper restriction endonucleases (Thermo Fisher Scientific) for 1 h at 37°C and purified using NucleoSpin® Gel and PCR Clean-up Kit. Ligations were performed for 1 h at room temperature using a T4 DNA ligase (Thermo Fisher Scientific). The constructions were transformed by heat shock into *E. coli* NZY5α competent cells

(NZYTech, Lisbon, Portugal). Super optimal broth with catabolite repression (SOC; NZYTech) was used for transformants recovery.

All plasmids herein constructed were verified by colony PCR using Dream *Taq* polymerase (Thermo Fisher Scientific) and digestion, and their sequences were further confirmed by sequencing (Eurofins) (Primers in Table A 4.1). After sequence confirmation the resulting plasmid with the assembled pathway, pRSFDuet_kfoCA_Zmugd, was transformed in *E. coli* K-12 MG1655 (DE3) and *E. coli* BL21 (DE3) strains for chondroitin production evaluation.

4.2.3. Mutant construction

The evaluated modifications to improve heterologous chondroitin production included membrane-bound lytic murein transglycosylase B (*mltB*), superoxide dismutase (*sodA*) and *N*-acetylglucosamine-1-phosphate uridyltransferase (*glmU*) overexpression, sugar phosphatase (*ybiV*) and lipid II flippase (*murJ*) underexpression, and β-*N*-acetylglucosaminidase (nagZ) deletion. *mltB*, *sodA* and *glmU* have been amplified from *E. coli* K-12 MG1655 (DE3) genome and cloned using the primers glmU_Fw, glmU_Rv, sodA_Fw, sodA_Rv, mltB_Fw and mltB_Rv (Table A 4.1) in the same conditions as previously described for chondroitin pathway construction. pETDuet-1 was used as vector to overexpress *glmU* and/or *sodA* while overexpression of *mltB* was conducted using pCDFDuet-1 as expression vector. pCRISPathBrick plasmid (Table 4.1) was used to implement the CRISPR interference (CRISPRi) system, by expressing and targeting a dCas9 (dead Cas9) [44]. The protospacer for *ybiV* and *murJ* targeting was inserted using Golden Gate assembly. For this procedure, annealing of pair of oligos (ybiV_UE_Fw, ybiV_UE_Rv for *ybiV*, Table A 4.1, and murJ_UE_Fw, murJ_UE_Rv or murJ2_Fw, murJ2_Rv for *murJ*, Table A 4.2) has been conducted by incubating 100 µmol of each oligo in T4 DNA ligase buffer at 95°C for 5 min and letting the temperature slowly decrease until room temperature was reached. Then the annealed oligo was mixed with the plasmid pCRISPathBrick, *Bsa*I (=*Eco*31I) and T4 DNA ligase. The mixture was incubated through 10 cycles of 5 min at 37°C and 10 min at 22°C, followed by 30 min at 37°C and 15 min at 75°C. All constructed plasmids were verified by colony PCR, digestion and their sequences were further confirmed by sequencing.

In the attempts to delete *nagZ* gene, the protocol for gene deletion using a CRISPR-Cas9 strategy was followed (Reisch and Prather, 2015). The primers used are described in Table A 4.2.

4.2.4. Flask fermentations

For chondroitin screening production tests in conical flask fermentations, *E. coli* K-12 MG1655 (DE3) and *E. coli* BL21 (DE3) harboring the heterologous chondroitin production pathway (pRSFDuet_kfoCA_Zmugd) have been used. Also, variations of these strains, with over and/or underexpression of genes, have been evaluated. Flasks of 250 mL with 50 mL lysogeny broth (LB) Miller (NZYTech) with proper antibiotics (i.e., depending on the plasmid(s) present in the strain, 50 μg/mL of kanamycin (Fisher), 100 μg/mL of ampicillin (Fisher Bioreagents), 100 μg/mL of spectinomycin (Alfa Aesar) and/or 34 μg/mL chloramphenicol (Alfa Aesar)) have been inoculated at 1% (v/v) with an overnight culture. The cultures were incubated at 37°C and 200 rpm until OD_{600nm} of 0.6-0.8 was reached. At this point, isopropyl β-D-1-thiogalactopyranoside (IPTG, 1 mM; NZYTech) was added to induce heterologous enzyme production and temperature was decreased to 30°C. The cultures were further incubated for 24 h. All assays were conducted in triplicate.

4.2.5. Bioreactor operation

The engineered *E. coli* K-12 MG1655 (DE3) strains were collected from LB plates supplemented with proper antibiotics and used to inoculate pre-inoculums of 10 mL of LB medium. The cultures were grown at 37°C and 200 rpm for about 22 h. Cells were harvested by centrifugation (4,000 x g, 15 min) and used to inoculate second phase pre-inoculums with 120 mL of defined medium (per liter: yeast extract 2 g (Labkem, Baldoyle, Ireland), K_2HPO_4 10 g (Panreac, Barcelona, Spain), KH_2PO_4 2 g (Panreac), $MgCl_2$ 0.1 g (VWR, Radnor, USA), sodium citrate 0.5 g (Panreac), $(NH_4)_2SO_4$ 1 g (Labkem), glucose 20 g (Acros Organics, Geel, Belgium) which were then shaken at 200 rpm and 37°C in 500 mL conical flasks for about 16 h. To inoculate 2-L bioreactors (DASGIP® Parallel Bioreactor System for Microbial Applications, Eppendorf, Hamburg, Germany) containing 400 mL of defined medium and proper antibiotics, pre-inoculum was used to start the inoculum with OD_{600nm} of 0.1. Batch fermentations were carried out operating the fermenters at constant dissolved oxygen feeding (DO-stat) mode, at 37°C, with initial agitation of 250 rpm, feeding 0.5 vvm of air (12 L/h), and maintaining pH at 7.0. The agitation was automatically regulated up to 800 rpm to maintain dissolved oxygen above 30%. A 2 M solution of NaOH (Sigma, St. Louis, USA) was used to control the pH. After 3 h (OD_{600nm} ~0.6), 1 mM IPTG was added to the culture and temperature was reduced to 30°C. The fermentation continued until glucose was completely consumed (until ~54 h). Samples were taken during the fermentation to monitor glucose concentration and OD_{600nm}.

For fed-batch fermentations, the batch phase was performed as previously described except that initial glucose concentration was 10 g/L. When glucose concentrations were reduced up to 2-3 g/L, the feeding phase was initiated by feeding a concentrated solution (80 g/L glucose and 18 g/L yeast extract) at the corresponding consumption rates calculated in the previous batch fermentations. All fermentations were conducted in duplicates.

4.2.6. Analytic methods

Samples from *E. coli* fermentations were centrifuged to separate cells from supernatant (4,000 x g, 15 min).

The supernatants from the end of flask fermentations were used to quantify extracellular chondroitin while the pellets were further processed to determine intracellular chondroitin.

Pellets (from ~50 mL of culture) were washed and then resuspended in deionized water (5 mL). The suspended cells were lysed by sonication with a microtip probe linked to Vibra-cell processor (Sonics, Newtown, CT, USA). Keeping the solution on ice during the procedure, short pulses of 3 s ON and 4 s OFF at 30% amplitude were performed until 5 min of active sonication was reached. The resulting lysate was centrifuged (12,000 x g 15 min) to remove insoluble material.

For intracellular chondroitin quantification, the lysates were treated with DNase (New England Biolabs, MA, USA) for 2 h at 37°C and then with proteinase K at final concentration of 2 mg/mL (NZYTech) for 2 h at 56°C. Then, the mixture was boiled for 5 min and centrifuged again (16,000 x g 20 min) to removed insoluble material.

Both extracellular and intracellular samples were precipitated by adding three volumes of cold ethanol and letting the mixture at 4°C overnight. The precipitate was collected by centrifugation at 4,000 x g for 10 min at 4°C. The precipitate was air-dried at room temperature overnight after which it was resuspended in deionized water. Insoluble material was removed by centrifugation (16,000 x g for 20 min). Chondroitin was quantified by uronic acid carbazole assay [45] using chondroitin sulfate (Biosynth, Staad, Switzerland) solutions as standards. Samples or standards (125 µL) were mixed with 750 µL sulfuric acid reagent (9.5 g/L sodium tetraborate, Supelco, Bellefonte, USA, dissolved in H_2SO_4 > 95%). The mixture was boiled for 20 min. Afterwards, 25 µL of carbazole reagent (1.25 g/L carbazole, Supelco, dissolved in absolute ethanol, Fisher) were added to the boiled samples. The mixture was boiled again for 15 min and cooled down for 15 min. The OD_{530nm} was then read in a microplate reader.

Protein expression in cultured *E. coli* K-12 MG1655 (DE3) carrying pETDuet-1, pETDuet_glmU, pETDuet_sodA, pCDFDuet-1, or pCDFDuet_mltB, was evaluated by sodium dodecyl sulfate

polyacrylamide gel electrophoresis (SDS–PAGE) [42] (4% stacking gel and 12% running gel). Samples (soluble and insoluble fractions of lysates) were mixed with 2x Laemmli Sample Buffer (65.8 mM Tris–HCl pH 6.8, 2.1% SDS, 26.3% glycerol, 0.01% bromophenol blue and 5% β-mercaptoethanol, from Fisher Scientific, JMGS, Sigma-Aldrich and AppliChem, respectively) and denatured at 95°C for 5 min. The protein marker used was Color Protein Standard—Broad Range (NEB, #77125). After electrophoresis, the gel was stained using Coomassie Blue R-250 (AppliChem) for 15 min and de-stained using distilled water. Samples from bioreactors were collected along the fermentations and analyzed in terms of glucose consumption and cellular growth. Final extracellular chondroitin was quantified using the carbazole method described above. Glucose consumption was monitored along the fermentations by quantifying glucose in supernatants through dinitrosalicylic acid (DNS) method [46]. 3,5-Dinitrosalicylic acid (Acros Organics), sodium potassium tartrate tetrahydrate (Panreac) and NaOH (Sigma) were used to prepare the DNS reagent, which was mixed with same volume of samples, boiled and cooled down by adding deionized water. The OD_{540nm} was then measured. The glucose concentrations were further confirmed by high performance liquid chromatography (HPLC) using a JASCO system associated with a refractive index (RI) detector (RI-2031), and an Aminex HPX-87H column from Bio-Rad kept at 60°C; the mobile phase used was 5 mM H_2SO_4 (Fisher) at a flow rate of 0.5 mL/min. All ODs were measured in a 96-well plate spectrophotometric reader Synergy HT (BioTek, Winooski, VT, USA), to establish the growth profile.

4.3. Results and Discussion

4.3.1. Bioinformatic results for mutant prediction

OptFlux was used to search for *E. coli* mutants that could produce higher amounts of chondroitin. Using this tool, it was not possible to predict the improvement of chondroitin production by combining gene knockouts (*data not shown*), probably because the competing pathways that use the intermediates are critical for cell growth. However, gene over- and underexpression search allowed the identification of several possible targets. The identified genes for under- and overexpression in the solutions of highest BPCY for each model are shown in Table 4.2.

Table 4.2. Optimization of chondroitin production using OptFlux. The optimization algorithm was run at least four times for each model. The predicted phenotype for the unmodified and modified strains (from the resulting solutions with highest biomass-product coupled yield (BPCY)) are shown. BPCY is calculated by OptFlux by multiplying biomass by product and then dividing by substrate consumed (in all cases being 10 mol/gDW/h), as predicted by pFBA simulation. Flux variability analysis (FVA) results are shown as minimum and maximum chondroitin obtained through pFBA for fixed biomass. Predicted biomass and chondroitin values are in units of mol/gDW/h.

Model	BPCY	Genes modified		Predicted phenotype (pFBA)		FVA	
		Underexpression	Overexpression	Biomass	Chondroitin	Min chondroitin	Max chondroitin
iB21_1397_c	-	-	-	0.9756	0.0000	-	-
iB21_1397_c	0.09607	*cmk*	*glmU*	0.3671	2.6168	2.6146	2.6345
	0.09607	*pyrH*	*glmM*	0.3671	2.6168	2.6146	2.6345
	0.09607	*pyrH*	*glmU*	0.3671	2.6168	2.6146	2.6345
	0.09200	*mltC*	*glmU*	0.6516	1.4119	1.4025	1.4401
	0.09200	*mltF*	*glmU*	0.6516	1.4119	1.4025	1.4401
	0.09200	*mltA*	*glmU*	0.6516	1.4119	1.4025	1.4401
	0.09200	*mltB*	*glmU*	0.6516	1.4119	1.4025	1.4401
	0.09200	*mltE*	*glmU*	0.6516	1.4119	1.4025	1.4401
	0.09200	*slt*	*glmU*	0.6516	1.4119	1.4025	1.4401
iECBD_1354_c	-	-	-	0.9756	0.0000	-	-
iECBD_1354_c	0.09200	*mltE*	*glmM*	0.6516	1.4119	1.4025	1.4401
	0.09200	*slt*	*glmM*	0.6516	1.4119	1.4025	1.4401
	0.09200	*mltC*	*glmM*	0.6516	1.4119	1.4025	1.4401
	0.09200	*mltA*	*glmM*	0.6516	1.4119	1.4025	1.4401
	0.09200	*amiABC, ampG*	*glmU*	0.6516	1.4119	1.4025	1.4401
	0.09200	*slt*	*glmU*	0.6516	1.4119	1.4025	1.4401
	0.09200	*mltB*	*glmM*	0.6516	1.4119	1.4025	1.4401

Table 4.3. Optimization of chondroitin production using OptFlux. The optimization algorithm was run at least four times for each model. The predicted phenotype for the unmodified and modified strains (from the resulting solutions with highest biomass-product coupled yield (BPCY)) are shown. BPCY is calculated by OptFlux by multiplying biomass by product and then dividing by substrate consumed (in all cases being 10 mol/gDW/h), as predicted by pFBA simulation. Flux variability analysis (FVA) results are shown as minimum and maximum chondroitin obtained through pFBA for fixed biomass. Predicted biomass and chondroitin values are in units of mol/gDW/h (*continuation*).

Model	BPCY	Genes modified		Predicted phenotype (pFBA)		FVA	
		Underexpression	Overexpression	Biomass	Chondroitin	Min chondroitin	Max chondroitin
iECBD_1354_c	0.09200	*mltF*	*glmM*	0.6516	1.4119	1.4025	1.4401
iEC1356_Bl21DE3_c	-	-	-	0.9767	0.000	-	-
iEC1356_Bl21DE3_c	0.09215	*mltE*	*glmU*	0.6519	1.4135	1.4015	1.4417
	0.09215	*oppC*	*glmU*	0.6519	1.4135	1.3268	1.4417
	0.09215	*mltC*	*glmU*	0.6519	1.4135	1.4015	1.4417
	0.09215	*oppB*	*glmU*	0.6519	1.4135	1.3268	1.4417
	0.09215	*oppF*	*glmU*	0.6519	1.4135	1.3268	1.4417
	0.09215	*amiA*	*glmU*	0.6519	1.4135	1.4015	1.4417
	0.09215	*amiB*	*glmU*	0.6519	1.4135	1.4015	1.4417
	0.09215	*oppD*	*glmU*	0.6519	1.4135	1.3268	1.4417
	0.09215	*mltA*	*glmU*	0.6519	1.4135	1.4015	1.4417
	0.09215	*mltB*	*glmU*	0.6519	1.4135	1.4015	1.4417
	0.09215	*amiC*	*glmU*	0.6519	1.4135	1.4015	1.4417
iJO1366_c	-	-	-	0.9824	0.000	-	-
iJO1366_c	0.09287	*mltB*	*glmM*	0.6531	1.4219	1.4000	1.4501
	0.09287	*mltE*	*glmM*	0.6531	1.4219	1.4000	1.4501

Most of the resulting solutions comprised two combined modifications, namely underexpression of one of the genes from cell wall biosynthesis and recycling pathways (such as: lytic transglycosylases MltABCEF and Slt; anhydromuropeptide permease AmpG; oligopeptide permeases *oppBCDF*; or *N*-acetylmuramoyl L-alanine amidases AmiABC) and overexpression of one of the genes responsible for the production of a chondroitin precursor (either glucosamine-1-phosphate *N*-acetyltransferase/ UDP-*N*-acetylglucosamine diphosphorylase GlmU or phosphoglucosamine mutase GlmM). Underexpression of cytidylate kinase *cmk* and UMP kinase *pyrH* genes were also identified, which catalyze reactions from the biosynthesis and salvage of pyrimidine ribonucleotides.

Overexpression of *glmM* and/or *glmU* are well reported metabolic engineering strategies for improving the production of chondroitin [47] or hyaluronic acid [48] in *E. coli*.

In Gram-negative bacteria, the cell-wall recycling is initiated in the periplasm by bacterial lytic transglycosylases such as Slt and MltABCEF, that degrade murein (peptidoglycan), the major cell wall component, at the glycosidic bond between *N*-acetylglucosamine (GlcNAc) and *N*-acetylmuramic acid (MurNAc), releasing a distinctive 1,6-anhydro-*N*-acetyl-β-D-muramyl (AnhydroMurNAc) product. The anhydromuropeptide permease AmpG specifically transports the resulting AnhydroMurNAc-containing muropeptides, from the periplasm into the cytoplasm [49–51]. MurNAc-L-Ala amidases AmiABC, are another class of peptidoglycan degrading enzymes in the periplasm, that cleave the amide bond between MurNAc and the stem peptide in peptidoglycan, releasing murein tri-, tetra-, and pentapeptides into the periplasm from which they diffuse out of the cell, or enter the cytoplasm via the MppA-OppBCDF permease system [52]. The membrane oligopeptide permeases OppBCDF and the murein tripeptide ABC transporter periplasmic binding protein MppA constitute an ABC transporter which is involved in the recycling of murein tripeptide from either exogenous sources or from amidases action [52,53]. Once in the cytoplasm, the muropeptides are degraded by amidase AmpD, β-*N*-acetylglucosaminidase NagZ, and LD-Carboxypeptidase LdcA, and their constituent components are used for Lipid II biosynthesis. The Lipid II assembled in the cytoplasm is delivered to the periplasm for the *de novo* synthesis of peptidoglycan [51]. Consequently, peptidoglycan biosynthesis is the main competing pathway that redirects precursors from chondroitin biosynthesis. Despite these reactions result in the production the chondroitin precursor glucosamine 6-phosphate (GlcN-6-P) that can also be produced by glucose, they consume the intermediate UDP-acetylglucosamine (UDP-GlcNAc) (Figure 4.1). Therefore, underexpressing cell wall recycling pathway genes might lead to more available precursors, as predicted by the models.

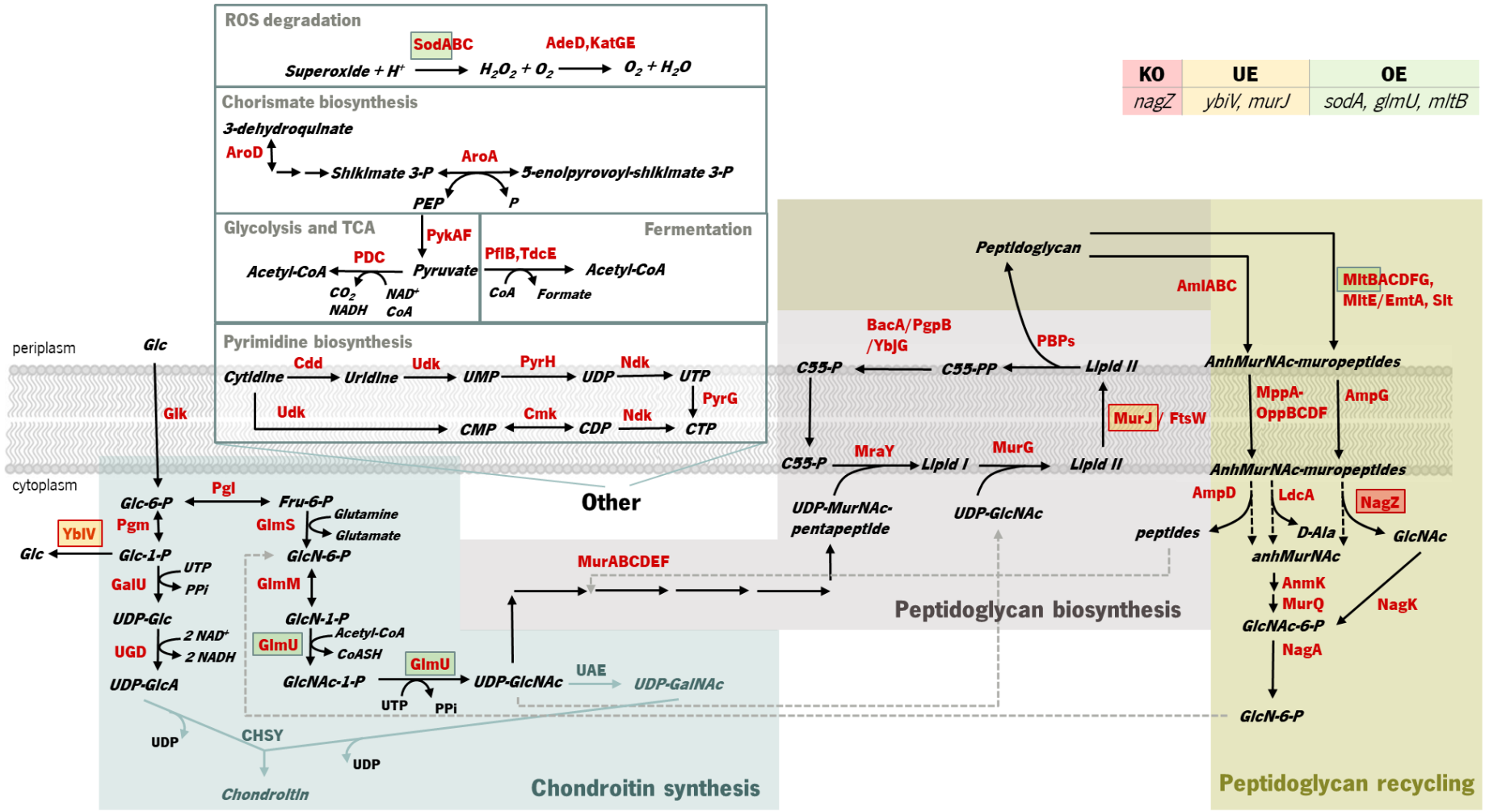


Figure 4.1. Identified targets for gene modification to potentially improve chondroitin heterologous production in *Escherichia coli* and their role in bacterial metabolism. The target for knock-out (KO) is marked in a red square, underexpressions (UE) are marked in orange squares, and overexpressions (OE) are marked in green squares. MltB, MurJ and NagZ are involved in cell wall biosynthesis and recycling, while GlmU produces the chondroitin precursor uridine-diphosphate (UDP)-*N*-acetylglucosamine and YbiV

dephosphorylates the precursor glucose 1-phosphate. SodA is not directly related with the metabolism of chondroitin synthesis intermediates. Enzyme and compounds abbreviations: AdeD: adenine deaminase; AmiABC and AmpD: *N*-acetylmuramoyl-*L*-alanine amidases; AmpG: anhydromuropeptide permease; AnmK: anhydro-N-acetylmuramic acid kinase; AroA: 3-phosphoshikimate 1-carboxyvinyltransferase; BacA: undecaprenyl pyrophosphate phosphatase; C55: *di-trans*,*octa-cis*-undecaprenyl; Cdd: cytidine/deoxycytidine deaminase; CHSY: chondroitin synthase; Cmk: cytidylate kinase; FtsW: peptidoglycan glycosyltransferase; GalU: uridine-triphosphate(UTP):glucose-1-phosphate uridylyltransferase; Glk: glucokinase; GlmU: glucosamine-1-phosphate *N*-acetyltransferase/ UDP-*N*-acetylglucosamine diphosphorylase; GlmM: phosphoglucosamine mutase; GlmS: glucosamine-6- phosphate synthase; KatGE: catalases; LdcA: murein tetrapeptide carboxypeptidase; MltABCDFG: membrane bound lytic transglycosylases; MppA-OppBCDF: oligopeptide permeases complex with muropeptide-binding protein; MraY: phospho-*N*-acetylmuramoyl-pentapeptide-transferase; MurG: *N*-acetylglucosaminyl transferase; MurJ: lipid II flippase; MurQ: *N*-acetylmuramic acid 6-phosphate etherase; NagA: *N*-acetylglucosamine-6-phosphate deacetylase; NagK: *N*-acetyl-D-glucosamine kinase; NagZ: β-*N*-acetylglucosaminidase; Ndk: nucleoside diphosphate kinase; PBPs: penicillin-binding proteins; PDC: pyruvate dehydrogenase complex encoded by genes *pdhA*, *pdhB* and *lpd*; PflB: pyruvate formate-lyase; Pgi: glucose-6-phosphate isomerase; Pgm: phosphoglucomutase; PgpB: phosphatidylglycerophosphatase B; PykAF: pyruvate kinases; PyrG: CTP synthetase; PyrH: UMP kinase; Slt: soluble lytic transglycosylase; SodA: superoxide dismutase; TdcE: 2-ketobutyrate formate-lyase/pyruvate formate-lyase 4; UAE: UDP-*N*-acetylglucosamine 4-epimerase; Udk: uridine/cytidine kinase; UGD: UDP-glucose 6-dehydrogenase; YbjG: undecaprenyl pyrophosphate phosphatase; YbiV: sugar phosphatase.

The pyrimidine nucleotide metabolic process was also identified as a target pathway to be modified to increase chondroitin production. Cmk rephosphorylates cytidylate (CMP) and deoxycytidylate (dCMP). Underexpression of *cmk* (included in an optimization solution using iB21_1397_c model) could lead to improved uridine-5′-triphosphate (UTP) pool as observed in a *cmk* deletion mutant [54]. This is beneficial for chondroitin production because UTP is a co-factor in chondroitin-precursors biosynthesis steps catalyzed by GalU and GlmU. Uridine monophosphate (UMP) kinase PyrH is also involved in the biosynthesis of pyrimidine nucleotides, specifically catalyzing the phosphorylation of UMP to UDP. In this case, targeting this gene for underexpression, as obtained in an optimization solution using iB21_1397_c model, has not an obvious explanation and no literature was found to support that the underexpression of *pyrH* would improve UTP pool, or in other manner be directly beneficial for chondroitin biosynthesis. However, unexpected targets can have other mechanisms for improving product titers. Targeting genes from purine and pyrimidine biosynthesis pathways for underexpression (including *cmk* and *pyrH*) has been applied for antibody production [55] as a growth decoupling strategy to increase product formation. Inhibiting excess biomass formation allows for carbon to be utilized efficiently for product formation instead of growth, resulting in increased product yields and titers [55]. Interestingly, in Landberg et al. (2020) study, from the reported 21 selected targets for underexpression, repression of *cmk* consistently resulted in higher product titers. On the other hand, *pyrH* underexpression did not affect production. It is also relevant to mention that the models herein used are stoichiometric models, therefore, they do not comprehend kinetic information that would be valuable for this analysis. The model iB21_1397_c has the UMP kinase reaction set as reversible and as being catalyzed by both *pyrH* and *cmk* encoded enzymes, although it is known that the specific favored reactions are, *in vivo*, the phosphorylation of UMP and phosphorylation of CMP (or dCMP), respectively [56,57]. Also, *pyrH* is a known essential gene for *E. coli* growth [58]. Therefore, despite *pyrH* has been identified *in silico* as a target for underexpression, this might not be the best target for improving chondroitin production.

Generally, solutions herein obtained with different models were composed by similar gene modifications which shows the robustness of the optimization. Despite a maximum of 10 modifications were allowed when setting the optimization, usually the solution comprised 2 gene modifications (the highest number of modifications was 5 with model iECBD_1354_c, Table 4.2), thus suggesting that the solutions can be further improved using different approaches besides the ones available in OptFlux.

The FVA analysis shows the range of flux distributions of chondroitin. To evaluate the robustness of a solution, the minimum and maximum chondroitin produced by the mutant, for the optimal biomass value, are calculated and the difference should be minimal. The solutions that included oligopeptide permeases

oppBCDF genes underexpression (iEC1356_Bl21DE3_c, Table 4.2) were the solutions that resulted in greater difference between minimum and maximum chondroitin production. This suggests that those solutions are less robust.

To seek for more robust mutants, a different computational approach was performed using the MEWpy tool developed for Python [35]. Using this approach, instead of maximizing one objective function (BPCY), two objective functions were defined, namely the BPCY and WYIELD, which allow to guide the evolutionary algorithm onto more robust solutions. BPCY was calculated by MEWpy by multiplying biomass by product, based on pFBA predictions. WYIELD is the weighed sum of the minimum and maximum product fluxes, constrained to a minimal growth.

Since the model iB21_1397_c achieved the best optimization results (highest BPCY of 0.09607 and highest chondroitin flux of 2.6168 mol/gDW/h), this model was selected to use in the MEWpy tool. The same environmental conditions were set, and the evolutionary optimization was run. This procedure resulted in 76 solutions, with 39 identified targets for gene modification. Type and frequency of each gene modification throughout all solutions was analyzed in Figure 4.2.

The results from the best solutions obtained using the MEWpy tool are shown in Table 4.4. Solution 1 and 2 were the ones with the highest BPCY score, while solutions 5, 6 and 7 exhibited the highest WYIELD. Phenotype simulation was then performed using OptFlux, with pFBA as simulation method, for each solution and those results are also shown in Table 4.4.

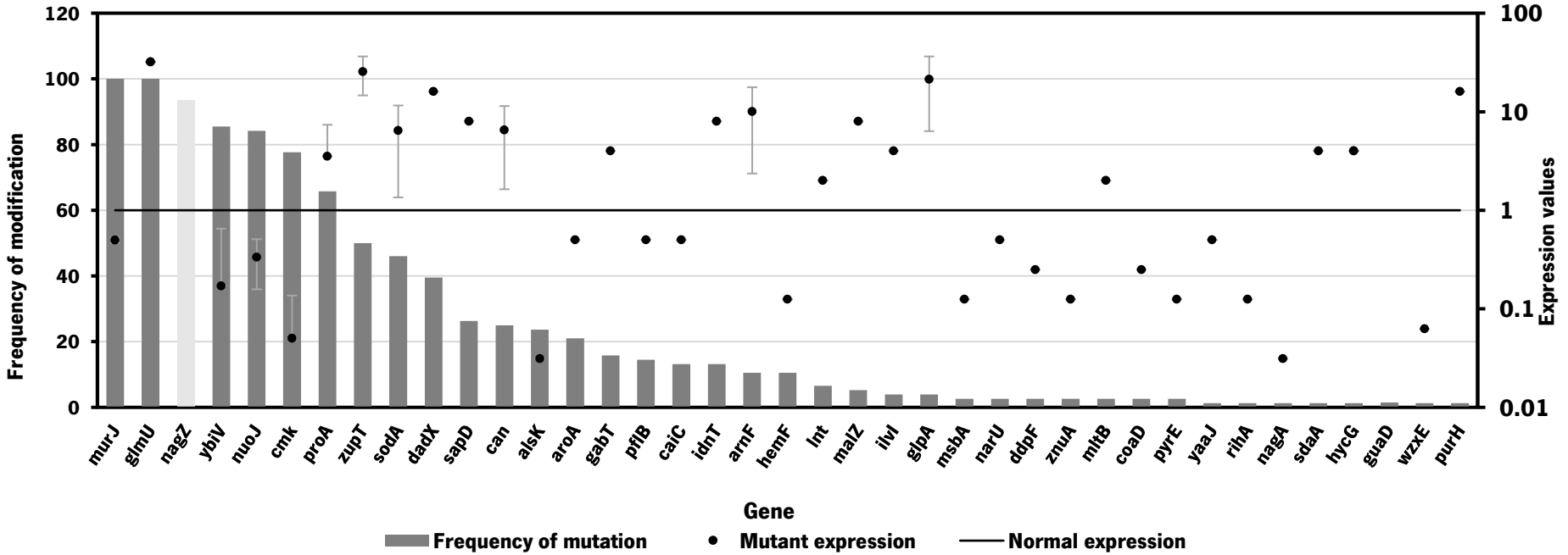


Figure 4.2. Frequency of gene modifications in the 76 solutions from strain optimization using MEWpy tool. The mutant expression (in dots) represents the average expression value. Mutant expressions higher than 1 represent overexpressions while values of expression lower than 1 represent underexpressions. Deletions are represented in light grey bars.

Table 4.4. Optimization results obtained for iB21_1397_c model using the MEWpy tool in Python, and the corresponding relevant fluxes according to phenotype simulations using parsimonious flux balance analysis (FBA) in OptFlux. BPCY was calculated by multiplying biomass by product and then dividing by substrate consumed, as predicted by pFBA. WYIELD is the weighed sum of the minimum and maximum product fluxes. Flux variability analysis (FVA) results are shown as minimum and maximum chondroitin obtained for fixed biomass. Predicted biomass and chondroitin values are in units of mol/gDW/h.

Model	BPCY	WYIELD	Genes modified			Predicted phenotype (pFBA)		FVA	
			Knock-out	Underexpression	Overexpression	Biomass	Chondroitin	Min chondroitin	Max chondroitin
1	0.08840	2.91104	*nagZ*	*ybiV, alsK, aroA, pflB, murJ, narU*	*sodA, glmU, mltB*	0.3040	2.9079	2.9034	2.9289
2	0.08887	2.90887	*ybiV*	*alsK, aroA, murJ, proA*	*idnT, lnt, glmU, gabT, malZ*	0.3056	2.9079	2.9033	2.9220
3	0.04316	3.14641	*nagZ, ybiV*	*pflB, murJ, msbA, cmk*	*sodA, sapD, glmU*	0.1221	2.9150	2.9150	3.6863
4	0.04316	3.14651	*nagZ*	*ybiV, pflB, murJ, msbA*	*sodA, sapD, idnT, glmU*	0.1221	2.9150	2.9150	3.6866
6	0.04340	2.95361	*nagZ*	*ybiV, alsk, aroA, murJ, pyrE*	*sodA, glmU, zupT, mltB*	0.1221	2.9354	2.6382	3.6896
Selected genes	0.08840	2.91104	*nagZ*	*ybiV, murJ*	*sodA, glmU, mltB*	0.3040	2.9079	2.9034	2.9289

The MEWpy method allowed to identify new mutants with improved chondroitin production, as high as 2.9150 mol/gDW/h, while the highest production obtained from OptFlux solutions was 2.6168 mol/gDW/h (Table 4.2).

As occurred in the optimization results obtained using OptFlux, solutions from MEWpy included targets from peptidoglycan biosynthesis and recycling pathways (*nagZ*, *murJ*, *mltB*) and pyrimidine ribonucleotides salvage pathway (*cmk*). Maltodextrin glucosidase (*malZ*) releases glucose from malto-oligosaccharides as part of the glycogen degradation pathway [59], and its suggested overexpression might be to improve glucose and glucose 6-phosphate availability through recycling of those carbon stocks. Many of these new solutions included varied transporters (*narU, znuA, sapD, msbA, idnT*) that were not found to be related with chondroitin production. Other genes that were not found to be related to chondroitin biosynthesis pathway, co-factor production or competing pathways were: D-allose kinase *alsK* that catalyzes the phosphorylation of D-allose to D-allose 6-phosphate; *aroA* that encodes 3-phosphoshikimate 1-carboxyvinyltransferase which is involved in chorismate pathway, leading to aromatic amino acids biosynthesis; enzyme apolipoprotein *N*-acyltransferase (*lnt*) that transfers a fatty acid from phospholipid to the amino terminus of a diacylglycerol prolipoprotein as part of lipoprotein posttranslational modification pathway; *purH* that encodes the bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/ inosinic acid cyclohydrolase involved in the *de novo* biosynthesis of purine nucleotides; glutamate-5-semialdehyde dehydrogenase encoded by *proA* that is from the L-proline biosynthesis; 4-aminobutyrate aminotransferase GabT that is involved in 4-aminobutyrate (GABA) degradation.

Overexpression of *glmU*, responsible to produce a chondroitin precursor UDP-GlcNAc, was included in most solutions, as occurred in those resulting from OptFlux.

Gene *nagZ* encodes β-*N*-acetylglucosaminidase (NagZ), an enzyme involved in peptidoglycan recycling. NagZ acts specifically by hydrolyzing the β-1,4 glycosidic bond, removing *N*-acetyl-glucosamine (GlcNAc) residues from peptidoglycan fragments that have been excised from the cell wall during growth [60]. Gene *murJ* encodes for lipid II flippase (MurJ) which flips lipid II from the inner face of the inner membrane to the outer face, being essential for peptidoglycan polymerization. The gene *mltB*, described above, was identified, in the best solution from MEWpy, as a target for slight overexpression (expression value: 2), contrarily to the solutions obtained in OptFlux, where it has been suggested for underexpression. This might be related with the combination and number of genes being different in the solutions, and *mltB* overexpression might be required to compensate for the negative effect on growth, caused by changes in the expression of several genes. The sugar phosphatase YbiV has been indicated in most solutions for

either underexpression or knockout. This is a logical solution as this enzyme redirects metabolic flux from chondroitin production by phosphating glucose-6-phosphate into glucose [61] (Figure 4.1). Pyruvate formate-lyase encoded by *pflB,* which catalyzes non-oxidative cleavage of pyruvate to acetyl-CoA and formate in anaerobically growing cells, was included in some solutions (14%) as underexpressed. The underexpression of *pflB* would decrease the flux through this reaction, leading to an increased availability of pyruvate, and consequently improved glucose-6-phosphate or fructose-6-phosphate levels. In fact, *pflB* deletion has been a common reported target for increased dicarboxylic acid production [62–64].

Other common gene identified for overexpression was the superoxide dismutase *sodA*. This gene is expressed in response to oxidative stress and acts by destructing toxic superoxide radicals that are naturally produced during respiratory growth [65]. No direct relationship with chondroitin production improvement was found.

Based on gene modification frequency, and after confirming that the phenotypes did not vary much from the originally proposed solution (Table 4.4), the gene selection was narrowed, from genes present in the best solution (Solution 1), for genes with more potential to engineer efficient *E. coli* strains. These selected modifications were: *nagZ* deletion, *ybiV* and *murJ* underexpressions and *sodA*, *glmU* and *mltB* overexpressions. A schematic representation of the affected pathways with these gene modifications is presented in Figure 4.1.

The individual modifications and cumulative modifications of the selected genes were experimentally implemented, to study which gene combinations could benefit chondroitin production the most without compromising *E. coli* growth.

4.3.2. *In vivo* validation of bioinformatic results

4.3.2.1. Chondroitin production using engineered *E. coli*

The biosynthetic pathway for heterologous production of chondroitin was constructed by cloning the genes *kfoC* and *kfoA* from *E. coli* O5:K4:H4 and *Zmugd* in the plasmid pRSFDuet-1. The amplification products are shown in Figure A 4.1. The assembled pathway was expressed in both *E. coli* K-12 MG1655 (DE3) and *E. coli* BL21 (DE3) to evaluate chondroitin production (Figure 4.3).

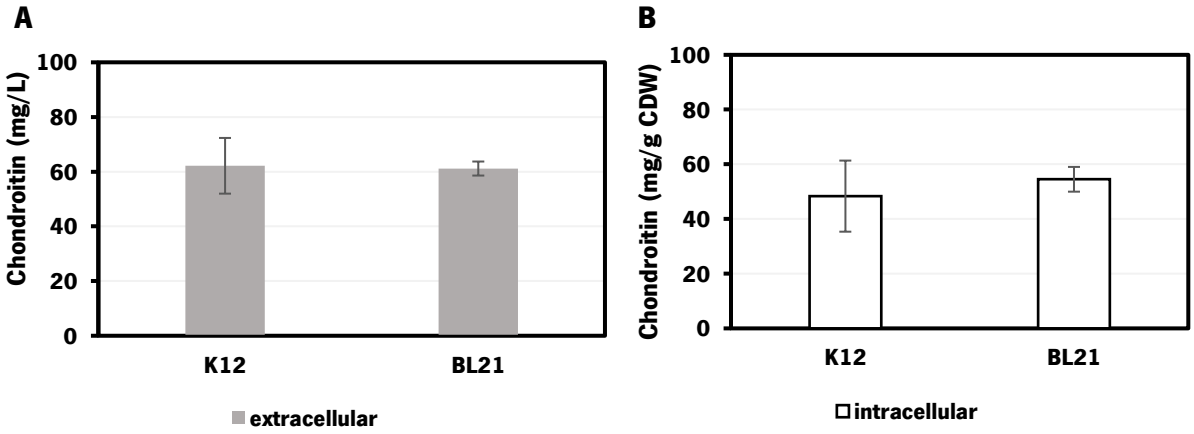


Figure 4.3. Chondroitin production in *Escherichia coli* engineered strains with chondroitin biosynthetic pathway in flask fermentation: A – extracellular chondroitin in mg/L and B – intracellular chondroitin in mg/g CDW, using K12 – *E. coli* K-12 MG1655 (DE3) and BL21 – *E. coli* BL21 (DE3), both carrying pRSFDuet_kfoCA_Zmugd. Assays were performed in triplicate. CDW – cell dry weight.

E. coli K-12 harboring the biosynthetic pathway for chondroitin production was able to produce 62 ± 10 mg/L of extracellular chondroitin and 48 mg/g of cell dry weight (CDW) of intracellular chondroitin. *E. coli* BL21 was able to produce 61 mg/L of extracellular chondroitin and 55 mg/g CDW of intracellular chondroitin. These results are higher than the ones obtained in a recent work [66] that reported a production of intracellular sulfated chondroitin of 126.64 µg/g CDW and 13.14 µg/g CDW using *E. coli* O5:K4:H4 and *E. coli* K-12 MG1655, respectively, in flasks. These strains harbored a biosynthetic pathway for chondroitin production similar to the one herein used (*kfoC, kfoA* and *kfoF*, naturally present in *E. coli* O5:K4:H4), but also expressed a chondroitin-4-*O*-sulfotransferase and lacked PAPS reductase *cysH* for the sulfation of chondroitin.

Extracellular chondroitin is more commonly measured, possibly because of its ease of purification and quantification, which makes it of greater interest for biotechnological production. Extracellular chondroitin production using engineered *E. coli* in shake flask cultivation has been reported to achieve concentrations from 0.01076 to 1739 mg/L [32,66–70], depending on the host, biosynthetic pathway, chassis optimizations and on culture conditions.

The highest chondroitin titers reported have been obtained using engineered *E. coli* O5:K4:H4, which naturally produces a fructosylated chondroitin, in defined medium, and in batch or fed-batch fermentations in bioreactors. For instance, a three-phase fermentation with pathogenic *E. coli* O5:K4:H4 overexpressing gene from transcription antitermination protein *rfaH* led to 9.2 g/L of chondroitin using glucose as substrate, and in a larger scale, the same strain produced 9 g/L using glycerol as substrate

[67]. To avoid the risks of using pathogenic bacteria, there have been efforts to construct alternative hosts to be efficient for chondroitin production using metabolic engineering strategies. A *Bacillus subtilis* 168, engineered with *kfoC*, *kfoA* expression and *tuaD* up-regulation, growing on sucrose, has achieved 5.22 g/L of chondroitin in fed-batch fermentation [39].

As the culture medium herein used for these initial screening tests in flasks was LB, the optimal production of chondroitin could be further optimized using different culture conditions. Based on work with highest reported chondroitin production [67], a defined medium was used for further fermentations in bioreactor (sections 3.2.4 and 3.2.5).

As the difference in chondroitin production between the two hosts was not significant (Figure 4.3), the *E. coli* K-12 MG1655 (DE3) was selected for further engineering strategies.

4.3.2.2. Construction of engineered *E. coli* strains based in bioinformatic optimizations

Based on the bioinformatics results, the selected modifications to be implemented towards an enhancement of chondroitin production were the overexpression of *sodA*, *glmU* and *mltB*; the deletion of *nagZ*; and the underexpression of *ybiV* and *murJ*.

The genes *sodA*, *glmU* and *mltB* for overexpressions were amplified (Figure A 4.2A) and cloned in pETDuet-1 or pCDFDuet-1. The plasmid with lower copy number (pCDFDuet-1) was used for *mltB* since, from the overexpressions predicted, it was the one with lower expression value (Figure 4.2). Their expression was confirmed by SDS-PAGE (Figure A 4.2B).

The *nagZ* knockout was attempted multiple times in *E. coli* K-12 MG1655 (and afterwards in *E. coli* BL21) using a CRISPR-Cas9 strategy [71] (schematized on Figure A 4.3), but it was unsuccessful. The primers used for this strategy are described in Table A 4.2. Although this gene is reported as non-essential for *E. coli* growth and has been previously deleted [72] using a different recombination-based strategy [73], it has a described role in cell wall biosynthesis as was previously mentioned. Therefore, it is possible that a strain lacking *nagZ* could be more susceptible to the antibiotics used as selective markers in the attempted CRISPR-Cas9 strategy (spectinomycin and chloramphenicol). A deletion of this gene could affect the cell wall integrity or permeability, that is maintained by the coordinated and regulated action of enzymes involved in the peptidoglycan synthesis and recycling [74]. In fact, the improved sensibility to β-lactam antibiotics in Gram-negative strains lacking *nagZ* compared to wild-type strains has been widely reported [60,74,75]. If the cells are more susceptible to the antibiotics used in the selection medium, then it could affect the growth and survival of the cells during the gene editing process. The selective pressure applied by the antibiotics could be too strong, resulting in a decrease in the number of cells that

survive and grow on the selection medium. One possible solution to this issue could be to use lower concentrations of the antibiotics to reduce the effect on growth of cells lacking *nagZ*. The integration of the chondroitin pathway genes in the genome without maintenance of antibiotic resistance markers can also be an efficient strategy to reduce the toxic effect of antibiotics.

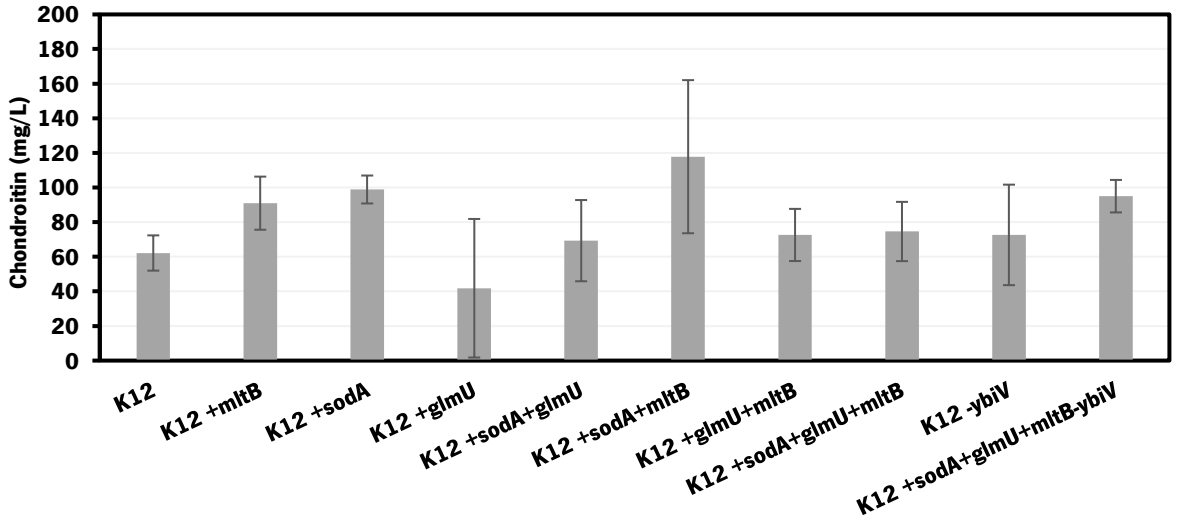
Regarding genes underexpressions, a CRISPRi system [44] was designed and constructed. In this strategy, a modified version of the caspase 9 protein commonly called as dead Cas9 (dCas9), which does not have the nuclease activity but maintains sequence-specific double stranded DNA-binding capability, is expressed to target the gene, ultimately repressing its expression. The CRISPRi system for *ybiV* underexpression was successfully constructed and evaluated. However, underexpression of *murJ* was not evaluated because construction of targeting protospacer has failed. A schematic representation of the strategy used can be found on Figure A 4.3. The primers used in the attempts of constructing the protospacer for *murJ* underexpression are described in Table A 4.2.

The individual and cumulative genetic modifications (*sodA*, *glmU*, and *mltB* overexpressions and/or *ybiV* underexpression) were further evaluated *in vivo*.

4.3.2.3. Chondroitin production using *E. coli* engineering strains based in bioinformatic optimizations

Chondroitin production in shake flasks

The solutions obtained though bioinformatics were further validated *in vivo*. Each modification was individually evaluated in the engineered *E. coli* strain harboring the chondroitin pathway already expressing 3 heterologous genes (*kfoC*, *kfoA* and *Zmugd*), through shake flasks fermentations (Figure 4.4) to seek for mutants with improved chondroitin production carrying as few modifications as possible.



	K12	K12 +*mltB*	K12 +*sodA*	K12 +*glmU*	K12 +*sodA* +*glmU*	K12 +*sodA* +*mltB*	K12 +*glmU* +*mltB*	K12 +*sodA* +*glmU* +*mltB*	K12 -*ybiV*	K12 +*sodA* +*glmU* +*mltB* -*ybiV*
Yield chondroitin/OD (mg/(L.OD))	72	95	192	47	106	259	133	149	93	190

Figure 4.4. Chondroitin production in *E. coli* K-12 MG1655 harboring the chondroitin biosynthetic pathway (K12, genes *kfoC*, *kfoA* and *Zmugd*) with additional modifications: *sodA*, *glmU* and *mltB* overexpressions and/or *ybiV* underexpression.

In these screening shake flasks experiments, *E. coli* K-12 mutants were able to produce extracellular chondroitin from 42 to 118 mg/L. Despite the differences in chondroitin production were not significant, mutants overexpressing *sodA* or *mltB,* or the one overexpressing both these two genes, seemed to be the most promising ones. As the cumulative effect of both overexpressions was not significantly better than the individual mutations, the two mutants containing individually overexpressed *sodA* or *mltB* were selected for further scale-up studies, in a more suitable culture medium for chondroitin production.

4.3.2.4. Chondroitin production in bioreactor (Batch experiments)

The two selected mutants were cultured in a bioreactor operated in batch starting with 20 g/L of glucose and compared to the control (*E. coli* K-12 harboring pRSFDuet_kfoCA_Zmugd). The strain growth and the glucose consumption were monitored (Figure 4.5).

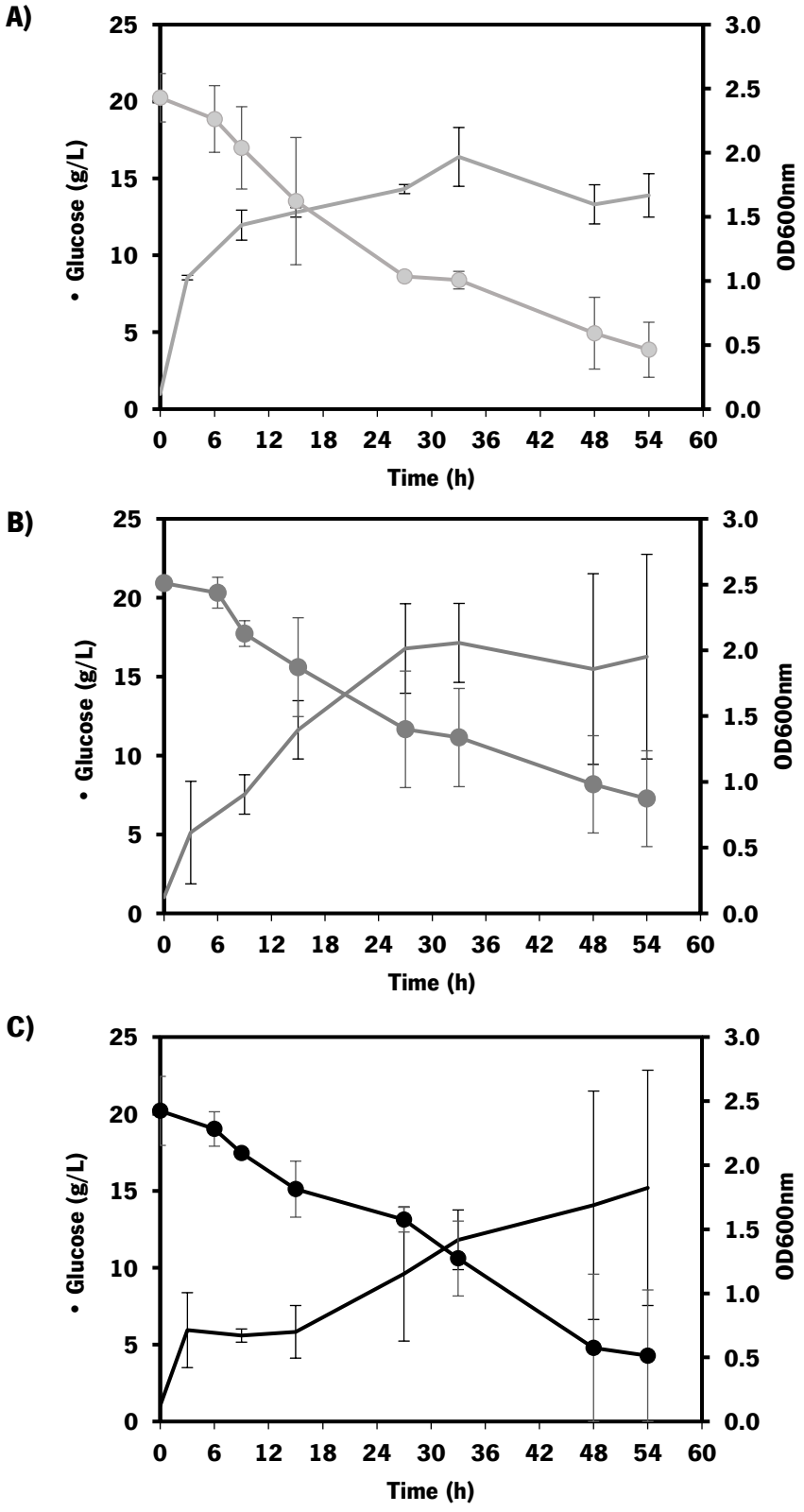
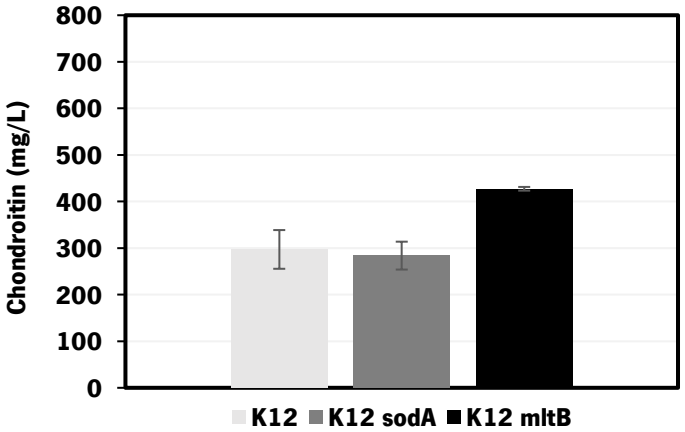


Figure 4.5. Growth and substrate consumptions curves of *E. coli* K-12 MG1655 DE3 mutants: **A)** carrying pRSFDuet_kfoCA_Zmugd (control - only containing chondroitin biosynthetic pathway); **B)** carrying

pRSFDuet_kfoCA_Zmugd and pETDuet_sodA; **C**) carrying pRSFDuet_kfoCA_Zmugd and pCDFDuet_mltB. Average of 2 Batch assays starting with 20 g/L glucose. Dots (•) indicate glucose concentration and lines (-) optical density at 600 nm (OD600nm).

The mutants with the selected overexpressions showed more variability between assays but the growth curve was similar to the control which indicates that the growth was not significantly affected by the additionally introduced modifications.

The chondroitin production in the end of fermentation (54 h) was evaluated for each engineered strain (Figure 4.6).



	K12	K12 *sodA*	K12 *mltB*
Yield chondroitin/OD (mg/(L.OD))	178	180	313
Yield chondroitin/glucose (mg/g)	19	22	32

Figure 4.6. Chondroitin production from cultured *E. coli* K-12 MG1655 (DE3) harboring pRSFDuet_kfoCA_Zmugd (K12) and its counterparts with additional *sodA* or *mltB* overexpressions, in bioreactors operated in batch mode.

The results obtained showed that *E. coli* K-12 MG1655 engineered with chondroitin production pathway and *mltB* overexpression performed better in terms of extracellular chondroitin production, achieving a concentration of 427 mg/L in 54 h. Regarding the yields per OD or substrate consumption, both mutants expressing *mltB* or *sodA* showed better results compared to the control host. Although the differences between the control and *sodA* mutant are not significant, *mltB*-overexpressing strain had improved yields of 1.7-fold.

4.3.2.5. Chondroitin production in bioreactor (Fed-batch experiments)

The most promising mutant (*E. coli* K-12 MG1655 overexpressing *kfoC*, *kfoA*, *Zmugd* and *mltB*) was further cultivated under fed-batch conditions and compared to the control (strain without *mltB* overexpression). The growth and substrate consumption are described on Figure 4.7.

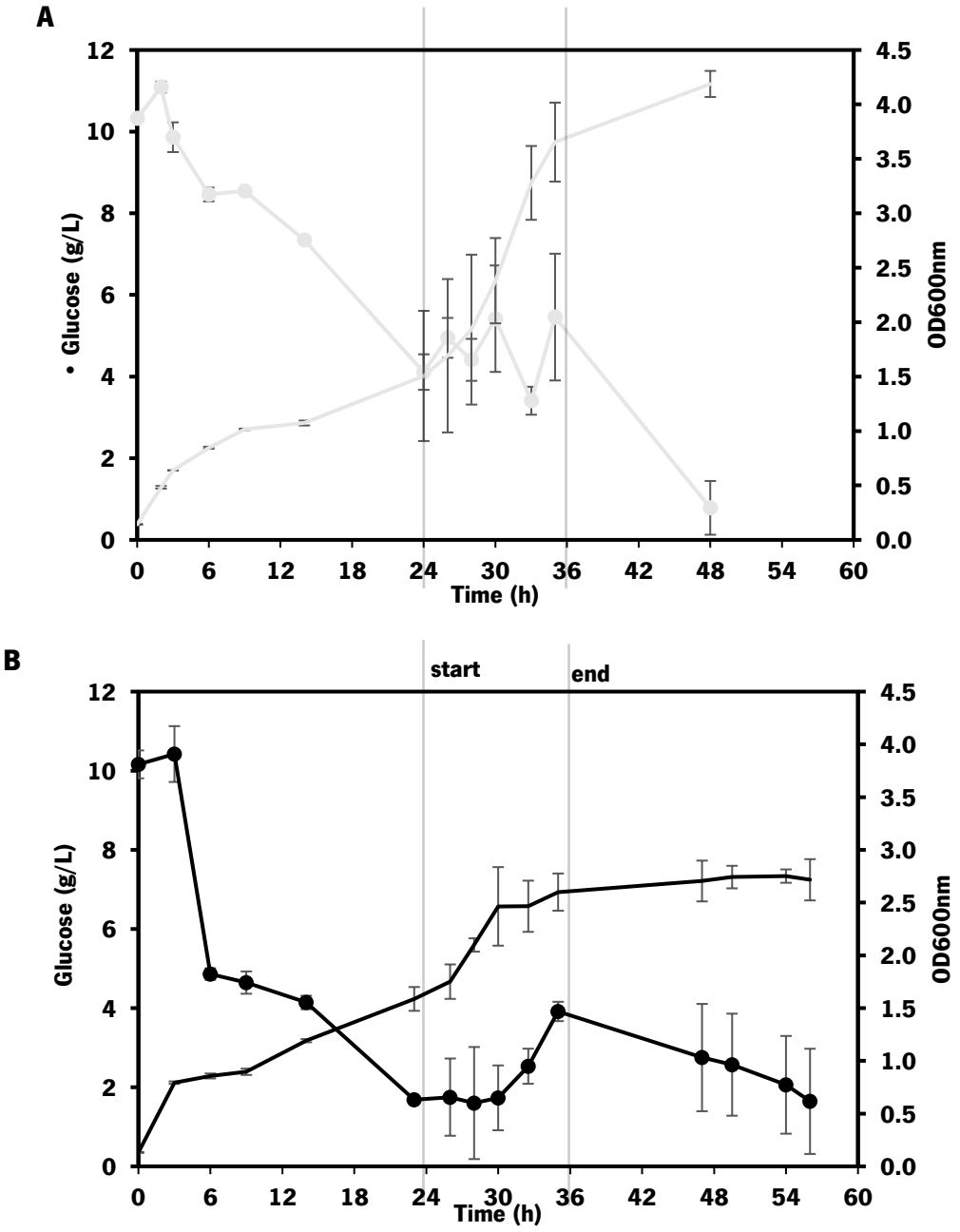
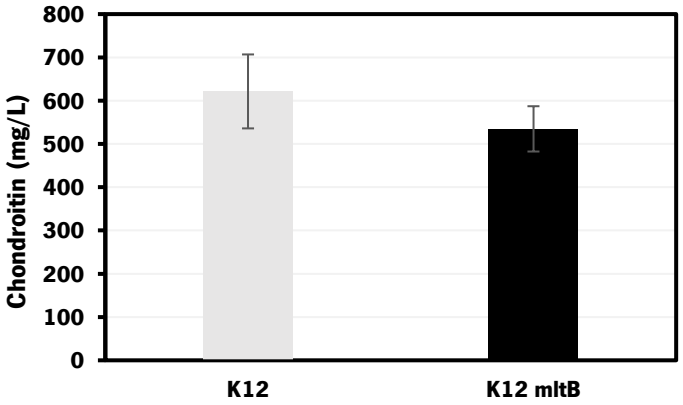


Figure 4.7. Growth and substrate consumptions curves of: **A:** *E. coli* K-12 MG1655 (DE3) control (only containing chondroitin biosynthetic pathway - pRSFDuet_kfoCA_Zmugd); **B:** *E. coli* K-12 MG1655 (DE3) carrying pRSFDuet_kfoCA_Zmugd and pCDFDuet_*mltB*. Average of 2 Fed-Batch assays starting with 10 g/L glucose. The feeding was started after approximately 24 h until 36 h after inoculum.

As expected, the growth of engineered *E. coli* K-12 in bioreactor was greatly improved by changing to fed-batch mode, comparing to the fermentations in batch operation mode. However, this effect was more evident on the control strain than on the one overexpressing *mltB*.

The chondroitin production in the end of fermentation was also evaluated (Figure 4.8).



	K12	K12 *mltB*
Yield chondroitin/OD (mg/L.OD)	159	196
Yield chondroitin/glucose (mg/g)	42	46

Figure 4.8. Chondroitin production from cultured *E. coli* K-12 MG1655 harboring pRSFDuet_kfoCA_Zmugd (K12) and its counterpart with additional *mltB* overexpression, in bioreactors operated in fed-batch mode.

Using the *E. coli* K-12 MG1655 engineered strains lacking or containing *mltB* overexpression, chondroitin concentration at the end of fed-bath fermentation achieved 621 mg/L and 535 mg/L, respectively. Despite the chondroitin production being lower in the mutant strain overexpressing the *mltB* gene, the yields based on growth (OD) and on glucose were higher compared to the parent mutant. This is justified with the lower growth of the *mltB*-expressing strain previously discussed (Figure 4.7).

Chondroitin production has benefited from changing the operation mode to the fed-batch for both strains, resulting in 2.1 and 1.3 -fold increase in chondroitin titers, for control strain and *mltB* strain, respectively. This is in accordance with other works in the literature that showed that fed-batch fermentations achieved the higher titers comparing to batch or shake flasks fermentations [32,67,68,76]. Although yields on biomass for both strains were lower than the ones obtained with batch, the yields on glucose consumption were 2.2 and 1.4 times improved, for control and *mltB* strain, respectively. It is common for the fed-batch fermentations to result in much higher cell density, which occurred in this work, and if the growth is not accompanied by product formation at the same range, the yields in biomass are naturally lower. Nevertheless, when comparing the yields, it is evident that *mltB*-overexpressing strain was consistently the best performer.

Although *mltB* overexpression was not a logical modification to improve chondroitin production, it has been predicted in some solutions from computational optimization, including the one with highest BPCY (Solution 1, Table 4.4). In the bioreactor batch experiments, the *mltB*-overexpressing mutant was indeed the best performing strain in terms of chondroitin production and its growth capability was similar to the strain containing only the pathway, without further modifications. When in fed-batch, the mutant with *mltB* up-regulation did not grow as much as the control. This can be due to the fact that the *E. coli* host used

was already expressing 3 heterologous genes for producing chondroitin, and further *mltB* overexpression might have constrained the bacterium growth as a result of an increase in the metabolic burden, which became more significant when higher cell densities were achieved (fed-batch). We believe that the chondroitin production by these strains could be further improved and become more reproducible with gene integration into *E. coli* genome, rather than being overexpressed using plasmids.

The strain overexpressing *mltB* presented the best chondroitin yields in bioreactors when operated in both batch and fed-batch modes, which suggests that a slight improvement in the peptidoglycan recycling can redirect the metabolic flux towards the chondroitin precursors production.

4.4. Conclusions

The current study used genome-scale model optimizations to identify genes for under- and overexpression, which allowed the selection of possible targets to improve chondroitin production. The suggested promising mutants were further validated *in vivo* by constructing the *E. coli* mutant strains containing the chondroitin heterologous pathway and the additionally selected modifications. In flask fermentation, *E. coli* harboring the biosynthetic pathway was able to produce 62 mg/L of chondroitin. The evaluated mutants with additional modifications on this engineered strain resulted in chondroitin titers from 42 to 118 mg/L. In bioreactor, batch fermentations led to improved chondroitin production, the highest achieved by *E. coli* K-12 overexpressing *mltB* (427 mg/L in 54 h). Further fed-batch assays resulted in an improvement up to 535 mg/L of chondroitin. This study highlights new possible metabolic engineering targets to improve chondroitin production which ultimately can contribute for the advances in the biotechnological production of this most sought glycosaminoglycan.

4.5. References

[1] E.J. O'Brien, J.M. Monk, B.O. Palsson, Using genome-scale models to predict biological capabilities, Cell. 161 (2015) 971–987. https://doi.org/10.1016/j.cell.2015.05.019.

[2] C. Gu, G.B. Kim, W.J. Kim, H.U. Kim, S.Y. Lee, Current status and applications of genome-scale metabolic models, Genome Biol. 20 (2019) 1–18. https://doi.org/10.1186/s13059-019-1730-3.

[3] A. Oliveira, J. Rodrigues, E.C. Ferreira, L. Rodrigues, O. Dias, A kinetic model of the central carbon metabolism for acrylic acid production in *Escherichia coli*, PLOS Comput. Biol. 17 (2021) e1008704. https://doi.org/10.1371/journal.pcbi.1008704.

[4] C.J. Norsigian, N. Pusarla, J.L. McConn, J.T. Yurkovich, A. Dräger, B.O. Palsson, Z. King, BiGG Models 2020: Multi-strain genome-scale models and expansion across the phylogenetic tree, Nucleic Acids Res. 48 (2020) D402–D406. https://doi.org/10.1093/nar/gkz1054.

[5] D. Segrè, D. Vitkup, G.M. Church, Analysis of optimality in natural and perturbed metabolic networks, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 15112–15117. https://doi.org/10.1073/pnas.232349399.

[6] S.A. Becker, A.M. Feist, M.L. Mo, G. Hannum, B. Palsson, M.J. Herrgard, Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox, Nat. Protoc. 2 (2007) 727–738. https://doi.org/10.1038/nprot.2007.99.

[7] J. Schellenberger, R. Que, R.M.T. Fleming, I. Thiele, J.D. Orth, A.M. Feist, D.C. Zielinski, A. Bordbar, N.E. Lewis, S. Rahmanian, J. Kang, D.R. Hyduke, B. Palsson, Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox v2.0, Nat. Protoc. 6 (2011) 1290–1307. https://doi.org/10.1038/nprot.2011.308.

[8] P. Maia, M. Rocha, I. Rocha, *In Silico* Constraint-based strain optimization methods: the quest for optimal cell factories, Microbiol. Mol. Biol. Rev. 80 (2016) 45–67. https://doi.org/10.1128/mmbr.00014-15.

[9] A.P. Burgard, P. Pharkya, C.D. Maranas, OptKnock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization, Biotechnol. Bioeng. 84 (2003) 647–657. https://doi.org/10.1002/bit.10803.

[10] P. Pharkya, A.P. Burgard, C.D. Maranas, OptStrain: A computational framework for redesign of microbial production systems, Genome Res. 14 (2004) 2367–2376. https://doi.org/10.1101/gr.2872004.

[11] K.R. Patil, I. Rocha, J. Förster, J. Nielsen, J. Forster, J. Nielsen, Evolutionary programming as a platform for *in silico* metabolic engineering, BMC Bioinformatics. 6 (2005) 1–12. https://doi.org/10.1186/1471-2105-6-308.

[12] M. Rocha, P. Maia, R. Mendes, J.P. Pinto, E.C. Ferreira, J. Nielsen, K.R. Patil, I. Rocha, Natural computation meta-heuristics for the *in silico* optimization of microbial strains, BMC Bioinformatics.

9 (2008) 1–16. https://doi.org/10.1186/1471-2105-9-499.

[13] G. Melzer, M.E. Esfandabadi, E. Franco-Lara, C. Wittmann, Flux Design: *In silico* design of cell factories based on correlation of pathway fluxes to desired properties, BMC Syst. Biol. 3 (2009) 1–16. https://doi.org/10.1186/1752-0509-3-120.

[14] J. Kim, J.L. Reed, OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains, BMC Syst. Biol. 4 (2010). https://doi.org/10.1186/1752-0509-4-53.

[15] S. Ranganathan, P.F. Suthers, C.D. Maranas, OptForce: An optimization procedure for identifying all genetic manipulations leading to targeted overproductions, PLoS Comput. Biol. 6 (2010). https://doi.org/10.1371/journal.pcbi.1000744.

[16] S. Ranganathan, T.W. Tee, A. Chowdhury, A.R. Zomorrodi, J.M. Yoon, Y. Fu, J. V. Shanks, C.D. Maranas, An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*, Metab. Eng. 14 (2012) 687–704. https://doi.org/10.1016/j.ymben.2012.08.008.

[17] J.M. Otero, D. Cimini, K.R. Patil, S.G. Poulsen, L. Olsson, J. Nielsen, Industrial systems biology of *Saccharomyces cerevisiae* enables novel succinic acid cell factory, PLoS One. 8 (2013) 1–10. https://doi.org/10.1371/journal.pone.0054144.

[18] N. Guan, B. Du, J. Li, H.D. Shin, R.R. Chen, G. Du, J. Chen, L. Liu, Comparative genomics and transcriptomics analysis-guided metabolic engineering of *Propionibacterium acidipropionici* for improved propionic acid production, Biotechnol. Bioeng. 115 (2018) 483–494. https://doi.org/10.1002/bit.26478.

[19] G. Xu, W. Zou, X. Chen, N. Xu, L. Liu, J. Chen, Fumaric acid production in *Saccharomyces cerevisiae* by *in silico* aided metabolic engineering, PLoS One. 7 (2012) 1–10. https://doi.org/10.1371/journal.pone.0052086.

[20] X. Chen, G. Xu, N. Xu, W. Zou, P. Zhu, L. Liu, J. Chen, Metabolic engineering of *Torulopsis glabrata* for malate production, Metab. Eng. 19 (2013) 10–16. https://doi.org/10.1016/j.ymben.2013.05.002.

[21] M. Kim, B.G. Park, E.J. Kim, J. Kim, B.G. Kim, *In silico* identification of metabolic engineering strategies for improved lipid production in *Yarrowia lipolytica* by genome-scale metabolic modeling, Biotechnol. Biofuels. 12 (2019) 1–14. https://doi.org/10.1186/s13068-019-1518-4.

[22] P. Mishra, N.R. Lee, M. Lakshmanan, M. Kim, B.G. Kim, D.Y. Lee, Genome-scale model-driven strain design for dicarboxylic acid production in *Yarrowia lipolytica*, BMC Syst. Biol. 12 (2018). https://doi.org/10.1186/s12918-018-0542-5.

[23] M. Kavščcek, G. Bhutada, T. Madl, K. Natter, Optimization of lipid production with a genome-scale model of *Yarrowia lipolytica*, BMC Syst. Biol. 9 (2015) 1–13. https://doi.org/10.1186/s12918-015-0217-4.

[24] J.E. Yang, S.J. Park, W.J. Kim, H.J. Kim, B.J. Kim, H. Lee, J. Shin, S.Y. Lee, One-step fermentative production of aromatic polyesters from glucose by metabolically engineered *Escherichia coli*

strains, Nat. Commun. 9 (2018). https://doi.org/10.1038/s41467-017-02498-w.

[25] K.R. Kildegaard, N.B. Jensen, K. Schneider, E. Czarnotta, E. Özdemir, T. Klein, J. Maury, B.E. Ebert, H.B. Christensen, Y. Chen, I.K. Kim, M.J. Herrgård, L.M. Blank, J. Forster, J. Nielsen, I. Borodina, Engineering and systems-level analysis of *Saccharomyces cerevisiae* for production of 3-hydroxypropionic acid via malonyl-CoA reductase-dependent pathway, Microb. Cell Fact. 15 (2016) 1–13. https://doi.org/10.1186/s12934-016-0451-5.

[26] I. Borodina, K.R. Kildegaard, N.B. Jensen, T.H. Blicher, J. Maury, S. Sherstyk, K. Schneider, P. Lamosa, M.J. Herrgård, I. Rosenstand, F. Öberg, J. Forster, J. Nielsen, Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in *Saccharomyces cerevisiae* via β-alanine, Metab. Eng. 27 (2015) 57–64. https://doi.org/10.1016/j.ymben.2014.10.003.

[27] J. Zhang, S.D. Petersen, T. Radivojevic, A. Ramirez, A. Pérez-Manríquez, E. Abeliuk, B.J. Sánchez, Z. Costello, Y. Chen, M.J. Fero, H.G. Martin, J. Nielsen, J.D. Keasling, M.K. Jensen, Combining mechanistic and machine learning models for predictive engineering and optimization of tryptophan metabolism, Nat. Commun. 11 (2020). https://doi.org/10.1038/s41467-020-17910-1.

[28] S. Ferreira, R. Pereira, S.A. Wahl, I. Rocha, Metabolic engineering strategies for butanol production in *Escherichia coli*, Biotechnol. Bioeng. 117 (2020) 2571–2587. https://doi.org/10.1002/bit.27377.

[29] P. Xu, S. Ranganathan, Z.L. Fowler, C.D. Maranas, M.A.G.G. Koffas, Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA, Metab. Eng. 13 (2011) 578–587. https://doi.org/10.1016/j.ymben.2011.06.008.

[30] F. Cheng, H. Yu, G. Stephanopoulos, Engineering *Corynebacterium glutamicum* for high-titer biosynthesis of hyaluronic acid, Metab. Eng. 55 (2019) 276–289. https://doi.org/10.1016/j.ymben.2019.07.003.

[31] M.R. Couto, J.L. Rodrigues, L.R. Rodrigues, Heterologous production of chondroitin, Biotechnol. Reports. 33 (2022). https://doi.org/10.1016/j.btre.2022.e00710.

[32] W. He, L. Fu, G. Li, J. Andrew Jones, R.J. Linhardt, M. Koffas, J.A. Jones, R.J. Linhardt, M. Koffas, J. Andrew Jones, R.J. Linhardt, M. Koffas, Production of chondroitin in metabolically engineered *E. coli*, Metab. Eng. 27 (2015) 92–100. https://doi.org/10.1016/j.ymben.2014.11.003.

[33] I. Rocha, P. Maia, P. Evangelista, P. Vilaça, S. Soares, J.P. Pinto, J. Nielsen, K.R. Patil, E.C. Ferreira, M. Rocha, OptFlux: an open-source software platform for *in silico* metabolic engineering, BMC Syst. Biol. 4 (2010) 1–12. https://doi.org/10.1186/1752-0509-4-45.

[34] N.E. Lewis, K.K. Hixson, T.M. Conrad, J.A. Lerman, P. Charusanti, A.D. Polpitiya, J.N. Adkins, G. Schramm, S.O. Purvine, D. Lopez-Ferrer, K.K. Weitz, R. Eils, R. König, R.D. Smith, B. Palsson, Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models, Mol. Syst. Biol. 6 (2010). https://doi.org/10.1038/msb.2010.47.

[35] V. Pereira, F. Cruz, M. Rocha, MEWpy: a computational strain optimization workbench in Python, Bioinformatics. (2021) 1–3. https://doi.org/10.1093/bioinformatics/btab013.

[36] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evol. Comput. 6 (2002) 182–197. https://doi.org/10.1109/4235.996017.

[37] R. Mahadevan, C.H. Schilling, The effects of alternate optimal solutions in constraint-based genome-scale metabolic models, Metab. Eng. 5 (2003) 264–276. https://doi.org/10.1016/j.ymben.2003.09.002.

[38] O.F. Restaino, I. di Lauro, R. Di Nuzzo, M. De Rosa, C. Schiraldi, R. Di Nuzzo, M. De Rosa, C. Schiraldi, R. Di Nuzzo, M. De Rosa, C. Schiraldi, New insight into chondroitin and heparosan-like capsular polysaccharide synthesis by profiling of the nucleotide sugar precursors, Biosci. Rep. 37 (2017) 1–12. https://doi.org/10.1042/BSR20160548.

[39] P. Jin, L. Zhang, P. Yuan, Z. Kang, G. Du, J. Chen, Efficient biosynthesis of polysaccharides chondroitin and heparosan by metabolically engineered *Bacillus subtilis*, Carbohydr. Polym. 140 (2016) 424–432. https://doi.org/10.1016/j.carbpol.2015.12.065.

[40] D. Cimini, M. De Rosa, A. Viggiani, O.F. Restaino, E. Carlino, C. Schiraldi, Improved fructosylated chondroitin production by *kfoC* overexpression in *E. coli* K4, J. Biotechnol. 150 (2010) 324–331. https://doi.org/10.1016/j.jbiotec.2010.09.954.

[41] H. Yu, G. Stephanopoulos, Metabolic engineering of *Escherichia coli* for biosynthesis of hyaluronic acid, Metab. Eng. 10 (2008) 24–32. https://doi.org/10.1016/j.ymben.2007.09.001.

[42] M.R. Couto, J.L. Rodrigues, L.R. Rodrigues, Cloning, expression and characterization of UDP-glucose dehydrogenases, Life. 11 (2021) 1201. https://doi.org/10.3390/life11111201.

[43] D.R. Nielsen, S.H. Yoon, C.J. Yuan, K.L.J. Prather, Metabolic engineering of acetoin and meso-2,3-butanediol biosynthesis in *E. coli*, Biotechnol. J. 5 (2010) 274–284. https://doi.org/10.1002/biot.200900279.

[44] B.F. Cress, O.D. Toparlak, S. Guleria, M. Lebovich, J.T. Stieglitz, J.A. Englaender, J.A. Jones, R.J. Linhardt, M.A.G.G. Koffas, CRISPathBrick: Modular combinatorial assembly of type II-A CRISPR arrays for dCas9-mediated multiplex transcriptional repression in *E. coli*, ACS Synth. Biol. 4 (2015) 987–1000. https://doi.org/10.1021/acssynbio.5b00012.

[45] T. Bitter, H.M. Muir, A modified uronic acid carbazole reaction, Anal. Biochem. 4 (1962) 330–334. https://doi.org/10.1016/0003-2697(62)90095-7.

[46] G.L. Miller, Use of dinitrosalicylic acid reagent for determination of reducing sugar, Anal. Chem. 31(3) (1959) 426–428.

[47] J. Wu, Q. Zhang, L. Liu, X. Chen, J. Liu, Q. Luo, Recombinant *Escherichia coli* for high efficiency production of fructosylated chondroitin and method for making thereof, US Patent 10,508,279 B2, 2019.

[48] J.E. Woo, H.J. Seong, S.Y. Lee, Y.S. Jang, Metabolic engineering of *Escherichia coli* for the production of hyaluronic acid from glucose and galactose, Front. Bioeng. Biotechnol. 7 (2019) 1–9. https://doi.org/10.3389/fbioe.2019.00351.

[49] Q. Cheng, J.T. Park, Substrate specificity of the AmpG permease required for recycling of cell wall

anhydro-muropeptides, J. Bacteriol. 184 (2002) 6434–6436. https://doi.org/10.1128/JB.184.23.6434.

[50] M. Lee, D. Hesek, L.I. Llarrull, E. Lastochkin, H. Pi, B. Boggess, S. Mobashery, Reactions of all *Escherichia coli* lytic transglycosylases with bacterial cell wall, J. Am. Chem. Soc. 135 (2013) 3311–3314. https://doi.org/10.1021/ja309036q.

[51] D.A. Dik, D.R. Marous, J.F. Fisher, S. Mobashery, Lytic transglycosylases: concinnity in concision of the bacterial cell wall, Crit. Rev. Biochem. Mol. Biol. 52 (2017) 503–542. https://doi.org/10.1080/10409238.2017.1337705.

[52] J.T. Park, T. Uehara, How bacteria consume their own exoskeletons (turnover and recycling of cell wall peptidoglycan), Microbiol. Mol. Biol. Rev. 72 (2008) 211–227. https://doi.org/10.1128/mmbr.00027-07.

[53] A. Maqbool, V.M. Levdikov, E. V. Blagova, M. Hervé, R.S.P. Horler, A.J. Wilkinson, G.H. Thomas, Compensating stereochemical changes allow murein tripeptide to be accommodated in a conventional peptide-binding protein, J. Biol. Chem. 286 (2011) 31512–31521. https://doi.org/10.1074/jbc.M111.267179.

[54] J. Fricke, J. Neuhard, R.A. Kelln, S. Pedersen, The *cmk* gene encoding cytidine monophosphate kinase is located in the *rpsA* operon and is required for normal replication rate in *Escherichia coli*, J. Bacteriol. 177 (1995) 517–523. https://doi.org/10.1128/jb.177.3.517-523.1995.

[55] J. Landberg, N.R. Wright, T. Wulff, M.J. Herrgård, A.T. Nielsen, CRISPR interference of nucleotide biosynthesis improves production of a single-domain antibody in *Escherichia coli*, Biotechnol. Bioeng. 117 (2020) 3835–3848. https://doi.org/10.1002/bit.27536.

[56] N. Bucurenci, H. Sakamoto, P. Briozzo, N. Palibroda, L. Serina, R.S. Sarfati, G. Labesse, G. Briand, A. Danchin, O. Bârzu, A.M. Gilles, CMP kinase from *Escherichia coli* is structurally related to other nucleoside monophosphate kinases, J. Biol. Chem. 271 (1996) 2856–2862. https://doi.org/10.1074/jbc.271.5.2856.

[57] P. Meyer, C. Evrin, P. Briozzo, N. Joly, O. Bârzu, A.M. Gilles, Structural and functional characterization of Escherichia coli UMP kinase in complex with its allosteric regulator GTP, J. Biol. Chem. 283 (2008) 36011–36018. https://doi.org/10.1074/jbc.M802614200.

[58] H. Sakamoto, S. Landais, C. Evrin, C. Laurent-Winter, O. Bârzu, R.A. Kelln, Structure-function relationships of UMP kinases from pyrH mutants of Gram-negative bacteria, Microbiology. 150 (2004) 2153–2159. https://doi.org/10.1099/mic.0.26996-0.

[59] J.T. Park, J.H. Shim, P.L. Tran, I.H. Hong, H.U. Yong, E.F. Oktavina, H.D. Nguyen, J.W. Kim, T.S. Lee, S.H. Park, W. Boos, K.H. Park, Role of maltose enzymes in glycogen synthesis by Escherichia coli, J. Bacteriol. 193 (2011) 2517–2526. https://doi.org/10.1128/JB.01238-10.

[60] G. Vadlamani, K.A. Stubbs, J. Désiré, Y. Blériot, D.J. Vocadlo, B.L. Mark, Conformational flexibility of the glycosidase NagZ allows it to bind structurally diverse inhibitors to suppress β-lactam antibiotic resistance, Protein Sci. 26 (2017) 1161–1170. https://doi.org/10.1002/pro.3166.

[61] A. Roberts, S.Y. Lee, E. McCullagh, R.E. Silversmith, D.E. Wemmer, YbiV from Escherichia coli

K12 is a HAD phosphatase, Proteins Struct. Funct. Genet. 58 (2005) 790–801. https://doi.org/10.1002/prot.20267.

[62] M. Zhao, Y. Li, F. Wang, Y. Ren, D. Wei, A CRISPRi mediated self-inducible system for dynamic regulation of TCA cycle and improvement of itaconic acid production in Escherichia coli, Synth. Syst. Biotechnol. 7 (2022) 982–988. https://doi.org/10.1016/j.synbio.2022.05.008.

[63] A.M. Olajuyin, M. Yang, Y. Liu, T. Mu, J. Tian, O.A. Adaramoye, J. Xing, Efficient production of succinic acid from Palmaria palmata hydrolysate by metabolically engineered Escherichia coli, Bioresour. Technol. 214 (2016) 653–659. https://doi.org/10.1016/j.biortech.2016.04.117.

[64] C.D. Lu, Y.Y. Xiong, J.F. Zhao, J.H. Wang, Fermentation Production of Succinic Acid from Sucrose by Engineered Escherichia Coli JH208, Adv. Mater. Res. 1088 (2015) 503–506. https://doi.org/10.4028/www.scientific.net/amr.1088.503.

[65] A. Negrete, J. Shiloach, Improving E. coli growth performance by manipulating small RNA expression, Microb. Cell Fact. 16 (2017) 1–12. https://doi.org/10.1186/s12934-017-0810-x.

[66] A. Badri, A. Williams, A. Awofiranye, P. Datta, K. Xia, W. He, K. Fraser, J.S. Dordick, R.J. Linhardt, M.A.G.G. Koffas, Complete biosynthesis of a sulfated chondroitin in Escherichia coli, Nat. Commun. 12 (2021) 1–10. https://doi.org/10.1038/s41467-021-21692-5.

[67] M. De Rosa, C. Schiraldi, D. Cimini, Biotechnological production of chondroitin, US Patent 8,592,186 B2, 2013.

[68] D. Cimini, M. De Rosa, E. Carlino, A. Ruggiero, C. Schiraldi, Homologous overexpression of rfaH in E. coli K4 improves the production of chondroitin-like capsular polysaccharide, Microb. Cell Fact. 12 (2013) 1. https://doi.org/10.1186/1475-2859-12-46.

[69] D. Cimini, R. Russo, S. D'Ambrosio, I. Dello Iacono, C. Rega, E. Carlino, O. Argenzio, L. Russo, B. D'Abrosca, A. Chambery, C. Schiraldi, Physiological characterization and quantitative proteomic analyses of metabolically engineered E. coli K4 strains with improved pathways for capsular polysaccharide biosynthesis, Biotechnol. Bioeng. 115 (2018) 1801–1814. https://doi.org/10.1002/bit.26597.

[70] Q. Wu, A. Yang, W. Zou, Z. Duan, J. Liu, J. Chen, L. Liu, Transcriptional engineering of Escherichia coli K4 for fructosylated chondroitin production, Biotechnol. Prog. 29 (2013) 1140–1149. https://doi.org/10.1002/btpr.1777.

[71] C.R. Reisch, K.L.J. Prather, The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in Escherichia coli, Sci. Rep. 5 (2015) 15096. https://doi.org/10.1038/srep15096.

[72] Q. Cheng, H. Li, K. Merdek, J.T. Park, Molecular characterization of the β-N-acetylglucosaminidase of Escherichia coli and its role in cell wall recycling, J. Bacteriol. 182 (2000) 4836–4840. https://doi.org/10.1128/JB.182.17.4836-4840.2000.

[73] D. Yu, H.M. Ellis, E.C. Lee, N.A. Jenkins, N.G. Copeland, D.L. Court, An efficient recombination system for chromosome engineering in Escherichia coli, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 5978–5983. https://doi.org/10.1073/pnas.100127597.

[74] S. Dhar, H. Kumari, D. Balasubramanian, K. Mathee, Cell-wall recycling and synthesis in Escherichia coli and Pseudomonas aeruginosa – Their role in the development of resistance, J. Med. Microbiol. 67 (2018) 1–21. https://doi.org/10.1099/jmm.0.000636.

[75] W. Vötsch, M.F. Templin, Characterization of a β-N-acetylglucosaminidase of Escherichia coli and elucidation of its role in muropeptide recycling and β-lactamase induction, J. Biol. Chem. 275 (2000) 39032–39038. https://doi.org/10.1074/jbc.M004797200.

[76] D. Cimini, S. Fantaccione, F. Volpe, M. De Rosa, O.F. Restaino, G. Aquino, C. Schiraldi, IS2-mediated overexpression of kfoC in E. coli K4 increases chondroitin-like capsular polysaccharide production, Appl. Microbiol. Biotechnol. 98 (2014) 3955–3964. https://doi.org/10.1007/s00253-014-5506-0.

CHAPTER 5

IN SILICO DESIGN OF *SACCHAROMYCES CEREVISIAE* STRAINS FOR IMPROVED PRODUCTION OF CHONDROITIN

Abstract

Chondroitin sulfate is a glycosaminoglycan that has gained widespread use in nutraceuticals and pharmaceuticals, mainly for treating osteoarthritis. Traditionally, it has been extracted from animal cartilage but recently, biotechnological processes have emerged as a commercial alternative to avoid the risk of viral or prion contamination and offer a vegan-friendly source. Typically, these methods involve producing the chondroitin backbone using pathogenic bacteria and then modifying it enzymatically through the action of sulfotransferases. Despite the challenges of expressing active sulfotransferases in bacteria, the use of eukaryotic microorganisms is still limited to a few works using *Pichia pastoris*. To create a safer and efficient biotechnological platform, we have constructed a biosynthetic pathway for chondroitin production in *S. cerevisiae*. Furthermore, as genome-scale models are valuable tools for identifying novel targets for metabolic engineering, a stoichiometric model of chondroitin-producing *S. cerevisiae* has been developed and used in optimization algorithms. Our research has yielded several novel targets, such as uridine diphosphate(UDP)-*N*-acetylglucosamine pyrophosphorylase (*QRI1*), glucosamine-6-phosphate acetyltransferase (*GNA1*) or *N*-acetylglucosamine-phosphate mutase *(PCM1*) overexpressions that might enhance chondroitin production, which will guide future experimental research to develop more efficient host organisms for the biotechnological production process.

Keywords: chondroitin; *S. cerevisiae*; metabolic models; metabolic engineering

This chapter is in preparation to be submitted for publication:

M.R. Couto, J.L. Rodrigues, Ó. Dias, L.R. Rodrigues, *In silico* design of *Saccharomyces cerevisiae* strains for improved production of chondroitin.

5.1. Introduction

Chondroitin sulfate is a glycosaminoglycan that naturally occurs in animals, in different concentrations, and in different sulfation patterns, percentage and structure ratios, according to the tissue where it is present [1]. Unsulfated or fructosylated forms of this compound also exist in some pathogenic bacteria as a capsular constituent [2]. Chondroitin sulfate has been mainly used in nutraceuticals, pharmaceuticals and veterinary supplements for osteoarthritis treatment and joint protection, but also in ophthalmological solutions and devices, as described in detail in Chapter 1. Its biological activity varies according to its sulfonation pattern, therefore making it versatile and useful in a wide range of other potential applications [1]. The biotechnological production of chondroitin is generally based on cultivating the pathogenic bacteria *Escherichia coli* O5:K4:H4, which naturally produces a fructosylated form of chondroitin [3]. Many efforts have been made to engineer chondroitin production using safer microorganisms [4–7], however the yields do not meet the growing demand. Furthermore, eukaryotic microorganisms remain relatively unexplored, with a singular work using *Pichia pastoris* for chondroitin production [8]. Nevertheless, eukaryotic microorganisms are particularly interesting for chondroitin production because of their ability to perform post-translational modifications, unlike prokaryotic organisms such as *E. coli*. In fact, glycosylation and correct folding are required for animal sulfotransferases to become active and perform the sulfonation of the chondroitin backbone [9], which makes the correct expression of these proteins challenging, especially in prokaryotic hosts [10]. *S. cerevisiae* has been one of the most widely used microbes for industrial biotechnological production of several compounds. Its broad use as host in metabolic engineer is related to its rapid and robust growth, its ease of genetic manipulation for recombinant protein expression, and its ability to perform post-translational modifications, including glycosylation, and to properly fold recombinant proteins [11–13]. Therefore, it can be an interesting host to be used in the production of chondroitin.

Genome-scale metabolic models (GEMs) are powerful resources that consist in the representation of the entire metabolic network of a biological system, including the enzymes, metabolites, reactions, genes, and their associations, containing information on stoichiometry, compartmentalization and biomass composition [14]. The use of these models to evaluate the organism biological capabilities requires the representation of the biochemical conversions following a stoichiometric matrix representation containing the stoichiometric coefficients for each metabolite in each reaction, where reactions are the columns and the metabolites the rows [14,15]. Constraint-based modelling assumes cells operate in a steady-state, meaning that the metabolites may not be accumulated, and by applying flux constrains through upper and lower bounds, this matrix is transformed into a system of linear equations which can be used to

calculate the flux of each reaction [14,15]. As this represents an undetermined system, a biological relevant reaction, usually biomass production, is used as the objective function to formulate a linear problem that can be solved using mathematical programming [14]. Manipulating the reaction bounds allows to simulate environmental conditions or genetic modifications such as knockouts [16]. Appling evolutionary algorithms is a common strategy in strain design for identifying targets for metabolic engineering. Additionally, other information can be integrated with GEMs such as regulatory, kinetics and omics data to improve the predictive power of these models in specific conditions [17].

As GEMs provide a systems biology framework for phenotype simulation, they have wide applications in metabolism studying, identification of novel targets for metabolic engineering, diseases understanding and drug target identification [14,18–22]. In particular, the industrial applications of GEMs are the most reported as they have been used for enhancing the biotechnological production of several compounds, either endogenous or heterologous, such as dicarboxylic acids [23–29], alcohols [30–32], amino acids [33–35], polymers [36–39], antibiotics [40,41] and polyphenols [42–44].

This study explores the potential of chondroitin sulfate production in *S. cerevisiae* using synthetic biology and metabolic engineering strategies. Furthermore, using a budding yeast GEM, *in silico* flux analysis was employed, as well as evolutionary algorithms, to identify novel targets for improving chondroitin titers.

5.2. Materials and Methods

5.2.1. Model construction

The consensus GEM of *Saccharomyces cerevisiae* yeast-GEM [45], version 8.4.2 (https://github.com/SysBioChalmers/yeast-GEM), composed by 4058 reactions, 2742 metabolites and 1150 genes, was used as template. From this, a new model containing the heterologous pathway for chondroitin production was constructed by including the reactions UGD (uridine diphosphate(UDP)-glucose 6-dehydrogenase), UAE (UDP-*N*-acetylglucosamine 4-epimerase), CHSY (chondroitin synthase/polymerase) and EX_chond (chondroitin drain reaction). The metabolites UDP-glucose (S_1543) and UDP-acetylglucosamine (S_1544) already existed in the template model. UDP-glucuronic acid (M_udpglcur), UDP-acetylgalactosamine (M_udpacgal) and chondroitin (M_chond) were included. The genes for the reactions UGD, UAE and CHSY were also included in the model. After preliminary tests, the biomass equation was adjusted to include 1% chitin in biomass composition. The final constructed model, named yeast-GEM_c, included 4062 reactions, 2745 metabolites and 1153 genes.

5.2.2. Conditions for *in silico* simulations and optimization

OptFlux software [46] version 3.3.5 was used to simulate the phenotype of *S. cerevisiae* engineered with chondroitin production pathway and further mutants, using parsimonious Flux Balance Analysis (pFBA) as simulation method [47]. The glucose uptake was set to 10 mol/gDW/h and oxygen was unlimited. Gene deletion and gene over and under expression identifications were performed by running optimization algorithms from the software. Strength Pareto Evolutionary Algorithm 2 (SPEA2) [48] was used as optimization strategy where the Biomass-Product Coupled Yield (BPCY) was set as the objective, and pFBA as the simulation method. pFBA used biomass reaction (r_2111) as the objective function to maximize. A maximum of 10 modifications were allowed. The maximum for evaluation functions was set to 50,000. CPLEX Optimization Studio version 12.9.0 (IBM) was used as Linear Programming solver.

Optimizations were also performed using MEWpy [49] under varied specified environmental conditions, with glucose as carbon source (10 mol/gDW/h) and unlimited oxygen (1000 mol/gDW/h) uptake. Optimizations were performed using the Evolutionary Algorithm Non-Dominated Sorting Genetic Algorithm II [50] as optimization strategy and pFBA as the simulation method. The evolutionary algorithm used two objective functions, BPCY and Weighted Yield (WYIELD, weighed sum of the minimum and maximum product fluxes). pFBA used the biomass reaction as the objective function to maximize. A maximum of 10 or 3 modifications were allowed.

Flux variability analysis (FVA) was performed to access the robustness of the optimization results. The phenotype was simulated for engineered strains with fixed biomass using pFBA and alternating the objective function between maximization and minimization of chondroitin production [51].

5.2.3. Strains and plasmids

The strains and plasmids used in this study are listed in Table 5.1. *E. coli* NZY5α (NZYTech, Lisbon, Portugal) competent cells were used for cloning procedures, vector propagation and storage. *E. coli* was cultured at 37°C and 200 rpm in lysogeny broth (LB) (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl; NZYTech) or on LB agar plates (20 g/L agar, JMGS, Odivelas, Portugal). Ampicillin (NZYTech) at a final concentration of 100 μg/mL was supplemented when necessary.

S. cerevisiae CEN.PK2-1C and *S. cerevisiae* BY4741 strains were obtained from Euroscarf (Oberursel, Germany). The plasmids pSP-GM1 (PGK1 promoter and TEF promoter; Addgene, Watertown, MA, USA) and pBEVY-L (GPD promoter; ATCC, Manassas, USA) were used as shuttle vectors. Wild-type yeast strains were cultivated at 30°C and 200 rpm in yeast extract peptone dextrose (YPD) media, composed by 20

g/L bacteriological peptone (HiMedia, Mumbai, India), 10 g/L yeast extract (Panreac AppliChem, Darmstadt, Germany), 20 g/L glucose (Acros Organics, New Jersey, USA), or in agar plates with the same composition.

The engineered yeast strains were grown in synthetic defined minimal media composed by 6.7 g L^{-1} of yeast nitrogen base (YNB) with ammonium sulfate without amino acids (Sigma Aldrich, Steinheim, Germany), supplemented with 20 g/L glucose and the required amino acids to compensate for auxotrophies, namely tryptophan or methionine (Panreac AppliChem), depending on the strain, and histidine (Panreac AppliChem) at final concentrations of 100 mg/L.

Table 5.1. Strains and plasmids used in this study.

Strains	**Relevant genotype**	**Source**
Escherichia coli NZY5α	*fhuA2* Δ(*argF lacZ*)U169 *phoA glnV44*Φ80 Δ(*lacZ*)M15 *gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	NZYTech (MB00401)
Saccharomyces cerevisiae CEN.PK2-1C	*MATa ura3-52 his3Δ1 leu2-3,112 trp1-289 MAL2-8c SUC2*	Euroscarf 30000A [52]
S. cerevisiae BY4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	Euroscarf Y00000 [53]
Plasmids	**Description**	**Source**
pSP-GM1	pUC *ori*, Amp^R, 2 μ *ori*, *URA3* $P_{TEF1}P_{PGK1}$	Addgene #64739 [54]
pBEVY-L	pUC *ori*, Amp^R, 2 μ *ori*, *LEU2* $P_{GPD}P_{ADH1}$	ATCC 51226
pUC57_Giuae	pMB1 *ori*, Amp^R; pUC57 carrying *Giardia intestinalis* uridine diphosphate-glucosamine-4-epimerase gene (*GiUAE*) codon-optimized for *S. cerevisiae*	NZYTech
pUC57_Btchsy1	pMB1 *ori*, Amp^R; pUC57 carrying *Bos taurus* chondroitin synthase 1 gene (*BtCHSY*) codon-optimized for *S. cerevisiae*	NZYTech
pETM6_kfoCA	pETM6 carrying chondroitin synthase, *kfoC,* and, uridine diphosphate-glucosamine-4-epimerase, *kfoA,* genes from *E. coli* O5:K4:H4	[5]
pSP-GM1_Zmugd	pSP-GM1 carrying *Zymomonas mobilis* uridine diphosphate glucose 6-dehydrogenase gene (*Zmugd*)	[55]
pSP-GM1_Giuae_Zmugd	pSP-GM1 carrying *GiUAE* and *Zmugd*	This study
pBEVY_Btchsy	pBEVY-L carrying *BtCHSY*	This study
pSP-GM1_kfoA_Zmugd	pSP-GM1 carrying *kfoA* and *Zmugd*	This study
pBEVY_kfoC	pBEVY-L carrying *kfoC*	This study

5.2.4 Biosynthetic pathway construction

Table A 5.1 compiles the primers used for cloning procedures. Two different biosynthetic pathways have been constructed and introduced in *S. cerevisiae* CEN.PK2-1C and *S. cerevisiae* BY4741 strains. Each pathway contained three genes for the expression of UGD, UAE and CHSY, which are absent in yeast metabolism. UAE gene from *Giardia intestinalis* (GenBank accession number AY187036.1, *GiUAE*) and CHSY gene from *Bos taurus* (GenBank accession number AF440749.1, *BtCHSY*) were codon-optimized for *S. cerevisiae* and synthesized by NZYTech (sequences can be seen in Table 5.2). Afterwards, *GiUAE* and *BtCHSY* were amplified from pUC57_Giuae and pUC57_Btchsy, respectively, using Gi_uae_Fw1 and Gi_uae_Rv1 or Bt_chsy1_Fw and Bt_chsy1_Rv as primers. After amplification, genes were first independently cloned in pSP-GM1 under PGK1 promoter. For further assembly of the entire chondroitin pathway, the *GiUAE* was cloned in pSP-GM1 under the TEF promoter regulation, using the primers Giuae_tefp_Fw and Giuae_tefp_Rv for gene amplification. Then, *Zmugd*, amplified from pSP-GM1_Zmugd [55] with primers Zm_psp_Fw and Zm_psp_Rv, was cloned in pSP-GM1_Giuae_tef, resulting in pSP-GM1_Giuae_Zmugd (*Zmugd* was cloned second as the restriction enzymes used to clone *GiUAE* cut *Zmugd,* that is not codon-optimized). The *BtCHSY* was amplified from pSP-GM1_Btchsy, using the primers Btchsy_pBEVY_Fw and Btchsy_pBEVY_Rv, and cloned in pBEVY-L generating pBEVY_Btchsy. The second pathway was composed by chondroitin-producing genes *kfoA* and *kfoC* (encoding UAE and CHSY, respectively) from *E. coli* K4 (serotype O5:K4(L):H4). *KfoA* and *kfoC* were amplified from pETM6_kfoCA [5], which was kindly provided by Dr. Mattheos Koffas (Rensselaer Polytechnic Institute, Troy, NY). For that, the primer pairs kfoA_psp_Fw/kfoA_psp_Rv and kfoC_pBEVY_Fw/kfoC_pBEVY_Rv were used, respectively. Then, *kfoA* was cloned into pSP-GM1_Zmugd [55] while *kfoC* was cloned in pBEVY-L, resulting in pSP-GM1_kfoA_Zmugd and pBEVY_kfoC, respectively.

All gene amplifications were performed through PCR (polymerase chain reaction) using phusion High Fidelity DNA Polymerase (Thermo Fisher Scientific, Wilmington, United States). The plasmids were extracted with Plasmid Miniprep Kit (Macherey-Nagel, Düren, Germany). PCR products from amplification were excised and purified from agarose gels using NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel). Quantification of plasmid DNA and PCR products was further achieved using NanoDrop One instrument (Thermo Fisher Scientific). Then, digestion was performed by incubating specific restriction endonucleases (Thermo Fisher Scientific) for 1 h at 37°C. The resulting digested DNA fragments were purified using NucleoSpin® Gel and PCR Clean-up Kit and used for ligations with T4 DNA ligase (Thermo Fisher Scientific) for 1 h at room temperature. The resulting mixture was transformed into *E. coli* NZY5α competent cells (NZYTech) by heat shock. Transformants were then recovered by adding super optimal

broth with catabolite repression (SOC; NZYTech) and incubating the mixture for 1 h at 37°C. Cells were then platted on agar plates containing selective medium. Finally, all construction sequences were verified by colony PCR using Dream Taq polymerase (Thermo Fisher Scientific), digestion, and further confirmation by sequencing (GATC Biotech, Konstanz, Germany). After sequence confirmation, transformations of the constructed plasmids into *S. cerevisiae* were performed by lithium acetate/single-stranded carrier DNA/polyethylene glycol method [56]. Lithium acetate, salmon sperm DNA and polyethylene glycol (PEG-3350) were obtained from Sigma-Aldrich. Selection of yeast transformants was performed in synthetic defined minimal media with the required amino acids.

5.2.5. Flask fermentation conditions

For each assay, a single *S. cerevisiae* colony was picked from the transformation plate and grown for 24 h at 30°C and 200 rpm in 8 mL of the synthetic defined minimal media supplemented with the required amino acids for pre-culture. Afterwards, 50 mL of medium with the same composition in 250 mL flasks were inoculated to an initial optical density at 600 nm (OD_{600nm}) of 0.1. Yeast cells were further cultured at 30°C and 200 rpm for 24 h.

5.2.6. Analytical methods

In the end of the fermentation, the culture of *S. cerevisiae* cells (~50 mL) was harvested by centrifugation (5,000 × g, 15 min). The supernatants were used to quantify extracellular chondroitin and glucose, while the pellets were further processed to determine intracellular chondroitin.

To obtain the intracellular fraction, cells were lysed. For each 0.1 g of wet cells, 0.2 g of glass beads (425–600 µm, Sigma- Aldrich) were added to the cells pellet, as well as 1 mL of deionized water. Cells were then lysed in FastPrep-24 (MP Biomedicals, Salon, USA) during 5 cycles of 1 min at 6–6.5 m/s interspersed with 1 min cooling on ice. Lysed samples were centrifuged (16,000 × g, 15 min). Afterwards, the lysates (supernatant) were treated with DNaseI (New England Biolabs, MA, USA) for 2 h at 37°C and then with proteinase K (2 mg/mL, NZYTech) for 2 h at 56°C. The mixture was further boiled for 5 min and centrifuged (16,000 x g, 20 min) to remove insoluble material.

To precipitate extracellular and intracellular chondroitin, three volumes of cold ethanol were added to the samples and the mixture was left at 4°C overnight. The resulting precipitate was collected through centrifugation (4,000 x g for 10 min at 4°C) and subsequently air-dried at room temperature overnight. The dried precipitate was resuspended in deionized water followed by removal of the insoluble material

by centrifugation (16,000 x g, 20 min). Uronic acid carbazole assay [57] was used to quantify chondroitin by using chondroitin sulfate (Biosynth, Staad, Switzerland) solutions as standards. Standards or samples with 125 µL were mixed with 750 µL sulfuric acid reagent (9.5 g/L sodium tetraborate, Supelco, Bellefonte, USA, dissolved in H_2SO_4 > 95%, Fisher Chemical, Hampton, USA) and boiled for 20 min. Then, 25 µL of carbazole reagent (1.25 g/L carbazole, Supelco, dissolved in absolute ethanol, Fisher Chemical) were added to the boiled samples, followed by an additional 15 min boil and a subsequent 15 min cooling period. The OD_{530nm} was measured using a 96-well plate spectrophotometric reader Synergy HT (BioTek, Winooski, VT, USA).

Yeast cell concentration was calculated by determining the OD_{600nm} in the end of culture, using a calibration with solutions of known biomass concentration.

Glucose samples at the end of fermentation were analyzed using high performance liquid chromatography (HPLC) with a JASCO system and a refractive index (RI) detector (RI-2031), employing an Aminex HPX-87H column from Bio-Rad maintained at 60°C. The mobile phase used was 5 mM H_2SO_4 at a flow rate of 0.5 mL/min.

5.3. Results and Discussion

5.3.1. Bioinformatic tool for identification of gene targets

A model of *S. cerevisiae* metabolism has been modified to include the heterologous reactions, intermediates and genes required for chondroitin production. At that stage, optimizations for improving chondroitin production could not find any solution, either searching for knockout or under- and overexpression targets. One possible hypothesis for this was that biomass growth was not being properly coupled with product formation. We then realized that the original biomass equation did not predict the inclusion of chitin. Even though *S. cerevisiae* is reported to have a minimal amount of chitin, its presence might still be necessary for essential functions related to cell wall integrity and other processes, as suggested by the finding that simultaneous knockout of all three chitin synthase genes is lethal in yeast [58]. Therefore, based on literature [59–61], the biomass equation was corrected to include 1% chitin, by adjusting the reaction stoichiometry in the model to maintain the stoichiometric coefficients of other compounds while including the necessary stoichiometric coefficient to achieve the desired percentage of chitin (Table 5.2). As chitin is an important intervenient in pathways related with chondroitin precursors, this adjustment could result in optimization results.

Table 5.2. Intermediates and their stoichiometric values in growth equation in the model yeast-GEM 8.4.2 before and after including 1% of chitin.

Compound ID code	Compound name	Template model		Corrected model 1% chitin	
		Stoichiometry	Percentage (%)	Stoichiometry	Percentage (%)
s_0001_ce	(1→3)-β-D-glucan [cell envelope]	0.748514964	33.88	0.748514964	33.54
s_0004_ce	(1→6)-β-D-glucan [cell envelope]	0.250091654	11.32	0.250091654	11.21
s_0773_c	glycogen [cytoplasm]	0.361414528	16.36	0.361414528	16.20
s_1107_c	mannan [cytoplasm]	0.710939625	32.18	0.710939625	31.86
s_1520_c	trehalose [cytoplasm]	0.138275712	6.26	0.138275712	6.20
s_0509_c	chitin [cytoplasm]	0	0.00	0.022313288	1.00

In fact, after performing these modifications, the optimization using evolutionary algorithms in OptFlux was able to find multiple solutions. The solutions with best BPCY are shown in Table 5.3.

Table 5.3. Optimization of chondroitin production in yeast-GEM_c model using OptFlux. The optimization algorithm for under and overexpression identification was run three times. The predicted phenotype for the unmodified and modified strains (from the resulting solutions with highest biomass-product coupled yield, BPCY) are shown. BPCY is calculated by OptFlux by multiplying biomass by product and then dividing by substrate consumed (in all cases being 10 mol/gDW/h), as predicted by pFBA simulation. Flux variability analysis (FVA) results are shown as minimum and maximum chondroitin obtained through pFBA for fixed biomass. Predicted biomass and chondroitin values are in units of mol/gDW/h.

Solution	BPCY	Genes modified		Predicted phenotype (pFBA)		FVA	
		Underexpression	Overexpression	Biomass	Chondroitin	Min chondroitin	Max chondroitin
-	-	-	-	0.8612	0.0000	-	-
1	0.04375	-	*QRI1*	0.7317	0.5980	0.5980	0.9358
2	0.04375	-	*GNA1*	0.7317	0.5980	0.5980	0.9358
3	0.04375	-	*PCM1*	0.7317	0.5980	0.5980	0.9358

Gene descriptions: *GNA1* - glucosamine-6-phosphate acetyltransferase; *PCM1* - *N*-acetylglucosamine-phosphate mutase; *QRI1* – uridine diphosphate-*N*-acetylglucosamine pyrophosphorylase

Despite ten modifications were allowed, the gathered solutions pointed to single modifications, namely the overexpression of one of the genes involved in the production of chondroitin precursors, *QRI1*, *GNA1*

or *PCM1* (expression values of 32). *GNA1* encodes glucosamine-6-phosphate acetyltransferase, which catalyzes *N*-acetylglucosamine 6-phosphate synthesis, from glucosamine 6-phosphate and acetyl-coenzyme A (acetyl-CoA). *PCM1*, encoding *N*-acetylglucosamine-phosphate mutase, is responsible for converting *N*-acetylglucosamine 6-phostate to *N*–acetylglucosamine 1-phostate. *QRI1*, encoding UDP-*N*-acetylglucosamine pyrophosphorylase, is responsible for the formation of UDP-*N*-acetylglucosamine. Figure 5.1 shows a schematic representation of the metabolism of *S. cerevisiae* that is involved in the biosynthetic production of chondroitin and possible competing pathways.

The overexpression of genes from synthesis pathway of precursors UDP-glucose and UDP-*N*-acetylglucosamine is a common strategy for improving the production of chondroitin and other glycosaminoglycans [3,7,62–64]. Interestingly, all the optimization results here obtained indicated genes that lead to UDP-*N*-acetylglucosamine production, suggesting this intermediate as the limiting precursor in *S. cerevisiae*.

Regarding FBA analysis, the difference between the predicted minimum and maximum chondroitin production shows the mutants are moderately robust.

In MEWpy, the optimization using evolutionary algorithms resulted in 75 solutions that included modifications in 53 different genes. The frequency and expression values of genes resulting from optimization are shown in Figure 5.2.

All the solutions from optimization presented a common modification, namely the overexpression of *PCM1*, a modification already identified by the OptFlux approach, which confirms it as a valuable strategy for improving chondroitin titers in engineered yeast cells. *QRI1* overexpression was also identified in the MEWpy approach, but only in 4 of the solutions. However, *GNA1* was not one of the identified targets in the MEWpy optimization. Instead, other gene commonly indicated for overexpression involved in the production of chondroitin precursors was *PGM2*. This gene encodes phosphoglucomutase, appearing in 54 solutions (in the fourth place, Figure 5.2). As it can be seen in Figure 5.1, this gene contributes to the production of UDP-glucose precursor.

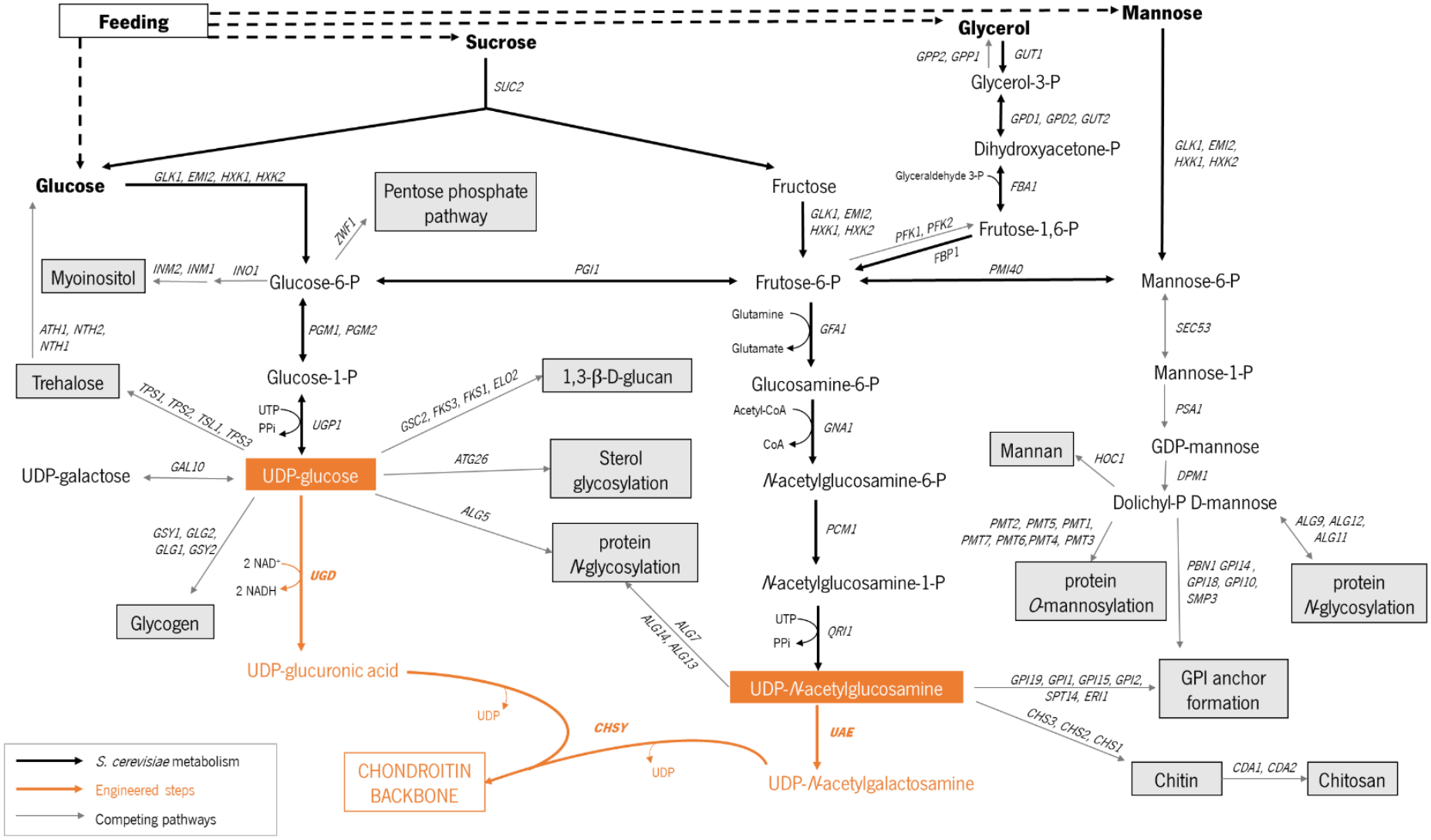

Feeding
Sucrose
Mannose
Glycerol
GPP2, GPP1
GUT1
Glycerol-3-P
GPD1, GPD2, GUT2
Dihydroxyacetone-P
Glyceraldehyde 3-P
FBA1
Frutose-1,6-P
PFK1, PFK2
FBP1
SUC2
Glucose
GLK1, EMI2, HXK1, HXK2
Pentose phosphate pathway
ZWF1
Fructose
GLK1, EMI2, HXK1, HXK2
GLK1, EMI2, HXK1, HXK2
Myoinositol
INM2, INM1
INO1
Glucose-6-P
PGI1
Frutose-6-P
PMI40
Mannose-6-P
ATH1, NTH2, NTH1
PGM1, PGM2
Glutamine
Glutamate
GFA1
SEC53
Trehalose
TPS1, TPS2, TSL1, TPS3
Glucose-1-P
1,3-β-D-glucan
GSC2, FKS3, FKS1, ELO2
Glucosamine-6-P
Mannose-1-P
UTP
PPi
UGP1
Acetyl-CoA
CoA
GNA1
PSA1
Mannan
HOC1
GDP-mannose
UDP-galactose
GAL10
UDP-glucose
ATG26
Sterol glycosylation
N-acetylglucosamine-6-P
DPM1
Dolichyl-P D-mannose
PCM1
PMT2, PMT5, PMT1, PMT7, PMT6,PMT4, PMT3
ALG9, ALG12, ALG11
GSY1, GLG2, GLG1, GSY2
ALG5
protein N-glycosylation
N-acetylglucosamine-1-P
protein O-mannosylation
PBN1 GPI14 , GPI18, GPI10, SMP3
protein N-glycosylation
2 NAD+
2 NADH
UGD
Glycogen
ALG14, ALG13
ALG7
UTP
PPi
QRI1
UDP-glucuronic acid
UDP-N-acetylglucosamine
GPI19, GPI1, GPI15, GPI2, SPT14, ERI1
GPI anchor formation
UDP
CHSY
UAE
CHS3, CHS2, CHS1
CHONDROITIN BACKBONE
UDP
UDP-N-acetylgalactosamine
Chitin
CDA1, CDA2
Chitosan
S. cerevisiae metabolism
Engineered steps
Competing pathways

Figure 5.1. Pathways involved in chondroitin production in engineered *Saccharomyces cerevisiae* cells, and competing pathways that redirect the metabolic flux from chondroitin precursors. **Compound abbreviations:** CoA – coenzyme A; NAD - nicotinamide adenine dinucleotide; PPi – diphosphate; GPI - glycosylphosphatidylinositol; UDP – uridine diphosphate; UTP - uridine triphosphate; **Gene descriptions:** *ALG11* - Alpha-1,2-mannosyltransferase; *ALG12* - Alpha-1,6-mannosyltransferase; *ALG9* – mannosyltransferase; *ALG5* - UDP-glucose:dolichyl-phosphate glucosyltransferase; *ATH1* - acid trehalase; *ATG26* - UDP-glucose:sterol glucosyltransferase; *CDA1*, *CDA2* – chitin deacetylase; *CHS1*, *CHS2*, *CHS3* - chitin synthases; *CHSY* - chondroitin synthase; *DPM1* - dolichol phosphate mannose synthase; *EMI2* - hexokinase; *ERI1* - endoplasmic reticulum-associated Ras Inhibitor; *FBA1* – fructose 1,6-bisphosphate aldolase; *FBP1* - fructose-1,6-bisphosphatase; *FKS1*, *FKS3* - 1,3-β-D-glucan synthase; *GAL10* – UDP-glucose-4-epimerase; *GLG1*, *GLG2* – glycogenin glucosyltransferase; *GLK1* - glucokinase; *GNA1* - glucosamine-6-phosphate *N*-acetyltransferase; *GPD1, GPD2* - glycerol-3-phosphate dehydrogenases; *GPI1*, *GPI10*, *GPI14*, *GPI15*, *GPI18*, *GPI19* – GPI anchor proteins; *GPP1*, *GPP2* - glycerol-3-phosphate phosphatases; *GSC2* - 1,3-β-glucan synthase; *GSY1*, *GSY2* - glycogen synthases; *GUT1* - glycerol kinase; *GUT2* - glycerol-3-phosphate dehydrogenase; *HXK1*, *HXK2* - hexokinases; *INO1* - inositol-3-phosphate synthase; *INM1*, *INM2* - inositol monophosphatases; *NTH1*, *NTH2* - neutral trehalases; *PBN1* - glycosylphosphatidylinositol-mannosyltransferase I; *PCM1* - *N*-acetylglucosamine-phosphate mutase; *PFK1, PFK2* - phosphofructokinase; *PGI1* - phosphoglucose isomerase; *PMT1*, *PMT2*, *PMT3*, *PMT4*, *PMT5*, *PMT6*, *PMT7* - protein *O*-mannosyltransferases; *PSA1* - guanosine diphosphate(GDP)-mannose pyrophosphorylase; *QRI1* - UDP-*N*-acetylglucosamine pyrophosphorylase; *SEC53* - Phosphomannomutase; *SMP3* - alpha 1,2-mannosyltransferase; *SPT14* - UDP-glycosyltransferase; *SUC2* – invertase; *TPS1* - trehalose-6-phosphate synthase; *TPS2* - trehalose-phosphatase; *TPS3* - trehalose-6-phosphatase; *UAE* - UDP-*N*-acetylglucosamine 4'-epimerase; *UGD* - UDP-glucose 6-dehydrogenase; *ZWF1* - glucose-6-phosphate dehydrogenase.

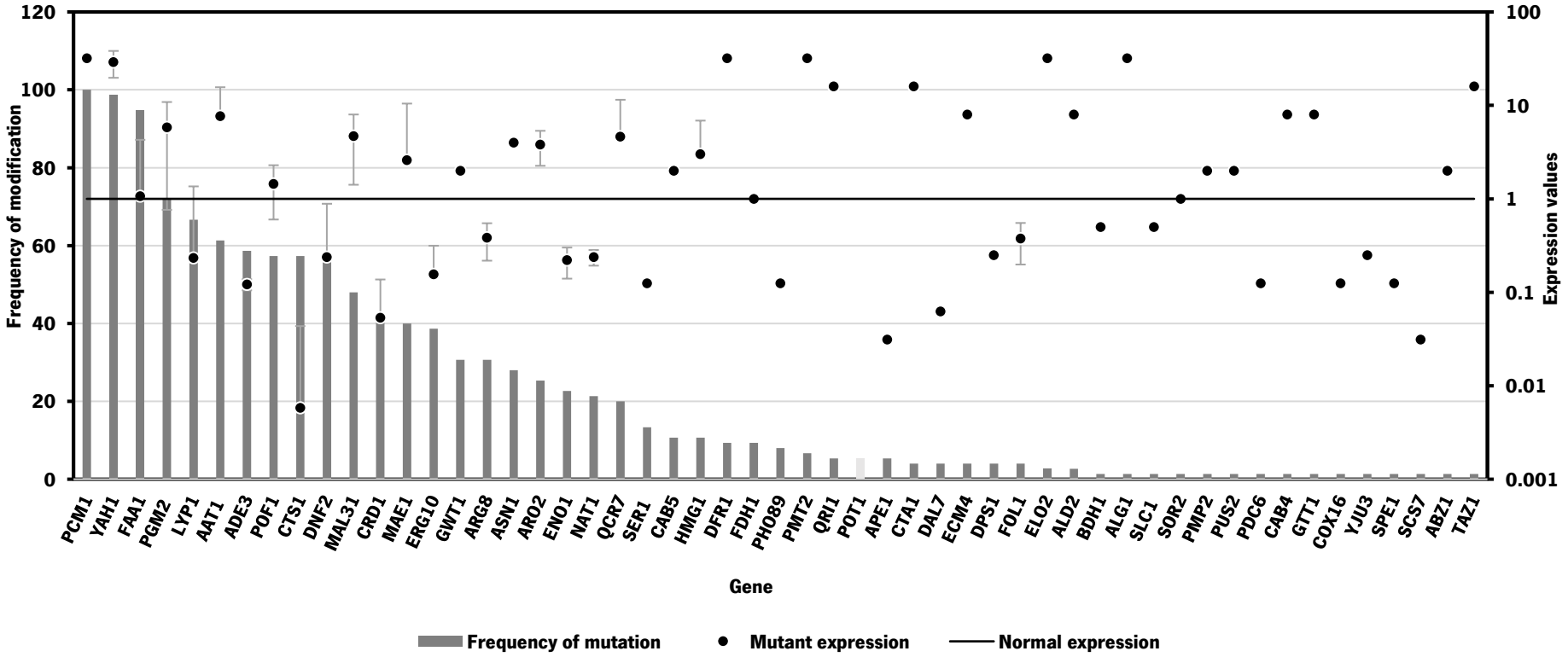


Figure 5.2. Frequency and expression values of genes in the solutions from optimization of *Saccharomyces cerevisiae* model for chondroitin production using MEWpy tool, limiting number of modifications to 10. The mutant expression (in dots) represents the average expression value. Mutant expressions higher than 1 represent overexpressions while values of expression lower than 1 represent underexpressions. Deletions are represented in light grey bars.

The second most common modification was the overexpression of *YAH1*, which encodes yeast adrenodoxin homolog, a ferredoxin involved in heme A biosynthesis by transferring electrons from nicotinamide adenine dinucleotide phosphate reduced form (NADPH) to heme O. The relationship between the overexpression of *YAH1* and the potential improvement of chondroitin production might not be immediately apparent. However, *YAH1* plays a crucial role in the electron transport chain and cellular redox balance within the mitochondria, and its overexpression leads to accumulation of heme A [65]. Consequently, the NAD^+ generated in this process could potentially be utilized in one of the reactions involved in chondroitin production, particularly the reaction catalyzed by UGD. This reaction requires NAD^+ as a co-factor, converting it to NADH during the transformation of UDP-glucose into UDP-glucuronic acid. The third most frequently identified gene target was the long chain fatty acyl-CoA synthetase gene (*FAA1*), which was observed either as overexpression or underexpression, depending on the proposed solution. Due to the inconsistency in the recommended gene expression for this gene, it can be inferred that its contribution to the enhancement of chondroitin production is not consistently significant.

The solutions with higher BPCY are described in Table 5.4.

Among the genes identified in the solutions with highest BPCY, only *QRI1* and *PCM1* were found to be directly involved in the pathways associated with chondroitin production (Figure 5.1).

However, there are several indirect relationships where modifications to other gene expressions may impact the *in silico* chondroitin production. For instance, the overexpression of the gene *POF1*, which encodes nicotinamide mononucleotide-specific adenylyltransferase, catalyzes the conversion of nicotinamide mononucleotide to nicotinamide adenine dinucleotide (NAD^+), an essential co-factor in chondroitin production, as discussed earlier. Therefore, the identification of *POF1* overexpression may be related with attempting to improve NAD^+ pool. Additionally, *CTS1*, which encodes endochitinase, was identified as a knockout target. As observed in Figure 5.1, chitin formation competes with chondroitin production pathway for UDP-acetylglucosamine substrate. Knocking out *CTS1* could redirect cellular resources and energy that would have been used for chitin breakdown towards the biosynthesis of chondroitin. This redirection could enhance the overall yield and efficiency of chondroitin production.

Table 5.4. Optimization results obtained for *Saccharomyces cerevisiae* model using MEWpy tool, allowing a maximum of 10 modifications. Growth rate and chondroitin production rate were predicted by phenotype simulations using parsimonious flux balance analysis (pFBA) and are presented in units of mol/gDW/h. BPCY was calculated by multiplying biomass by product and then dividing by substrate consumed, as predicted by pFBA. WYIELD is the weighed sum of the minimum and maximum product fluxes. Flux variability analysis (FVA) results are shown as minimum and maximum chondroitin obtained for fixed biomass.

Solution	BPCY	WYIELD	Genes modified			Predicted phenotype (pFBA)		FVA	
			Knock-out	Underexpression	Overexpression	Biomass	Chondroitin	Min chondroitin	Max chondroitin
1	0.04375	0.60872	*DNF2, CTS1*	*CRD1, LYP1, FAA1*	*YAH1, QRI1, PCM1, CAB5*	0.7317	0.5980	0.5980	0.9358
2	0.04374	0.60872	*FAA1, POT1, CTS1*	*MAL31, ARG8*	*YAH1, PCM1, ALD2*	0.7316	0.5980	0.5980	0.9358
3	0.04335	0.60876	*ENO1, CTS1*	*ERG10*	*FAA1, YAH1, AAT1, POF1, CTA1, PCM1, MAL31*	0.7248	0.5981	0.5981	0.9365

Gene descriptions: *AAT1* - aspartate aminotransferase; *ALD2* - aldehyde dehydrogenase; *ARG8* - acetylornithine aminotransferase; *CAB5* - subunit of the CoA-Synthesizing Protein Complex; *CRD1* - cardiolipin synthase; *CTA1* - catalase A; *CTS1* - endochitinase; *DNF2* – phospholipid-transporting ATPase; *ENO1* - enolase I; *ERG10* - acetyl-CoA C-acetyltransferase; *FAA1* - long chain fatty acyl-CoA synthetase; *LYP1* - lysine permease; *MAL31* - maltose permease; *PCM1* - *N*-acetylglucosamine-phosphate mutase; *POF1* - nicotinamide mononucleotide-specific adenylyltransferase, *POT1* - 3-ketoacyl-CoA thiolase; *QRI1* - UDP-*N*-acetylglucosamine pyrophosphorylase; *YAH1* - yeast adrenodoxin homolog.

The size of the resulting solutions was between 8 to 10 gene modifications. However, the resulting BPCY was not higher than the one obtained in the OptFlux solutions where only one gene expression was altered. In terms of FVA analysis, the robustness from MEWpy solutions was neither higher nor lower than the ones from OptFlux approach. Also, changing the gene expression of 8 to 10 genes would be difficult to implement and would possibly significantly affect the *S. cerevisiae* growth. Therefore, the optimization was again run now limiting the number of modifications to 3. The new optimization using MEWpy led to 28 solutions. These solutions included modifications in 14 different genes. The frequency and expression value of each gene throughout the solutions is presented in Figure 5.3.

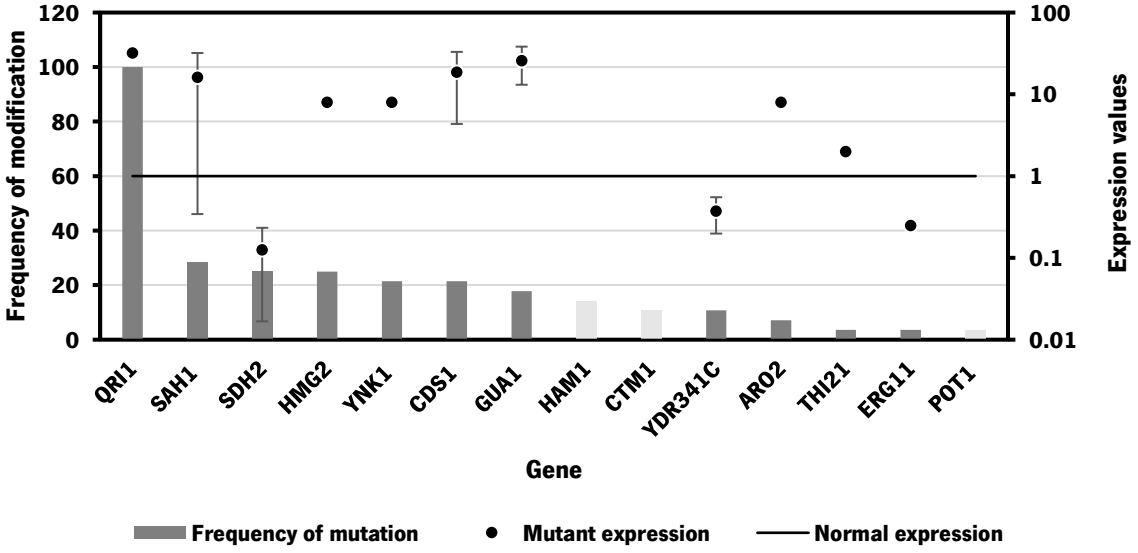


Figure 5.3. Frequency and expression values of genes in the solutions from optimization of *Saccharomyces cerevisiae* model using MEWpy tool, limiting number of modifications to 3. The mutant expression (in dots) represents the average expression value. Mutant expressions higher than 1 represent overexpressions while values of expression lower than 1 represent underexpressions. Deletions are represented in light grey bars.

In this case, all solutions included *QRI1* overexpression, which was also predicted in the above-mentioned approaches (Table 5.3 and Figure 5.2). The best solutions in terms of BPCY are described in Table 5.5.

Table 5.5. Optimization results obtained for *Saccharomyces cerevisiae* model using MEWpy tool, allowing a maximum of 3 modifications. Growth rate and chondroitin production rate were predicted by phenotype simulations using parsimonious flux balance analysis (pFBA) and are presented in units of mol/gDW/h. BPCY was calculated by multiplying biomass by product and then dividing by substrate consumed, as predicted by pFBA. WYIELD is the weighed sum of the minimum and maximum product fluxes. Flux variability analysis (FVA) results are shown as minimum and maximum chondroitin obtained for fixed biomass.

Solution	BPCY	W YIELD	Genes modified			Predicted phenotype (pFBA)		FVA	
			Knock-out	Under expression	Over expression	Biomass	Chondroitin	Min chondroitin	Max chondroitin
1	0.04375	0.83445	-	-	*QRI1*	0.7317	0.5980	0.5980	0.9358
2	0.01488	2.80306	*CTM1*	*CDS1*	*QRI1*	0.0541	0.6131	0.6131	3.7416
3	0.04359	0.83489	*HAM1*	*SDH2*	*QRI1*	0.7288	0.5980	0.5980	0.9364

Gene descriptions: *CDS1* - phosphatidate cytidylyltransferase; *CTM1* - cytochrome c lysine methyltransferase; *HAM1* - Nucleoside triphosphate pyrophosphohydrolase; *QRI1* – uridine diphosphate-*N*-acetylglucosamine pyrophosphorylase; SDH2 - iron-sulfur protein subunit of succinate dehydrogenase.

As the BPCY values were still low, attempts to find more efficient mutants were made by combining several of the identified promising modifications *in silico* (Table 5.6). However, the obtained phenotypes of engineered strains with cumulative mutations did not exhibit improved BPCY nor chondroitin production compared to the ones with single modifications.

For future perspectives, other types of models should be explored for more meaningful results on identification of targets for metabolic engineering. While GEMs can give insights into novel metabolic engineering targets, the phenotype prediction could be more accurate if kinetic data, enzyme usage-constraints and regulatory information were included in the model. For example, GECKO is a method that enhances a GEM to account for enzymes as part of reactions and has been applied to a *S. cerevisiae* model [66].

Table 5.6. Attempts to obtain higher chondroitin production *in silico* by combining promising gene overexpressions. Results from simulations in OptFlux using parsimonious flux balance analysis (pFBA) as simulation method for growth rate (biomass) and chondroitin production rate, presented in units of mol/gDW/h. BPCY was calculated by multiplying biomass by product and then dividing by substrate consumed, as predicted by pFBA.

Overexpressions	Biomass	Chondroitin	BPCY
PCM1, QRI1	0.7317	0.5980	0.04375
PCM1, GNA1	0.7317	0.5980	0.04375
QRI1, GNA1	0.7317	0.5980	0.04375
PCM1, QRI1, GNA1	0.7317	0.5980	0.04375

Gene descriptions: *GNA1* - glucosamine-6-phosphate acetyltransferase; *PCM1* - *N*-acetylglucosamine-phosphate mutase; *QRI1* – uridine diphosphate-*N*-acetylglucosamine pyrophosphorylase

5.3.2. Flasks fermentations

The reports on the use of eukaryotic microorganisms for chondroitin or chondroitin sulfate production are still very limited. The first study on this area used *P. pastoris* to express the sulfotransferases that were then used to sulfate a chondroitin backbone produced by an engineered *Bacillus subtilis* strain [7]. The same group then engineered a *P. pastoris* strain containing both chondroitin biosynthesis and sulfonation modules, that was able to reach a production of 2.1 g/L chondroitin sulfate with a sulfonation degree of 4.0%. [8]. That work is the only one using a eukaryotic microorganism for chondroitin production. More recently, the same group also engineered a *P. pastoris* strain to produce another complex glycosaminoglycan, namely heparin [67].

To evaluate the ability of a widely used eukaryotic microorganism to potentially host industrial biotechnological process of chondroitin production, we applied efforts to produce chondroitin in *S. cerevisiae*. However, after transforming the plasmids carrying the designed pathways for chondroitin production, the transformants were rare, and after picking colonies from agar plates, some colonies were not able to grow on pre-inoculum liquid medium. Figure 5.4 shows the performance on chondroitin production by the few transformants we were able to test.

S. cerevisiae is often used as a host organism for the expression of heterologous genes and can carry multiple plasmids simultaneously. However, introducing multiple plasmids into a yeast cell can have various effects on cell growth and physiology. Some potential problems that may arise include: (a) metabolic burden - the presence of multiple plasmids and the expression of heterologous genes can impose an additional metabolic burden on the host yeast cell, resulting in reduced growth rates and compromised cell viability; (b) competitive replication - plasmids often compete for limited cellular

resources during replication, leading to instability and loss of one or both plasmids over time, leading to a heterogeneous population of cells with varying plasmid content; and (c) induced stress responses - the expression of foreign genes may induce stress responses in the host cell, triggering various regulatory mechanisms that can affect cellular homeostasis and growth.

To overcome these potential issues, several strategies can be employed including: (a) balanced expression of genes - fine-tuning the expression levels of multiple genes can help alleviate the metabolic burden and minimize adverse effects on cell growth and physiology; (b) strain engineering - using engineered yeast strains with improved capabilities for handling metabolic stress or expressing foreign genes can help mitigate the negative impacts on cell growth; and (c) adaptative laboratory evolution - improving the performance of microbial strains under specific conditions by subjecting a population of microorganisms to prolonged periods of growth under controlled selective pressure, allows the natural selection of beneficial mutations that may result in yeast strains adapted to efficiently manage the additional genetic load. The combination of these strategies might result in robust yeast strains capable of efficiently carrying multiple plasmids and expressing heterologous genes without compromising growth or productivity.

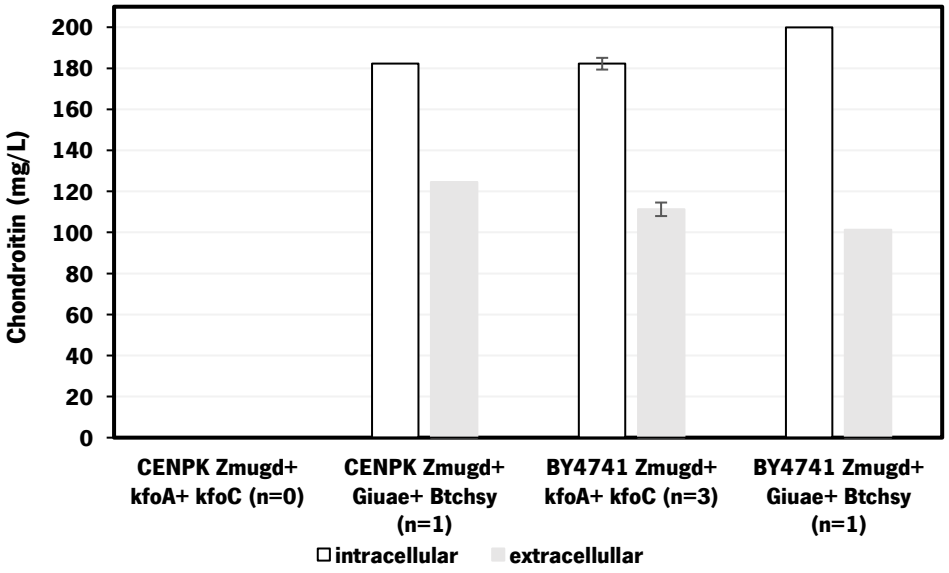


Figure 5.4. Chondroitin production in engineered *Saccharomyces cerevisiae* CEN.PK2-1C and BY4741 strains.

Despite few viable colonies were obtained in transformations, the strains herein constructed were able to produce chondroitin between 182 and 200 mg/L, without significant differences between the different constructs and strains.

Comparing to the other work describing chondroitin production using *P.* pastoris [67], the genes used for the chondroitin production module were *kfoC*, *kfoA* (from *E. coli* K4), and *tuaD* (UDP-glucose dehydrogenase from *B. subtilis*) which yielded only 5.5 mg/L chondroitin. After the genes were further codon-optimized, and the bioprocess was changed to a fed-batch fermentation, the chondroitin production increased to 2.1 g/L. This suggests that the chondroitin production herein obtained can also be further improved using some of the techniques discussed above.

Furthermore, since the robustness of the engineered strains, culture medium and operational mode need additional optimizations, further genetic modifications on the host harboring the heterologous pathway were not performed. Once a robust host strain is constructed, further modifications, such as *QRI1, GNA1 or PCM1* overexpression, as suggested by the bioinformatic approaches herein described, should be performed.

5.4. Conclusions

Chondroitin sulfate stands as a valuable natural compound, with its distinct sulfation patterns and properties leading to a wide range of practical uses. Its biotechnological production presents an intriguing opportunity. As *S. cerevisiae* is a robust, fast-growing, and easily mutated host, it has been selected to be engineered with chondroitin biosynthetic pathways. Also, this is an alternative host that should be considered for chondroitin production due to its ability to perform post-translational modifications. Since the application of computational-aided metabolic engineering might help discover critical bottlenecks in its heterologous biosynthesis, a metabolic model and flux analysis has been herein used for strain design. The model yeastGEM_c has been constructed using yeast GEM as scaffold in which chondroitin production reactions and genes have been included. Using evolutionary algorithms, several promising novel targets, such as *QRI1, GNA1* or *PCM1* overexpression, have been identified as promising to improve chondroitin titers of engineered *S. cerevisiae* strains. Thus, the strategy reported here should be useful in developing industrial *S. cerevisiae* strains capable of efficient chondroitin production. Furthermore, novel biosynthetic pathways have been evaluated for producing chondroitin in this host, and 125 mg/L of extracellular chondroitin have been obtained through flask fermentation. The use of these pathways to construct the predicted engineered strains, in combination with other methods, such as integration of genes in yeast genome and adaptative lab evolution, can lead to further improved chondroitin yields.

5.5. References

[1] M.R. Couto, J.L. Rodrigues, L.R. Rodrigues, Heterologous production of chondroitin, Biotechnol. Reports. 33 (2022). https://doi.org/10.1016/j.btre.2022.e00710.

[2] B.F. Cress, J.A. Englaender, W. He, D. Kasper, R.J. Linhardt, M.A.G. Koffas, Masquerading microbial pathogens: Capsular polysaccharides mimic host-tissue molecules, FEMS Microbiol. Rev. 38 (2014) 660–697. https://doi.org/10.1111/1574-6976.12056.

[3] D. Cimini, E. Carlino, A. Giovane, O. Argenzio, I. Dello Iacono, M. De Rosa, C. Schiraldi, Engineering a branch of the UDP-precursor biosynthesis pathway enhances the production of capsular polysaccharide in *Escherichia coli* O5:K4:H4, Biotechnol. J. 10 (2015) 1307–1315. https://doi.org/10.1002/biot.201400602.

[4] P. Jin, L. Zhang, P. Yuan, Z. Kang, G. Du, J. Chen, Efficient biosynthesis of polysaccharides chondroitin and heparosan by metabolically engineered *Bacillus subtilis*, Carbohydr. Polym. 140 (2016) 424–432. https://doi.org/10.1016/j.carbpol.2015.12.065.

[5] W. He, L. Fu, G. Li, J. Andrew Jones, R.J. Linhardt, M. Koffas, J.A. Jones, R.J. Linhardt, M. Koffas, J. Andrew Jones, R.J. Linhardt, M. Koffas, Production of chondroitin in metabolically engineered *E. coli*, Metab. Eng. 27 (2015) 92–100. https://doi.org/10.1016/j.ymben.2014.11.003.

[6] F. Cheng, S. Luozhong, H. Yu, Z. Guo, Biosynthesis of chondroitin in engineered *Corynebacterium glutamicum*, J. Microbiol. Biotechnol. 29 (2019) 392–400. https://doi.org/10.4014/JMB.1810.10062.

[7] Z. Zhou, Q. Li, H. Huang, H. Wang, Y. Wang, G. Du, J. Chen, Z. Kang, A microbial–enzymatic strategy for producing chondroitin sulfate glycosaminoglycans, Biotechnol. Bioeng. 115 (2018) 1561–1570. https://doi.org/10.1002/bit.26577.

[8] X. Jin, W. Zhang, Y. Wang, J. Sheng, R. Xu, J. Li, G. Du, Z. Kang, Biosynthesis of non-animal chondroitin sulfate from methanol using genetically engineered *Pichia pastoris*, Green Chem. 23 (2021) 4365–4374. https://doi.org/10.1039/d1gc00260k.

[9] M.M. Desko, D.A. Gross, J.J. Kohler, Effects of N-glycosylation on the activity and localization of GlcNAc-6-sulfotransferase 1, Glycobiology. 19 (2009) 1068–1077. https://doi.org/10.1093/glycob/cwp092.

[10] A. Badri, A. Williams, A. Awofiranye, P. Datta, K. Xia, W. He, K. Fraser, J.S. Dordick, R.J. Linhardt, M.A.G.G. Koffas, Complete biosynthesis of a sulfated chondroitin in *Escherichia coli*, Nat. Commun. 12 (2021) 1–10. https://doi.org/10.1038/s41467-021-21692-5.

[11] J. Nielsen, Yeast systems biology: Model organism and cell factory, Biotechnol. J. 14 (2019) 1–9. https://doi.org/10.1002/biot.201800421.

[12] M. Li, I. Borodina, Application of synthetic biology for production of chemicals in yeast *Saccharomyces cerevisiae*, FEMS Yeast Res. 15 (2015) 1–12. https://doi.org/10.1111/1567-1364.12213.

[13] S.K. Gupta, P. Shukla, Sophisticated cloning, fermentation, and purification technologies for an

enhanced therapeutic protein production: A review, Front. Pharmacol. 8 (2017) 1–17. https://doi.org/10.3389/fphar.2017.00419.

[14] H. Lopes, I. Rocha, Genome-scale modeling of yeast: chronology, applications and critical perspectives, FEMS Yeast Res. 17 (2017) 1–14. https://doi.org/10.1093/femsyr/fox050.

[15] V. Vieira, P. Maia, M. Rocha, I. Rocha, Comparison of pathway analysis and constraint-based methods for cell factory design, BMC Bioinformatics. 20 (2019) 1–15. https://doi.org/10.1186/s12859-019-2934-y.

[16] P. Maia, M. Rocha, I. Rocha, *In silico* constraint-based strain optimization methods: the quest for optimal cell factories, Microbiol. Mol. Biol. Rev. 80 (2016) 45–67. https://doi.org/10.1128/mmbr.00014-15.

[17] X. Bi, Y. Liu, J. Li, G. Du, X. Lv, L. Liu, Construction of multiscale genome-scale metabolic models: Frameworks and challenges, Biomolecules. 12 (2022) 1–21. https://doi.org/10.3390/biom12050721.

[18] R.H. Ng, J.W. Lee, P. Baloni, C. Diener, J.R. Heath, Y. Su, Constraint-based reconstruction and analyses of metabolic models: Open-source Python tools and applications to cancer, Front. Oncol. 12 (2022) 1–23. https://doi.org/10.3389/fonc.2022.914594.

[19] A. Mardinoglu, J. Nielsen, Systems medicine and metabolic modelling, J. Intern. Med. 271 (2012) 142–154. https://doi.org/10.1111/j.1365-2796.2011.02493.x.

[20] C. Zhang, Q. Hua, Applications of genome-scale metabolic models in biotechnology and systems medicine, Front. Physiol. 6 (2016) 1–8. https://doi.org/10.3389/fphys.2015.00413.

[21] X. Fang, C.J. Lloyd, B.O. Palsson, Reconstructing organisms *in silico*: genome-scale models and their emerging applications, Nat. Rev. Microbiol. 18 (2020) 731–743. https://doi.org/10.1038/s41579-020-00440-4.

[22] C. Gu, G.B. Kim, W.J. Kim, H.U. Kim, S.Y. Lee, Current status and applications of genome-scale metabolic models, Genome Biol. 20 (2019) 1–18. https://doi.org/10.1186/s13059-019-1730-3.

[23] X. Chen, G. Xu, N. Xu, W. Zou, P. Zhu, L. Liu, J. Chen, Metabolic engineering of *Torulopsis glabrata* for malate production, Metab. Eng. 19 (2013) 10–16. https://doi.org/10.1016/j.ymben.2013.05.002.

[24] G. Xu, W. Zou, X. Chen, N. Xu, L. Liu, J. Chen, Fumaric Acid Production in *Saccharomyces cerevisiae* by *In Silico* Aided Metabolic Engineering, PLoS One. 7 (2012) 1–10. https://doi.org/10.1371/journal.pone.0052086.

[25] J.M. Otero, D. Cimini, K.R. Patil, S.G. Poulsen, L. Olsson, J. Nielsen, Industrial Systems Biology of *Saccharomyces cerevisiae* enables novel succinic acid cell factory, PLoS One. 8 (2013) 1–10. https://doi.org/10.1371/journal.pone.0054144.

[26] C.W. Song, D.I. Kim, S. Choi, J.W. Jang, S.Y. Lee, Metabolic engineering of *Escherichia coli* for the production of fumaric acid, Biotechnol. Bioeng. 110 (2013) 2025–2034.

https://doi.org/10.1002/bit.24868.

[27] X. Chen, J. Wu, W. Song, L. Zhang, H. Wang, L. Liu, Fumaric acid production by *Torulopsis glabrata*: Engineering the urea cycle and the purine nucleotide cycle, Biotechnol. Bioeng. 112 (2015) 156–167. https://doi.org/10.1002/bit.25334.

[28] B.J. Harder, K. Bettenbrock, S. Klamt, Model-based metabolic engineering enables high yield itaconic acid production by *Escherichia coli*, Metab. Eng. 38 (2016) 29–37. https://doi.org/10.1016/j.ymben.2016.05.008.

[29] P. Mishra, N.R. Lee, M. Lakshmanan, M. Kim, B.G. Kim, D.Y. Lee, Genome-scale model-driven strain design for dicarboxylic acid production in *Yarrowia lipolytica*, BMC Syst. Biol. 12 (2018). https://doi.org/10.1186/s12918-018-0542-5.

[30] C. Bro, B. Regenberg, J. Förster, J. Nielsen, *In silico* aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production, Metab. Eng. 8 (2006) 102–111. https://doi.org/10.1016/j.ymben.2005.09.007.

[31] H. Yim, R. Haselbeck, W. Niu, C. Pujol-Baxley, A. Burgard, J. Boldt, J. Khandurina, J.D. Trawick, R.E. Osterhout, R. Stephen, J. Estadilla, S. Teisan, H.B. Schreyer, S. Andrae, T.H. Yang, S.Y. Lee, M.J. Burk, S. Van Dien, Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol, Nat. Chem. Biol. 7 (2011) 445–452. https://doi.org/10.1038/nchembio.580.

[32] C. Ng, M. Jung, J. Lee, M.-K. Oh, Production of 2,3-butanediol in *Saccharomyces cerevisiae* by *in silico* aided metabolic engineering, Microb. Cell Fact. 11 (2012) 68. https://doi.org/10.1186/1475-2859-11-68.

[33] J. Becker, O. Zelder, S. Häfner, H. Schröder, C. Wittmann, From zero to hero-Design-based systems metabolic engineering of *Corynebacterium glutamicum* for l-lysine production, Metab. Eng. 13 (2011) 159–168. https://doi.org/10.1016/j.ymben.2011.01.003.

[34] J.H. Park, T.Y. Kim, K.H. Lee, S.Y. Lee, Fed-batch culture of *Escherichia coli* for L-valine production based on *in silico* flux response analysis, Biotechnol. Bioeng. 108 (2011) 934–946. https://doi.org/10.1002/bit.22995.

[35] J. van Ooyen, S. Noack, M. Bott, A. Reth, L. Eggeling, Improved L-lysine production with *Corynebacterium glutamicum* and systemic insight into citrate synthase flux and activity, Biotechnol. Bioeng. 109 (2012) 2070–2081. https://doi.org/10.1002/bit.24486.

[36] Y.K. Jung, T.Y. Kim, S.J. Park, S.Y. Lee, Metabolic engineering of Escherichia coli for the production of polylactic acid and its copolymers, Biotechnol. Bioeng. 105 (2010) 161–171. https://doi.org/10.1002/bit.22548.

[37] I. Poblete-Castro, D. Binger, A. Rodrigues, J. Becker, V.A.P. Martins Dos Santos, C. Wittmann, In-silico-driven metabolic engineering of *Pseudomonas putida* for enhanced production of poly-hydroxyalkanoates, Metab. Eng. 15 (2013) 113–123. https://doi.org/10.1016/j.ymben.2012.10.004.

[38] K.R. Kildegaard, N.B. Jensen, K. Schneider, E. Czarnotta, E. Özdemir, T. Klein, J. Maury, B.E. Ebert, H.B. Christensen, Y. Chen, I.K. Kim, M.J. Herrgård, L.M. Blank, J. Forster, J. Nielsen, I.

Borodina, Engineering and systems-level analysis of *Saccharomyces cerevisiae* for production of 3-hydroxypropionic acid via malonyl-CoA reductase-dependent pathway, Microb. Cell Fact. 15 (2016) 1–13. https://doi.org/10.1186/s12934-016-0451-5.

[39] J.E. Yang, S.J. Park, W.J. Kim, H.J. Kim, B.J. Kim, H. Lee, J. Shin, S.Y. Lee, One-step fermentative production of aromatic polyesters from glucose by metabolically engineered *Escherichia coli* strains, Nat. Commun. 9 (2018). https://doi.org/10.1038/s41467-017-02498-w.

[40] M. Kim, J. Sang Yi, J. Kim, J.N. Kim, M.W. Kim, B.G. Kim, Reconstruction of a high-quality metabolic model enables the identification of gene overexpression targets for enhanced antibiotic production in *Streptomyces coelicolor* A3(2), Biotechnol. J. 9 (2014) 1185–1194. https://doi.org/10.1002/biot.201300539.

[41] D. Huang, J. Wen, G. Wang, G. Yu, X. Jia, Y. Chen, *In silico* aided metabolic engineering of *Streptomyces roseosporus* for daptomycin yield improvement, Appl. Microbiol. Biotechnol. 94 (2012) 637–649. https://doi.org/10.1007/s00253-011-3773-6.

[42] J.A. Chemler, Z.L. Fowler, K.P. McHugh, M.A.G. Koffas, Improving NADPH availability for natural product biosynthesis in *Escherichia coli* by metabolic engineering, Metab. Eng. 12 (2010) 96–104. https://doi.org/10.1016/j.ymben.2009.07.003.

[43] N. Bhan, P. Xu, O. Khalidi, M.A.G. Koffas, Redirecting carbon flux into malonyl-CoA to improve resveratrol titers: Proof of concept for genetic interventions predicted by OptForce computational framework, Chem. Eng. Sci. 103 (2013) 109–114. https://doi.org/10.1016/j.ces.2012.10.009.

[44] P. Xu, S. Ranganathan, Z.L. Fowler, C.D. Maranas, M.A.G.G. Koffas, Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA, Metab. Eng. 13 (2011) 578–587. https://doi.org/10.1016/j.ymben.2011.06.008.

[45] H. Lu, F. Li, B.J. Sánchez, Z. Zhu, G. Li, I. Domenzain, S. Marcišauskas, P.M. Anton, D. Lappa, C. Lieven, M.E. Beber, N. Sonnenschein, E.J. Kerkhoven, J. Nielsen, A consensus *S. cerevisiae* metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism, Nat. Commun. 10 (2019). https://doi.org/10.1038/s41467-019-11581-3.

[46] I. Rocha, P. Maia, P. Evangelista, P. Vilaça, S. Soares, J.P. Pinto, J. Nielsen, K.R. Patil, E.C. Ferreira, M. Rocha, OptFlux: an open-source software platform for *in silico* metabolic engineering, BMC Syst. Biol. 4 (2010) 1–12. https://doi.org/10.1186/1752-0509-4-45.

[47] N.E. Lewis, K.K. Hixson, T.M. Conrad, J.A. Lerman, P. Charusanti, A.D. Polpitiya, J.N. Adkins, G. Schramm, S.O. Purvine, D. Lopez-Ferrer, K.K. Weitz, R. Eils, R. König, R.D. Smith, B. Palsson, Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models, Mol. Syst. Biol. 6 (2010). https://doi.org/10.1038/msb.2010.47.

[48] E. Zitzler, M. Laumanns, L. Thiele, SPEA2: Improving the Strength Pareto Evolutionary Algorithm, Evol. Methods Des. Optim. Control with Appl. to Ind. Probl. (2001) 95–100. https://doi.org/10.3929/ethz-a-004284029.

[49] V. Pereira, F. Cruz, M. Rocha, MEWpy: a computational strain optimization workbench in Python, Bioinformatics. (2021) 1–3. https://doi.org/10.1093/bioinformatics/btab013.

[50] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evol. Comput. 6 (2002) 182–197. https://doi.org/10.1109/4235.996017.

[51] R. Mahadevan, C.H. Schilling, The effects of alternate optimal solutions in constraint-based genome-scale metabolic models, Metab. Eng. 5 (2003) 264–276. https://doi.org/10.1016/j.ymben.2003.09.002.

[52] K.-D. Entian, P. Kötter, 25 Yeast Genetic Strain and plasmid collections, in: I. Stansfield, M.J.R. Stark (Eds.), Methods Microbiol., 2nd ed., Elsevier, 2007: pp. 629–666. https://doi.org/10.1016/S0580-9517(06)36025-4.

[53] C.B. Brachmann, A. Davies, G.J. Cost, E. Caputo, J. Li, P. Hieter, J.D. Boeke, Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications, Yeast. 14 (1998) 115–132. https://doi.org/10.1002/(SICI)1097-0061(19980130)14:2<115::AID-YEA204>3.0.CO;2-2.

[54] Y. Chen, S. Partow, G. Scalcinati, V. Siewers, J. Nielsen, Enhancing the copy number of episomal plasmids in *Saccharomyces cerevisiae* for improved protein production, FEMS Yeast Res. 12 (2012) 598–607. https://doi.org/10.1111/j.1567-1364.2012.00809.x.

[55] M.R. Couto, J.L. Rodrigues, L.R. Rodrigues, Cloning, expression and characterization of UDP-glucose dehydrogenases, Life. 11 (2021) 1201. https://doi.org/10.3390/life11111201.

[56] R.D. Gietz, R.A. Woods, Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method, in: Methods Enzymol., 2002: pp. 87–96. https://doi.org/10.1016/S0076-6879(02)50957-5.

[57] T. Bitter, H.M. Muir, A modified uronic acid carbazole reaction, Anal. Biochem. 4 (1962) 330–334. https://doi.org/10.1016/0003-2697(62)90095-7.

[58] J.A. Shaw, P.C. Mol, B. Bowers, S.J. Silverman, M.H. Valdivieso, A. Duran, E. Cabib, The function of chitin synthases 2 and 3 in the *Saccharomyces cerevisiae* cell cycle, J. Cell Biol. 114 (1991) 111–123. https://doi.org/10.1083/jcb.114.1.111.

[59] G. Lesage, H. Bussey, Cell wall assembly in *Saccharomyces cerevisiae* , Microbiol. Mol. Biol. Rev. 70 (2006) 317–343. https://doi.org/10.1128/mmbr.00038-05.

[60] P. Orlean, Architecture and biosynthesis of the *Saccharomyces cerevisiae* cell wall, Genetics. 192 (2012) 775–818. https://doi.org/10.1534/genetics.112.144485.

[61] F.M. Klis, P. Mol, K. Hellingwerf, S. Brul, Dynamics of cell wall structure *in Saccharomyces cerevisiae*, FEMS Microbiol. Rev. 26 (2002) 239–256. https://doi.org/10.1016/S0168-6445(02)00087-6.

[62] Q. Zhang, R. Yao, X. Chen, L. Liu, S. Xu, J. Chen, J. Wu, Enhancing fructosylated chondroitin production in *Escherichia coli* K4 by balancing the UDP-precursors, Metab. Eng. 47 (2018) 314–322. https://doi.org/10.1016/j.ymben.2018.04.006.

[63] J. Wu, Q. Zhang, L. Liu, X. Chen, J. Liu, Q. Luo, Recombinant *Escherichia coli* for high efficiency production of fructosylated chondroitin and method for making thereof, US Patent 10,508,279

B2, 2019.

[64] S. D'ambrosio, A. Alfano, E. Cassese, O.F. Restaino, S. Barbuto Ferraiuolo, R. Finamore, M. Cammarota, C. Schiraldi, D. Cimini, Production and purification of higher molecular weight chondroitin by metabolically engineered *Escherichia coli* K4 strains, Sci. Rep. 10 (2020) 1–10. https://doi.org/10.1038/s41598-020-70027-9.

[65] M.H. Barros, A. Tzagoloff, Regulation of the heme A biosynthetic pathway in *Saccharomyces cerevisiae*, FEBS Lett. 516 (2002) 119–123. https://doi.org/10.1016/S0014-5793(02)02514-0.

[66] B.J. Sánchez, C. Zhang, A. Nilsson, P. Lahtvee, E.J. Kerkhoven, J. Nielsen, Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints, Mol. Syst. Biol. 13 (2017) 935. https://doi.org/10.15252/msb.20167411.

[67] Y. Zhang, Y. Wang, Z. Zhou, P. Wang, X. Xi, S. Hu, R.R. Xu, G. Du, J. Li, J. Chen, Z. Kang, Synthesis of bioengineered heparin by recombinant yeast *Pichia pastoris*, Green Chem. (2022) 3180–3192. https://doi.org/10.1039/d1gc04672a.

CHAPTER 6

MAIN CONCLUSIONS AND FUTURE PERSPECTIVES

Abstract

The biotechnological production of chondroitin sulfate, and other glycosamingoglycans (GAGs), is still very challenging. This chapter presents the main conclusions of this thesis. Additionally, future perspectives in this research area are discussed, and the next steps for improving the knowledge on chondroitin biotechnological production.

Main Conclusions and Future Perspectives

The main goal of this work was to explore alternative enzymes and alternative hosts to produce chondroitin. This goal was accomplished by ultimately introducing heterologous genes in *Saccharomyces cerevisiae* strains, although the chondroitin titers obtained were modest and further improvement of the engineered strains are still needed.

The main conclusions that can be drawn from this thesis are further summarized.

The first experimental work from this thesis was to study the potential of different novel uridine diphosphate (UDP)-glucose 6-dehydrogenase (UGD) enzymes for the biotechnological production of chondroitin. Both *Escherichia coli* and *S. cerevisiae* were used for the expression of these enzymes from the GAG production pathway. Three UGD enzymes (from *Zymomonas mobilis*, *Lactobacillus johnsonii* and *Capra hircus*), each one being a representative of a different evolutionary group of these enzymes, were evaluated in terms of kinetics and activity in *in vivo* and *in vitro* conditions. Interestingly, the best performing enzyme varied according to the conditions, emphasizing the importance of choosing the right enzyme for the intended production strategy. *E. coli* was found to be a useful system for protein expression, with the recombinant proteins being clearly overexpressed in a soluble form. Furthermore, significant increases in the UDP-glucuronic acid (UDP-GlcA) pool in cell extracts of engineered *E. coli* were observed compared to the negative control. *S. cerevisiae* was also able to express active UGDs as shown by the improved UDP-GlcA pools. *Zm*UGD generally exhibited the highest activity *in vivo* in both *E. coli* and *S. cerevisiae* mutants. The UGD from animal (*C. hircus*) origin tested was codon-optimized for *S. cerevisiae* and therefore only expressed in this system. Although it was active, it did not result in as much UDP-GlcA production as *Zm*UGD did.

As our goal was to construct a microbial cell factory for chondroitin production, the best performing enzyme regarding *in vivo* activity, *Zm*UGD, was used, in combination with other required enzymes for the complete assembly (UAE, UDP-N-acetylglucosamine 4'-epimerase, and CHSY, chondroitin synthase, encoded by *E. coli* K4 genes *kfoA* and *kfoC*, respectively) of chondroitin biosynthetic pathway. This construct was able to produce up to 62 mg/L of chondroitin. Additionally, genome-scale model optimization was applied to identify potential targets for improving chondroitin production. The solutions varied according to the type of computational approach. Generally, genes identified for repression were often from the cell wall-related pathways, such as *nagZ* (β-*N*-acetylglucosaminidase) or amidases *amiABC,* which compete for precursors with chondroitin synthesis. Regarding predicted overexpressions, targets were usually from the pathways from chondroitin precursors production, including *glmM*

(phosphoglucosamine mutase) and *glmU* (glucosamine-1-phosphate *N*-acetyltransferase/ *N*-acetylglucosamine-1-phosphate uridyltransferase). Based on the predicted *in silico* mutations, *E. coli* mutants with *mltB* (lytic murein transglycosylase), *sodA* (superoxide dismutase), or *glmU* overexpressions, and *ybiV* (sugar phosphatase) underexpression were evaluated for chondroitin production in preliminary flask fermentation assays. The evaluated mutants with additional modifications on the strain harboring the chondroitin biosynthetic pathway resulted in chondroitin titers from 42 to 118 mg/L. Therefore, some of these mutants demonstrated increased chondroitin production, highlighting the effectiveness of metabolic engineering in enhancing chondroitin yields. Scale up assays showed that chondroitin production could be greatly improved in *E. coli* engineered strains. Batch fermentations in bioreactor improved chondroitin production, with the highest titers being achieved by *E. coli* K-12 overexpressing *mltB*. Using fed-batch fermentation, chondroitin titers with this mutant improved to 535 mg/L.

S. cerevisiae showed promise as an alternative host for chondroitin production, due to its wide use, robustness, and capability for post-translational modifications. Heterologous biosynthetic pathways to produce chondroitin were constructed in *S. cerevisiae* strains by introducing *Zm*UGD (the best performing UGD) assembled with *kfoA* and *kfoC*, or with animal-sourced genes codon-optimized for *S. cerevisiae* (UAE from *Giardia intestinalis* and CHSY from *Bos taurus*), which led to extracellular chondroitin production up to 125 mg/L. Computational-aided metabolic engineering was key to identify critical bottlenecks and potential genetic targets for improving chondroitin titers. Commonly identified targets included the overexpression of genes responsible for the production of UDP-acetylglucosamine, a precursor for chondroitin. For example, genes *QRI1, GNA1* or *PCM1* (encoding uridine diphosphate-*N*-acetylglucosamine pyrophosphorylase, glucosamine-6-phosphate acetyltransferase and *N*-acetylglucosamine-phosphate mutase, respectively) were commonly identified by different bioinformatic approaches herein used. These genes are potential targets to overexpress to improve chondroitin production in the future.

Despite the conclusions herein obtained comprise an advancement in this research field, there are still many questions that remain unsolved, so some future perspectives can be suggested:

- Genetic and metabolic engineering to enhance precursor availability, enzyme activity, and overall chondroitin production efficiency:
    - Enzyme engineering strategies to improve enzyme solubility and activity include random or directed mutagenesis;

  - Continued efforts in the evaluation and optimization of novel biosynthetic pathways, with alternative genes, for chondroitin production, potentially leading to even higher yields;
  - Implementation of the gene modifications herein predicted for *E. coli* and *S. cerevisiae* that were not yet performed should be considered in future work.
- Enhancement of host strains robustness and consistency:
  - The integration of genes into the yeast genome can make the chondroitin production process more reproductible and efficient;
  - Making host cells more adaptable through adaptive lab evolution is an interesting strategy for further improving chondroitin yields.
- Evaluation of alternative hosts for chondroitin production, considering their specific advantages and capabilities:
  - Particularly eukaryotic hosts, that might offer more soluble and active enzymes as most of the genes required for sulfated GAGs production are from animals.
- Development and optimization of integrated approaches combining microbial production and post-polymerization processes, leading to more efficient and cost-effective production of chondroitin sulfate (CS):
  - Post-polymerization processes, such as different sulfonations performed through the action of sulfotransferases, are required to obtain different chondroitin-derived structures that lead to different biological activities;
  - The expression of sulfotransferases for further sulfation of chondroitin is still on its infancy and many challenges can be expected regarding both expressing them in a soluble, active manner, and providing the required amounts of the sulfotransferase co-factor 3'-phosphoadenylyl-sulfate (PAPS).
- Further utilization of different computational tools to identify targets and optimize metabolic pathways for enhanced chondroitin production:
  - As highlighted in this thesis, leveraging computational approaches for the identification of metabolic engineering targets can provide valuable insights into designing and constructing optimized microbial cell factories;
  - For example, as previously discussed, GECKO method can be applied to genome-scale metabolic models to enhance its predictive capabilities by accounting for enzymes as part of reactions, thereby incorporating enzymatic constraints.

- Improvement of the methods for purification, quantification and characterization, towards more cost-effective processes. Expensive and complex purification and analysis techniques of these compounds are some of the current impairments in the study and use of GAGs produced by microbial cell factories.
- Optimization of the culture conditions can significantly impact the production process, ensuring better yield and quality of chondroitin.
- Industrial-scale implementation and optimization of the developed strategies for chondroitin production:
    - Improving the process to an industrial-scale production usually allows for the optimization of processes, resulting in higher yields of chondroitin, reduced costs per unit, better quality control and standardized production processes;
    - Specifically, for *S. cerevisiae*, its fermentative capabilities are known to be robust and efficient, therefore, the use of large-scale bioreactors can certainly improve chondroitin production for the strains herein constructed.

Despite the advancements on chondroitin biotechnological production, there is still a gap of knowledge on the responsible enzymes and on the actual mechanisms of action for chondroitin polymerization and sulfation especially in the animals, that could give insights into finding new enzymes. Advancements on the enzymes for polymerization, transport, and post-polymerization, coupled with the development of efficient technologies for gene editing, such as clustered regularly interspaced short palindromic repeats-Caspase 9 (CRISPR-Cas9), to efficiently engineer strains, are expected to pave the way for the establishment of efficient microbial cell factories for CS production. The potential for the large-scale industrial production of CS through heterologous microbes is becoming closer, offering a safer and more sustainable alternative to conventional production methods. Ultimately, this could contribute to wider availability of CS and improve its applications in various fields.

APPENDIX

FIGURES

CHAPTER 3

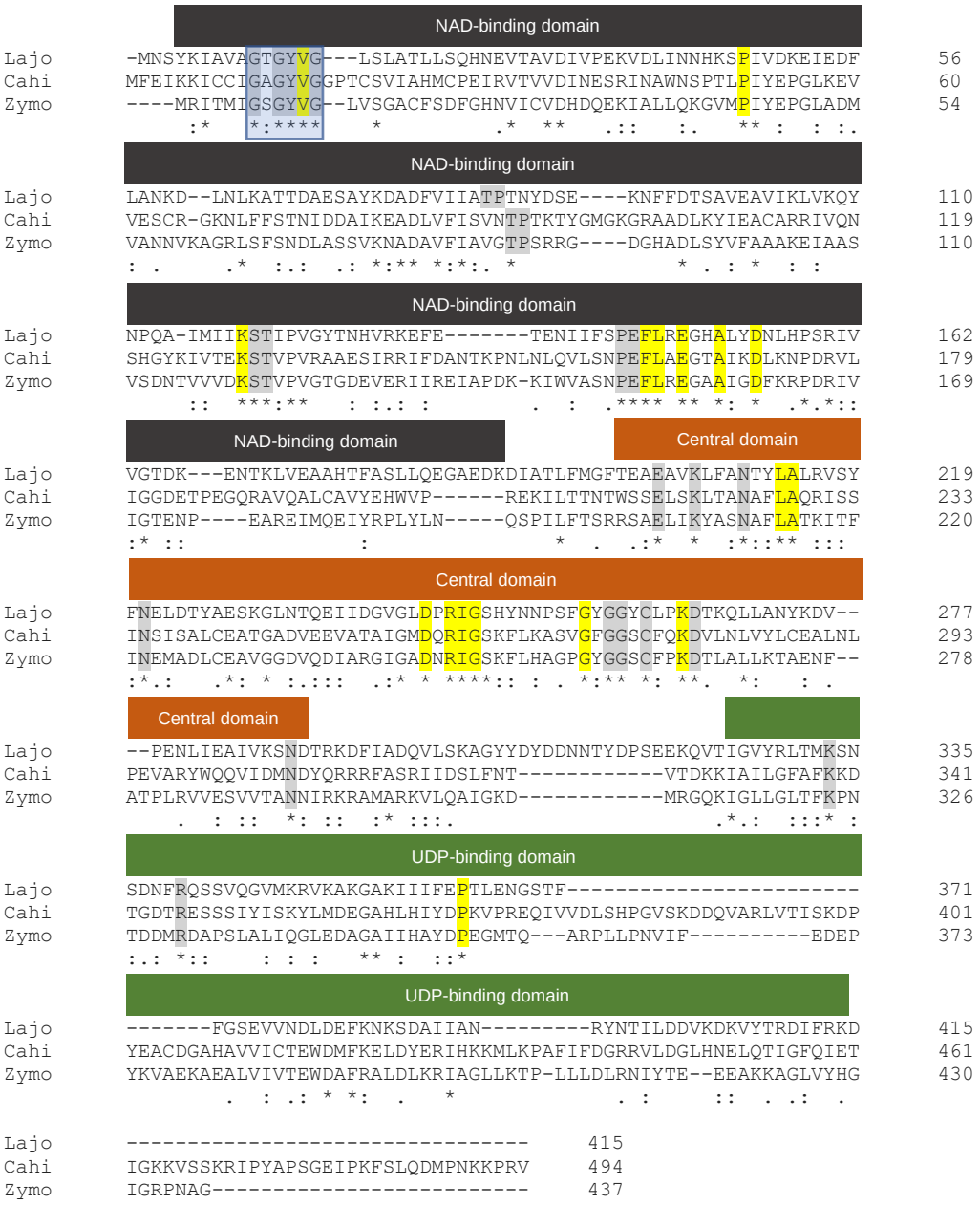


Figure A 3.1. Multiple amino acid sequence alignment of the three uridine diphosphate (UDP)-glucose dehydrogenases used in this study generated by Clustal Omega (1.2.4) (EMBL-EBI). Species are abbreviated as the first two letters of the genus followed by the two first letters of the species: Lajo, *Lactobacillus johnsonii*; Cahi, *Capra hircus*; and Zymo, *Zymomonas mobilis*. An asterisk (*) marks the residues that are conserved in the three sequences; a colon (:) indicates conservation between groups of strongly similar properties (scoring > 0.5 in the Gonnet Point Accepted Mutation (PAM) 250 matrix); and a period (.) indicates conservation between groups of weakly similar properties (scoring =< 0.5 in the Gonnet PAM 250 matrix). Conserved amino acids within 63 members of UDP-glucose 6-dehydrogenase/ guanosine diphosphate (GDP)-mannose 6-dehydrogenase

(UGD/GMD) superfamily are highlighted in grey while conserved sites within only UGD enzymes are highlighted in yellow. The nicotinamide adenine dinucleotide (NAD)-binding domain contains a characteristic dinucleotide binding Rossman fold, indicated in a blue box, composed by a conserved glycine-rich phosphate-binding loop with the pattern GXGXXG linking the first β-sheet to the following α-helix that, in the case of UGD/GMD dehydrogenase enzymatic family, has a strictly conserved tyrosine so that the region becomes GXGYXG. The central domain is a long α-helical region that serves as the dimerization interface for these enzymes. The C-terminal substrate-binding domain forms an incomplete dinucleotide binding fold, primarily responsible for binding the UDP moiety of the UDP-sugar in a deep pocket. The described active site of UGD contains residues from the three domains that are highly conserved within the superfamily and include: N-terminal threonine (T_{122}), central α-helix lysine, (K_{207}), and asparagine, (N_{211}) and C-terminal contributes with cysteine (C_{263}) and aspartic acid (D_{267}) (reference sites from *Z. mobilis* sequence).

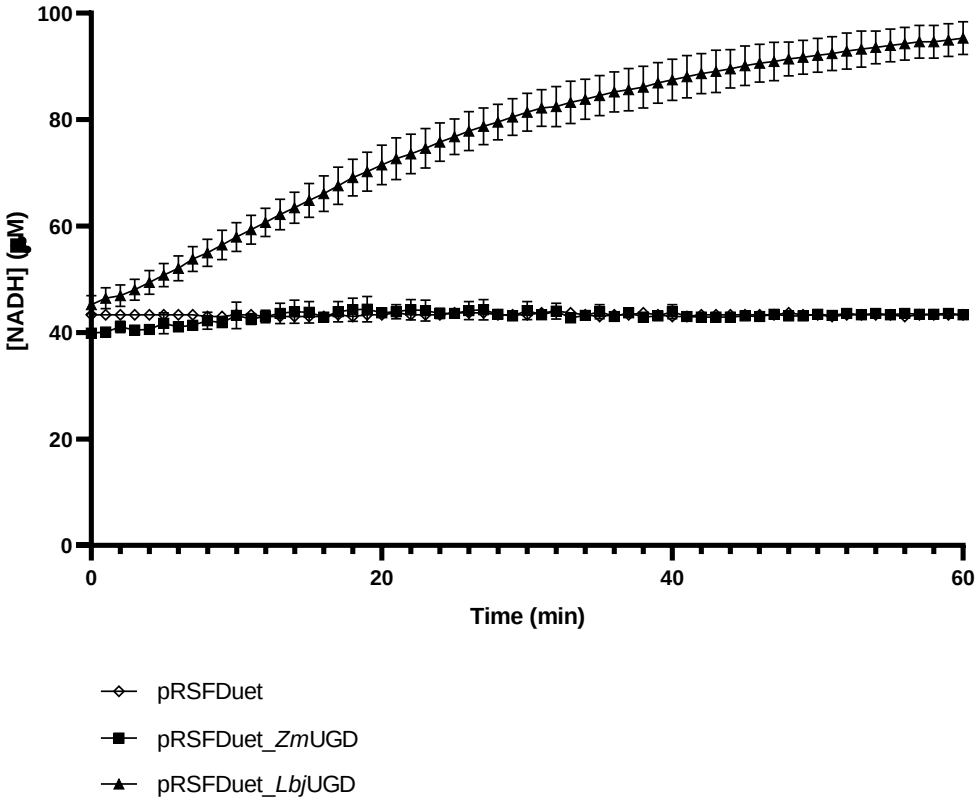


Figure A 3.2. *In vitro* reduced nicotinamide adenine dinucleotide (NADH) production over time in engineered *Escherichia coli* at pH 7.2 as obtained by monitoring absorbance at 340 nm.

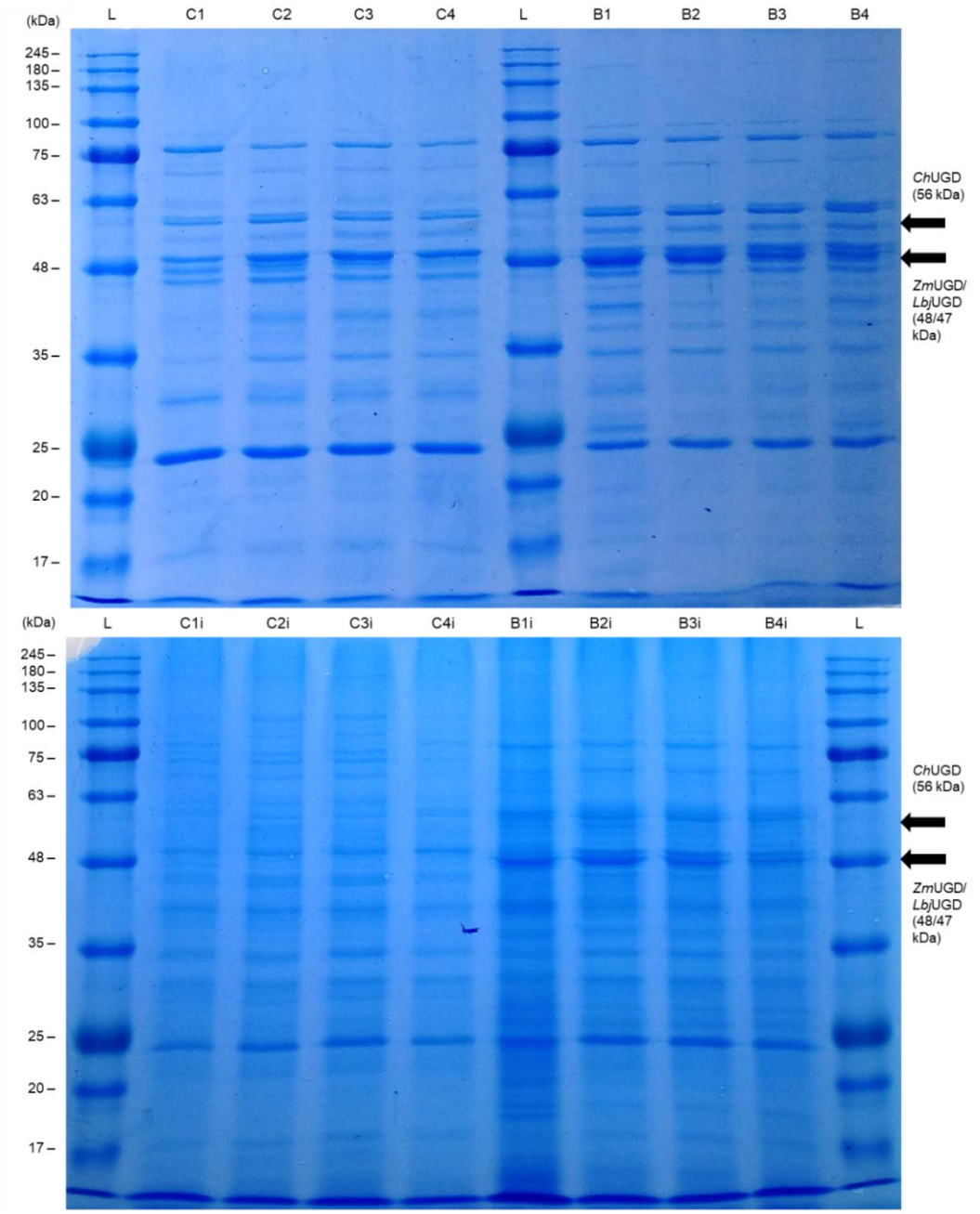


Figure A 3.3. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels from *Saccharomyces cerevisiae* cell extracts (15 µg). In the upper gel, the samples loaded were: cell extract from *S. cerevisiae* CEN.PK2-1C wild-type (C1), cell extract from *S. cerevisiae* CEN.PK2-1C expressing *Ch*UGD (C2) cell extract from *S. cerevisiae* CEN.PK2-1C expressing *Zm*UGD (C3), cell extract from *S. cerevisiae* CEN.PK2-1C expressing *Lbj*UGD (C4), *S. cerevisiae* BY4741 wild-type (B1), *S. cerevisiae* BY4741 expressing *Ch*UGD (B2), *S. cerevisiae* BY4741 expressing *Zm*UGD (B3) and *S. cerevisiae* BY4741 expressing *Lbj*UGD (B4). The lower gel shows the proteins in the corresponding insoluble phases. L: NZYColour Protein Marker II ladder.

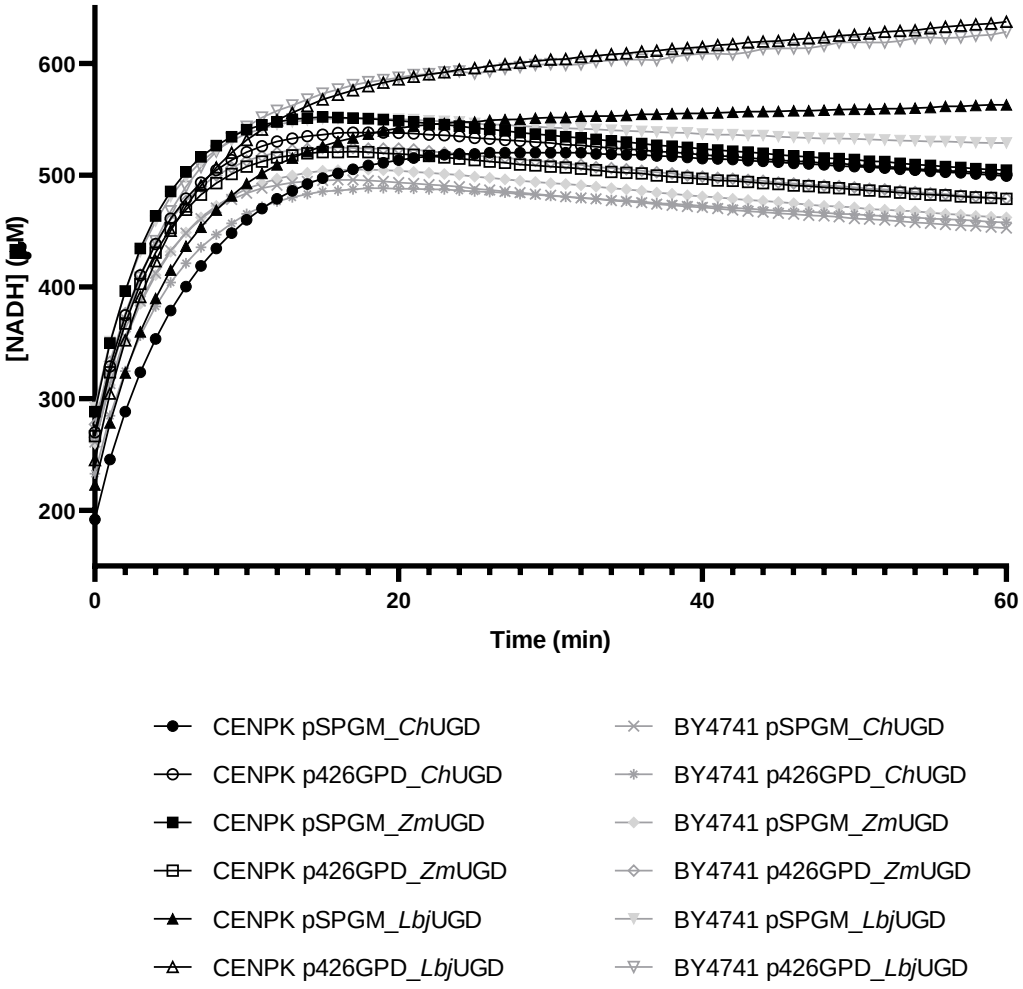


Figure A 3.4. *In vitro* reduced nicotinamide adenine dinucleotide (NADH) production over time in engineered *S. cerevisiae* yeasts at pH 9.2 obtained by monitoring absorbance at 340 nm.

CHAPTER 4

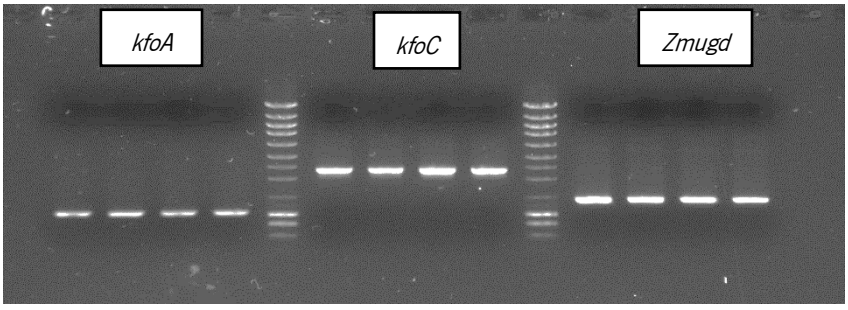


Figure A 4.1. Agarose gel 0.7% showing polymerase chain reaction (PCR) amplification results of genes for chondroitin biosynthetic pathway construction. Ladder: NZYDNA Ladder III, NZYTech.

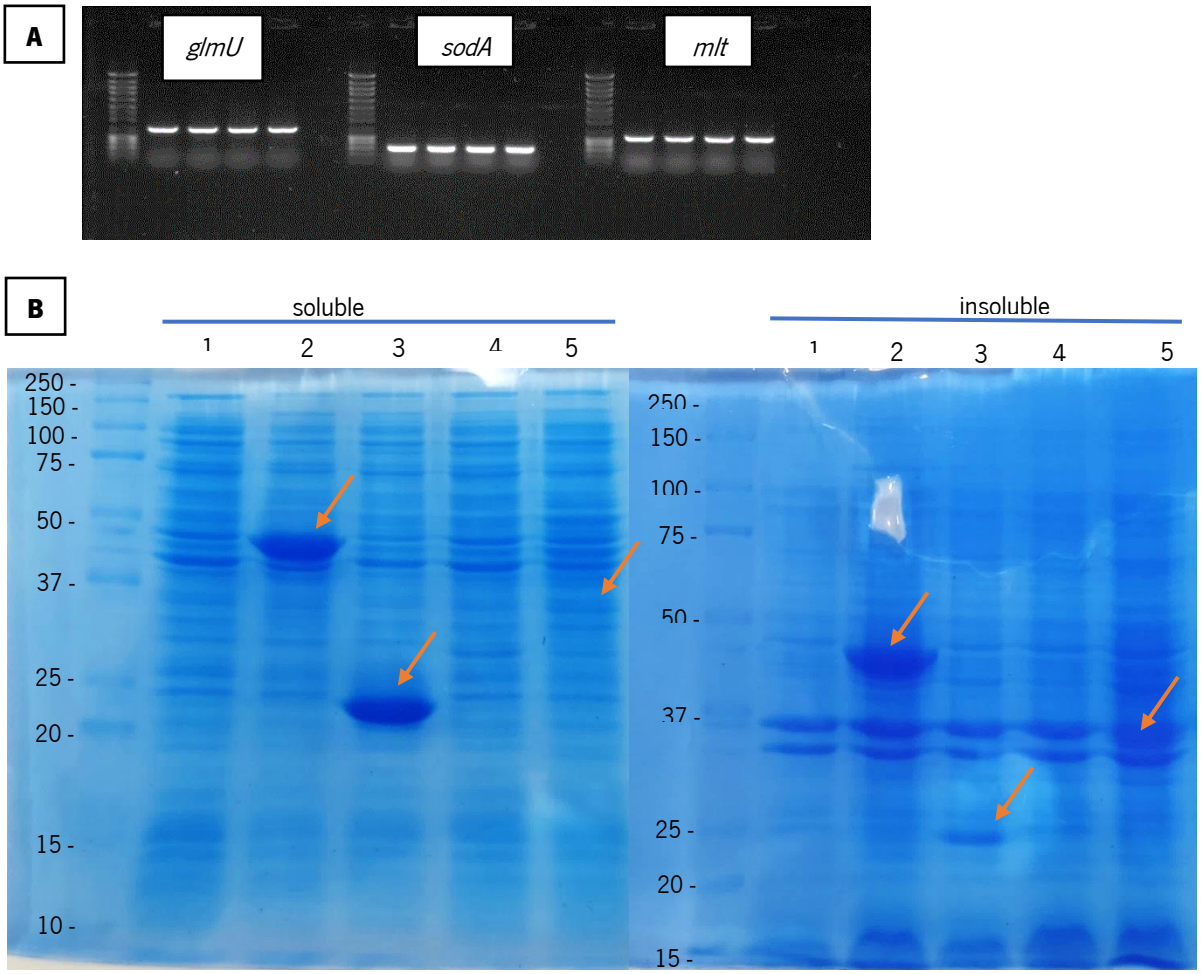


Figure A 4.2. Cloning and expression of identified targets for chondroitin production enhancement: *glmU*, glucosamine-1-phosphate acetyltransferase/ *N*-acetylglucosamine-1-phosphate uridyltransferase; *mltB*, membrane-bound lytic murein transglycosylase; *sodA*, superoxide dismutase. **A.** Agarose gel 0.7% showing PCR result of amplification of genes *glmU*, *sodA* and *mltB* from *E. coli* K-12 MG1655 (DE3) genome. **B.** SDS-PAGE gel showing overexpression of genes *glmU*, *sodA* and *mltB* in *E. coli* K-12 MG1655 (DE3). 1 – pETDuet-1, 2 - pETDuet_glmU, 3 – pETDuet_sodA, 4 – pCDFDuet-1, 5 – pCDFDuet_mltB. The predicted sizes were: GlmU 49.2 kDa; SodA 24.91 kDa; MltB 41.2 kDa (Slt35 36 kDa).

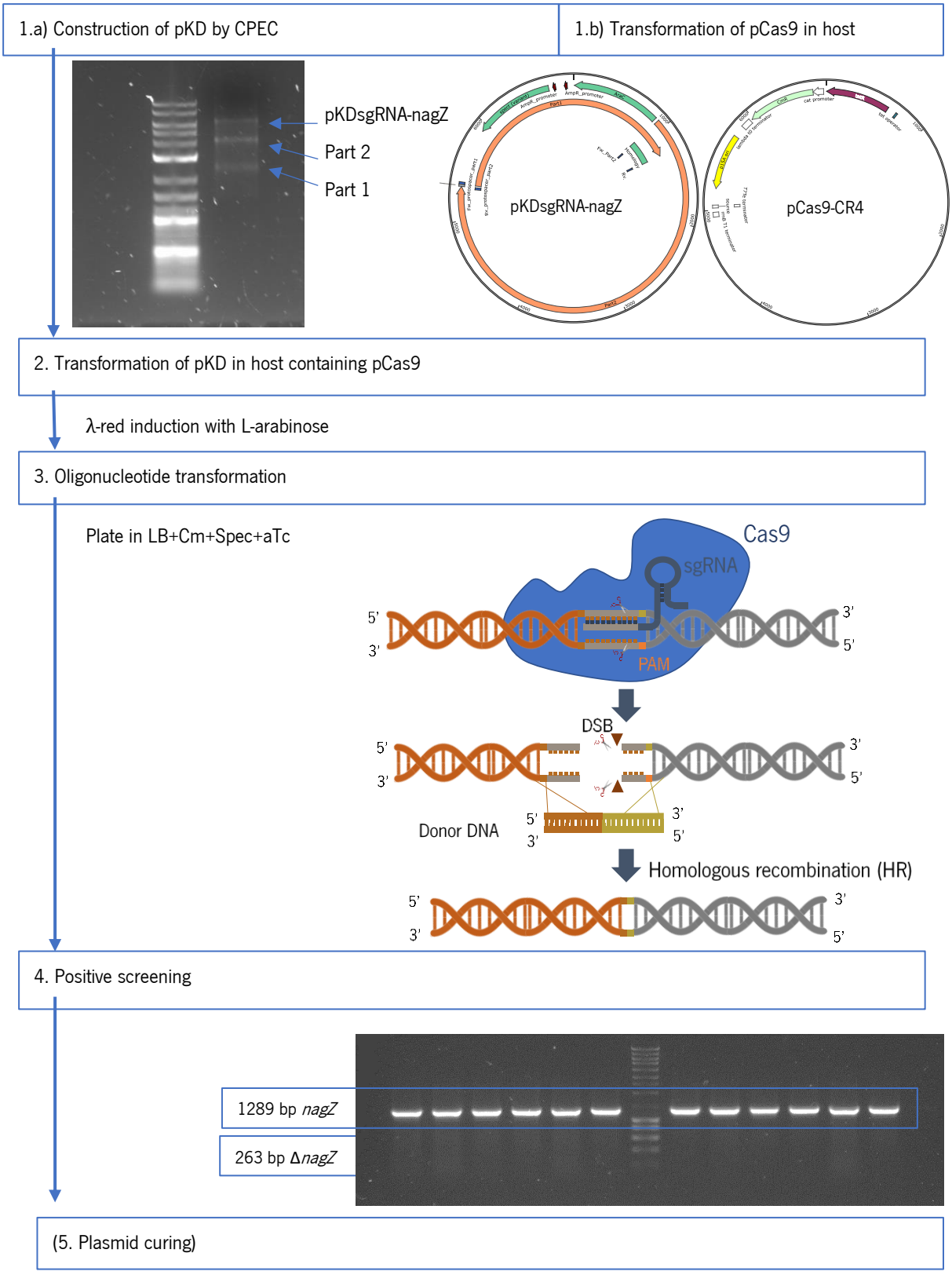


Figure A 4.3. Attempted CRISPR-Cas9 strategy for *nagZ* deletion. This methodology consists in the cleavage of double stranded DNA by Cas9 and λ-Red recombinase system-facilitated genomic integration of donor DNA. The single-guide RNA (sgRNA) encoding plasmid pKDsgRNA-nagZ to target the *nagZ* has been constructed. The plasmids for the expression of Cas9 and of the sgRNA were transformed. Then, the transformation of the

oligonucleotide (donor DNA) that should induce the gene deletion was performed. During the subsequent screenings, no positive colonies were found in the several attempts of this procedure.

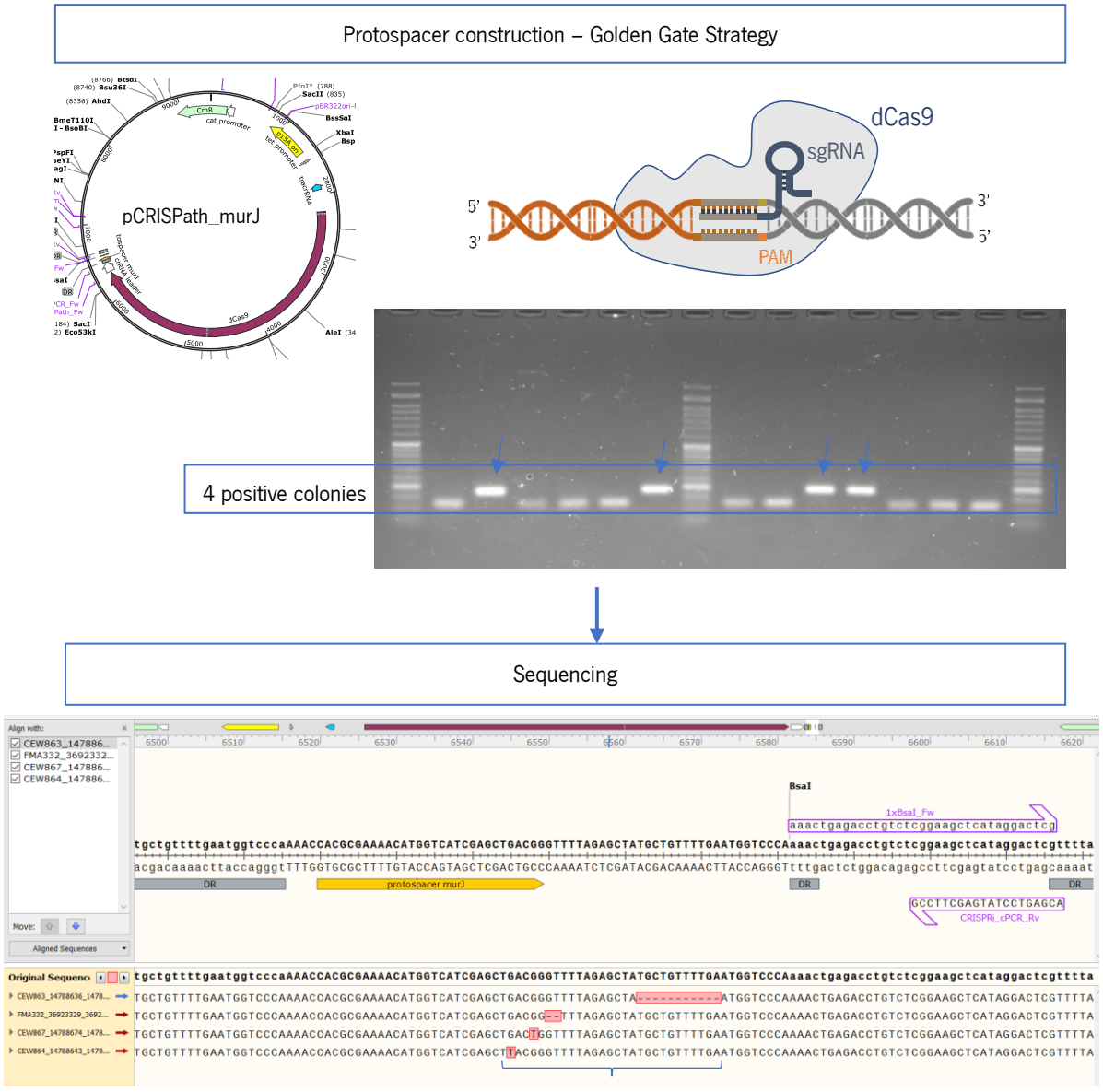


Figure A 4.4. Attempted CRISPR interference (CRISPRi) strategy to underexpress *murJ*. In this strategy, a modified version of the caspase 9 protein commonly called as dead Cas9 (dCas9), which does not have the nuclease activity but maintains sequence-specific double stranded DNA-binding capability, is expressed to target the *murJ* gene, ultimately repressing its expression. The cloning of the protospacer in the pCRISPathBrick was performed by Golden Gate Strategy. Sequencing of resulting pCRISPath_murJ showed that constructed plasmids constantly had errors in or next to the protospacer.

APPENDIX

TABLES

CHAPTER 2

Table A 2.1. Reported applications of genome-scale metabolic modeling in identification of targets for metabolic engineering.

Microorganism - model	Compound	Computational procedure	Results	Reference
Escherichia coli - EcoMBEL979	Lycopene	FSEOF and MOMA	Overexpression of *dxs*, *idi*, and *mdh* led to 2.7- and 3.2-fold increase in lycopene concentration and yield on biomass, respectively.	[1]
S. cerevisiae - iFF708	Vanillin	OptGene with MoMA	Knockouts of *PDC1*, *GDH1*, and *GDH2+* resulted in 1.5-fold increase in vanillin β-D-glucoside yield in batch mode. Single knockout of *PDC1* showed 2-fold productivity improvement in continuous culture.	[2]
E. coli - EcoM-BEL97	Polylactic acid and its copolymers	MOMA and flux response analysis	Triple knockout of the *ackA*, *ppc* and *adhE*, with *ldhA* and *acs* up-regulations, increased polymer content and lactate fraction up to 3.7-fold.	[3]
E. coli - EciJR904	NADPH, leucocyanidin and (+)-catechin	CiED	*pgi*, *ppc* and *pldA* deletions resulted in 4- and 2-fold increase in leucocyanidin and (+)-catechin, respectively.	[4]
E. coli - iJR904	1,4-Butanediol	OptKnock	Knockouts of *adhE*, *pflB*, *ldhA*, and *mdh* genes were identified, and with additional rational modifications, the engineered strain resulted in 3-fold increase on 1,4-butanediol titer.	[5]
E. coli - iAF1260	Malonyl-CoA and Naringenin	Optforce	Knockout of *fumC* and *sucC*, and overexpression of *accA*, *pgk*, *gapA* and PDH led to 5.6-fold increase on naringenin concentration.	[6]
E. coli - iAF1260	6-Deoxyery-thronolide B	FDCA and LMOMA	Several targets identified, for example *ackA* or *sucCD* knockouts, *sdhABCD* or *pgi* underexpressions, and *fbaAB* up-regulations. No *in vivo* data.	[7]
E. coli - Ec_iAF1260	Taxadiene	FBA	Overexpression of *dxs*, *idi, ispB, ispDF, crtE* and *txs* knockout led to taxadiene yields of 876 mg L^{-1} and 8.9 mg L^{-1} h^{-1}.	[8]
Acinetobacter baylyi – iAbaylyi[v4]	Triacylglycerol	FBA	Strain with deletions of *acr1*, *poxB*, *dgkA*, and a triacylglycerol lipase produced 5.6-fold more triacylglycerol.	[9]

Table A 2.1. Reported applications of genome-scale metabolic modeling in identification of targets for metabolic engineering (*continuation*).

Microorganism - model	Compound	Computational procedure	Results	Reference
Corynebacterium glutamicum -N/A	L-Lysine	Constraint-based MFA	Citrate synthase flux was diminished with different promoter strengths to enhance L-lysine titer up to 13.19 g L^{-1}., 1.8-fold higher than the strain having citrate synthase at 100% activity.	[10]
Streptomyces roseosporus - N/A	Daptomycin	MFA	Overexpression of *zwf2*, *dptI* and *dptJ* resulted in 34.4% higher daptomycin concentration.	[11]
E. coli - iAF1260	Taxadiene	FBA and MoMA	Overexpression of *idi* and *ppk*, *sthA*, and *purN* resulted in over 12-fold improvement.	[12]
E. coli - iAF1260	Fatty acids	OptForce	*fabZ* and acyl-ACP thioesterase were upregulated and *fadD* was deleted resulting in a strain able to produce 1.70 g L^{-1} and 0.14 g g^{-1} of C_{14-16} fatty acids.	[13]
S. cerevisiae - iND750	Fumaric acid	Literature mining + FBA	*FUM1* was deleted and *RoPYC* and *SFC1* genes were introduced which resulted in a final engineered strain able to produce 1675 mg L^{-1} fumaric acid in batch culture.	[14]
S. cerevisiae - iMM904	2,3-Butanediol	Optknock	Triple knockout of *ADH1, ADH3* and *ADH5* strain produced highest 2,3-butanediol titer (2.29 g L^{-1}) and yield (0.113 g g^{-1}) under anaerobic condition.	[15]
E. coli - EcoMBEL979	Shikimic acid and putrescine	FVSEOF	Prediction of *glk*, *acnA*, *acnB*, *ackA* and *ppc* genes as overexpression targets led to 20.5% increase in production of putrescine.	[16]
Candida glabrata - iNX804	Malate	FBA	Overexpression of pyruvate carboxylase, malate dehydrogenase, *RoPYC*, *RoMDH* and *SpMAE1* led to 10-fold malate accumulation.	[17]
E. coli iAF1260	Resveratrol	OptForce	Overexpression of PDH, *pgk*, *gapA* and deletion of *fumC* led to 60% increase in the yield.	[18]
S. cerevisiae - iFF708	Succinate	OptGene	Deletion of *SDH3p*, *SER3p* and *SER33p* resulted in 13- and 14-fold improvement in titer and biomass yield, respectively. Further strain improvement by directed evolutions led to final 30- and 43-fold improvements in titer and yield on biomass, respectively.	[19]
Pseudomonas putida – N/A	Poly-hydroxyalkanoates	Elementary flux mode analysis	*gcd* deletion strain showed an increase of the PHA yield (80%), the PHA titer (100%) and cellular PHA content (50%).	[20]
E. coli - EcoMBEL979	Fumaric acid	Flux response analysis	Overexpression of the native *ppc* gene resulted in the increased 2.8-fold titer of fumaric acid.	[21]

Table A 2.1. Reported applications of genome-scale metabolic modeling in identification of targets for metabolic engineering (*continuation*).

Microorganism - model	Compound	Computational procedure	Results	Reference
E. coli - *N/A*	Fatty acids	Bi-level optimization framework	Deletion of *cyoA, adhE, nuoA, ndh, pta* and *dld*, with overexpression of *accBCDA* and thioesterase increased to 202 mg/g dry cell weight, ~250% of that in the wild-type strain.	[22]
E. coli - iAF1260	Lycopene	FDCA	Knockout targets (*deoB*, *yhfw*, *yahI*, *pta* and *eutD*) and up-regulated targets (*ompN*, *ompE*, *ndk* and *cmk*) led to 10–45 % increases of lycopene yield.	[23]
Streptomyces coelicolor - iMK1208	Antibiotic actinorhodin	FSEOF	BCDH overexpression mutant showed a 52-fold increase in actinorhodin production.	[24]
Pichia pastoris - PpaMBEL1254	Recombinant protein hSOD	MOMA and FSEOF	Five out of 9 tested targets led to enhanced production of hSOD yield up to 40%, namely: overexpressions: *SOL3*, *GDH3*, *MDH1*; knock-outs: *ADH2*, *GUT2*.	[25]
S. cerevisiae - iMM904	Amorphadiene	FBA, MOMA and FDCA	Single knockouts resulted in greater amorphadiene yield, being the highest improvement of 12-fold with *HXK2* knockout strain.	[26]
C. glabrata - iNX804	Acetoin	FBA	Inactivation of *ADH* and expression of heterologous pathway (*BsALS* and *BaALDC*) led to a final acetoin titer of 3.67 g L^{-1}.	[27]
C. glabrata - iNX804	Fumaric acid	FBA and FSEOF	Overexpressing *ASL*, *ADSL*, and *SpMAE1* resulted in a final fumarate titer of 8.83 g L^{-1}.	[28]
S. cerevisiae - iTO977	3-Hydroxypropionic acid	pFBA	Multiple copy integration of *MCR* and *ACC1*, overexpression of *PDC1*, *ALD6* and acetyl-CoA synthase and engineering of cofactor specificity of the glyceraldehyde-3-phosphate dehydrogenase led to 3HP titer of 9.8 g $L^{-1}$and yield of 13 % C-mol C-mol^{-1} glucose.	[29]
S. cerevisiae - iLL672 (extended version)	β-Farnesene	pFBA	Engineered strain with overexpressed *eutE*, *PK*, PTA, and NADH-HMGr, and deleted *ALD4*, *ALD6* and *RHR2*, produced 25% more farnese, requiring 75% less oxygen.	[30]
E. coli - *N/A*	Itaconic acid	MCSs	Overexpression of *cadA* and *gltA*, deletion of *aceA*, *sucCD*, *pykA*, *pykF* and *pta* and downregulation of *icd* led to 32 g L^{-1} itaconic acid with an overall yield of 0.68 mol mol^{-1} and a peak productivity of 0.45 g L^{-1} h^{-1}.	[31]

Table A 2.1. Reported applications of genome-scale metabolic modeling in identification of targets for metabolic engineering (*continuation*).

Microorganism - model	Compound	Computational procedure	Results	Reference
Y. lipolytica - iYLI647	Dicarboxylic acids	GDLS, Flux activity analysis, tSOT and CMA	Overexpression of malate dehydrogenase and malic enzyme genes and overexpression of glutamate dehydrogenase gene generated additional NADPH required for fatty acid synthesis, leading to the increased dodecanedioic acid *in silico* fluxes by 48% and 22% higher, respectively. No *in vivo* data.	[32]
E. coli - iJO1366	Aromatic polyesters	Flux response analysis	The *tyrB* and *aspC* genes were deleted from an engineered strain resulting in D-phenyllactic acid titer increase by 3-fold.	[33]
P. putida – iJN1463 and iMATD1b539	Indigoidine	cMCS, EMA and FBA	Six predicted deletions plus three targets from rational strain design, and adaptative lab evolution resulted in a final strain producing 7.3 g/L indigoidine.	[34]

Abbreviations: 3HP, 3-hydroxypropionic acid; hSOD, human superoxide dismutase; NADPH, nicotinamide adenine dinucleotide phosphate; PHA, polyhydroxyalkanoates; **Computational methods:** CMA, cofactor modification analysis; cMCS, constrained minimal cut set; CiED, cipher for evolutionary design; EMA, elementary mode analysis; FBA, flux balance analysis; FDCA, flux distribution comparison analysis; FSEOF, flux scanning based on enforced objective flux; FVSEOF, flux variability scanning based on enforced objective flux; GDLS, genetic design by local search, LMOMA, linear minimization of metabolic adjustment; MFA, metabolic flux analysis; pFBA, parsimonious enzyme usage FBA; tSOT, transcriptomics-based strain optimization tool. **Enzymes and genes:** *ACC1*, *accBCDA*, acetyl-CoA carboxylase; *aceA*, isocitrate lyase; *ackA*, acetate kinase; *acna*, aconitate hydratase A; *acnB*, aconitate hydratase B; *acr1*, fatty acyl-CoA reductase; *acs*, acetyl-CoA synthetase; *ADH, ADH1, ADH2, ADH3, ADH5,* alcohol dehydrogenase*; adhE*, acetaldehyde/alcohol dehydrogenase; *ALD4*, *ALD6*, aldehyde dehydrogenase; *ALS*, acetolactate synthase; *ALDC*, acetolactate decarboxylase; *ASL*, argininosuccinate lyase, *ADSL*, adenylosuccinate lyase; *aspC*, aspartate aminotransferase; BCDH, branched-chain α-keto acid dehydrogenase; *cadA*, lysine decarboxylase; *cmk*, cytidylate kinase; *crtE*, geranylgeranyl pyrophosphate synthase; *cyoA*, cytochrome bo3 subunit 2; *deoB*, phosphopentomutase; *dgkA*, diacylglycerol kinase; *dld*, quinone-dependent D-lactate dehydrogenase; *dptI*,α-ketoglutarate methyltransferase; *dptJ*, tryptophan oxygenase; *dxs,* 1-deoxy-D-xylulose-5-phosphate synthase *; eutD*, phosphate acetyltransferase; *eutE,* acetaldehyde dehydrogenase (acetylating);*fabZ*, 3-hydroxy-acyl-[acyl-carrier-protein] dehydratase; *fadD*, long-chain-fatty-acid—CoA ligase; *fbaAB*, fructose-bisphosphate aldolase; *FUM1, fumC*, fumarase; *gapA*, glyceraldehyde-3-phosphate dehydrogenase; *gcd*; quinoprotein glucose dehydrogenase ; *GDH1, GDH2*, *GDH3* glutamate dehydrogenase; *glk*; glucokinase; *gltA*, citrate synthase; *GUT2,* glycerol-3-phosphate dehydrogenase; *HXK2,* hexokinase isoenzyme 2; *icd*, isocitrate dehydrogenase; *idi*, isopentenyl-diphosphate isomerase; *ispB*, octaprenyl-diphosphate synthase*; ispDF*, 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase and 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; *ldhA*, lactate dehydrogenase; *MAE1*, malic enzyme; MCR, malonyl-CoA reductase; *mdh*, *MDH1,* malate dehydrogenase; NADH-HMGr, NADH-consuming hydroxymethylglutaryl(HMG)-CoA reductase; *ndh*, NADH:quinone oxidoreductase II; *ndk*, nucleoside diphosphate kinase; *nuoA*, NADH:quinone oxidoreductase subunit A; *ompE*, *ompN*, outer membrane porin; *PDC1*, pyruvate decarboxylase; PDH, pyruvate dehydrogenase multi-enzyme complex; *pflB*, pyruvate formate lyase; *pgi*, glucose-6-phosphate isomerase; PGK, phosphoglycerate kinase; *PK*, phosphoketolase; *pldA*, outer membrane phospholipase A; *poxB*, pyruvate dehydrogenase (cytochrome); *ppc*, phosphoenolpyruvate carboxylase; *ppk*, polyphosphate kinase; *pta*, phosphate acetyltransferase; *purN*, phosphoribosylglycinamide formyltransferase; *PYC,* pyruvate carboxylase; *pykA*, *pykF*, pyruvate kinase; *RHR2*, glycerol 3-phosphate phosphatase; *SDH3*, cytochrome b subunit of the succinate dehydrogenase complex; *sdhABCD*, succinate:quinone oxidoreductase; *SER3*p/ *SER33*p (isoenzymes), 3-phosphoglycerate dehydrogenase; *SFC1*, succinate–fumarate transporter; *SOL3*, 6-phosphogluconolactonase; *sthA*, soluble pyridine nucleotide transhydrogenase; *sucCD*, succinyl-CoA synthetase; *txs*, taxadiene synthase; *tyrB*, tyrosine aminotransferase; *yahI,* carbamate kinase; *yhfw*, phospho- pentomutase; *zwf2*, glucose-6-phosphate hydrogenase.

References

[1] H.S. Choi, S.Y. Lee, T.Y. Kim, H.M. Woo, *In silico* identification of gene amplification targets for improvement of lycopene production, Appl. Environ. Microbiol. 76 (2010) 3097–3105. https://doi.org/10.1128/AEM.00115-10.

[2] A.R. Brochado, C. Matos, B.L. Møller, J. Hansen, U.H. Mortensen, K.R. Patil, Improved vanillin production in baker's yeast through *in silico* design, Microb. Cell Fact. 9 (2010) 1–15. https://doi.org/10.1186/1475-2859-9-84.

[3] Y.K. Jung, T.Y. Kim, S.J. Park, S.Y. Lee, Metabolic engineering of *Escherichia coli* for the production of polylactic acid and its copolymers, Biotechnol. Bioeng. 105 (2010) 161–171. https://doi.org/10.1002/bit.22548.

[4] J.A. Chemler, Z.L. Fowler, K.P. McHugh, M.A.G. Koffas, Improving NADPH availability for natural product biosynthesis in *Escherichia coli* by metabolic engineering, Metab. Eng. 12 (2010) 96–104. https://doi.org/10.1016/j.ymben.2009.07.003.

[5] H. Yim, R. Haselbeck, W. Niu, C. Pujol-Baxley, A. Burgard, J. Boldt, J. Khandurina, J.D. Trawick, R.E. Osterhout, R. Stephen, J. Estadilla, S. Teisan, H.B. Schreyer, S. Andrae, T.H. Yang, S.Y. Lee, M.J. Burk, S. Van Dien, Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol, Nat. Chem. Biol. 7 (2011) 445–452. https://doi.org/10.1038/nchembio.580.

[6] P. Xu, S. Ranganathan, Z.L. Fowler, C.D. Maranas, M.A.G.G. Koffas, Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA, Metab. Eng. 13 (2011) 578–587. https://doi.org/10.1016/j.ymben.2011.06.008.

[7] H. Meng, Z. Lu, Y. Wang, X. Wang, S. Zhang, *In silico* improvement of heterologous biosynthesis of erythromycin precursor 6-deoxyerythronolide b in *Escherichia coli*, Biotechnol. Bioprocess Eng. 16 (2011) 445–456. https://doi.org/10.1007/s12257-010-0321-7.

[8] H. Meng, Y. Wang, Q. Hua, S. Zhang, X. Wang, *In silico* analysis and experimental improvement of taxadiene heterologous biosynthesis in *Escherichia coli*, Biotechnol. Bioprocess Eng. 16 (2011) 205–215. https://doi.org/10.1007/s12257-010-0329-z.

[9] S. Santala, E. Efimova, V. Kivinen, A. Larjo, T. Aho, M. Karp, V. Santala, Improved triacylglycerol production in *Acinetobacter baylyi* ADP1 by metabolic engineering, Microb. Cell Fact. 10 (2011) 1–10. https://doi.org/10.1186/1475-2859-10-36.

[10] J. van Ooyen, S. Noack, M. Bott, A. Reth, L. Eggeling, Improved L-lysine production with *Corynebacterium glutamicum* and systemic insight into citrate synthase flux and activity, Biotechnol. Bioeng. 109 (2012) 2070–2081. https://doi.org/10.1002/bit.24486.

[11] D. Huang, J. Wen, G. Wang, G. Yu, X. Jia, Y. Chen, *In silico* aided metabolic engineering of *Streptomyces roseosporus* for daptomycin yield improvement, Appl. Microbiol. Biotechnol. 94 (2012) 637–649. https://doi.org/10.1007/s00253-011-3773-6.

[12] B.A. Boghigian, J. Armando, D. Salas, B.A. Pfeifer, Computational identification of gene over-expression targets for metabolic engineering of taxadiene production, Appl. Microbiol. Biotechnol. 93 (2012) 2063–2073. https://doi.org/10.1007/s00253-011-3725-1.

[13] S. Ranganathan, T.W. Tee, A. Chowdhury, A.R. Zomorrodi, J.M. Yoon, Y. Fu, J. V. Shanks, C.D. Maranas, An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*, Metab. Eng. 14 (2012) 687–704. https://doi.org/10.1016/j.ymben.2012.08.008.

[14] G. Xu, W. Zou, X. Chen, N. Xu, L. Liu, J. Chen, Fumaric acid production in *Saccharomyces cerevisiae* by *in silico* aided metabolic engineering, PLoS One. 7 (2012) 1–10. https://doi.org/10.1371/journal.pone.0052086.

[15] C. Ng, M. Jung, J. Lee, M.-K. Oh, Production of 2,3-butanediol in *Saccharomyces cerevisiae* by *in silico* aided metabolic engineering, Microb. Cell Fact. 11 (2012) 68. https://doi.org/10.1186/1475-2859-11-68.

[16] J.M. Park, H.M. Park, W.J. Kim, H.U. Kim, T.Y. Kim, S.Y. Lee, Flux variability scanning based on enforced objective flux for identifying gene amplification targets, BMC Syst. Biol. 6 (2012). https://doi.org/10.1186/1752-0509-6-106.

[17] X. Chen, G. Xu, N. Xu, W. Zou, P. Zhu, L. Liu, J. Chen, Metabolic engineering of *Torulopsis glabrata* for malate production, Metab. Eng. 19 (2013) 10–16. https://doi.org/10.1016/j.ymben.2013.05.002.

[18] N. Bhan, P. Xu, O. Khalidi, M.A.G. Koffas, Redirecting carbon flux into malonyl-CoA to improve resveratrol titers: Proof of concept for genetic interventions predicted by OptForce computational framework, Chem. Eng. Sci. 103 (2013) 109–114. https://doi.org/10.1016/j.ces.2012.10.009.

[19] J.M. Otero, D. Cimini, K.R. Patil, S.G. Poulsen, L. Olsson, J. Nielsen, Industrial systems biology of *Saccharomyces cerevisiae* enables novel succinic acid Cell Factory, PLoS One. 8 (2013) 1–10. https://doi.org/10.1371/journal.pone.0054144.

[20] I. Poblete-Castro, D. Binger, A. Rodrigues, J. Becker, V.A.P. Martins Dos Santos, C. Wittmann, *In-silico*-driven metabolic engineering of *Pseudomonas putida* for enhanced production of poly-hydroxyalkanoates, Metab. Eng. 15 (2013) 113–123. https://doi.org/10.1016/j.ymben.2012.10.004.

[21] C.W. Song, D.I. Kim, S. Choi, J.W. Jang, S.Y. Lee, Metabolic engineering of *Escherichia coli* for the production of fumaric acid, Biotechnol. Bioeng. 110 (2013) 2025–2034. https://doi.org/10.1002/bit.24868.

[22] F. Lin, Y. Chen, R. Levine, K. Lee, Y. Yuan, X.N. Lin, Improving fatty acid availability for bio-hydrocarbon production in *Escherichia coli* by metabolic Engineering, PLoS One. 8 (2013) 1–14. https://doi.org/10.1371/journal.pone.0078595.

[23] J. feng Wang, H. lin Meng, Z. qiang Xiong, S. liang Zhang, Y. Wang, Identification of novel knockout and up-regulated targets for improving isoprenoid production in *E. coli*, Biotechnol. Lett. 36 (2014) 1021–1027. https://doi.org/10.1007/s10529-014-1460-2.

[24] M. Kim, J. Sang Yi, J. Kim, J.N. Kim, M.W. Kim, B.G. Kim, Reconstruction of a high-quality metabolic model enables the identification of gene overexpression targets for enhanced antibiotic production in *Streptomyces coelicolor* A3(2), Biotechnol. J. 9 (2014) 1185–1194. https://doi.org/10.1002/biot.201300539.

[25] J. Nocon, M.G. Steiger, M. Pfeffer, S.B. Sohn, T.Y. Kim, M. Maurer, H. Rußmayer, S. Pflügl, M. Ask, C. Haberhauer-Troyer, K. Ortmayr, S. Hann, G.

Koellensperger, B. Gasser, S.Y. Lee, D. Mattanovich, Model based engineering of *Pichia pastoris* central metabolism enhances recombinant protein production, Metab. Eng. 24 (2014) 129–138. https://doi.org/10.1016/j.ymben.2014.05.011.

[26] Z. Sun, H. Meng, J. Li, J. Wang, Q. Li, Y. Wang, Y. Zhang, Identification of novel knockout targets for improving terpenoids biosynthesis in *Saccharomyces cerevisiae*, PLoS One. 9 (2014) 9–14. https://doi.org/10.1371/journal.pone.0112615.

[27] S. Li, X. Gao, N. Xu, L. Liu, J. Chen, Enhancement of acetoin production in *Candida glabrata* by *in silico*-aided metabolic engineering, Microb. Cell Fact. 13 (2014) 1–11. https://doi.org/10.1186/1475-2859-13-55.

[28] X. Chen, J. Wu, W. Song, L. Zhang, H. Wang, L. Liu, Fumaric acid production by *Torulopsis glabrata*: Engineering the urea cycle and the purine nucleotide cycle, Biotechnol. Bioeng. 112 (2015) 156–167. https://doi.org/10.1002/bit.25334.

[29] K.R. Kildegaard, N.B. Jensen, K. Schneider, E. Czarnotta, E. Özdemir, T. Klein, J. Maury, B.E. Ebert, H.B. Christensen, Y. Chen, I.K. Kim, M.J. Herrgård, L.M. Blank, J. Forster, J. Nielsen, I. Borodina, Engineering and systems-level analysis of *Saccharomyces cerevisiae* for production of 3-hydroxypropionic acid via malonyl-CoA reductase-dependent pathway, Microb. Cell Fact. 15 (2016) 1–13. https://doi.org/10.1186/s12934-016-0451-5.

[30] A.L. Meadows, K.M. Hawkins, Y. Tsegaye, E. Antipov, Y. Kim, L. Raetz, R.H. Dahl, A. Tai, T. Mahatdejkul-Meadows, L. Xu, L. Zhao, M.S. Dasika, A. Murarka, J. Lenihan, D. Eng, J.S. Leng, C.L. Liu, J.W. Wenger, H. Jiang, L. Chao, P. Westfall, J. Lai, S. Ganesan, P. Jackson, R. Mans, D. Platt, C.D. Reeves, P.R. Saija, G. Wichmann, V.F. Holmes, K. Benjamin, P.W. Hill, T.S. Gardner, A.E. Tsong, Rewriting yeast central carbon metabolism for industrial isoprenoid production, Nature. 537 (2016) 694–697. https://doi.org/10.1038/nature19769.

[31] B.J. Harder, K. Bettenbrock, S. Klamt, Model-based metabolic engineering enables high yield itaconic acid production by *Escherichia coli*, Metab. Eng. 38 (2016) 29–37. https://doi.org/10.1016/j.ymben.2016.05.008.

[32] P. Mishra, N.R. Lee, M. Lakshmanan, M. Kim, B.G. Kim, D.Y. Lee, Genome-scale model-driven strain design for dicarboxylic acid production in *Yarrowia lipolytica*, BMC Syst. Biol. 12 (2018). https://doi.org/10.1186/s12918-018-0542-5.

[33] J.E. Yang, S.J. Park, W.J. Kim, H.J. Kim, B.J. Kim, H. Lee, J. Shin, S.Y. Lee, One-step fermentative production of aromatic polyesters from glucose by metabolically engineered *Escherichia coli* strains, Nat. Commun. 9 (2018). https://doi.org/10.1038/s41467-017-02498-w.

[34] T. Eng, D. Banerjee, J. Menasalvas, Y. Chen, J. Gin, H. Choudhary, E. Baidoo, J.H. Chen, A. Ekman, R. Kakumanu, Y.L. Diercks, A. Codik, C. Larabell, J. Gladden, B.A. Simmons, J.D. Keasling, C.J. Petzold, A. Mukhopadhyay, Maximizing microbial bioproduction from sustainable carbon sources using iterative systems engineering, Cell Rep. 42 (2023) 113087. https://doi.org/10.1016/j.celrep.2023.113087.

CHAPTER 3

Table A 3.1. Primers used in this study for gene amplification, colony PCR and sequencing. Histidine Tag sequence is underlined.

Primers	Template	Sequence (5′ → 3′)	Restriction enzyme
Gene amplification for cloning			
Zm_Fw	*Zymomonas mobilis* gDNA	AAAGGATCCAAAAAAATGCATCACCATCACCATCAC CGCATTACCATGATCGGTTCC	*Bam*HI
Zm_Rv		AAAACTCGAGTTACCCTGCATTGGGTCGAC	*Xho*I
Lbj_Fw	*Lactobacillus. johnsonii* gDNA	AAAGGATCCAAAAAAATGCATCACCATCACCATCAC AATTCATATAAAATTGCAGTTG	*Bam*HI
Lbj_Rv		AAAACTCGAGCTAATCTTTTCTAAAAATATCTCTTG	*Xho*I
Zm_prsf_Fw	p426GPD_*Zm*UGD	AAAGAATTCTATGCGCATTACCATGATCGGTTC	*Eco*RI
Zm_prsf_Rv		AAAGCGGCCGCTTACCCTGCATTGGGTCGAC	*Not*I
Lbj_prsf_Fw	pGEM_*Lbj*UGD	AAAAAGAATTCTATGAATTCATATAAAATTGCAGTTG	*Eco*RI
Lbj_prsf_Rv		AAAGCGGCCGCCTAATCTTTTCTAAAAATATCTCTTG	*Not*I
Ch_Fw	pUC19_*Ch*UGD	AAAGGATCCAAAAAAATGCATCACCATCACCATCAC TTTGAGATTAAAAAAATCTGC	*Bam*HI
Ch_Rv		AAAACTCGAGCTATACTCTCGGTTTCTTATTC	*Xho*I
Colony PCR and Sequencing			
M13_Fw	pGEM®-T Easy and derivatives	GTAAAACGACGGCCAGT	
M13_Rv		GCGGATAACAATTTCACACAGG	
pRSFDuet_Fw	pRSFDuet-1 and derivatives	GGATCTCGACGCTCTCCCTT	
T7 terminator_Rv		CTAGTTATTGCTCAGCGGT	
PGK1p_Fw	pSP-GM1 and derivatives	GTGATCTCCAGAGCAAAGT	
CYCt_Rv		CTTCGAGCGTCCCAAAAC	
p426seq_Fw	p426GPD and derivatives	ACTAAAGGGAACAAAAGCTG	
CYC 1_Rv		ACTTCAGGTTGTCTAACTCC	

Table A 3.2. Gene sequence of UGD from *Capra hircus* codon optimized (NZYTech) for *Saccharomyces cerevisiae.*

>ugd1 (UDP-glucose 6-dehydrogenase) from *C. hircus* (1485 bp; GC%: 41.89; CAI: 0.74)

ATGTTTGAGATTAAAAAAATCTGCTGCATCGGTGCGGGTTATGTCGGTGGTCCTACTTGTAGCGTTATTGCACACATGT
GTCCTGAAATTCGTGTCACCGTCGTTGACATCAATGAGTCAAGAATAAATGCTTGGAATAGTCCGACTCTACCAATTTAT
GAACCCGGTTTGAAAGAAGTAGTTGAATCTTGCAGAGGTAAGAACTTGTTTTTCTCTACAAACATCGACGACGCGATAA
AGGAAGCTGATCTGGTATTCATCTCTGTTAATACACCCACCAAAACGTACGGTATGGGCAAGGGTAGAGCTGCTGATCT
AAAGTACATCGAAGCGTGTGCACGTCGAATAGTTCAGAACTCTCATGGTTACAAGATCGTGACAGAAAAGTCCACAGTT
CCAGTTAGAGCAGCTGAATCCATAAGGAGGATCTTCGATGCTAACACAAAGCCAAACTTGAACTTACAGGTATTGTCAA
ACCCGGAATTTTTAGCTGAAGGGACGGCCATCAAGGATTTGAAGAACCCAGACCGAGTTCTAATCGGTGGTGACGAGA
CACCTGAAGGTCAAAGAGCTGTCCAAGCTCTCTGTGCAGTTTACGAGCATTGGGTGCCCCGTGAAAAAATTCTTACAA
CCAACACTTGGTCTTCAGAGTTGTCGAAACTAACCGCTAATGCATTTTTAGCTCAAAGAATTTCTAGCATCAATTCAATT
AGTGCCTTGTGTGAAGCAACCGGCGCAGATGTAGAGGAAGTCGCTACAGCTATTGGAATGGATCAAAGAATAGGTTCT
AAATTTCTAAAAGCTAGCGTAGGATTTGGTGGTTCTTGCTTTCAAAAGGATGTCCTTAATTTAGTGTATTTGTGTGAAGC
CCTTAATTTACCAGAAGTCGCTAGATACTGGCAACAGGTCATAGATATGAACGACTATCAACGTAGAAGATTCGCTTCG
AGGATAATTGATTCTCTATTTAACACAGTTACTGATAAAAAGATTGCCATATTGGGATTCGCCTTTAAGAAAGACACTGG
CGACACAAGAGAGTCATCATCAATATACATTTCTAAGTATTTGATGGATGAAGGTGCTCACTTGCATATTTATGATCCAA
AGGTGCCACGTGAACAGATTGTGGTAGATTTATCGCATCCCGGCGTTAGCAAGGATGATCAAGTTGCTCGTCTTGTGA
CAATCTCCAAAGACCCATACGAAGCATGTGATGGTGCTCACGCAGTGGTTATTTGTACAGAATGGGACATGTTTAAAGA
ACTGGATTACGAGAGAATACACAAAAAAATGTTGAAGCCCGCTTTTATTTTTGATGGTAGGAGAGTGTTGGACGGGCTA
CATAACGAATTGCAAACCATAGGCTTTCAAATTGAGACAATTGGGAAGAAGGTTTCCAGTAAAAGAATCCCTTATGCGC
CGTCAGGGGAGATACCAAAGTTTAGTCTACAAGACATGCCGAATAAGAAACCGAGAGTATAG

Table A 3.3. Results of BLASTp sequences search against UniProtKB/Swiss-Prot database ordered by the identity percentages obtained. Results with score <50 were removed to exclude enzymes that showed similarity with only one of the proteins domains.

Query	Organism	Protein	Access Number	Identity (%)
*Zm*UGD	*Sinorhizobium meliloti* 1021	UGD	O54068.2	67.05
	Bacillus subtilis subsp. *subtilis* str. 168	UGD (YwqF)	P96718.1	48.51
	Bacillus subtilis subsp. *subtilis* str. 168	UGD (TuaD)	O32271.1	46.67
	Pseudomonas aeruginosa PAO1	UGD	O86422.2	45.86
	Rickettsia bellii RML369-C	UGD	Q1RKF8.1	44.93
	Methanocaldococcus jannaschii DSM 2661	UGD	Q58454.1	44.04
	Rickettsia conorii str. Malish 7	UGD	Q92GB1.1	44.01
	Rickettsia felis URRWXCal2	UGD	Q4UK39.1	43.91
	Rickettsia typhi str. Wilmington	UGD	Q68VX0.1	43.78
	Rickettsia prowazekii str. Madrid E	UGD	O05973.1	43.09
	Bacillus subtilis subsp. *subtilis* str. 168	Putative UGD (YtcA)	O34862.1	40.50
	Haloferax volcanii DS2	UGD (AglM)	D4GYH5.1	39.00
	Drosophila melanogaster	UGD	O02373.1	37.53
	Pseudomonas aeruginosa PAO1	GMD	P11759.2	36.91
	Homo sapiens	UGD	O60701.1	36.85

Table A 3.3. Results of BLASTp sequences search against UniProtKB/Swiss-Prot database ordered by the identity percentages obtained. Results with score <50 were removed to exclude enzymes that showed similarity with only one of the proteins domains (*continuation*).

Query	Organism	Protein	Access Number	Identity (%)
*Zm*UGD (cont.)	*Pongo abelii*	UGD	Q5R7B3.1	36.85
	Bos taurus	UGD	P12378.2	36.64
	Mus musculus	UGD	O70475.1	36.64
	Oryza sativa Japonica Group	UGD1	Q75GS4.1	36.46
	Glycine max	UGD1	Q96558.1	36.23
	Gallus gallus	UGD	Q5F3T9.1	36.23
	Oryza sativa Japonica Group	UGD2	B7F958.1	36.09
	Rattus norvegicus	UGD	O70199.1	36.01
	Arabidopsis thaliana	UGD2	Q9LIA8.1	35.96
	Arabidopsis thaliana	UGD3	Q9LF33.1	35.73
	Arabidopsis thaliana	UGD1	Q9FZE1.1	35.39
	Caenorhabditis elegans	UGD	Q19905.1	35.36
	Arabidopsis thaliana	UGD4	Q9FM01.1	35.31
	Azotobacter vinelandii	GMD	P51585.1	35.08
	Oryza sativa Japonica Group	UGD5	Q2QS13.1	35.02
	Pseudomonas syringae pv. Tomato str. DC3000	GMD	Q887P8.1	34.99
	Oryza sativa Japonica Group	UGD3	Q9AUV6.1	34.75
	Oryza sativa Japonica Group	UGD4	Q2QS14.1	34.53
	Pseudomonas putida KT2440	GMD	Q88NC4.1	33.18
	Pseudomonas syringae pv. Syringae	GMD	P59793.1	32.60
	Escherichia coli K5	UGD	Q47329.1	32.39
	Pseudomonas savastanoi pv. Phaseolicola	GMD	O07299.1	31.56
	Escherichia coli K-12	UAMAD	P27829.4	31.01
	Escherichia coli O8:K40	UGD	O33952.1	30.79
	Salmonella enterica subsp. *Enterica serovar Typhi*	UAMAD	Q8Z389.1	30.66
	Salmonella enterica subsp. *Enterica serovar Typhimurium* str. LT2	UAMAD	Q9L6R4.1	30.39
	Streptococcus pneumoniae	UGD	Q57346.2	29.92
	Salmonella enterica subsp. *Enterica serovar Typhimurium* str. LT2	UGD	Q04873.1	29.86
	Shigella flexneri	Putative UGD	P37791.1	29.56
	Escherichia coli O157:H7	UGD	Q7DBF9.1	29.21
	Escherichia coli CFT073	UGD	Q8FG45.1	29.21
	Escherichia coli K-12	UGD	P76373.1	29.21
	Escherichia coli O111:H-	UGD	Q04872.1	29.13
	Methanococcus maripaludis S2	UAMAD	Q6LZC3.1	28.70
	Streptococcus pyogenes	UGD	P0C0F4.1	28.57
	Streptococcus pyogenes serotype M1	UGD	P0C0F5.1	28.57
	Methanococcus maripaludis C7	UAMAD	A6VK13.1	28.47

Table A 3.3. Results of BLASTp sequences search against UniProtKB/Swiss-Prot database ordered by the identity percentages obtained. Results with score <50 were removed to exclude enzymes that showed similarity with only one of the proteins domains (*continuation*).

Query	Organism	Protein	Access Number	Identity (%)
*Zm*UGD (cont.)	*Streptococcus pyogenes* SSI-1	UGD	P0DG69.1	28.30
	Streptococcus pyogenes MGAS315	UGD	P0DG68.1	28.30
	Methanococcus maripaludis C5	UAMAD	A4FY94.1	28.02
	Ralstonia solanacearum GMI1000	NAGAAD	P58591.1	27.87
	Yersinia pestis	UAMAD	Q8ZAE4.1	27.85
	Methanocaldococcus jannaschii DSM 2661	UAMAD	Q57871.1	27.52
	Ralstonia solanacearum	NAGAAD	Q45410.1	27.17
	Methanococcus vannielii SB	UAMAD	A6USK4.1	26.90
	Methanococcus aeolicus Nankai-3	UAMAD	A6UU98.1	26.09
	Pseudomonas aeruginosa PAO1	UAGAD	G3XD94.1	25.74
*Lbj*UGD	*Escherichia coli* O111:H-	UGD	Q04872.1	61.80
	Escherichia coli O8:K40	UGD	O33952.1	61.56
	Escherichia coli CFT073	UGD	Q8FG45.1	60.34
	Escherichia coli O157:H7	UGD	Q7DBF9.1	60.10
	Escherichia coli K-12	UGD	P76373.1	59.85
	Salmonella enterica subsp. *enterica serovar Typhimurium* str. LT2	UGD	Q04873.1	59.61
	Streptococcus pyogenes serotype M1	UGD	P0C0F5.1	57.77
	Streptococcus pyogenes WF50	UGD	P0C0F4.1	57.77
	Streptococcus pneumoniae	UGD	Q57346.2	57.66
	Escherichia coli K5	UGD	Q47329.1	57.63
	Streptococcus pyogenes MGAS315	UGD	P0DG68.1	57.52
	Streptococcus pyogenes SSI-1	UGD	P0DG69.1	57.52
	Shigella flexneri	Putative UGD	P37791.1	57.42
	Methanocaldococcus jannaschii DSM 2661	UGD	Q58454.1	32.85
	Rickettsia prowazekii str. Madrid E	UGD	O05973.1	30.42
	Bacillus subtilis subsp. *subtilis* str. 168	UGD (YwqF)	P96718.1	30.03
	Rickettsia bellii RML369-C	UGD	Q1RKF8.1	30.02
	Rickettsia typhi str. Wilmington	UGD	Q68VX0.1	29.71
	Bacillus subtilis subsp. *subtilis* str. 168	UGD (TuaD)	O32271.1	29.64
	Rickettsia felis URRWXCal2	UGD	Q4UK39.1	29.10
	Sinorhizobium meliloti 1021	UGD	O54068.2	29.06
	Bacillus subtilis subsp. *subtilis* str. 168	Putative UGD (YtcA)	O34862.1	28.76
	Rickettsia conorii str. Malish 7	UGD	Q92GB1.1	28.46
	Azotobacter vinelandii	GMD	P51585.1	28.39
	Pseudomonas syringae pv. Syringae	GMD	P59793.1	28.00
	Pseudomonas syringae pv. tomato str. DC3000	GMD	Q887P8.1	27.93

Table A 3.3. Results of BLASTp sequences search against UniProtKB/Swiss-Prot database ordered by the identity percentages obtained. Results with score <50 were removed to exclude enzymes that showed similarity with only one of the proteins domains (*continuation*).

Query	Organism	Protein	Access Number	Identity (%)
*Lbj*UGD (cont.)	*Pseudomonas savastanoi pv. phaseolicola*	GMD	O07299.1	27.76
	Arabidopsis thaliana	UGD4	Q9FM01.1	27.11
	Caenorhabditis elegans	UGD	Q19905.1	27.03
	Pseudomonas aeruginosa PAO1	GMD	P11759.2	26.77
	Bacillus licheniformis DSM 13 = ATCC 1458	GPD	Q65I16.1	26.53
	Rattus norvegicus	UGD	O70199.1	26.51
	Homo sapiens	UGD	O60701.1	26.51
	Bos taurus	UGD	P12378.2	26.51
	Pseudomonas aeruginosa PAO1	UGD	O86422.2	26.37
	Haloferax volcanii DS2	UGD (AglM)	D4GYH5.1	26.32
	Pongo abelii	UGD	Q5R7B3.1	26.25
	Mus musculus	UGD	O70475.1	26.25
	Pseudomonas putida KT2440	GMD	Q88NC4.1	25.85
	Drosophila melanogaster	UGD	O02373.1	25.39
	Staphylococcus aureus	UGD/GMD superfamily (CapL)	P39861.1	25.24
*Ch*UGD	*Bos taurus*	UGD	P12378.2	99.80
	Homo sapiens	UGD	O60701.1	98.38
	Pongo abelii	UGD	Q5R7B3.1	98.18
	Mus musculus	UGD	O70475.1	96.36
	Rattus norvegicus	UGD	O70199.1	96.15
	Gallus gallus	UGD	Q5F3T9.1	92.51
	Drosophila melanogaster	UGD	O02373.1	68.55
	Caenorhabditis elegans	UGD	Q19905.1	65.87
	Glycine max	UGD1	Q96558.1	61.75
	Arabidopsis thaliana	UGD4	Q9FM01.1	61.75
	Oryza sativa Japonica Group	UGD5	Q2QS13.1	61.44
	Arabidopsis thaliana	UGD2	Q9LIA8.1	61.32
	Arabidopsis thaliana	UGD1	Q9FZE1.1	61.31
	Arabidopsis thaliana	UGD3	Q9LF33.1	61.11
	Oryza sativa Japonica Group	UGD4	Q2QS14.1	61.02
	Oryza sativa Japonica Group	UGD2	B7F958.1	60.98
	Oryza sativa Japonica Group	UGD3	Q9AUV6.1	60.81
	Oryza sativa Japonica Group	UGD1	Q75GS4.1	59.75
	Sinorhizobium meliloti 1021	UGD	O54068.2	37.96
	Bacillus subtilis subsp. *subtilis* str. 168	UGD (YwqF)	P96718.1	35.42
	Methanocaldococcus jannaschii DSM 2661	UGD	Q58454.1	35.38
	Bacillus subtilis subsp. *subtilis* str. 168	UGD (TuaD)	O32271.1	34.77

Table A 3.3. Results of BLASTp sequences search against UniProtKB/Swiss-Prot database ordered by the identity percentages obtained. Results with score <50 were removed to exclude enzymes that showed similarity with only one of the protein domains (*continuation*).

Query	Organism	Protein	Access Number	Identity (%)
*Ch*UGD (cont.)	*Haloferax volcanii* DS2	UGD (AglM)	D4GYH5.1	34.68
	Pseudomonas aeruginosa PAO1	UGD	O86422.2	32.70
	Rickettsia conorii str. Malish 7	UGD	Q92GB1.1	32.54
	Rickettsia felis URRWXCal2	UGD	Q4UK39.1	32.47
	Rickettsia typhi str. Wilmington	UGD	Q68VX0.1	32.47
	Rickettsia prowazekii str. Madrid E	UGD	O05973.1	32.40
	Rickettsia bellii RML369-C	UGD	Q1RKF8.1	32.26
	Bacillus subtilis subsp. *subtilis* str. 168	Putative UGD (YtcA)	O34862.1	31.29
	Pseudomonas syringae pv. syringae	GMD	P59793.1	29.17
	Pseudomonas syringae pv. tomato str. DC3000	GMD	Q887P8.1	29.09
	Pseudomonas savastanoi pv. phaseolicola	GMD	O07299.1	28.91
	Azotobacter vinelandii	GMD	P51585.1	28.68
	Pseudomonas aeruginosa PAO1	GMD	P11759.2	27.98
	Shigella flexneri	Putative UGD	P37791.1	27.82
	Salmonella enterica subsp. *enterica serovar Typhimurium* str. LT2	UGD	Q04873.1	27.72
	Methanococcus maripaludis C7	UAMAD	A6VK13.1	27.47
	Escherichia coli O8:K40	UGD	O33952.1	27.37
	Methanococcus maripaludis S2	UAMAD	Q6LZC3.1	27.25
	Escherichia coli O111:H-	UGD	Q04872.1	27.25
	Methanocaldococcus jannaschii DSM 2661	UAMAD	Q57871.1	27.21
	Pseudomonas putida KT2440	GMD	Q88NC4.1	27.08
	Methanococcus aeolicus Nankai-3	UAMAD	A6UU98.1	26.87
	Methanococcus vannielii SB	UAMAD	A6USK4.1	26.82
	Escherichia coli K5	UGD	Q47329.1	26.53
	Escherichia coli O157:H7	UGD	Q7DBF9.1	26.36
	Methanococcus maripaludis C5	UAMAD	A4FY94.1	26.18
	Escherichia coli CFT073	UGD	Q8FG45.1	26.09
	Escherichia coli K-12	UGD	P76373.1	26.09
	Pyrococcus horikoshii OT3	UAMAD	O59284.1	25.74
	Streptococcus pneumoniae	UGD	Q57346.2	25.68
	Salmonella enterica subsp. *enterica serovar Typhi*	UGD/GMD superfamily (VipA/TviB)	Q04972.1	23.79

Enzyme abbreviations: UGD, uridine diphosphate(UDP)-glucose 6-dehydrogenase; GMD, guanosine diphosphate(GDP)-mannose 6-dehydrogenase; UAMAD, UDP-*N*-acetyl-D-mannosamine dehydrogenase; UAGAD, UDP-*N*-acetylglucosamine 6-dehydrogenase; and NAGAAD, nucleoside diphosphate(NDP)-*N*-acetyl-D-galactosaminuronic acid dehydrogenase.

CHAPTER 4

Table A 4.1. Primers used in this study.

Primers	Template	Sequence (5′ → 3′)	Restriction enzyme
Zmugd_MCS2_Fw	pRSFDuet_Zmugd	GGAATCATATGCGCATTACCATGATCG	*Nde*I
Zmugd_MCS2_Rv		AACCGCTCGAGTTACCCTGCATTGGGTCG	*Xho*I
kfoC_MCS1_Fw	pETM6_PCA	ACATGCCATGGGCATGAGTATTCTTAATCAAGC	*Nco*I
KfoC_MCS1_Rv		AACGCGGATCCTGCCATATGTATATCTCCTT	*BamH*I
KfoA_MCS1_Fw	pETM6_PCA	AACGCGGATCCATGAATATATTAGTTACAGGTG	*BamH*I
KfoA_MCS1_Rv		AGAATGCGGCCGCTTAAATATAACCATTTGGGTTT	*Not*I
glmU_Fw	*Escherichia coli* K-12 MG1655 (DE3) gDNA	GGAATCATATGTTGAATAATGCTATGAGC	*Nde*I
glmU_Rv		GGAATGATATCACTTTTTCTTTACCGGAC	*EcoRV*=Eco32I
sodA_Fw		GGAATGAATTCTATGAGCTATACCCTGCCA	*EcoR*I
sodA_Rv		GGAATGTCGACTTATTTTTTCGCCGCAAAACG	*Sal*I
mltB_Fw		GGAATGATATCTATGTTCAAGCGTCGTTATG	*EcoRV*
mltB_Rv		GGAATCTCGAGCTACTGTACTCGCGCC	*Xho*I
ybiV_UE_Fw	-	AAACGAGCGTAAAAGTTATCGTCACAGACATGGAGTTTTAGAGCTATGCTGTTTTGAATGGTCCCA	-
ybiV_UE_Rv		GTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACTCCATGTCTGTGACGATAACTTTTACGCTC	-
Colony PCR and Sequencing			
MCS2_Fw	pRSFDuet-1 and derivatives	GTACACGGCCGCATAATCG	
MCS1_Fw		GGATCTCGACGCTCTCCCTT	
T7 terminator_Rv		CTAGTTATTGCTCAGCGGT	
CRISPRi_cPCR_Fw	pCRISPathBrick	AGCTGAGACAAATAGTGCGA	
CRISPRi_cPCR_Rv		ACGAGTCCTATGAGCTTCCG	
pCRISPath_Fw		CGCATTGATTTGAGTCAGCTAG	
pAC_seq_Rv		ATCGGTGATGTCGGCGATA	

Table A 4.2. Primers used in the attempts of *nagZ* deletion and *murJ* underexpression.

Purpose	Name	Sequence (5′ → 3′)
To clone the protospacer (20 bp) into the pKDsgRNA plasmid using CPEC	nagZ_CPEC_Fw	GGATGTCGGGCATATCAGCGGTTTTAGAGCTAGAAATAGCAAG
	K_Rv	TTTATAACCTCCTTAGAGCTCGA
	nagZ_CPEC_Rv	CGCTGATATGCCCGACATCCGTGCTCAGTATCTCTATCACTGA
	K_Fw	CCAATTGTCCATATTGCATCA
To verify deletion	nagZ_seq_Fw	TTGGTCAATGACTATGCCACTC
	nagZ_seq_Rv	ACCCGGACTGTTAGAGTCAAAA
Oligo with desired mutations / modifications: oligo_nagZ: = 37+36=73 mer oligo	oligo_nagZ	TGGCGGCAGCTATTAATAAAACAATAAGGAGAGCAGT CCCTGGCTTATGTGAGGAAGCGATGATTATCTATTT
new oligo nagZ	oligo_nagZ_2	AATGTTGGATGTCGAAGGTTACGAACTGGACGCGGAA CGATCAGCACCCGTCTGAATCAGTTACATGAACGCT
Primer to verify protospacer in pKD plasmid by sequencing	pKD_Seq5	CAGTGAATGGGGGTAAATGG
For murJ underexpression	murJ_UE_Fw	AAACCACGCGAAAACATGGTCATCGAGCTGACGGGTTTTAGAGCTATG CTGTTTTGAATGGTCCCA
	murJ_UE_Rv	GTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACCCGTCAGCTCGA TGACCATGTTTTCGCGTG
For murJ underexpression, new primers	murJ2_Fw	AAACTAAATTCATCGGTGTTCTAATCCAGACCCAGTTTTAGAGCTATGC TGTTTTGAATGGTCCCA
	murJ2_Rv	GTTTTGGGACCATTCAAAACAGCATAGCTCTAAAACTGGGTCTGGATTA GAACACCGATGAATTTA

CHAPTER 5

Table A 5.1. Primers used in this study.

Purpose	Name	Template	Sequence (5′ → 3′)	Restriction enzyme
Cloning of *Giuae* in pSP-GM1 (PGK1)	Gi_uae_Fw1	pUC57_Giuae	AAAGGATCCAAAAAAATGCATCACCATCACCATCACCAGAATCATTCCTTCTTGG	*Bam*HI
	Gi_uae_Rv1		AAAACTCGAGCTATAATCCAATACCTTTTACATC	*Xho*I
Cloning of *Btchsy* in pSP-GM1 (PGK1)	Bt_chsy1_Fw	pUC57_Btchsy1	AAAGGATCCAAAAAAATGCATCACCATCACCATCACGCTGTCCGAGGTCGAAGAGC	*Bam*HI
	Bt_chsy1_Rv		AAAACTCGAGTTAAGCCGTGCGGACTGAGC	*Xho*I
Cloning of *Giuae* in pSP-GM1_(PTEF)	Giuae_tefp_Fw	pSP-GM1_Giuae	ATTTGCGGCCGCAAAACAATGCATCACCATCACCAT	*Not*I
	Giuae_tefp_Rv		AACGAGCTCCTATAATCCAATACCTTTTACATC	*Sac*I
Zmugd cloning in pSP-GM1_tefp_*Giuae*	Zm_psp_Fw	pSP-GM1_Zmugd	AAAGGATCCAAAAAAATGCATCACCATCACCATCACCGCATTACCATGATCGGTTCC	*Bam*HI
	Zm_psp_Rv		AAAGCTAGCTTACCCTGCATTGGGTCGAC	*Nhe*I
Btchsy cloning in pBEVY	Btchsy_pBEVY_Fw	pSP-GM1_Btchsy	AACGCGGATCCAAAACAATGCATCACCATCAC	*Bam*HI
	Btchsy_pBEVY_Rv		AACGCCGTCGACTTAAGCCGTGCGGACTGAGC	*Sal*I
kfoA cloning in pSP-GM1_Zmugd	kfoA_psp_Fw	pETM6_kfoCA	ATTTGCGGCCGCAAAACAATGAATATATTAGTTACAGGTG	*Not*I
	kfoA_psp_Rv		AACGAGCTCTTAAATATAACCATTTGGGTTT	*Sac*I
kfoC cloning in pBEVY-L	kfoC_pBEVY_Fw	pETM6_kfoCA	AACGCGGATCCAAAACAATGAGTATTCTTAATCAAGC	*Bam*HI
	kfoC_pBEVY_Rv		AACGCCGTCGACTTATAAATCATTCTCTATTTTTTCCC	*Sal*I
Colony PCR and sequencing				
	tADH1_Fw	pSP-GM1_tefp_Giuae / pSP-GM1_Zmugd_kfoA	CCTAAGAGTCACTTTAAAATTTGT	
	TEF1p_Rv		CGTACCACTTCAAAACACC	
	PGK1p_FW	pSP-GM1_Zmugd / pSP-GM1_Giuae_Zmugd	GTGATCTCCAGAGCAAAGT	
	CYCt_REV		CTTCGAGCGTCCCAAAAC	
	tADH1_Fw	pBEVY_Btchsy / pBEVY_kfoC	CCTAAGAGTCACTTTAAAATTTGT	
	pGAL_middle_Rv		ATTATTCCCCTACTTGACTAATAAG	
	Bt_chsy1_middle	pBEVY_Btchsy	CGGTAAATGTTTAAGAGAGATG	
	kfoC_pBEVY_Rv	pBEVY_kfoC	AACGCCGTCGACTTATAAATCATTCTCTATTTTTTCCC	

Table A 5.2. Codon-optimized gene sequences (5′ → 3′) for *Saccharomyces cerevisiae.*

>Chondroitin synthase (chsy1) (2412 bp GC%: 41,54; CAI: 0,74)

ATGGCTGTCCGAGGTCGAAGAGCATGGTTAAGCGTTTTGCTGGGTTTAGTTCTTGGTTTTGTCTTAGCCTCTAGACTCGTTTT
GCCGAGGGCTTCTGAGTTAAAAAGAGCAGGCCCACGGAGAAGAGCTAGTCCAGAGGCTTGTAGGCCGGGTCAGGCAGCAG
CCGTACCATTAGCTGGAGGTGCTAGAGGAGATGCGAGAGGCCAAAGATTATGGCCCCACGGGGAAGCTCAAGACGGTGTTC
CCCGCGACCGCAATTTCTTATTCGTAGGAGTCATGACCGCTCAAAAGTATCTACAAACGAGAGCTGTGGCTGCTTTCAGAAC
TTGGTCAAAAACTATCCCAGGCAAGGTCGAATTTTTCTCATCTGAAGGTTCTGATACCAGTATTAGTATTCCTGTGGTCCCATT
GCGAGGTGTTGATGATTCATATCCGCCTCAGAAAAAATCATTCATGATGTTAAAATATATGCATGATCATTACTTGGATAAGTA
CGAATGGTTCATGAGAGCTGACGATGATGTTTATATTAAAGGTGACAGATTGGAATCATTCTTAAGATCGCTCAACAGCAGCG
AGCCTTTGTTCTTGGGACAGACCGGCTTAGGAACCACTGAAGAAATGGGAAAACTTGCTCTAGAACCCGGTGAAAACTTTTG
CATGGGAGGCCCCGGCGTTATTATGAGCAGAGAAGTATTGAGAAGAATGGTACCTCACATCGGTAAATGTTTAAGAGAGATGT
ACACTACTCACGAGGATGTAGAGGTGGGTCGCTGCGTTAGAAGATTTGCTGGTGTACAATGCGTCTGGTCGTACGAAATGCA
ACAACTCTTTTATGAAAATTATGAACAAAATAAAAAAGGTTACATACGTGATTTACACAATTCCAAGATACATAGAGCCATCAC
CTTACATCCTAACAAAAATCCTCCATATCAATATCGTTTGCACAGCTATATGCTTTCTAGAAAAATCGCTGAACTGAGACATCG
TACTATCCAACTGCATCGCGAGATTGTCCTTATGTCCAAGTACAGTAATACGGAAGTTCATAAAGAGGATTTGCAATTAGGTAT
TCCACCGTCTTTCATGAGATTTCAACCAAGACAAAGAGAAGAAATTTTGGAATGGGAATTTCTAACCGGTAAGTATTTATATAG
TGCAGCCGATAGCCAGCCCCCACGTAGAGGTATGGACAGCGCCCAAAGAGAAGCATTGGATGATATTGTTATGCAAGTTATG
GAAATGATAAATGCCAATGCCAAAACTAGAGGGAGAATTATAGATTTCAAAGAAATTCAATACGGTTATCGCCGTGTTAATCCG
ATGTACGGGGCTGAATATATACTGGATTTGTTGTTGCTTTACAAGAAACATAAAGGTAAGAAAATGACAGTTCCAGTTAGAAGA
CATGCCTATTTGCAACAAACTTTTTCAAAAATTCAGTTTGTTGAACATGAAGAACTTGATGCAAAGGAATTGGCGAATAAAATT
AATCAAGAATCCGGCTCCCTGTCTTTCCTGTCCAATTCATTGAAAAAGTTGGTGCCATTTCAACTTCCCGGGAGTAAAAACGA
ACACAAAGAACCTAAGGAGAAAAAGATTAATATTTTAATCCCCCTGAGCGGAAGATTTGACATGTTTGTCCGTTTCATGGGAAA
CTTTGAAAAGACATGTTTAATTCCGAATCAAAACGTGAAACTGGTTGTCTTATTGTTTAATTCTGATTCGAACCCTGACAAAGC
CAAGCAAGTGGAACTTATGCGTGATTACAGAATTAAATATCCCAAGGCCGATATGCAAATTTTGCCGGTGTCAGGTGAGTTTT
CCAGAGCCTTGGCATTGGAGGTGGGCAGTTCTCAATTTAACAATGAGAGTCTTTTGTTTTTCTGTGATGTTGACTTAGTCTTCA
CGGCCGAGTTTTTGCAACGTTGTAGAGCTAACACTGTCTTAGGCCAACAAATTTACTTCCCTATTATCTTCTCACAATATGATC
CTAAAATTGTCTACTCCGGCAAAGTTCCAAGTGATAATCACTTTGCTTTTACTCAGAAAACTGGTTTTTGGAGGAATTATGGAT
TTGGAATTACGTGTATATACAAAGGGGATTTAGTGAGGGTGGGTGGCTTTGATGTGTCCATCCAGGGTTGGGGTCTGGAAGAT
GTCGATTTATTCAACAAAGTCGTTCAAGCAGGGCTTAAAACTTTCAGGAGCCAGGAAGTTGGCGTGGTCCACGTACACCATC
CTGTTTTTTGTGATCCGAACTTAGATCCTAAGCAATACAAAATGTGTTTGGGCAGTAAAGCCTCCACTTATGGTAGCACTCATC
AATTGGCAGAAATGTGGTTAGAAAAAAACGATCCCAATTACTCTAAGTCGAGTAACAACAACGGCTCAGTCCGCACGGCTTAA

>UDP-N-acetylglucosamine 4-epimerase (UAE) (1158 bp GC%: 42,31; CAI: 0,77)

ATGCAGAATCATTCCTTCTTGGGTAAAACGGTCTTGATCACAGGCGGTTGTGGTTTCATTGGATCTCATTTTGTAGAGGCTTG
TCATGTTCTCGGTATGACCGTCTATGTTTTAGATAACCTCAGTTCTGGTAAGAATGTATTTAAGACAACATCTGATTGTTCTTC
ATCATTGGTTTACACTATCGGTGACATTAGAGACAAGGCTATCTTCTCTAGACTTCCGCAGAAAATTGACTTCGTCATTCATTT
GGCCGCAGCTGTTTCTGTTGCCGAATCAGTAACAAATCCGCAGAAATATATGTTAACCAATGTTGAGGGTAGTAGAAACGTTT
TCCAGTATGCTGTTGACGCAAAGGCTTCAGCCGTCTTAAGTGCTTCTACTGCTGCATATTACGGCGATTGTGGTAAGTCAGCT
ATTACCGAGGCTTTCCCCTATGGTGGCATTTCTCCATACGCTGAAAGTAAGATGGAGATGGAGAGATTAGGAGCCGAATTTCA
AAAGACATCTCGTTGCAGATTCATTTTCTGCAGATTTTTCAATGTCTACGGCCCTAGACAGGACCCCTCCTCGCCCTACACTG
GAGTTATGAGTATTTTTATGGACAGGTGCGCCGCTAGAAAACCGATTACAATCTTTGGTACCGGCGAGCAAACAAGAGATTTT
GTTTTCATCAAGGACTTAATTGTCGCTGCCATTAACTTATTAGGTCAATTGGATAAGTTTCCTATTGGAGCAGATGCCGTGCAA
CAGAACGATCCAGAAGAAGTCCAAAGGTCCGCATATACCGGTGAAGGTGTGTACCCAACTGTTTTCAACATCGGATCGGGAA
TTTCGATCAGTGTCAATGAGCTTGCTGAATTAGCCAAAATTGTTTCTGGTAGACATGAAGTGGAAATTGTTCATGGAGAACCAC
GTAGTGGAGACATATTGCATTCATTATCCGATTGCACTAGGATTAGAAACGCTACTGGATGGAGCGCATCTACTACTCTTAGA
GTTGGTATGTCTGAAACTTGGGGATGGGCTGCTGGCGAAATTAGTTACTTGTCTGGTGACTTGGTAAGAGTGTTGGAAAACGA
ATTAAAGATTGATGGGGTTTCAGTAGCAAAATCTCTATGTGGTAAAGATGCAGATGTAAAAGGTATTGGATTATAG