



HAL
open science

Biosynthesis, Depolymerization, Detection, and Analysis of Glycosaminoglycans with a focus on heparosan

Malgorzata Sulewska

► **To cite this version:**

Malgorzata Sulewska. Biosynthesis, Depolymerization, Detection, and Analysis of Glycosaminoglycans with a focus on heparosan. Biochemistry [q-bio.BM]. Université Grenoble Alpes [2020-..]; Medizinische Hochschule Hannover, 2023. English. NNT : 2023GRALV001 . tel-04484917

HAL Id: tel-04484917

<https://theses.hal.science/tel-04484917>

Submitted on 1 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES
et de HANNOVER MEDICAL SCHOOL**

École doctorale : Chimie et Sciences du Vivant ; Hannover Biomedical Research School

Spécialité : Biochimie

Unité de recherche : Cermav-CNRS ; Institute of Clinical Biochemistry

**Biosynthèse, dépolymérisation, détection
et analyse des glycosaminoglycanes
et de l'héparosane en particulier**

*Biosynthesis, Depolymerization, Detection,
and Analysis of Glycosaminoglycans
with a Focus on Heparosan*

Présentée par :

Małgorzata Sulewska

Direction de thèse :

Dr. Bernard Priem

Université Grenoble Alpes

Directeur de thèse

Prof. Dr. Rita Gerardy-Schahn

Hannover Medical School

Co-directrice de thèse

Rapporteurs :

Prof. Dr. Lothar Elling

RWTH Aachen University

Dr. Régis Daniel

Université d'Évry Val-d'Essonne

Thèse soutenue publiquement le 2023/01/13, devant le jury composé de :

Dr. Bernard Priem

Université Grenoble Alpes

Directeur de thèse

Prof. Dr. Rita Gerardy-Schahn

Hannover Medical School

Co-directrice de thèse

Prof. Dr. Lothar Elling

RWTH Aachen University

Rapporteur

Dr. Régis Daniel

Université d'Évry Val-d'Essonne

Rapporteur

Prof. Dr. Corinne Mercier

Université Grenoble Alpes

Examinatrice/Président du jury

Dr. Mélanie Leroux

HTL Biotechnology

Examinatrice

**Biosynthesis, Depolymerization, Detection,
and Analysis of Glycosaminoglycans
with a Focus on Heparosan**

*Biosynthèse, dépolymérisation, détection
et analyse des glycosaminoglycans
et de l'héparosan en particulier*

A thesis submitted for the degree of
Doctor rerum naturalium (Dr. rer. nat.)
in the subject of Biochemistry
by

Małgorzata Sulewska

November 2022



This doctoral dissertation was prepared in dual affiliation with the Université Grenoble Alpes and Hannover Medical School as part of a Joint PhD Agreement signed by all parties on November 10, 2018. All presented data were generated at Cermav-CNRS and the Institute of Clinical Biochemistry at Hannover Medical School.

Cette thèse de doctorat a été préparée en double affiliation avec l'Université Grenoble Alpes et la Faculté de Médecine de Hanovre dans le cadre d'une convention de thèse conjointe signée par toutes les parties le 10 novembre 2018. Toutes les données présentées ont été générées au Cermav-CNRS et à l'Institut de Biochimie Clinique de la Faculté de Médecine de Hanovre.



The PhD Thesis Committee:

Co-Supervisor: Prof. Dr. Rita Gerardy-Schahn
Hannover Medical School

Co-Supervisor: Dr. Bernard Priem
The Université Grenoble Alpes

Internal expert: Prof. Dr. Corrine Mercier
The Université Grenoble Alpes

External expert: Prof. Dr. Lothar Elling
RWTH Aachen University

External expert: Dr. Daniel Régis
The Université d'Évry Val-d'Essonne

External expert: Dr. Mélanie Leroux
HTL Biotechnology

The PhD Thesis Reviewers:

- Prof. Dr. Lothar Elling, *RWTH Aachen University*
- Dr. Daniel Régis, *The Université d'Évry Val-d'Essonne*

Date of public defence: January 13, 2023

Acknowledgements

Getting a PhD degree has been my dream since many years, with ups and downs on the way. Meanwhile, some people became my inspiration and support; some provided the fundamentals to make it possible. I would like to express my gratitude to all of them.

During the PhD program, I have been lucky to be advised and guided by many outstanding scientists. My deepest appreciation goes to a true role model of a scientist and a leader, Prof. Dr. Rita Gerardy-Schahn. I am thankful for all the kindness and support I received over the last years on all levels. I am equally grateful to Dr. Bernard Priem for choosing me to be a member of this international collaboration and for the subsequential guidance and help. I could not have undertaken this journey without Dr. Timm Fiebig who has been constantly providing me with excellent scientific support, and giving me the confidence to ask questions anytime. I am also grateful to Prof. Dr. Falk Buettner for all the instructions. For all the help, especially at the beginning of my work, great atmosphere and organization of the laboratory, special thanks should also go to the Kapselgroup and the CBO team. I am peculiarly grateful to Monika Berger, who taught me a lot, and was always able to instantly find chemical reagents that I could not. Moreover, I thank the whole team of the Clinical Biochemistry Institute at Hannover Medical School and Cermav-CNRS, especially all the unnamed heroes who used to leave cakes in the common room with a note "For all". I would also like to recognize the examiners of this dissertation, Prof. Dr. Lothar Elling and Dr. Daniel Régis, for acceptance of the assignment. I would like to extend my sincere gratitude to Dr. Serge Perez for trusting me and giving me the opportunity to present my work during the InnoGly event in Greece.

It would be much more difficult to achieve this without the constant support of my friends, especially Paula, Sylwia and Yair. Thank you for always believing in and encouraging me. For giving me the necessary foundations and support, I am forever grateful to my parents and sisters, Diana, Kamila, and Emila. To my parents, Kochani Rodzice dziękuję za całe wsparcie i możliwość edukacji, którą mi daliście.

Finally yet importantly, I am truly indebted to all, who provided me with guidance in preparing this thesis, especially the co-authors; without their help, I would not have been able to create this dissertation in this form.

“Sometimes you're ahead, sometimes you're behind.
The race is long and, in the end, it's only with yourself”.

Mary T. Schmich, 1997

Abstract

Glycosaminoglycans (GAGs) are linear anionic polymers built of repeating heterogeneously modified disaccharide units. As GAGs have a massive impact on various biological processes, including cell differentiation and cancer progression, there is a significant interest in developing GAG-based biomarkers. Some bacteria synthesize GAGs. For instance, *E. coli* K5 produces heparosan, a nonimmunogenic and biocompatible GAG against which antibody production is challenging. Consequently, reagents specifically recognizing heparosan are limited.

Work in the course of my PhD thesis addresses the development of tools detection of heparosan and other GAGs in biological samples and, secondly, the study of heparosan biosynthesis.

The development of analytical tools started with engineering the heparosan-specific tailspike protein (KflB). Using a structure-guided mutational approach, the enzymatic function of KflB was inactivated and the protein transformed into a heparosan-binding protein (an artificial lectin). To fully characterize the functional mechanism of KflB, we designed two HPLC-based assays: (i) Heparosan was digested with KflB, and the products were visualized at $\lambda = 250$ nm. (ii) Fluorescently labeled heparosan nonamers (see **Chapter 2**) were used as KflB substrates. Our results suggest that KflB is a lyase and breaks glycosidic bonds with the formation of a double bond in the uronic acid residue. Moreover, product profiles reveal an endolytic, processive mode of action. Notably, a C-terminal chaperone domain (CTD), present in KflB, was shown to be essential for proper folding but autocatalytically released from the mature protein. Using homology modeling on the highly related KflA and site-directed mutagenesis, the amino acids F202 and E206 were identified as critical for catalytic activity. A mutant with alanine in these positions was inactive but retained heparosan binding function, thus turning into a new heparosan detection tool.

Next (**Chapter 3**), we focused on developing an analysis technology suited to all GAGs. Therefore, we adapted multiplexed capillary gel electrophoresis with laser-induced fluorescence detection (xCGE-LIF) for GAGs analysis. The prerequisite for this type of analysis was establishing a reference database (migration time database). We used commercially available GAG-disaccharides (diGAGs), labeled them with AMAC (2-aminoacridone), and applied the established xCGE-LIF analysis. We show

that diGAG-isomers can be detected with great sensitivity and reliability. For the measurement of GAGs isolated from biological samples, we used GAG-specific lyases to generate diGAGs, labeled them with AMAC, and recorded migration times in electropherograms. Peak assignment was achieved using the established migration time database. After the proof-of-concept studies, we examined GAGs expression during cell differentiation. Preliminary results with hESCs indicate that the diGAGs composition alters during cell differentiation. Our protocol revealed the promising potential to be a powerful tool in GAGs analytics.

The *E. coli* K5 capsule gene cluster contains four genes, *kfiA-D*. While KfiA and KfiC are confirmed heparosan synthases, the role of KfiB is unknown. Identifying KfiB isoforms with varying numbers of a spatially defined sequence, designated as KfiB-repeat, prompted the third study (**Chapter 4**). To explore the role of the KfiB-repeat, we generated KfiB isoforms with one, two, and four repeats (KfiB_{1r}, KfiB_{2r}, and KfiB_{4r}). Synthetic genes were expressed in *E. coli* BL21-*kfiCAD* derivative and we evaluated their impact on heparosan synthesis and distribution. Although a tendency towards longer chains was visible if KfiB_{1r} and KfiB_{4r} were co-expressed, the impact of KfiB could only partially be determined.

Résumé

Les glycosaminoglycannes (GAGs) sont des polymères linéaires constitués d'une répétition d'unités disaccharidiques (diGAGs) présentant une forme hétérogénéité. Ils sont impliqués dans un grand nombre de processus biologiques, tels que la différenciation cellulaire et la cancérisation, où ils peuvent faire office de biomarqueurs spécifiques de ces états physiologiques. Certaines bactéries synthétisent des GAGs. Par exemple, *E. coli* K5 produit de l'héparosan, un GAG non immunogène et biocompatible contre lequel la production d'anticorps est difficile. Par conséquent, les réactifs reconnaissant spécifiquement l'héparosan sont limités.

Dans cette thèse, nous décrivons d'une part le développement d'outils de détection de l'héparosan et autres GAGs dans des extraits biologiques; et d'autre part une étude de la biosynthèse bactérienne de l'héparosan.

Le développement d'outils analytiques a commencé par l'ingénierie de la protéine tailspike spécifique à l'héparosan (KflB). A l'aide d'une approche de mutagenèse dirigée basée sur une analyse structurale, l'activité enzymatique de KflB est neutralisée tout en conservant ses propriétés de fixation à l'héparosan (lectine artificielle). Afin de caractériser le mécanisme catalytique de KflB, deux approches basées sur l'analyse HPLC sont mises en œuvre: (i) la digestion de l'héparosan par KflB et la détection des produits de dégradation à $\lambda = 250$ nm; (ii) Des nonamères d'héparosan marqués à la fluorescence (**chapitre 2**) sont utilisés comme substrats de KflB. Nos résultats montrent que KflB est une lyase qui génère une double liaison au niveau des résidus glucuronyles. Cette activité est de nature endolytique et processive. Un domaine C-terminal chaperone (CTD) est nécessaire au repliement correct de l'enzyme; celui-ci est détaché de façon autolytique pour libérer l'enzyme mature. Par alignement avec la protéine homologue KflA et mutagenèse dirigée, les acides aminés F202 et E206 sont identifiés comme essentiels à la catalyse. Un mutant alanine au niveau de ces positions se révèle inactif mais garde ses propriétés de liaison ce qui en fait un agent de détection de l'héparosan.

Ensuite (**chapitre 3**), nous développons une technique d'analyse applicable aux GAGs en général. Pour cela, nous adaptons une analyse multiplexe en gel d'électrophorèse capillaire avec détection de fluorescence laser (xCGE-LIF).

Ceci nécessite la construction d'une bibliothèque de référence basée sur les temps de rétention, réalisée à partir de diGAGs commerciaux marqués à l'AMAC et analysés par xCGE-LIF. Nous montrons que les isomères diGAGs peuvent être détectés avec une grande fiabilité et sensibilité. Afin d'analyser les GAGs dans des extraits biologiques, nous utilisons des lyases spécifiques afin de générer des diGAGs, marqués à l'AMAC puis analysés en comparaison des différents standards. Nous appliquons ensuite cet étalonnage à l'étude des variations de la composition en GAGs au cours de la différenciation cellulaire. Les résultats préliminaires obtenus avec des hESCs indiquent que la composition en diGAGs est altérée. Il s'en dégage que notre protocole d'analyse s'avère particulièrement intéressant comme outil d'analyse des GAGs.

Le cluster des gènes de capsule d'*E. coli* K5 est composé de *kfiA-D*. Alors que KfiAC sont des synthases confirmées, le rôle de KfiB est méconnu. A partir d'une analyse des banques de données, nous identifions une série d'isoformes de KfiB qui diffèrent au niveau d'une séquence répétée, ce qui constitue la base du **chapitre 4**. Afin d'explorer l'importance de ce motif, nous entreprenons la génération des isoformes contenant un, deux, et quatre motifs (KfiB_{1r}, KfiB_{2r} et KfiB_{4r}). Les gènes de synthèse sont exprimés dans *E. coli* BL21-*kfiCAD* et nous avons évalué leur impact sur la synthèse et la distribution de l'héparosan. Bien qu'une tendance vers des chaînes plus longues était visible si KfiB_{1r} et KfiB_{4r} étaient co-exprimés, l'impact de KfiB n'a pu être déterminé que partiellement.

Table of Contents

ACKNOWLEDGEMENTS	IV
ABSTRACT.....	VI
RESUME	VIII
LIST OF FIGURES	XIII
ABBREVIATIONS	XV
CHAPTER 1 – GENERAL INTRODUCTION	1
1.1. Bacteriophages.....	1
1.2. Tailspike protein	1
1.3. Capsular polysaccharides	2
1.4. Glycosaminoglycans.....	3
1.4.1. Chondroitin sulfate/Dermatan sulfate	5
1.4.2. Keratan sulfate	6
1.4.3. Heparin/Heparan sulfate	6
1.4.4. Heparosan	8
1.4.5. Hyaluronic acid	10
1.5. Techniques used and challenges associated with the analysis of glycosaminoglycans.....	10
1.6. Research aims	12
CHAPTER 2 – EXTENDING THE ENZYMATIC TOOLBOX FOR HEPAROSAN POLYMERIZATION, DEPOLYMERIZATION, AND DETECTION	13
Preface	14
CHAPITRE 2 – DEVELOPPEMENT D’OUTILS POUR LA POLYMERISATION, LA DEPOLYMERISATION ET LA DETECTION DE L’HEPAROSAN	15
Préface	16
2.1. Introduction	17
2.2. Results and discussion	19
2.2.1. Generation of fluorescently labeled heparosan oligomer	19
2.2.2. Investigation of the elongation mechanism of untagged KfiA and KfiC	21

2.2.3.	Evaluating tagged and purified KfiA and KfiC constructs for the build-up of heparosan	22
2.2.4.	Expression and purification of the Φ K5B tailspike protein KflB.....	25
2.2.5.	KflB is a heparosan lyase.....	27
2.2.6.	Analyzing the degradation mode of KflB using 2AB-K5 heparosan as substrate 29	
2.2.7.	Characterization of KflB substrate specificity	32
2.2.8.	Identification of KflB's active center and development of a heparosan detection agent	34
2.2.9.	Development of a heparosan detection agent	35
2.3.	Conclusions	38
2.4.	Materials and methods.....	39
2.4.1.	Materials	39
2.4.2.	HPLC-based anion-exchange chromatography (HPLC-AEC)	39
2.4.3.	Cloning.....	39
2.4.4.	Heparosan preparation	41
2.4.6.	Protein Expression and Purification.....	42
2.4.7.	Elongation of the K5 oligosaccharide acceptor – fluorescent heparosan	43
2.4.8.	K5 lyase B activity assays.....	44
2.4.9.	Labeling of an inactive KflB.....	45
2.4.10.	Dot blots	45
2.5.	Footnotes	46
	SUPPLEMENTARY INFORMATION	47

**CHAPTER 3 – ANALYSIS OF GLYCOSAMINOGLYCAN-DERIVATIVES
FROM MULTIPOTENT PANCREATIC PROGENITOR CELLS BY
XCGE-LIF**

	Preface	67
--	---------------	----

**CHAPITRE 3 – ANALYSE DES DÉRIVÉS DE GLYCOSAMINOGLYCANES
DE CELLULES PROGÉNITRICES PANCRÉATIQUES MULTIPOTENTES
PAR LA TECHNIQUE DE XCGE-LIF**

	Préface	69
3.1.	Introduction	70
3.2.	Results and discussion	72
3.2.1.	Standards library development	72
3.2.2.	Analysis of diGAGs from purified GAGs	75
3.2.3.	Characterization of diGAGs derived from hiPSCs.....	77
3.2.4.	Determining changes in diGAGs during cell differentiation	79
3.3.	Conclusions	80
3.4.	Materials and methods.....	81

3.4.1.	hPSC culture and differentiation	81
3.4.2.	Extraction and enzymatic digestion of GAGs	81
3.4.3.	Fluorescent labeling and HILIC-solid phase extraction (SPE).....	82
3.4.4.	xCGE-LIF	82
3.4.5.	Data processing and statistics	83
3.5.	Footnotes	83
SUPPLEMENTARY INFORMATION		84
CHAPTER 4 – EFFECT OF KFiB ON HEPAROSAN SYNTHESIS IN RECOMBINANT <i>ESCHERICHIA COLI</i> BL-21		89
Preface		90
CHAPITRE 4 – EFFET DE KFiB SUR LA SYNTHÈSE D’HEPAROSAN RECOMBINANT DANS <i>ESCHERICHIA COLI</i> BL-21		91
4.1.	Introduction	93
4.2.	Results and discussion	95
4.2.1.	Design of recombinant strains carrying different isoforms of KfiB	95
4.2.2.	In vitro assays with heparosan synthases and KfiB isoforms	97
4.2.3.	Influence of KfiB isoforms in vivo	97
4.2.4.	Influence of Influence of KfiB isoforms on the size of Heparosan	101
4.3.	Conclusions	102
4.4.	Materials and methods	103
4.4.1.	Bacterial strains, plasmids, application	103
4.4.2.	<i>In vitro</i> assay	103
4.4.3.	Fed-batch culture	104
4.4.4.	Heparosan quantification	104
4.4.5.	Heparosan extraction.....	104
4.4.6.	Heparosan purification	105
4.5.	Footnotes	105
CHAPTER 5 – GENERAL DISCUSSION.....		106
5.1.	Extending the enzymatic toolbox for heparosan polymerization, depolymerization, and detection	106
5.2.	Analysis of glycosaminoglycan-derivatives from multipotent pancreatic progenitor cells by xCGE-LIF	109
5.3.	Effect of KfiB on heparosan synthesis in recombinant <i>Escherichia coli</i> BL-21	110
REFERENCES.....		113
CURRICULUM VITAE.....		135

List of Figures

Chapter 1 – General Introduction

Figure 1. Bacteriophages infecting bacteria.....	2
Figure 2. Structures of glycosaminoglycans.	4
Figure 3. Heparosan synthesis and transport.....	9

Chapter 2 – Extending the enzymatic toolbox for heparosan polymerization, depolymerization, and detection

Figure 4. Generation of a fluorescently labeled heparosan nonamer.....	20
Figure 5. Elongation of the 2AB labeled K5 nonamer.....	23
Figure 6. Recombinant expression of KflB.....	26
Figure 7. Enzymatic activity of KflB.....	28
Figure 8. The minimal substrate for KflB.....	29
Figure 9. Depolymerization of fluorescently labeled heparosan by KflB.....	32
Figure 10. KflB substrate specificity..	33
Figure 11. Identification of active site residues in KflB.	35
Figure 12. Engineering an iKflB construct for the detection of heparosan.....	37
Figure 13. KflB as detection agent for heparosan on bacterial surfaces.....	38

Chapter 3 – Analysis of glycosaminoglycan-derivatives from multipotent pancreatic progenitor cells by xCGE-LIF

Figure 14. Labeling of diGAGs with AMAC and pic-BH3 complex on the example of the CS-derived disaccharide.	72
Figure 15. xCGE-LIF analysis of diGAGs-AMAC.....	74
Figure 16. Reproducibility of xCGE-LIF migration times.	75

Figure 17. xCGE-LIF analysis of AMAC-labeled products of GAG digestions and peaks’ assignment using our established database.	77
Figure 18. Workflow of AMAC-labeled diGAGs preparation from the cell lysate.	78
Figure 19. xCGE-LIF analysis of diGAG-AMAC from hiPSCs.	79
Figure 20. Comparison of GAG composition of hESCs and hESC-derived MPPCs by xCGE-LIF.	80
Chapter 3 – Effect of KfiB on heparosan synthesis in recombinant <i>Escherichia coli</i> BL-21	
Figure 21. Analysis of the predicted KfiB structure.	95
Figure 22. Schematic representation of the generated vectors.	96
Figure 23. Elongation of K5 nonamer-2AB catalyzed by KfiCA/and KfiB and its variants...	97
Figure 24. Heparosan production in <i>E. coli</i> BL21-Gold (DE3).	98
Figure 25. Heparosan production in recombinant <i>E. coli</i> BJ..	99
Figure 26. Detection of heparosan produced by engineered strains	100
Figure 27. Comparison of heparosan molecular weight produced by engineered strains.	101
Figure 28. SEC-MALS analysis of heparosan.	102

Abbreviations

Φ	bacteriophage, phage
Δ	unsaturated uronic acid at the non-reducing end
λ	wavelength
\mathfrak{D}	dispersity
A	alanine
2AB	2-aminobenzamide
AcOH	acetic acid
AEC	anion-exchange chromatography
AMAC	2-aminoacridone
AMAC-diGAGs	AMAC-labeled glycosaminoglycan-derived disaccharides
Asp	aspartic acid
Bis-Tris	bis-tris methane
caHep	cell-associated heparosan
cmR	chloramphenicol resistance
CPS	capsular polysaccharide
CS	chondroitin sulfate
CTD	C-terminal chaperone domain
diGAGs	glycosaminoglycan-derived disaccharides
DNA	deoxyribonucleic acid
DP	degree of polymerization
DMSO	dimethyl sulfoxide
DS	dermatan sulfate
DTT	dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
E	glutamic acid
endoN	endosialidase, endo-N-acetylneuraminidase
eGFP	enhanced green fluorescent protein
EXT	extosin
F	phenylalanine
GAG	glycosaminoglycan
Gal	galactose

GlcA	glucuronic acid
GlcNAc	<i>N</i>-acetylglucosamine
GalNAc	<i>N</i>-acetylgalactosamine
GSL	glycosphingolipid
HA	hyaluronic acid
HAS	hyaluronan synthases
hESCs	human embryonic stem cells
HILIC	hydrophilic interaction liquid chromatography
hiPSCs	human induced pluripotent stem cells
hPSCs	human pluripotent stem cells
HPSEC	high-performance size-exclusion chromatography
HPLC	high-performance liquid chromatography
HS	heparan sulfate
IdoA	iduronic acid
IPTG	isopropyl β-D-1-thiogalactopyranoside
K	lysine
kanR	kanamycin resistance
kDa	kilo Dalton
KfiA	<i>N</i>-acetylglucosaminyltransferase
KfiC	D-glucuronyltransferase
KfiD	UDP-Glc-dehydrogenase
KfiA	K5 lyase A
KfiB	K5 lyase B
KS	keratan sulfate
LC/ESI	liquid chromatography/electrospray ionization
M	mol per liter
MALS	multi-angle light scattering
Man	mannose
MBP	maltose-binding protein
MPPCs	multipotent pancreatic progenitor cells
MPS	mucopolysaccharidoses
MS	mass spectrometry
MS/MS	tandem MS
MTUs	migration time units
MW	molecular weight

MWCO	molecular weight cut-off
NA	highly sulfated domain of heparan sulfate
NA/NS	intermediately sulfated domain of heparan sulfate
NS	nonsulfated domain of heparan sulfate
nRFU	normalized RFU
OD₆₀₀	optical density measured at a $\lambda = 600$ nm
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PGs	proteoglycans
pic-BH₃	2- picoline borane complex
RFU	relative fluorescence units
SDS	sodium dodecyl sulfate
SEC	size-exclusion chromatography
Ser	serine
sHep	supernatant heparosan
tetR	tetracycline resistance
TF	trigger factor
TRs	tandem repeats
TSP	tailspike protein
Thr	threonine
Tris	tris(hydroxymethyl)aminomethane
Tween[®] 20	polyethylene glycol sorbitan monolaurate
UDP	uridine diphosphate
UDP-GlcA	uridine diphosphate glucuronic acid
UDP-GlcNAc	uridine diphosphate N-acetylglucosamine
UV	ultraviolet
wt	wildtype
Xyl	xylose
xCGE-LIF	multiplexed capillary gel electrophoresis with laser-induced fluorescence detection
Y	tyrosin

Chapter 1 – General introduction

1.1. Bacteriophages

Life on the Earth probably started with undifferentiated prokaryotes that were divided into archaea, eubacteria, and eukaryotic ancestors (Ackermann, 1998). Bacteriophages, in short, 'phages' (Φ), are viruses that infect bacteria. Their rise may have occurred 3.5 to 3.7 billion years ago, simultaneously with bacteria, making them the first existing viral group (Ackerman, 1998). Nowadays, billions of years later, with an estimated number of 10^{31} , bacteriophages represent the most abundant entities, colonizing the most diverse habitats on Earth. In the early 20th century, Frederick William Twort and Felix Hubert d'Hérelle independently discovered 'ultra-microscopic viruses' as 'invisible, antagonistic microbe of the dysentery bacillus' (Twort, 1915; d'Hérelle, 1917). In 1922, F. H. d'Hérelle (d'Hérelle, 1922) coined the term 'bacteriophages', and the first image of a bacteriophage was published in 1940, after the invention of the electron microscope (Ackermann, 2011).

After infection of bacteria, phages can replicate in the lytic cycle leading to the lysis of the host cell or, in the so-called lysogenic cycle, integrate into the host genome and replicate as part of the host genome (Ackermann, 1998).

1.2. Tailspike protein

Bacteriophages are the most common of all virus groups (Ackermann, 1998). The order of *Caudovirales*, referred to as tailed phages, accounts for 96% of all known bacteriophages. Structurally, tailed phages comprise the head (capsid) of an icosahedron-like structure and the tail of fixed length and width. These morphological features are easily detectable by electron microscopy (Ackermann, 1998).

In the first stage of the infection process, this is adsorption; a phage binds to specific bacterial surface receptors (Samson et al., 2013). Bacterial cell receptors involved in adsorption are often carbohydrates or surface-exposed proteins (Samson et al., 2013). In the evolutionary arms race, some bacteria developed protective polysaccharide capsules that sterically mask their receptors, thus preventing adsorption (Hampton et al., 2020; Knecht et al., 2020). For example, encapsulated *E. coli* K5 is resistant to T4 phage because the lipopolysaccharide of the bacterial cell

surface is concealed by the capsular polysaccharide (Soundararajan et al., 2019). To overcome this barrier, some phages, especially members of the *Podoviridae* family of the order *Caudovirales*, developed tailspike proteins (TSPs) equipped with enzymatic functions (Knecht, 2020). For instance, TSPs can be highly specific depolymerases that bind and break down the capsule polysaccharides of the host to bring phages in close contact with the bacterial outer membrane (**Fig. 1B**) (Hughes et al., 1998). A well-studied example to illustrate this function is Φ K1F, the bacteriophage infecting *E. coli* K1F (Mühlenhoff et al., 2003). The TSPs of Φ K1F comprise an endosialidase (endo-*N*-acetylneuraminidase, endoN) highly specific for binding and degradation of the *E. coli* K1 capsular polysaccharide named polysialic acid and consisting of α -2,8-linked sialic acid residues (Mühlenhoff et al., 2003). A second example is bacteriophage K5A (Φ K5A) infecting *E. coli* K5. The capsule surrounding *E. coli* K5 consists of the polysaccharide heparosan, a glycosaminoglycan also found in humans (see the paragraph below). Φ K5A is equipped with TSPs comprising the K5 lyase A. This enzyme degrades heparosan, allowing the virus to penetrate the capsule and initiate the infection process (**Fig. 1B**) (Thompson et al., 2010).

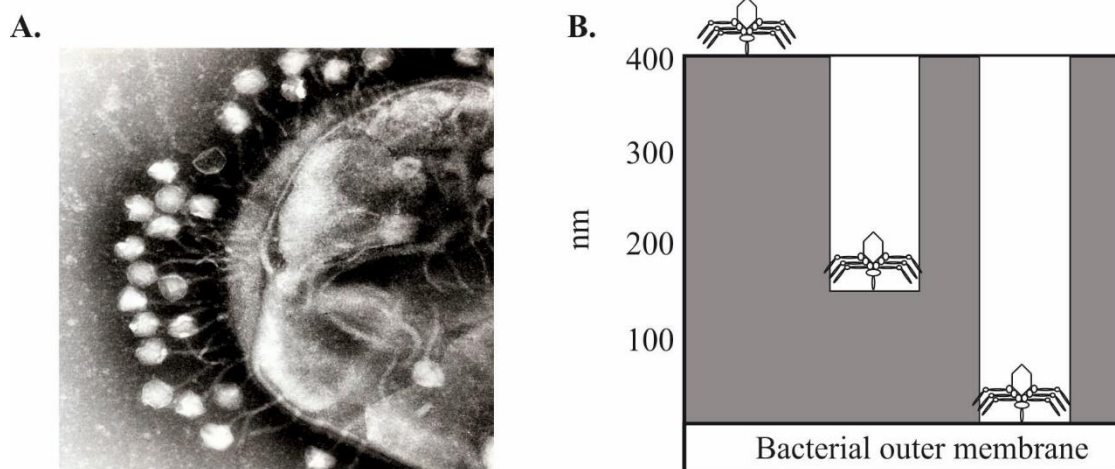


Figure 1. Bacteriophages infecting bacteria. **A.** Transmission electron micrograph image of bacteriophages attached to a bacterial cell wall (in approximately 200,000 magnification). The image was obtained by Professor Graham Beards and is presented without any changes (Link to the license: <https://creativecommons.org/licenses/by-sa/3.0/deed.en>). **B.** The scheme of the capsular polysaccharide degradation (grey) by podoviruses (adapted from Hughes et al., 1998).

1.3. Capsular polysaccharides

The survival of microbes is under the constant threat due to the complexity and variability of many environments (Rittershaus et al., 2013). Developing survival

strategies is the key to evolutionary success (Rittershaus et al., 2013). Many bacteria produce thick layers of capsular polysaccharides (CPSs) that stick firmly to the cell surface and extend up to 400 nm. CPSs, also known as the K antigens (Whitfield, 2006), protect bacteria against environmental threats and host immune defense mechanisms. Moreover, because CPSs often mimic host structures (see *E. coli* K5), CPSs impact host-microbe interactions and support immune escape (Buffet et al., 2020; Cress et al., 2014). The reported nearly 80 serologically distinct CPSs in *E. coli* show this great variety (Whitfield, 2006; K. Jann & B. Jann, 1983). *E. coli* capsules are classified into four groups based on serological, genetic, and biochemical properties (Whitfield, 2006). Although less extensively studied, CPSs can also be present in Gram-positive bacteria, e.g., *Staphylococcus aureus* (Willis et al., 2013). Next to their role as virulence factors (DeAngelis, 2002), CPSs offer great potential as basis material for the production of medical implants (Guzzo et al., 2020) and in the development of glycoconjugate vaccines (Fiebig et al., 2014; Sokaribo et al., 2021).

1.4. Glycosaminoglycans

Glycosaminoglycans (GAGs) are negatively charged polysaccharides composed of repeating disaccharide units. GAGs are produced in every mammalian tissue, and their functions are manifold, reaching from cell signaling to structural scaffolding (L. Schaefer & R. Schaefer, 2010; Mikami & Kitagawa, 2016; Nikitovic & Pérez, 2021). Based on the structure of the repeating disaccharides, GAGs are subdivided into chondroitin sulfate (CS)/dermatan sulfate (DS), heparin/heparan sulfate (HS)/heparosan, keratan sulfate (KS), and hyaluronic acid (HA), whereby differences between those separated by backslash concern post-synthetic enzymatic modifications like variant decoration with sulfate groups and variant degrees of epimerization of D-glucuronic acid (GlcA) to L-iduronic acid (IdoA) (**Fig. 2A**). GAGs, except HA, are present as side chains of proteoglycans (PGs) described in the next paragraph.

Many tissues are rich in PGs, a specialized form of heavily glycosylated proteins. The core proteins in PGs carry GAGs covalently attached to serine or in the case of KS, also threonine residues (**Fig. 2B**). Although there is a limited number of core proteins, the enormous combinatorial freedom inherent in attached GAGs (types, chain length, a pattern of modifications) and a far from being understood interactome with soluble proteins (Couchman & Pataki, 2012; Schaefer L. & Schaefer R., 2010). I will discuss some examples concerning the GAGs further.

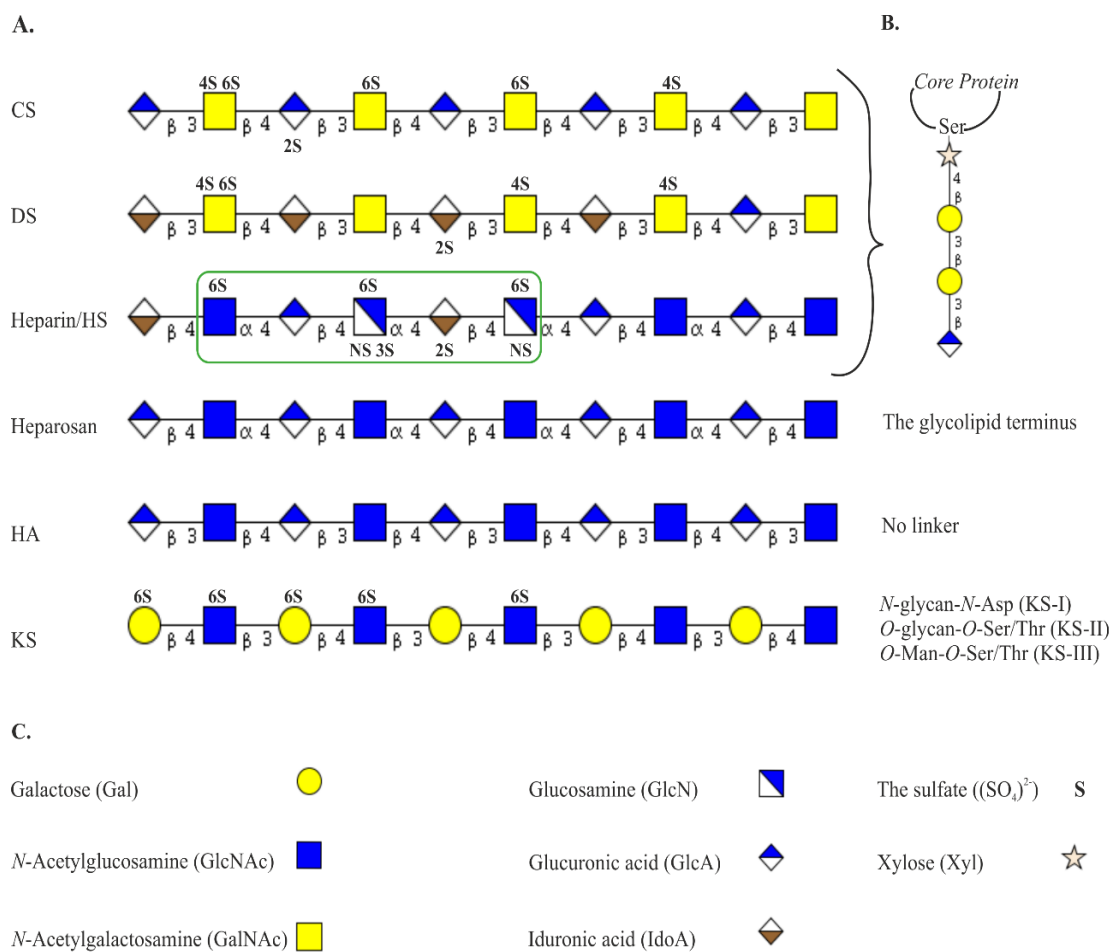


Figure 2. Structures of glycosaminoglycans. **A.** Structure of CS, DS, heparin/HS, heparosan, HA, KS; CS, DS, HS, heparin, and KS are decorated with sulfate groups on hexuronic acid (CS, DS, heparin, HS: 2S) and hexosamine (CS, DS, heparin, HS: NS, 4S, 6S; KS: 6S), while heparosan and HA are unsulfated. DS, HS, and heparin are modified by partial epimerization of GlcA to IdoA. The structure of the heparin antithrombin-binding pentasaccharide is highlighted in green **B.** GAGs' corresponding linkers. **C.** Description of used building blocks. GAGs structures were drawn with the GlycanBuilder software (Damerell et al., 2012).

Except for HA, which is produced on the inner side of the cell membrane (Prehm, 1984), the complex GAGs are synthesized in the Golgi apparatus, whereby enzymes act in a template-free process. Regimens controlling the molecular composition of GAGs are barely understood. However, the entire process of chain elongation and subsequent modifications is very fast, estimated to only last 1 minute (Carlsson & Kjellén, 2011). The ubiquitous expression of GAGs and, not least, their essential function in the shaping of the extracellular matrix of all animal cells explains that disturbances in the biosynthesis and degradation of GAGs cause a large variety of pathological processes. For instance, mucopolysaccharidoses (MPS), a genetically

inherited group of metabolic disorders, are caused by defects in GAG-degrading enzymes. Seven MPS subtypes have been identified that result in varying severity of clinical symptoms (Fenzi et al., 2015).

1.4.1. Chondroitin sulfate/Dermatan sulfate

Chondroitin sulfate and DS are structurally similar, and their biosynthesis proceeds in a similar pathway. Consequently, for years DS was classified as a subtype of CS.

The biosynthesis of CS and DS begins with the stepwise build-up of a tetrasaccharide (-GlcA- β 3-Gal- β 3-Gal- β 4-Xyl- β -) on serine present in the core protein. Upon the linker formation, the first GlcNAc residue is added to the non-reducing end by GalNAc transferase I. This is the critical step determining the synthesis of CS/DS or heparin/HS. (Carlsson & Kjellén, 2011; Mikami & Kitagawa, 2013). Next, the CS/DS backbone is synthesized by the joint action of GlcA transferase II and GalNAc transferase II and further partially modified by sulfation at the C-2 position of GlcA and the C-4, C-6 positions of GalNAc. Depending on the sulfation pattern, the heterogeneous disaccharides of CS chains were assigned into five subtypes O, A, C, D, and E (Mikami & Kitagawa, 2013). The DS-epimerases, Dse and DseI, catalyze multiple epimerization reactions of GlcA to IdoA. These modifications are of great importance for functionality, and the synthesized chains are considered a new molecule - DS. The isomeric equivalents of CS-derived disaccharides found in DS were classified as iO, iA, and iE. Besides the predominance of the characteristic epimeric state of hexuronic acid, the two GAGs differ in the presence of unique disaccharides; GlcA-GalNAc6S (CS-C), GlcA2S-GalNAc6S (CS-D) and IdoA2S-GalNAc4S (DS-iB), in CS and DS, respectively (Mikami & Kitagawa). Chondroitin sulfate is especially abundant in cartilage, while DS is in the skin, blood vessels, and aorta (Mizumoto & Yamada, 2021). The mechanism of CS degradation remains unclear (Mikami & Kitagawa, 2013). Presumably, it occurs in lysosomes by endo-type hydrolases leading to chain fragmentation. The produced oligomers are further degraded by exolytic glycosidases and sulfatases (Mikami & Kitagawa, 2013).

Chondroitin sulfate functions are associated with the sulfation pattern (Mikami & Kitagawa, 2013). Several pathogens have been reported that are exploiting cell surface CS chains in the infection process. For instance, after infection of erythrocytes,

malaria parasites (*Plasmodium falciparum*) require low-sulfated CS-A chains for the adhesion to endothelial cells (Mikami & Kitagawa, 2013). Another interesting example is the *Herpes simplex* virus utilizing sulfate-rich CS chains (CS-E) (Mikami & Kitagawa, 2013). CS-E chains are essential in bone development, and their deficiency contributes to osteoporosis. A well-known function of DS is the formation of the extracellular matrix and participation in cell signaling (Mizumoto & Yamada, 2022). Moreover, recent studies in mice deficient in DS-epimerases suggest a significant role of DS in embryonic development and perinatal survival (Stachtea et al., 2015).

1.4.2. Keratan sulfate

Keratan sulfate from a cornea extract was first described in 1939 (Suzuki, 1939). KS is classified into different forms depending on the type of glycosidic linkage to the protein. The initially described corneal KS (KS-I) is attached to *N*-linked glycans. KS-II is found on *O*-linked glycans of the mucin-type (*O*-GalNAc to threonine or serine). In addition, one report describes KS to occur on *O*-mannose-linked glycans as well (Krusius et al., 1986). KS biosynthesis proceeds by extension of an existing *N*-acetyllactosamine (Gal-GlcNAc) backbone, appearing in both *N*- and *O*-glycans, in the alternating action of glycosyltransferases adding GlcNAc and Gal (Funderburgh, 2002). Sulfation can occur at the C-6 position of both GlcNAc and Gal (Caterson & Melrose, 2018).

KS is widely spread in the cartilage, meniscus, and cornea, where it is involved in tissue hydration, weight-bearing compensation, and cell signaling processes (Caterson & Melrose, 2018). Zhang *et al.* reported the participation of KS in the glial scar formation upon injury in the central nervous system (H. Zhang et al., 2006). Abnormal KS expression has been linked to Alzheimer's disease (Lindahl et al., 1996). Recent studies have demonstrated the potential of KS for use as a biomarker in pancreatic cancer progression and metastasis (Leiphrakpam et al., 2019).

1.4.3. Heparin/Heparan sulfate

Heparin and heparan sulfate are two distinct GAGs, but a remarkable similarity allows presenting them together. Heparin and HS share the same backbone composed of the repeating units of (-GlcA- β 1,4-GlcNAc- α 1,4) partially modified by decoration with sulfate groups, deacetylation, and epimerization of GlcA to IdoA. However, the

level of sulfation distinguishes these two very similar polymers. The sulfate modifications to HS are performed in clusters resulting in domains of varying degrees of sulfation within one chain: highly sulfated, intermediately sulfated, and nonsulfated domain. In heparin, the entire polymer chain is highly sulfated (Carlsson & Kjellén, 2011). Due to the high degree of sulfation, heparin is believed to be the biomolecule with the highest negative charge density. Another difference between HS and heparin is the size of the polymers. Newly synthesized heparin chains are in the range of molecular weight (MW) = 60-100 kDa, depending on the source, while HS chains are shorter with MW = 22-45 kDa. It is noteworthy that commercial-grade heparins are processed and significantly shorter (7-25 kDa) (Carlsson & Kjellén, 2011).

The assembly of heparin and HS chains occurs in the Golgi apparatus and starts with the same GAG-linker as for CS/DS. The subsequent addition of GlcNAc initiates the synthesis of heparin/HS. Although the synthesis of both polymers follows the same pathway, heparin is produced exclusively by mast cells (Carlsson & Kjellén, 2011) and basophils (Khurana, 2008), while HS production takes place in all animal tissues. The heparin/HS chains are assembled by Exostosin-1 (EXT1) and EXT2 polymerases with UDP-sugars as substrates (Carlsson & Kjellén, 2011). The *in vitro* studies have shown that EXT1 and EXT2 have a dual enzyme activity – GlcA transferase II and GlcNAc transferase II. However, the EXT2 activity is low and seems to have a chaperone-like role for EXT1 in the HS/heparin synthesis (Busse & Kusche-Gullberg, 2003). The HS/heparin chain modifications occur simultaneous with the elongation process (Carlsson & Kjellén, 2011).

Heparin is well known for its anticoagulant activity and has been used in medicine for almost a century (Jin et al., 1997; Keire et al., 2015). Recent studies suggest a beneficial role of heparin in the survival of cancer patients. However, the underlying mechanism remains unknown (Ma et al., 2020). HS has a tremendous number of functions in mammals (Rodgers et al., 2018). Intensively studied are its roles in neural development (Kamimura & Maeda, 2021) and thymus growth (Hsu et al., 2020). HS can interact with more than 300 secreted proteins, such as chemokines and growth factors (Xu & Esko, 2014). HS modulates skin repair, and a low molecular weight heparan sulfate was proven to have therapeutic potential on mature and damaged skin (Bucay et al., 2020). Alterations in HS synthesis have been linked to

several disorders, including Alzheimer's disease (Pérez-López et al., 2021) and Autism spectrum disorders (Risi et al., 2021).

1.4.4. Heparosan

Heparosan is the least investigated molecule within the family of GAGs. Like the backbone of heparin/HS, heparosan consists of repeating disaccharides (-GlcA- β 1,4-GlcNAc- α 1,4-), and its biosynthesis in mammals is briefly described in subchapter 1.4.3. In bacteria such as *E. coli* K5 and *E. Nissle* 1917, heparosan exists as a CPS that serves as a protective barrier against the environment. Moreover, since heparosan is ubiquitous in the animal body, the capsule drastically reduces the immunogenicity of the bacteria (DeAngelis, 2015). Secondly, as explained in an earlier paragraph, the stereochemical properties of the capsule protect the bacterium against phage infection (Soundararajan et al., 2019).

Heparosan biosynthesis in bacteria is illustrated in **Fig. 3**. The process begins with a glycolipid acceptor consisting of several 3-deoxy-D-manno-oct-2-ulosonic acid moieties, phosphatidylglycerol, and mono- or diacyl fatty acids (Yan et al. 2020). Next, chain elongation is catalyzed at the inner membrane by the coordinated action of KfiA (*N*-acetylglucosaminyltransferase) (Hodson et al., 2000; M. Chen et al., 2006) and KfiC (D-glucuronyltransferase) (Sugiura et al., 2010). The genes expressing these enzymes are located in the serotype-specific region 2 along with *kfiD* (KfiD, UDP-Glc-dehydrogenase) and *kfiB* (Hodson et al., 2000). So far, the function of KfiB is not fully understood, but existing data suggest that the gene product is important for the stability of heparosan synthases *in vivo* (Leroux & Priem, 2016). Region 2 (**Fig. 3A**), which harbors the heparosan synthase complex, is part of the capsule gene cluster comprising three regions present in all *E. coli* strains expressing group 2 capsules (Whitfield, 2006). Regions 1 and 3 contain genes involved in synthesizing the glycolipid acceptor (Willis et al., 2013; Whitfield, 2006) and proteins involved in the extracellular transport of polysaccharides (Silver et al., 2001; Whitfield, 2006). Both regions are conserved in *E. coli* strains expressing group 2 capsules (Silver et al., 2001).

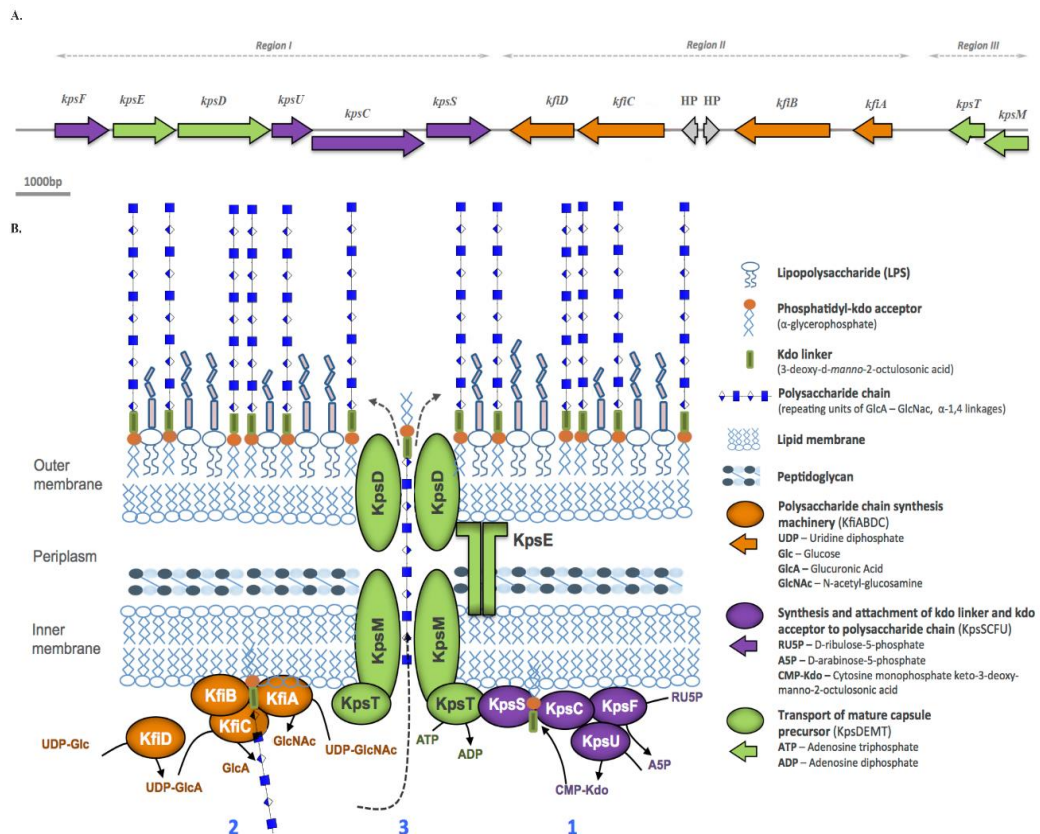


Figure 3. Heparosan synthesis and transport. A. *E. coli* Nissle 1917 capsular polysaccharide gene cluster consisting of three regions encoding for KpsFEDUCS, KfiABCD, and KpsMT proteins. B. Overview of K5 polysaccharide biosynthesis and exportation. The numbers represent the order of events (adapted from Nzakizwanayo *et al.*, 2015). Symbols, as well as colors, are explained in the added legend.

The so-called 'heparin crisis' that occurred in 2008, when the administration of contaminated heparin to patients resulted in approximately 350 adverse events and 150 deaths in the United States (Keire *et al.*, 2015), generated new momentum for the production of heparin from heparosan and thus avoid animal sources. Moreover, chemoenzymatic synthesis starting from heparosan would overcome concerns with respect to heparin shortages (Keire *et al.*, 2015). Further arguments in favor of an increased interest in heparosan come from the proven low immunogenicity, stability in extracellular spaces of mammalian tissues, its biodegradability in lysosomes, and the fact, that the blood-brain barrier is impermeable for heparosan (DeAngelis, 2015). These biological properties make heparosan a promising candidate as a drug delivery vehicle (DeAngelis, 2015; Rippe *et al.*, 2019; Mao *et al.*, 2021; X.Yang *et al.*, 2021).

1.4.5. Hyaluronic acid

Hyaluronic acid is widespread in all vertebrates. Exceptionally high concentrations are found in the umbilical cord and vitreous body, a transparent gel in the eye. In addition, some bacteria produce HA as a capsular polysaccharide. One example is the mucoid strain of group A *Streptococcus*. This strain causes acute rheumatic fever and HA is a proven virulence factor (Wessels et al., 1994).

In vertebrates, HA is produced by unique plasma membrane glycosyltransferases secreting this glycosaminoglycan directly to the extracellular space. Hyaluronan synthases (HAS) are three isoenzymes (HAS1, HAS2, and HAS3) of different enzymatic properties (Itano et al., 1999), generating well-defined HA chains of varying MW (Itano et al., 1999). Recent studies show an association between low molecular weight HA and aggressive solid tumors (Abatangelo et al., 2020). HA degradation in humans occurs rapidly, and its half-life depends on its location (Fraser, 2003). In the bloodstream, HA is catabolized within a few minutes in lymph nodes and the liver. The degradation of the HA in the skin occurs within 12-24 hours with the assistance of hyaluronidases, superoxide, nitric oxide, or lysosomes (Abatangelo et al., 2020). Interestingly, the HA synthesis and degradation relationship impact its properties, as differing molecular weights can evoke contradictory actions (Abatangelo et al., 2020).

Hyaluronic acid plays a key role during embryonic development, wound healing, and cancer progression. Moreover, due to its biocompatibility (Abatangelo et al., 2020), HA is a fascinating molecule for medical applications. Indeed, HA has been used in manufacturing wound healing devices (Winstein-Oppenheim et al., 2017), tear supplements (Hynnekleiv et al. 2022), and other medicinal products.

1.5. Techniques used and challenges associated with the analysis of glycosaminoglycans

The analysis of GAGs is highly challenging because of size and zonal variation concerning modifications and because GAGs exhibit significant structural and chemical similarities. However, easy to understand based on the above-described GAGs functions, there is an urgent need for efficient analytical procedures. Moreover, patients affected by MPS would greatly benefit from continuous GAGs analysis for diagnostic purposes (Khan et al., 2020). Ideally, a developed method would allow for

quick, sensitive analysis at a low cost and without the high-end infrastructure of specialized laboratories.

Several methods for the detection and quantification of GAGs have been reported. The initial approaches were based on electrophoretic separation, allowing MW comparison and subsequent staining with dyes like Alcian Blue (Khan et al., 2020), which is positively charged and therefore can interact with negatively charged GAGs (Scott et al., 1964). The drawback of this method is that different GAGs cannot be distinguished (Khan et al., 2020). High-performance liquid chromatography (HPLC) was used alone and in combination with other methods. Studelska *et al.* employed reversed-phase (RP) HPLC to quantify GAGs containing glucosamine and galactosamine (Studelska et al., 2016). In this method, the polysaccharides underwent acidic hydrolysis, generating single monomers (glucosamine and galactosamine), which were then reduced and fluorescently labeled. The synthesized derivatives were separated by RP-HPLC and quantified. The developed method is very sensitive and easily applicable in laboratories with RP-HPLC equipment. However, the method is not suitable for GAGs differentiation. Volpi established a protocol for the analysis of CS/HA disaccharide derivatives with 2-aminoacridone (AMAC) by RP-HPLC electron spray ionization mass spectrometry (Volpi, 2010). The high sensitivity allows this method's use in the GAG analysis from biological samples. Two other groups analyzed AMAC-labeled GAGs disaccharides using capillary electrophoresis with laser-induced fluorescence (LIF) detection (Hitchcock et al., 2008; Chang et al., 2012). Already in 1977, Radhakrishnamurthy *et al.* studied the composition of aorta GAGs by ion-exchange chromatography (Radhakrishnamurthy et al., 1977). Tomatsu *et al.* successfully applied one of the elementary biochemical assays – enzyme-linked immunosorbent assay - for KS and HS detection in blood and urine samples (Tomatsu et al., 2014). Several protocols utilizing mass spectrometry (MS) combined with chromatographic techniques were reported to be successful in GAG analysis from urine (Auray-Blais et al., 2012) and tissue samples (Oguma et al., 2001). MS methods are rapid, sensitive, and highly accurate (Szekeres et al., 2020). They provide information on oligomers' composition, length, and modifications (Pepi & Amster, 2021). However, identifying the location of the modifications requires tandem MS (MS/MS) (Pepi & Amster, 2021). Pepi and Amster developed protocols utilizing electron-activated tandem MS and successfully sequenced heparin/HS and CS/DS

oligomers of DP = 4-10 (degree of polymerization) (Pepi & Amster, 2021). Szekeres *et al.* wrote a comprehensive review of current advances in spectroscopic methods used to analyze GAG-cell interactions (Szekeres *et al.*, 2022). Osago *et al.* developed a method for analyzing GAGs disaccharides by LC/ESI/MS/MS (liquid chromatography/electrospray ionization) upon enzymatic digestion of GAGs polymers from cartilage. The group successfully identified and quantified 23 disaccharides showing GAGs composition in cartilages (Osago *et al.*, 2014).

Tremendous progress has been made in GAG analysis and sequencing. While not all techniques are listed, more can be found in Khan *et al.*, a comprehensive review of advances in GAG detection (Khan *et al.*, 2020). Currently, methods based on LC-MS/MS, despite some limitations, are considered the best and widely used for clinical and research purposes (Khan *et al.*, 2020).

1.6. Research aims

GAGs belong to a relatively small subcategory of carbohydrates. However, their complex, diverse structures, and similar physico-chemical properties make them most challenging molecules to analyze and differentiate. Therefore, I aimed to broaden the toolbox for GAG analysis and detection. The focus was set on heparosan, the CPS of bacteria like *E. coli* K5 and *E. coli* Nissle 1917. The high level of interest existing in heparosan owns to the fact that heparosan is a precursor for the chemoenzymatic synthesis of medically important heparin. Despite of considerable investments made towards understanding the heparosan biosynthesis machinery in bacteria, there are aspects left that are still puzzling. I aimed to address open questions during my doctoral thesis. In addition, an objective of my study was the generation of a reagent for the specific detection of heparosan. Towards this goal, I followed the concept, of designing an artificial lectin starting with an enzyme that specifically recognizes heparosan. Based on experience in the laboratory, we decided for the heparosan specific K5 lyase B. After intensive mechanistic characterization and guided by structural data that exist for the homologueous K5 lyase A, a site specific mutational approach was used to develop a heparosan detection tool.

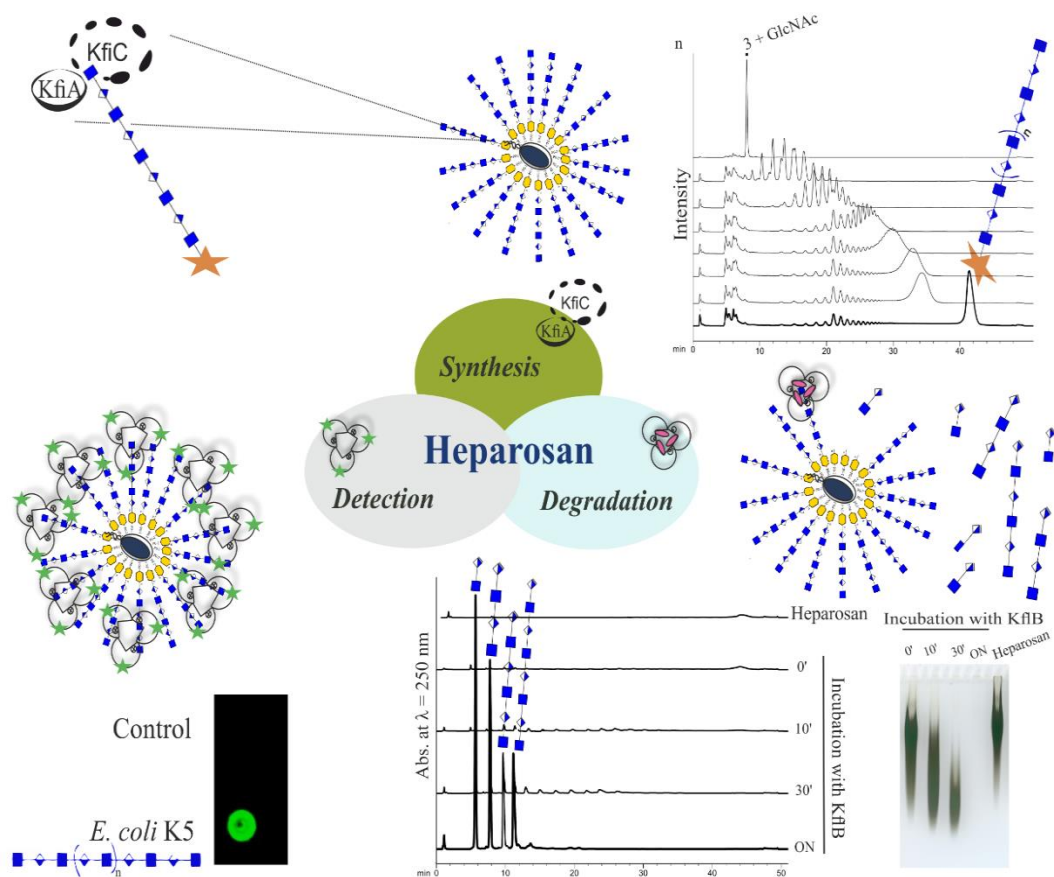
Chapter 2 – Extending the enzymatic toolbox for heparosan polymerization, depolymerization, and detection

Małgorzata Sulewska^{ab}, Monika Berger^a, Manuela Damerow^a, David Schwarzer^a,
Martina Mühlenhoff^a, Rita Gerardy-Schahn^a, Bernard Priem^b, Timm Fiebig^a

^a*Institute of Clinical Biochemistry, Hannover Medical School, Hannover, Germany*

^b*Centre de Recherche sur les Macromolécules Végétales, Groupe Chimie et
Biotechnologie des Oligosaccharides, 601 rue de la Chimie, BP 53X, 38041
Grenoble, Cedex 09, France*

Manuscript in preparation



Preface

The study aims to extend the enzymatic toolbox for heparosan polymerization, depolymerization, and detection and was conceptualized by Dr. Timm Fiebig, Prof. Rita Gerardy-Schahn, Dr. Bernard Priem, and me. Supervised by Dr. Bernard Priem, I isolated heparosan, produced defined oligomers, and fluorescently labeled heparosan oligomers (**Fig. 4**). These oligomers were then used as acceptors to monitor the *in vitro* polymerization reaction catalyzed by heparosan synthases.

The obtained fluorescently labeled heparosan chains as well as unlabeled heparosan, were used as substrates for the K5 lyase B. Dr. Manuela Damerow, supervised by Dr. David Schwarzer and PD Dr. Martina Mühlenhoff, designed and performed the initial experiments, including the cloning of primary constructs, the expression, and the purification methodology regarding the K5 lyase B. I carried out the cloning and purifications partially based on their work and supervised by Dr. Timm Fiebig and Prof. Rita Gerardy-Schahn performed all the experiments with respect to the characterization of the K5 lyase B enzymatic mechanism and its engineering to an artificial lectin, i.e., an enzymatically inactive K5 lyase B mutant that preserved its affinity for heparosan binding (**Fig. 13**). Finally, I evaluated and interpreted all the presented data together with Dr. Timm Fiebig, Prof. Rita Gerardy-Schahn, and Dr. Bernard Priem.

This manuscript was conceptualized and written by me and Dr. Timm Fiebig. Prof. Rita Gerardy-Schahn and Dr. Bernard Priem discussed the data and edited the manuscript.

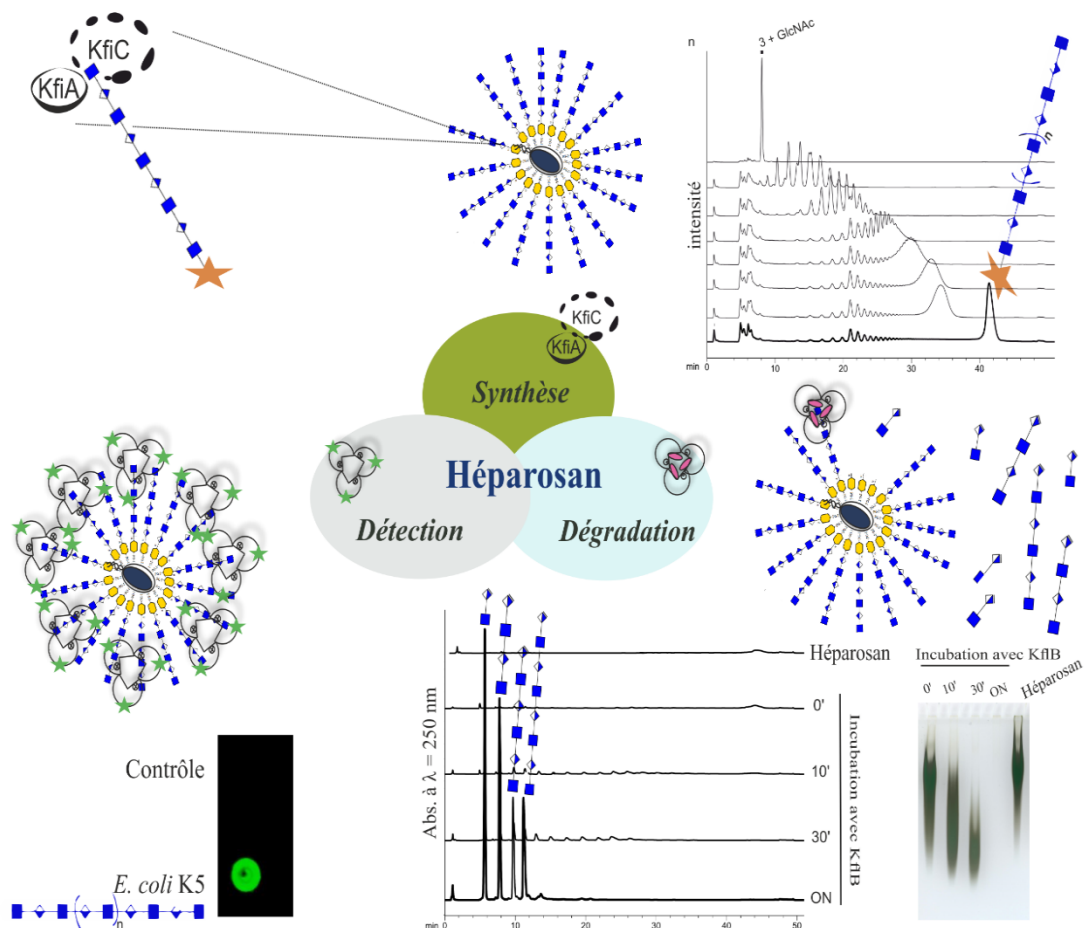
Chapitre 2 – Développement d’outils pour la polymérisation, la dépolymérisation et la détection de l’héparosan

Małgorzata Sulewska^{ab}, Monika Berger^a, Manuela Damerow^a, David Schwarzer^a,
Martina Mühlenhoff^a, Rita Gerardy-Schahn^a, Bernard Priem^b, Timm Fiebig^a

^a*Institute of Clinical Biochemistry, Hannover Medical School, Hannover, Germany*

^b*Centre de Recherche sur les Macromolécules Végétales, Groupe Chimie et
Biotechnologie des Oligosaccharides, 601 rue de la Chimie, BP 53X, 38041
Grenoble, Cedex 09, France*

Manuscrit en préparation



Préface

L'objet de ce travail ambitionne le développement d'outils pour la polymérisation, la dépolymérisation et la détection de l'héparosan, il a été conçu par les Dr Timm Fiebig, Prof Rita Gerardy-Schahn, Dr Bernard Priem, et moi même.

Sous la direction du Dr Bernard Priem, j'ai produit, isolé et marqué des oligosaccharides d'héparosan (**Fig. 4**). Ceux-ci ont été utilisés comme accepteurs afin de suivre l'activité de polymérisation *in vitro* catalysée par les héparosan synthases.

Les chaînes d'héparosan obtenues ainsi que l'héparosan non marqué ont été utilisés comme substrats de la K5 lyase. Dr Manuela Damerow, sous l'encadrement des Dr David Schwarzer et PD Dr Martina Mühlenhoff, a réalisé les premières expériences, à savoir les premières constructions, l'expression et la méthodologie de purification de la K5 lyase B. J'ai ensuite réalisé les clonages et purifications sur la base de ce travail préliminaire sous l'encadrement du Dr Timm Fiebig et de la Prof Rita Gerardy-Schahn, ainsi que les analyses du mécanisme d'action de la K5 lyase B et de son mécanisme d'action, son ingénierie en lectine artificielle, à savoir l'inactivation de ses propriétés catalytiques avec le maintien de ses propriétés d'adhésion à l'héparosan (**Fig. 13**). Finalement, j'ai évalué et interprété tous les résultats présentés ici avec l'aide du Dr Timm Fiebig, Prof Rita Gerardy-Schahn et du Dr Bernard Priem.

Ce manuscrit a été conçu et écrit par moi-même, revu et édité par Dr Timm Fiebig, Prof Rita Gerardy-Schahn et le Dr Bernard Priem.

2.1. Introduction

Heparosan is a natural, acidic polysaccharide belonging to the glycosaminoglycans family. Its dimeric repeating unit is built of β -*N*-acetylglucosamine and α -D-glucuronic acid (DeAngelis, 2015). Heparosan is expressed by pathogenic bacteria like *Escherichia coli* K5 and *Pasteurella multocida* and by non-pathogenic bacteria like *E. coli* Nissle 1917 (Cress et al., 2014). As a natural precursor for synthesizing heparin and heparan sulfate, heparosan is recognized as “self”, which facilitates evasion of the host immune response (Cress et al., 2014). Furthermore, due to its stability in extracellular spaces (Jing et al., 2017), high biocompatibility, and lack of immunogenicity, heparosan can be used to facilitate drug delivery (Lane et al., 2017). Consequently, developing a toolbox for the synthesis and engineering of heparosan has received much attention (X. Zhang et al., 2019a). So far, this toolbox contains, e.g., engineered bacteria for the production of functionalized heparosan (Leroux & Priem, 2016; Priem et al., 2017), bacterial polymerases for the enzymatic built-up of heparosan (Na et al., 2020; Otto et al., 2012; Sismey-Ragatz et al., 2007), heparosan lyases (Clarke et al., 2000; Thompson et al., 2010) and other means to disassemble heparosan (Higashi et al., 2011), as well as reagents for the detection of the polymer (Kizer et al., 2018). However, due to its lack of immunogenicity, producing antibodies against heparosan is difficult (DeAngelis, 2015).

The *in vitro* synthesis of heparosan has been most successful using bacterial enzymes. The best-studied and biotechnologically most relevant heparosan synthases, PmHS1 and PmHS2, are derived from *Pasteurella multocida* (Chavaroche et al., 2012; DeAngelis & White, 2002, 2004). These enzymes are two-domain polymerases that transfer GlcNAc or GlcA alternatingly from UDP-GlcNAc or UDP-GlcA to the non-reducing end of the nascent chain (Otto et al., 2012). They are part of a group 2 capsule expression system (Sande & Whitfield, 2021) and assemble the polymer using a distributive mechanism (Sismey-Ragatz et al., 2007). Distributive polymerases display weak affinity towards the growing polymer, frequently dissociate, and thus allow control over the product length by adjusting the donor (nucleotide sugar) to acceptor (fragmented polymer) ratio when used *in vitro* (Sismey-Ragatz et al., 2007; Yakovlieva & Walvoort, 2020). Thus, PmHS1 and PmHS2 have been successfully

used for biotechnological purposes (Chavaroche et al., 2012; Jing et al., 2017; Na et al., 2020; X. Zhang et al., 2019).

Interestingly, *E. coli* K5 expresses a heparosan synthase consisting of two GT-A folded domains that share similarities with the *Pm* enzymes. However, the *E. coli* K5 enzymes are expressed as separate polypeptides, KfiA and KfiC (Sugiura et al., 2010). They are co-expressed with (i) KfiB, which is speculated to promote membrane localization and is supposed to have a scaffolding function (Cress et al., 2014; Hodson et al., 2000), and (ii) KfiD, a UDP-glucose dehydrogenase responsible for providing UDP-GlcA, the substrate of KfiC (Sieberth et al., 1995). *In vitro* and *in vivo* studies showed that heparosan synthases activity can be observed without KfiB (Leroux & Priem, 2016; Sugiura et al., 2010; C. Zhang et al., 2012). However, KfiC is only active in the presence of KfiA, whereas KfiA is active alone (Sugiura et al., 2010). Interestingly, only KfiA has been used for *in vitro* biotechnological purposes (Li et al., 2014; Xu et al., 2011), whereas the homologous domain from PmHS1/2 usually substitutes KfiC, reportedly because of insufficient expression (X. Zhang et al., 2019a). Since information about the elongation mechanism utilized by KfiC and KfiA (from here on referred to as KfiCA if used together) is missing, one aim of this study was to provide this information and to evaluate how suitable KfiCA are for building up heparosan of different sizes.

Available tools for the depolymerization of heparosan include two heparosan lyases, the phage tailspike protein KflA (Clarke et al., 2000) and the putative pro-phage (ϕ) derived eliminase ElmA (Legoux et al., 1996). Both enzymes have been studied biochemically (Schwarzer et al., 2007; Thompson et al., 2010), and a crystal structure shows that KflA adopts a trimeric single-stranded β -helix fold (Thompson et al., 2010). The structure demonstrates extensive contacts between the protomers, which might explain the high thermostability of KflA (Thompson et al., 2010), a feature observed for many tailspike proteins (TSPs) (Mühlenhoff et al., 2003). The correct folding of TSPs can be assisted by a C-terminal chaperone domain (CTD), which is subsequently cleaved by an autocatalytic reaction involving a conserved serine residue (Schwarzer et al., 2007). Some TSPs, like the polysialic acid-degrading endosialidase endoNF of Φ K1F, have different polymer binding sites located at a distance from the catalytic center (Stummeyer et al., 2005). This architecture might allow the phage to tightly bind to the host capsule while simultaneously degrading the polymer, ultimately

reaching the bacterial membrane where infection occurs (Leiman et al., 2007). Studies on the endosialidase NF have shown that the catalytic center can be deactivated without destroying the polymer binding site, which led to the development of a detection agent for polysialic acid (Stummeyer et al., 2005). Since the structure of KflA was obtained without polymer, it is unknown if bacterial heparosan lyases exploit the same multivalent binding mode and if they can be transformed into a detection agent for heparosan.

This study is based on the hypothesis that *E. coli* K5-based polymerases and lyases can broaden the toolbox for heparosan synthesis, depolymerization, and detection. First, we generated a fluorescently labeled oligosaccharide that can be used to characterize the elongation mechanism of KfiA and KfiC. Moreover, these enzymes can be exploited to synthesize fluorescently labeled heparosan of different sizes. We show that the tailspike protein KflB from the as yet undescribed *E. coli* K5-specific Φ K5B has heparosan lyase activity. Additionally, we present an extensive biochemical characterization of KflB using natural and fluorescently labeled heparosan as a substrate. Finally, we demonstrate that the enzyme can be engineered to become a detection agent for heparosan.

2.2. Results and discussion

2.2.1. Generation of fluorescently labeled heparosan oligomer

Our first aim was to investigate if the K5 heparosan synthases, KfiA and KfiC, can generate fluorescently labeled heparosan of low dispersity and different sizes. To enable these experiments, we wanted to analyze the elongation mechanism of KfiA and KfiC starting from a fluorescently labeled oligomeric acceptor. To obtain such an acceptor, heparosan oligomers were harvested from a previously described *E. coli* K-12 derivative, expressing K5 capsule biosynthesis genes and the heparosan lyase ElmA (Barreteau et al., 2012) (Supplementary **Table S1**). ElmA leads to the glycosidic bond breaking in a β -elimination reaction with the simultaneous formation of a double bond at the non-reducing end of GlcA. This structural feature enables the detection of the oligosaccharides at $\lambda = 232\text{-}250$ nm after purification by SEC (**Fig. 4A-C**). Then, we used mass spectrometry (Supplementary **Fig. S1**) to determine the chain length of the generated K5 oligosaccharides (**Fig. 4C**). The K5 Δ -decamer (Δ = unsaturated uronic acid at the non-reducing end) was the dominant species in the

mixture (see **Fig. 4C**) and likely to be of sufficient size for enzymatic elongation (M. Chen et al., 2006; Sugiura et al., 2010). Therefore, this compound was chosen as a substrate for subsequent reactions to generate a fluorescently labeled K5 acceptor. First, the unsaturated uronic acid had to be removed to free the non-reducing end for polymerase elongation. Thus, the K5 Δ -decamer **1** (**Fig. 4D**) was treated with mercury (II) acetate (Ludwigs et al., 1987), yielding the K5 nonamer **2** with a terminal GlcNAc available for elongation. The K5 nonamer **2** was labeled using 2-aminobenzamide (2AB), resulting in the K5 nonamer **3**. The identity of **3** was confirmed by mass spectrometry (Supplementary **Fig. S1F**).

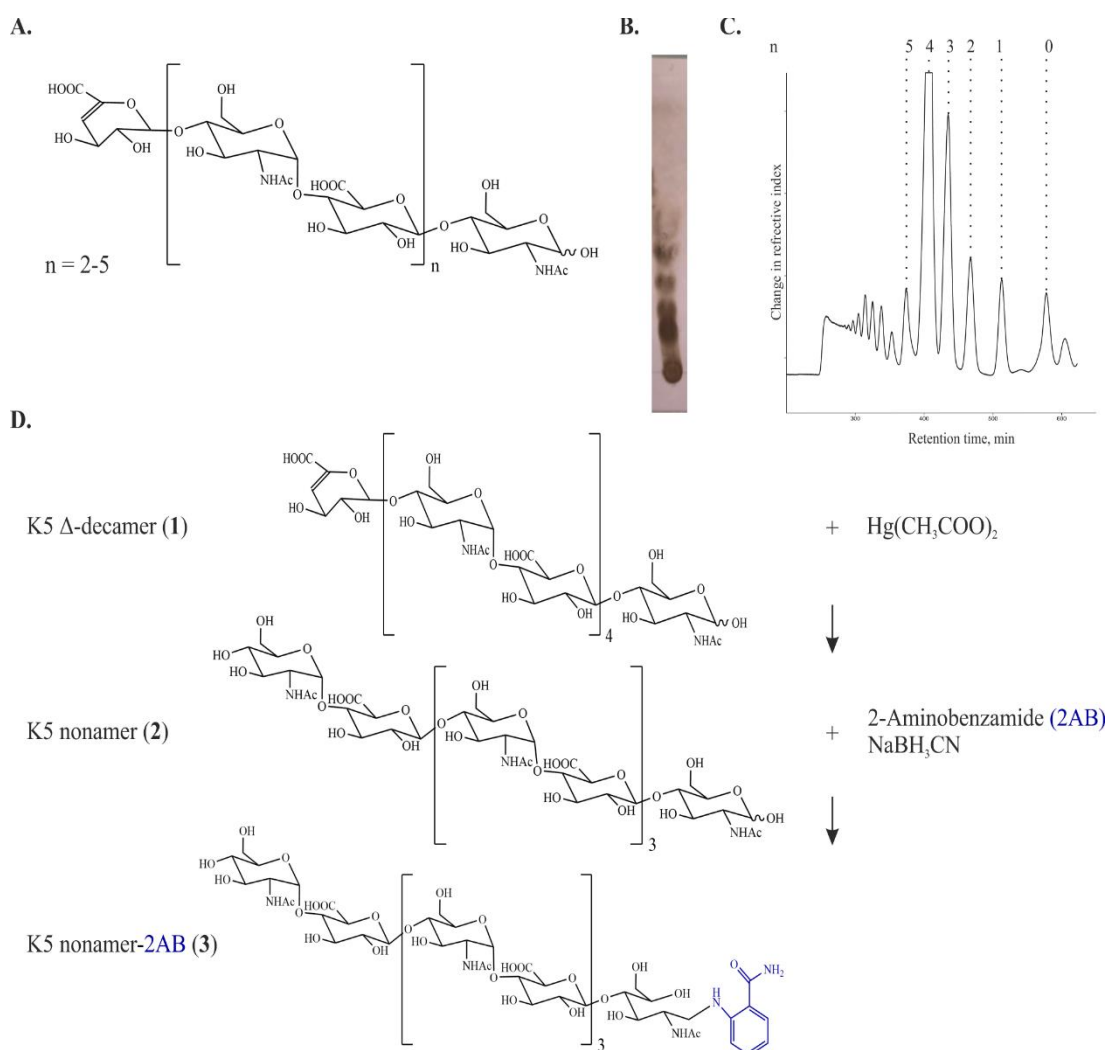


Figure 4. Generation of a fluorescently labeled heparosan nonamer. **A.** Structure of heparosan (K5) oligosaccharides harvested from bacterial cultures (Barreteau et al., 2012), n represents the number of internal GlcNAc-GlcA repeating units. Note the double bond introduced by ElmA (Barreteau et al., 2012). **B.** Thin layer chromatogram of the harvested K5 oligosaccharides and **C.** corresponding size exclusion chromatogram. **D.** Schematic representation of the strategy utilized to produce the 2AB labeled K5 nonamer **3**.

2.2.2. Investigation of the elongation mechanism of untagged KfiA and KfiC

To utilize KfiA and KfiC for biotechnological purposes, the enzymes should ideally elongate **3** in a distributive way. A hallmark of distributive elongation is a product growing in size over time with a low dispersity profile. Contrastly, in processive elongation, long products arise at early time points of the reaction and the dispersity is high (Budde et al., 2020; Fiebig et al., 2018; Yakovlieva & Walvoort, 2020). KfiC could successfully be purified if fused to the chaperone trigger factor (Sugiura et al., 2010). Since fusion tags might interfere with the elongation mode of the polymerase enzymes, KfiA and KfiC were cloned without tags into an arabinose inducible (pBAD33) vector (Supplementary **Table S2** and **S3**) and recombinantly expressed to analyze their activity from supernatants of bacterial lysates. Elongation of **3** in the presence of UDP-GlcA, UDP-GlcNAc (**Fig. 5A**), and lysates of the expression cultures was observed in a time-dependent manner using HPLC-based anion-exchange chromatography (AEC) with fluorescent detection. As expected, the product profile grew in size over time, indicating distributive elongation. The high molecular weight heparosan produced overnight and the populations observed in between showed low dispersity (**Fig. 5B**). Interestingly, in most reactions, a few chains of intermediate size (RT = 20-30 min) were neglected by the enzymes. The reason for this observation is currently unclear (e.g., steric hindrance caused by the label, formation of a secondary structure that is dominant at a certain chain length as reported for other group 2 polymers (Hlozek et al., 2018), or a preference of KfiCA for the elongation of longer chains). Noteworthy, short fluorescently labeled acceptors (e.g., DP = 2 or 4) have been successfully used in the elongation mode studies of group 2 polymerases, indicating that the label on **3** should not influence the elongation mode (Keys, Berger, et al., 2012; Keys, Freiburger, et al., 2012; Jing & DeAngelis, 2004). Although neglected chains can indicate processivity (Fiebig et al., 2018; Yakovlieva & Walvoort, 2020), the continuous growth of the chains over time, the Gaussian-shaped product profile in the absence of long chains at early reaction times argues that KfiA and KfiC sequentially build up the heparosan chain in a distributive manner. This agrees with reports describing homologous GAG polymerases from *Pasteurella multocida* as distributive (Jing & DeAngelis, 2004; Sismey-Ragatz et al., 2007). The processivity of one-domain and two-domain group 2 polymerases is usually mediated by extended polymer binding sites that are sometimes located in separate domains

(Budde et al., 2020; Fiebig et al., 2018; Keys et al., 2014). In line with our results, no indication of these binding sites has been found in bacterial GAG polymerases so far (Osawa et al., 2009), even though they have been reported for viral enzymes with homologous function (Maloney et al., 2022). In summary, the above-presented results indicate that KfiA and KfiC can be used to synthesize heparosan of different sizes *in vitro* for biotechnological applications.

2.2.3. Evaluating tagged and purified KfiA and KfiC constructs for the build-up of heparosan

Besides size control, one of the most important advantages of the enzyme-based *in vitro* production of heparosan is the reduced effort to purify the polymer (DeAngelis et al., 2013). Both enzymes were cloned as tagged fusion constructs to enable the controlled *in vitro* use of KfiA and KfiC. KfiA was cloned with an N-terminal hexa-His-tag, and KfiC (Supplementary **Fig. S2**) was cloned as a fusion protein with the chaperone trigger factor (TF), forming the previously described constructs (Sugiura et al., 2010) His₆-KfiA and His₆-TF-KfiC (hereafter referred to as pKfiA (5644) and pKfiC (5822), see Supplementary **Table S2**). pKfiA and pKfiC could be purified in sufficient yield (1.2 mg/l culture and 12.8 mg/l culture, respectively), but the purity was not ideal (Supplementary **Fig. S3A**, Supplementary **Table S4**). Using a radioactive incorporation assay, Sugiura *et al.* demonstrated that pKfiC was only active in the presence of pKfiA (Sugiura et al., 2010). The acceptor **3**, ending with GlcNAc, allowed us to confirm these results (Supplementary **Fig. S3B**). We could confirm that both enzymes were active as polymerases (Supplementary **Fig. S3C**). However, a direct comparison between the chain length achieved with pKfiCA, and the products synthesized by non-tagged constructs from culture lysates, demonstrated that the latter produced considerably larger chain lengths (**Fig. 5C**). Additionally, the reproducibility of chain lengths achieved with purified constructs was low, potentially indicating issues with enzyme stability (compare **Fig. 5B** and Supplementary **Fig. S3C**).

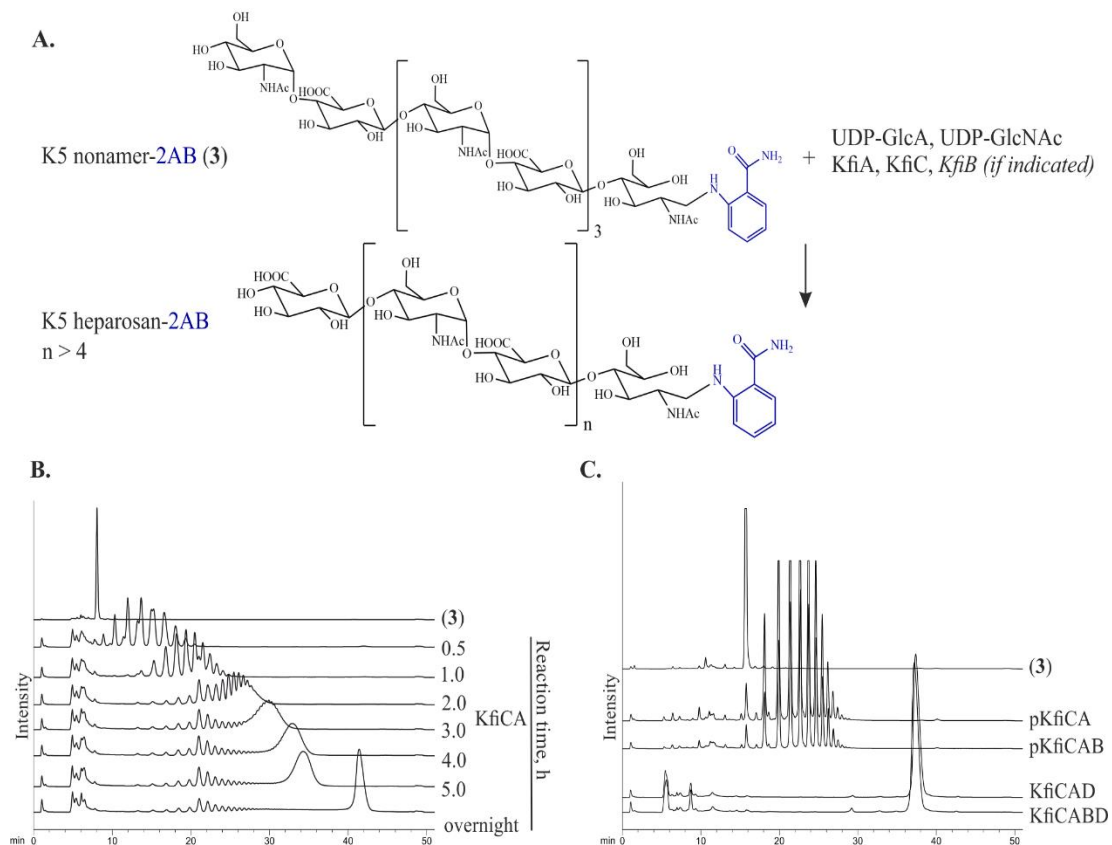


Figure 5. Elongation of the 2AB labeled K5 nonamer. **A.** Schematic representation of the reaction. **B.** Time course of the elongation of K5 nonamer-2AB by non-tagged KfiCA obtained from expression culture lysates. Please note that the gradient was optimized to allow a better separation of intermediate-sized chains (if compared to C). **C.** Elongation of the K5 nonamer-2AB 3 by tagged and purified pKfiCA, pKfiCAB, non-tagged KfiCAD, or KfiCABD obtained from expression culture lysates. The reactions were incubated overnight. Note that long polymers elute between 450-470 mM of NaCl in both gradients used in B and C, indicating that the co-expression of KfiD or KfiBD has no considerable effect on the chain length produced under the investigated *in vitro* conditions.

Since KfiB has previously been reported as a scaffolding factor (Cress et al., 2014; Hodson et al., 2000; C. Zhang et al., 2012), we hypothesized that it might stabilize KfiA and/or KfiC, promoting the production of longer heparosan chains. Therefore, we cloned MBP-KfiB-His₆, in which KfiB was fused to maltose binding protein and which could be expressed and purified by affinity chromatography (5.2 mg/l) (Supplementary **Fig. S3A**). As a control, we further cloned a non-tagged version of KfiB into the pBAD33-KfiCA expression vector used in **Fig. 5A**. Interestingly, the influence of KfiB on the chain length was not reproducible, neither for the tagged and purified construct (**Fig. 5C** and Supplementary **Fig. S3C**) nor for an untagged version co-expressed with untagged KfiCA and tested from lysates (**Fig. 5C**). Noteworthy, the expression level of KfiCAB from pBAD33 has not been

determined due to the absence of affinity tags. While KfiCA are clearly active and thus sufficiently expressed for the purpose of the experiment, KfiB expression might very well be insufficient under the chosen conditions. Thus, the above-presented results are not excluding the influence of KfiB on the stability of KfiCA *in vivo* (Cress et al., 2014; Hodson et al., 2000). It was beyond the scope of the current study to further analyze the function of KfiB.

While KfiA has been successfully used for biotechnological purposes *in vitro* (Xu et al., 2011; X. Zhang et al., 2019), KfiC expression has been considered insufficient for biotechnological applications (X. Zhang et al., 2019a). Despite considerable efforts for rational construct optimization (e.g., truncation studies) and reported successful purification of the enzyme (Sugiura et al., 2010). In our study, pKfiC could be produced with a satisfactory yield that exceeded the yield obtained for pKfiA. However, both the purity and the activity of pKfiCA remained an issue. The low molecular weight (MW) heparosan produced by pKfiCA could be exploitable as a substrate for heparin synthesis, while the production of high MW heparosan (> 200 kDa), e.g., for the production of biodegradable materials (Mills, 2002; Williams et al., 2019) could not be achieved. It is currently unclear if the impaired ability of pKfiCA to produce long chains results from the introduced tags, suboptimal buffer conditions, stability issues, or a combination of the above. Both stability and activity issues were reported for the polysialyltransferase from *Neisseria meningitidis* serogroup B and could be improved by optimizing the tagged enzyme by directed evolution (Keys, Berger et al., 2012; Keys et al., 2014). Unfortunately, these kinds of studies would require an antibody for the screening of large enzyme libraries using intra- and/or extracellular heparosan expression as readout. To the best of our knowledge, such antibodies are unavailable due to the low immunogenicity of heparosan (DeAngelis, 2015), and antibodies against *E. coli* K5 that were available to us cross-react with ubiquitous surface molecules (Peters et al., 1985). Since bacteriophage TSPs can be engineered to detect bacterial polysaccharides (Stummeyer et al., 2005), our next step was to investigate if an inactive heparosan lyase could be used to detect heparosan, paving the way toward directed evolution experiments with KfiCA.

2.2.4. Expression and purification of the Φ K5B tailspike protein KflB

We decided to explore if the enzymatic machinery of the as yet undescribed *E. coli* K5-specific Φ K5B would contain a heparosan lyase that could be exploited as a heparosan detection agent. The open reading frame (ORF) gp46 (Supplementary **Fig. S4**) of the Φ K5B showed homology to known heparosan lyases. In analogy to the heparosan lyase KflA from the *E. coli* K5-specific Φ K5A, gp46 was termed KflB. KflB shares 97.9% sequence identity with the heparosan lyase from Φ K1-5 (GenBank AAG59821.1 (Scholl et al., 2001)), 98.3% sequence identity to the revised sequence of KflA (CAA71133.2, (O’Leary et al., 2013)) and 56.6% sequence identity to ElmA (CAA65353 (Legoux et al., 1996)). To investigate if KflB can be expressed as a recombinant protein in *E. coli*, the ORF was amplified and cloned with N-terminal Strep II and C-terminal His₆-tag (**Fig. 6A**). In parallel, we designed a construct to investigate if KflB contains a C-terminal chaperone domain (CTD). CTDs have been described, e. g., for ElmA (Schwarzer et al., 2007) and have been suspected of playing a role in the folding of KflA (Thompson et al., 2010). Thus, we performed a multiple sequence alignment with KflB and TSPs with characterized CTDs, demonstrating that S505 is a putative candidate for the catalytic serine involved in the autocatalytic cleavage of the CTD of KflB (**Fig. 6B**). Consequently, KflB-S505A was generated. KflB wt and KflB-S505A were expressed in *E. coli* BL21(DE3), purified via affinity chromatography, and analyzed using SDS-PAGE and western blotting (**Fig. 6C, D**). Since TSPs can be SDS resistant (Mühlenhoff et al., 2003), samples were analyzed with (+) and without (-) denaturation (95°C for 5 min) prior to SDS-PAGE. The band of the denatured KflB wt corresponds to the expected molecular weight of a mature monomer (54 kDa). In contrast, the band corresponding to KflB-S505A was shifted towards higher molecular weights, with the shift being consistent with a retained CTD (16 kDa). In addition, the His₆-tag (**Fig. 6D**, red channel) located at the C-terminus of the CTD was only detectable by Western Blot in KflB-S505A. Both observations confirm that S505 is required for the autocatalytic loss of the CTD. When the denaturation step was omitted, a higher molecular weight band was revealed at ~150 kDa and ~ 130 kDa for KflB wt and KflB-S505A, respectively. This suggests the formation of a homotrimer for KflB wt, which agrees with the trimer formation observed in the crystal structure of KflA (Thompson et al., 2010).

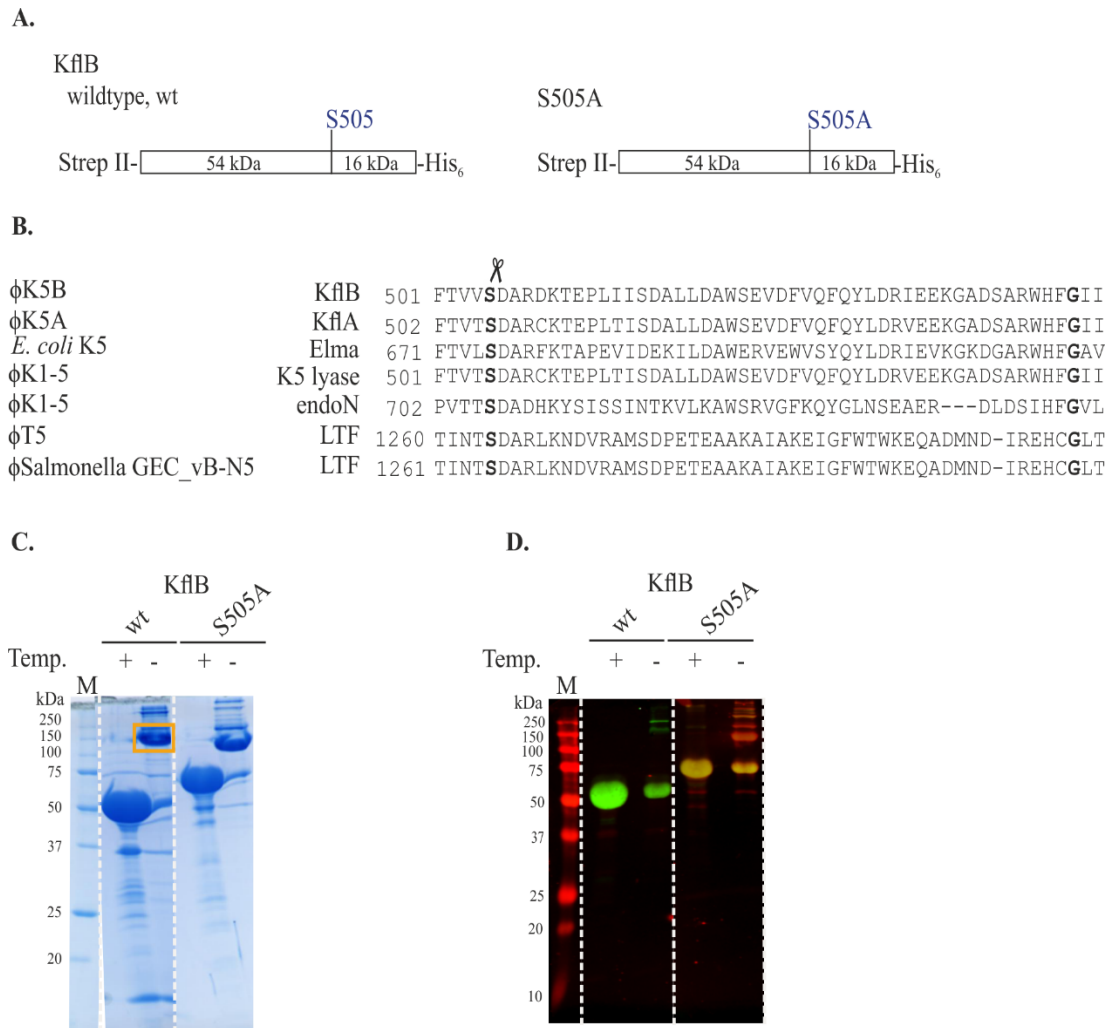


Figure 6. Recombinant expression of KfIB. **A.** Schematic representation of KfIB wild-type and KfIB-S505A. **B.** Partial multiple sequence alignment of tailspike proteins harboring a C-terminal chaperone domain (CTD). From the left: the origin of the protein, name of the protein, number of the first residue included in the alignment, and amino acid sequence. The conserved serine residue is marked in bold. The scissors (Ser^XAsp) indicate the proteolytic cleavage site. The multiple sequence alignment was generated by Clustal 2.1 using proteins of the following accession numbers (European Nucleotide Archive database): CAA71133 (KfIA), CAA65353 (Elma), AAG59821 (K5 lyase), AAG59822 (endoN), CAJ29339 (L-shaped tail fibre protein LTF of coliphage T5), QPI15176.1 (Long tail fibre LTF of the salmonella phage GEC_vB_N5). **C.** Gel from SDS-PAGE stained with Coomassie Brilliant Blue G-250 (Roth) showing purified Strep II-KfIB wt and Strep II-KfIB-S505A. The whole gel is shown in Supplementary Fig S5. The SDS-resistant complex of Strep II-KfIB wt is marked by the orange box. **D.** Corresponding Western Blot analysis. Strep II-tag (green channel), His₆-tag (red channel). M – marker, T - temperature, „+“ – denatured samples.

In contrast, the apparent MW of KfIB-S505A is more consistent with the formation of a homodimer. Interestingly, and in contrast to KfIB-S505A, a corresponding mutant (endoNF-S991A) of the endosialidase endoNF (Schwarzer et al., 2007) showed

decreased instead of increased electrophoretic mobility in an SDS-PAGE. This observation led to the proposal of a folding mechanism involving a partially folded intermediate that eventually leads to the formation of a stable protimer from which the CTD is released. It is beyond the scope of this study to further investigate the folding mechanism of KfIB. Interestingly, although SDS-resistant complexes are visible in the Coomassie-stained SDS-PAGE, the Strep-tag II is not detected (compare **Fig. 6C**, yellow box, with **Fig. 6D**, green channel, 150 kDa) in the corresponding Western Blot. This indicates that the N-terminal tag might be buried inside the trimeric complex, suggesting that at least this N-terminal part of KfIB wt and KfIB-S505A is folded.

2.2.5. *KfIB is a heparosan lyase*

In the next step, we wanted to investigate if KfIB has heparosan lyase activity. Heparosan from metabolically engineered *E. coli* BJ (Leroux & Priem, 2016) was purified, fractionated (Supplementary **Fig. S6**), incubated with KfIB, and samples were taken at different time points. To see if the CTD's presence impacts enzyme activity, we included KfIB S505A. Products were separated and visualized on a PA (polyacrylamide) gel stained with (**Fig. 7A, C**). Additionally, an HPLC-based assay was developed, allowing separation via anion-exchange chromatography (AEC) (**Fig. 7B, D**) coupled to UV detection at $\lambda = 250$ nm. Both assays clearly showed that both constructs are active. Overnight (ON) incubation of high molecular weight (HMW) heparosan with both constructs led to the complete disappearance of the substrate on the PA gel (**Fig. 7A, C**). This finding is in line with earlier studies which reported that the loss of the CTD of endosialidases is not required to produce an active enzyme (Mühlenhoff et al., 2003; Schwarzer et al., 2007). Interestingly, KfIB leaves a mixture of four short oligosaccharide products as detected by HPLC-AEC (**Fig. 7B, D**).

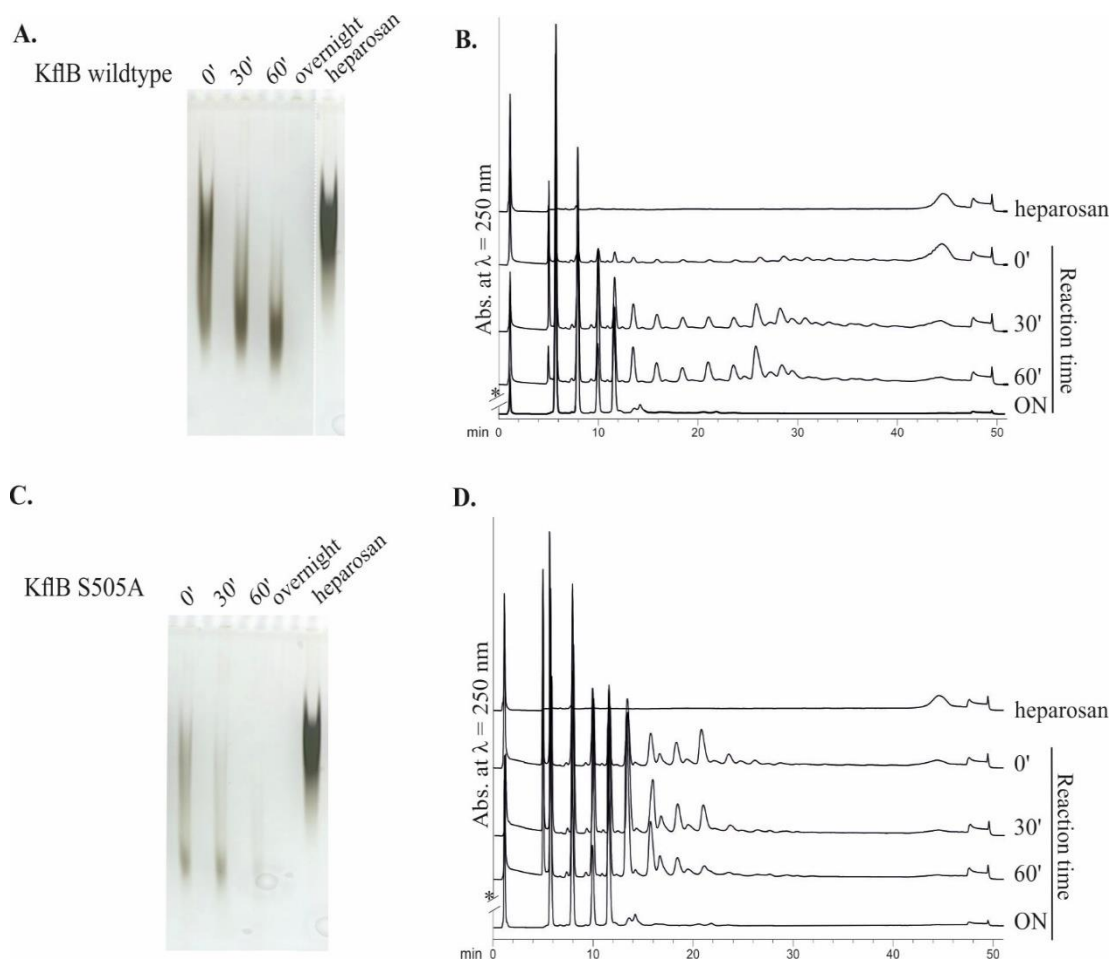


Figure 7. Enzymatic activity of KfIB. Heparosan degradation catalyzed by KfIB wt (**A, B**) and KfIB S505A (**C, D**). Products were visualized after (**A, C**) PAGE followed by alcian blue/silver staining or (**B, D**) UV detection after separation by HPLC-AEC. *The signal intensity of the overnight (ON) sample was decreased for illustrative purposes.

Moreover, we investigated the minimal substrate size of KfIB using the oligosaccharides shown in **Fig. 4C**. K5-derived tetramer ($n = 1$, see **Fig. 4A**), hexamer ($n = 2$), octamer ($n = 3$) and decamer ($n = 4$) were incubated in the presence of KfIB wt. HPLC with detection at $\lambda = 232$ nm (**Fig. 8**) demonstrated degradation of the deca-, octa-, and hexamer into mixtures of shorter oligomers, while the tetramer remained unchanged. Therefore, we concluded that the minimal length of the substrate for KfIB would be a hexamer ($n = 2$, see **Fig. 4A**). These results are in agreement with published literature that reported hexa-octamers of heparosan as a minimal substrate for KfIA (Murphy et al., 2004; O’Leary et al., 2013).

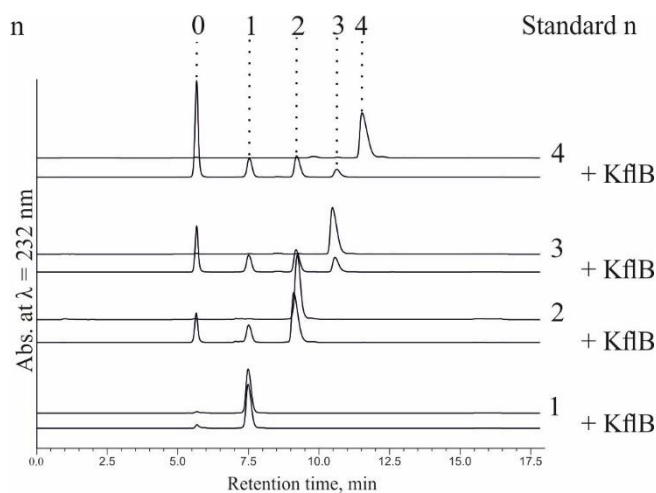


Figure 8. The minimal substrate for KflB. Heparosan oligomers in the presence and absence of KflB after overnight incubation. Products were analyzed via HPLC-AEC coupled to UV detection.

2.2.6. Analyzing the degradation mode of KflB using 2AB-K5 heparosan as substrate

The mode of action of heparosan lyases is still a matter of debate (Murphy et al., 2004; O’Leary et al., 2013; Thompson et al., 2010). This part of our study aimed to analyze if the product pool of KflB is consistent with an exo- or an endo-mode of degradation. We used 2AB-labeled heparosan for these experiments. An ideal exo lyase would cleave off small unlabeled fragments of the non-reducing end of 2AB-K5 heparosan, resulting in a population of 2AB-K5 heparosan that slowly decreases in size over time. It is reasonable to assume that this exo activity would favor the non-reducing end of heparosan, which is pointing away from the bacterial surface in the *in vivo* situation, while the reducing end is coupled to a lipid carrier (Sande & Whitfield, 2021; Willis et al., 2013; Yan et al., 2020). However, in the unlikely event of degradation initiated at the reducing end (and not being hindered by the label), one would expect a progressive disappearance of the starting material with the simultaneous formation of short, final products of very low dispersity. In contrast, an ideal endo lyase would cut the polymer at random locations without bias, leading to a broad product spectrum of labeled and unlabeled products with short oligomers present from the start and accumulating over time. To analyze the degradation mode of KflB, 2AB-K5 heparosan was produced with KfiCA from a bacterial lysate, as shown in (Fig. 5B) and the enzymes were heat-inactivated. KflB was supplemented, and the product profile was further analyzed by HPLC-AEC with fluorescence detection (Fig. 9B). Fig. 9B shows that the intensity but not the retention time of the peak corresponding to long 2AB-K5 heparosan (46 min) decreases, clearly arguing against

an exo mode of degradation. The fact that comparably short (eluting between 20-30 min) 2AB-labeled fragments start accumulating early in the reaction and keep accumulating over time (rather than at the end of the observed reaction period) again argues against an exo mode. The high dispersity of the product spectrum maintained over a long period is in favor of an endo mode (**Fig. 9C**). Interestingly, the major 2AB-labeled products accumulating during the first 120 minutes are of intermediate size (app. 10-20 repeating units) and reminiscent of the species that were not further elongated by KfiCA (**Fig. 5B**; RT = 20-30 min). The fact that these species are observed in the majority of both polymerization and depolymerization reactions suggests that their production is not due to inherent enzyme properties (e.g., binding sites) but rather due to the properties of the polymer (induced by the label or a secondary structure). Importantly, these species are absent when non-labeled heparosan is digested (**Fig. 7B, D**). After overnight incubation, 2AB-K5 heparosan is fully degraded to small oligosaccharides (**Fig. 9B**).

Studies analyzing the mode of degradation of KflA, which has the highest sequence similarity to KflB of the as yet described heparosan lyases, have analyzed the reaction endpoint of the degradation of non-labeled heparosan or the digest of chemically synthesized fragments (Murphy et al., 2004; O'Leary et al., 2013). In this study, the enzymatic built-up of labeled heparosan provided HMW material that allowed the analysis of the depolymerization in a time-dependent manner by HPLC-AEC. The 2AB labeled heparosan has the advantage that only the reducing end product of the chain digested by KflB is detected. In contrast, both reaction products are observable when the double bond is used for UV detection, especially in the later stages of the experiment, when every chain has been cleaved at least once, making the product spectrum harder to interpret. Although the accumulation of labeled oligosaccharides of intermediate size prevented analyzing the time-dependent degradation of smaller oligosaccharides, a considerable range of reaction intermediates could be visualized, strongly suggesting an endolytic mechanism for KflB.

In previous studies, (i) the fact that a mixture (di-decamers) of products has been observed after complete heparosan digest by KflA or the Φ K5A (Hänfling et al., 1996; Murphy et al., 2004), and (ii) the observation that an octamer is the minimal substrate for the enzyme, have been considered an indication for an endo mode of KflA (Murphy

et al., 2004; O’Leary et al., 2013). In contrast, the predominance of dimers in the final product pool of KflA has been interpreted as an indication of exo cleavage, while products larger than a hexamer would be expected for the endo mode (O’Leary et al., 2013). A model of processive endolytic cleavage has been suggested to explain this observation (O’Leary et al., 2013), but no experimental evidence has been presented. Interestingly, Schwarzer *et al.* reported that the wild-type endosialidase NF displays processive degradation of polysialic acid, whereas a serine mutant retaining the CTD was more distributive (Schwarzer et al., 2009). In their study, the distributive mutant degraded the long polymer early in the reaction. In contrast, the processive wild-type enzyme required a longer time, presumably, because the pool was degraded “polymer by polymer” rather than in a random fashion. Unmasking a processivity-mediating polysialic acid binding site in this mushroom-shaped enzyme’s stalk domain was suggested as a molecular basis for the different modes of action. In line with that, the more distributive Φ phi92 endosialidase lacks such a binding site (Schwarzer et al., 2015).

Since a distributive KflB would yield short to intermediate products early in the reaction, 2AB labeled HMW heparosan might not be the suitable substrate to further analyze the depolymerization mode because the accumulation of intermediate-sized products might interfere with the interpretation of the results. Noteworthy, the differences between the product profiles produced by KflB and KflB-S505A during the digest of non-labeled heparosan are comparable to the differences observed between endoNF and endoNF-S911A (Schwarzer et al., 2009). This indicates that KflB-S505A might also be more distributive than KflB wt, or in other words, KflB wt displays some degree of processivity (**Fig. 7B** and **D**, Supplementary **Fig. S7**). In summary, our results support a processive and endolytic mode for KflB. However, since no structural data is available for KflB, and since the homologous KflA was crystallized without substrate (Thompson et al., 2010), it is currently unclear if a processivity-mediating polymer binding site is present in heparosan lyases.

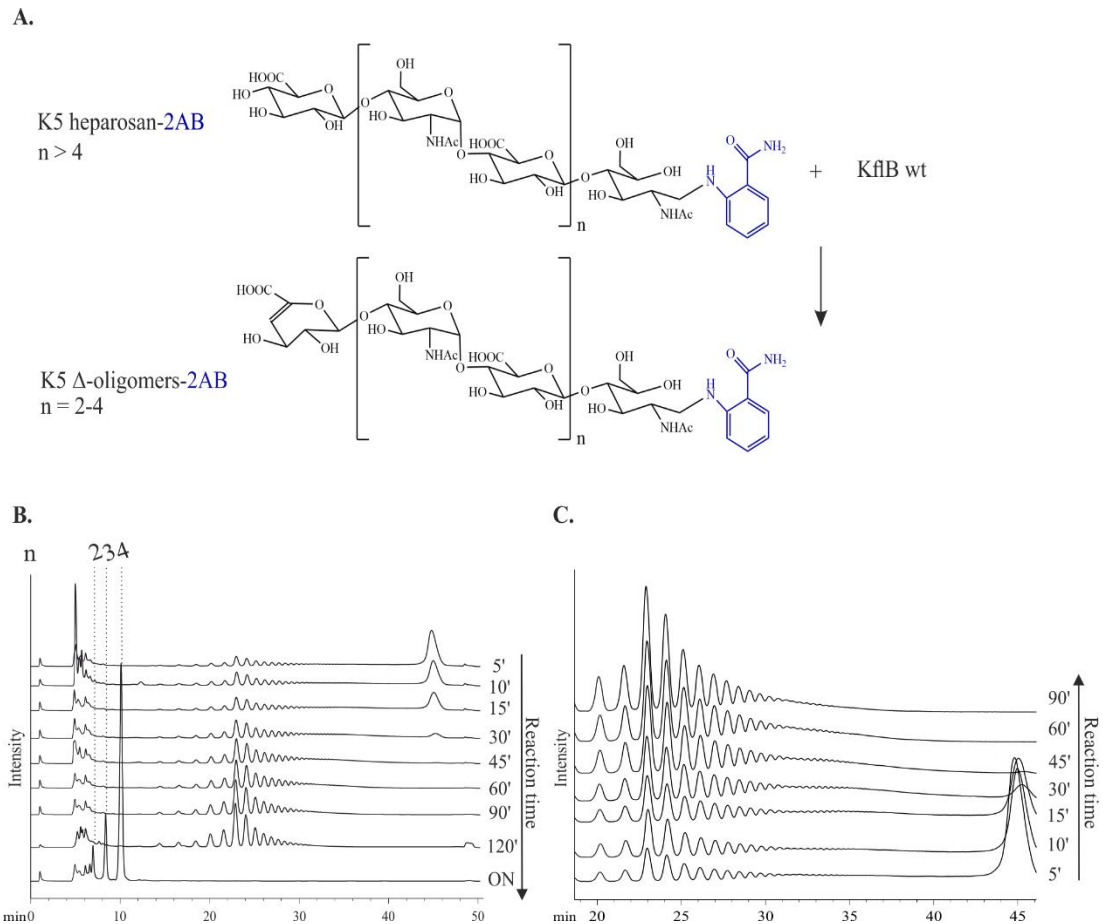


Figure 9. Depolymerization of fluorescently labeled heparosan by KflB. **A.** Schematic representation of the reaction. **B.** Time-course of the reaction with separation and detection of 2AB labeled substrates and products by HPLC-AEC. **C.** Zoomed-in section of B. Note that the reaction time is reversed to improve the visualization of the accumulation of small-intermediate products.

2.2.7. Characterization of KflB substrate specificity

The development of molecular tools for engineering glycosaminoglycans (GAGs) is highly interesting (DeAngelis et al., 2013; Miller et al., 2014). Glycosaminoglycans are very stable and require harsh conditions to undergo chemical degradation, whereas enzymatic degradation is performed under mild conditions, thus representing an appealing alternative (Bohlmann et al., 2015; Miller et al., 2014; Murphy et al., 2004). Consequently, we wanted to investigate if KflB could degrade other GAGs. KflB wt was incubated with hyaluronic acid, heparan sulfate, heparosan, chondroitin sulfate, and dermatan sulfate. Products were analyzed using PAGE combined with alcian blue/silver staining (**Fig. 10**). As expected, heparosan was completely degraded already after 0.5 h. Hyaluronic acid, chondroitin sulfate, and dermatan sulfate, structurally different from heparosan (DeAngelis et al., 2013), were resistant to KflB digestion. Heparan sulfate was slightly degraded within half an hour

of incubation, but longer incubation times did not change the product profile. It is tempting to speculate that this is because the precursor of heparan sulfate is structurally similar to heparosan (Y. Chen et al., 2013; DeAngelis et al., 2013). The majority of heparan sulfate is *N*- or *O*-sulfated, and glucuronic acid is partly epimerized to iduronic acid. The modified domains (NS domains) are alternating with less processed (NA/NS) domains and non-modified (N-acetylated or NA) domains that are still identical to heparosan (Sarrazin et al., 2011). We hypothesize that the partial degradation indicates that Kf1B digests heparan sulfate in its non- or less-modified domains. Like Kf1A, which has been shown to cleave heparan sulfate in non-sulfated regions or even in the vicinity of isolated *N*-sulfated residues (Murphy et al., 2004; O’Leary et al., 2013), Kf1B could be an attractive tool for the enrichment of NS and NA/NS domains structural analysis.

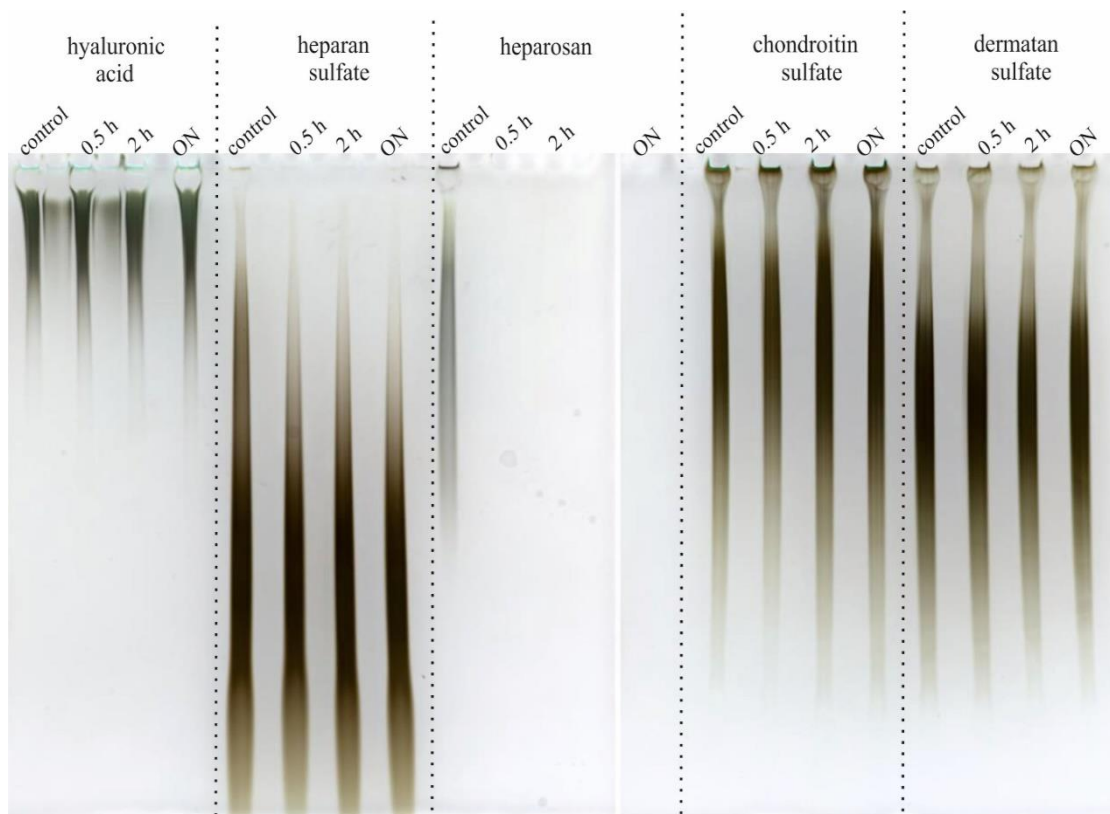


Figure 10. Kf1B substrate specificity. Different glycosaminoglycans were incubated with Kf1B and analyzed after the indicated time points using PAGE, followed by alcian blue/silver staining. A control reaction in the absence of Kf1B was incubated overnight to document the stability of the respective polymer in the reaction buffer.

2.2.8. Identification of KflB's active center and development of a heparosan detection agent

Stummeyer *et al.* demonstrated the possibility of transforming the tailspike endosialidase endoNF into a polysialic acid detection agent (Stummeyer *et al.*, 2005). Active site mutations led to the loss of endosialidase activity, while the resulting mutants retained their affinity towards polysialic acid, most likely via the above-discussed binding site present in the stalk domain. Therefore, we wanted to investigate if KflB can be transformed into a heparosan detection agent. Thompson *et al.* determined catalytically essential residues in KflA based on the apo structure and corroborated them using the kinetic measurements (Thompson *et al.*, 2010). A multiple sequences alignment identified F202, E206, K208, and Y229 as corresponding residues in KflB (**Fig. 11A**). Using site-directed mutagenesis, we introduced single or double amino acid exchanges to alanine in KflB, purified the resulting constructs (Supplementary **Fig. S5**) and analyzed their activity towards heparosan in an overnight reaction. Heparosan was included as a control in the absence of the enzyme. The reactions were analyzed by alcian blue/silver stained PAGE (**Fig. 11B**). As expected, incubation with KflB wt and KflB-S505A led to the complete degradation of heparosan. Interestingly, all single amino acid exchange mutants showed residual activity toward high MW heparosan (**Fig. 11B**, compare product profile to heparosan control). Only those constructs in which two amino acids were exchanged for alanine showed a complete activity loss.

A.

KflB	191	IDNVIAKNYPQ F GAV EL KGTASYNIVSNVIGADCQHVT Y NGT
KflA	191	IDNVIAKNYPQ F GAV EL KGTASYNIVSNVIGADCQHVT Y NGT
K1-5	191	IDNVIAKNYPQ F GAV EL KGTASYNIVSNVIGTDCQHVT Y NGT
ElmA	361	IDGVIARNYPQ F GAV EL KTAAKYNI VSNVIGEECQHVV Y NGT

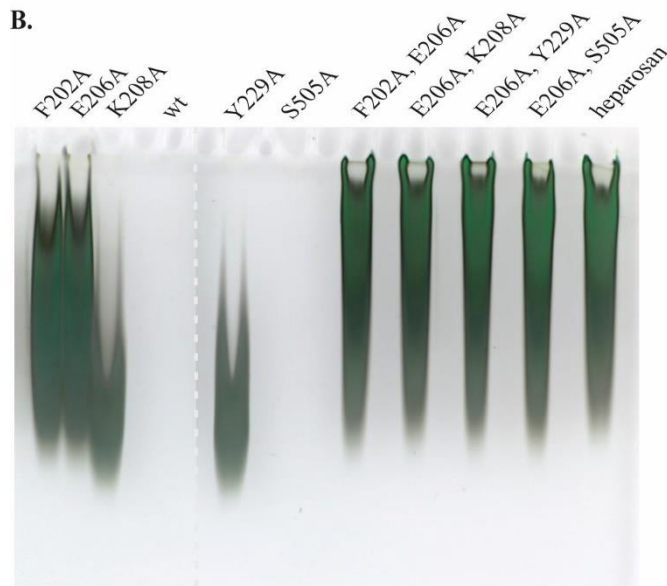


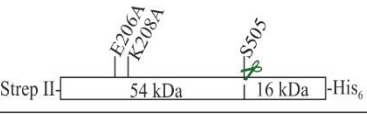
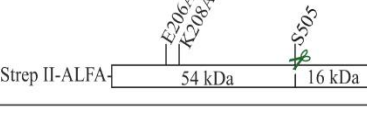
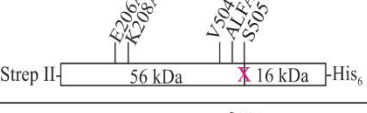
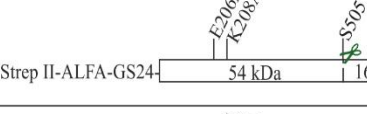
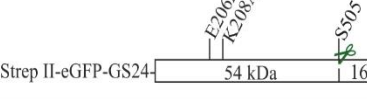
Figure 11. Identification of active site residues in KflB. **A.** Section of a multiple sequence alignment of heparosan lyases using the sequences as indicated in **Fig. 3**. From left to right: name of the protein, number of the first residue included in the alignment, amino acid sequence. Residues crucial for the activity of KflA (Thompson et al., 2010) are marked in bold. Residues marked with an asterisk were replaced by alanine resulting in the mutants shown in **B**. Effect of recombinant KflB and its mutants on heparosan in overnight reactions. Products were visualized with alcian blue/silver staining after PAGE. Heparosan incubated in the absence of the enzyme was used as a control. Recombinant KflB constructs and their identifiers: F202A (6112); E206 (6113); K208A (6114); wt (6118); Y229A (6119); S505A (6122); F206A, E206A (6167); E206A, K208A (6168); E206A, Y229A (6169); E206A, S505A (6170) (see Supplementary **Table S2**).

2.2.9. Development of a heparosan detection agent

All double amino acid exchange mutants were initially screened for their suitability as detection agents on *E. coli* K5 cells. However, without considerable success, most likely due to the poor accessibility of the Strep-tag II on a complex bacterial surface (see also **Fig. 4D**). As a minor degree of heparosan detection was possible using inactive KflB-E206A, K208A (hereafter referred to as iKflB), we decided to optimize the N-terminus of this construct for detection, by introducing various linkers and/or fusion tags. Constructs were generated according to the following rationale (see also **Fig. 12**): (I) an ALFA-tag (Götzke et al., 2019) was introduced between Strep-tag II and KflB's N-terminus to expose the Strep-tag II and

allow an additional possibility of detection via the anti-ALFA antibody (6223); (II) an ALFA-tag was introduced near the S505 cleavage site to have a C-terminal tag after the loss of the CTD (6224); (III) a flexible GS24 linker (X. Chen et al., 2013) was introduced in 6223 between the ALFA-tag and the enzyme's N-terminus to expose further the tags (6336); (IV) KfIB was N-terminally fused to Strep II-eGFP-GS24 (enhanced green fluorescent protein) for better exposure of the Strep-tag II and to enable detection via green fluorescent protein (GFP) if required; (V) iKfIB (construct 6168) was directly labeled with a fluorescent dye (Strep-iKfIB-IR800). All constructs except for 6224 could be purified. Interestingly, none of the ALFA-tags could be detected in a Western Blot using nano-anti-ALFA (not shown). Only the Strep-tag II of the eGFP fusion protein could be visualized without prior denaturation using Strep-Tactin IR800 (**Fig. 12B**).

A.

Protein	Protein design	Soluble protein?	Processed protein?	Is the homotrimer detected in WB?
6168 Strep II-iKfIB		yes	yes	no
6223 Strep II-ALFA-iKfIB		yes	yes	no
6224 Strep II-iKfIB-504-ALFA		no	no	no
6336 Strep II-ALFA-GS24-iKfIB		yes	yes	no
6338 Strep II-eGFP-GS24-KfIB		yes	yes	yes

B.

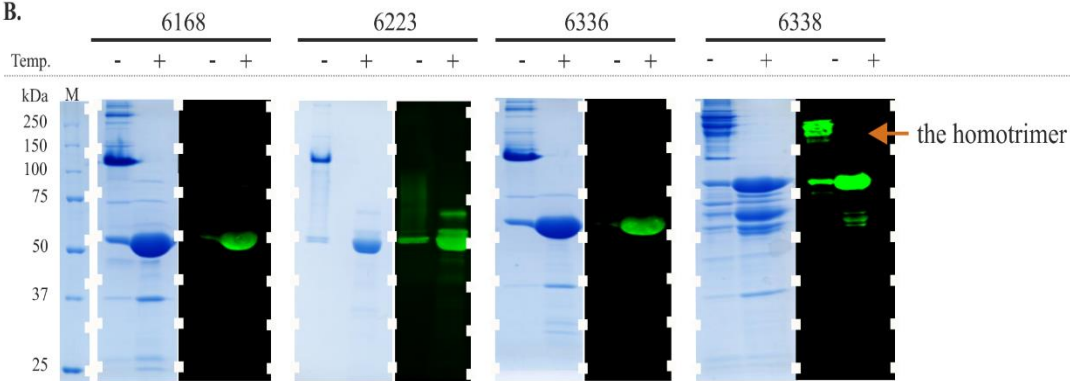


Figure 12. Engineering an iKfIB construct for the detection of heparosan. **A.** Overview of the designed construct. Green scissors indicate that the constructs show the expected molecular weight after CTD cleavage. Construct 6224 was insoluble and appeared to have a non-processed CTD (magenta x). **B.** SDS-PAGE of purified 6168, 6223, 6336, and 6338 and corresponding Western Blot analysis. Strep-tag II (green channel), M – marker, T – temperature, „+” – denatured samples.

We performed dot blots with strains obtained from liquid bacterial cultures to determine the applicability of 6338 and Strep II-iKfIB IR800 for detecting heparosan on bacterial surfaces. More specifically, we included *E. coli* K5 as a natural heparosan-producing strain and an engineered *E. coli* BJ derivative (Leroux & Priem, 2016) carrying pBAD33-kfiCAB (5744, see Supplementary **Table S2**) and pBBR3-kfiD (5691-tetR) to simulate a directed evolution experiment (Keys, Berger, et al., 2012). *E. coli* BJ was included as a control, and *E. coli* K1, a polysialic acid-producing strain, was used to exclude the possibility of unspecific binding to a negatively charged

capsule polymer. Both constructs 6338 and Strep II-iKfIB IR800 could be exploited to detect heparosan on *E. coli* harvested from bacterial cultures (**Fig. 13**), while both controls clearly remained negative.

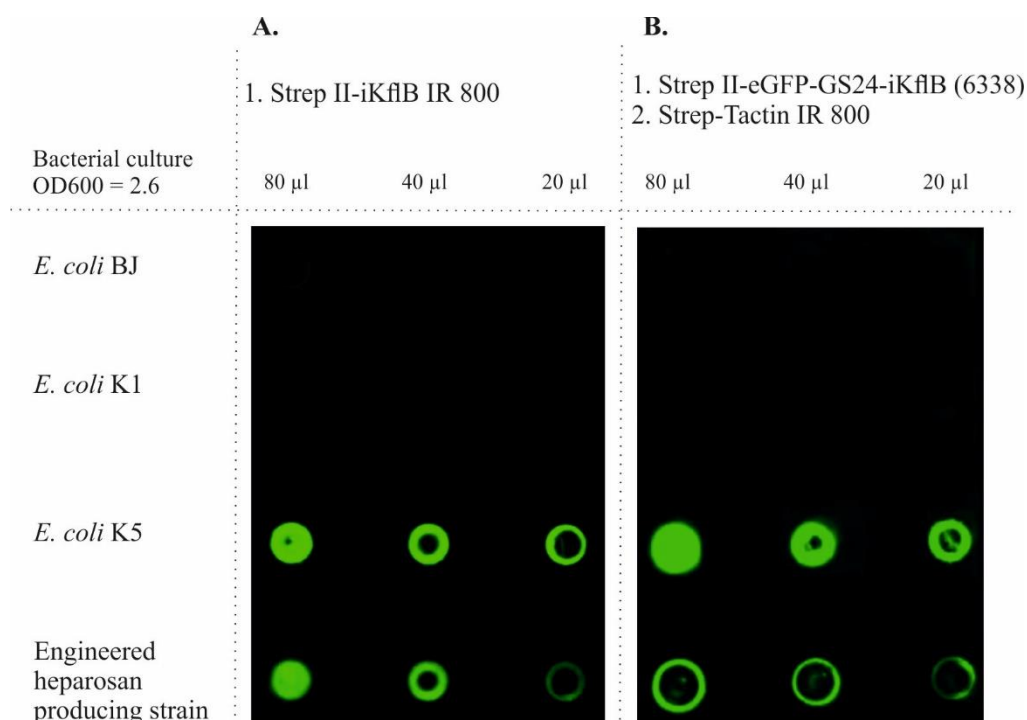


Figure 13. KfIB as detection agent for heparosan on bacterial surfaces. A. Direct detection of heparosan using fluorescently labeled KfIB. B. Indirect detection via Strep-Tactin.

2.3. Conclusions

The presented comprehensive study confirmed that *E. coli* K5 heparosan synthases and the phage lyase can broaden the tool for heparosan synthesis, depolymerization, and detection. Using a fluorescently labeled K5 acceptor, we characterized the elongation mechanism of heparosan synthases. Moreover, the same acceptor was used to prepare a substrate, leading to the KfIB degradation mode elucidation. Finally, we have transformed KfIB into heparosan-specific lectins. This precious agent lays the foundations for directed evolution studies with KfiCA and bacterial serotyping.

Interestingly, KfIB catalyzes partial degradation of heparan sulfate. Further studies are needed to confirm that it takes place in a non-sulfated domain of a chain. This could extend the use of KfIB-based artificial lectins as recognition and quantification agents for unsulfated domains, for example in tissues.

2.4. Materials and methods

2.4.1. Materials

Enzymes were purchased from Thermo Fisher Scientific and New England Biolabs GmbH. Recombinant DNA was isolated with QiAprep Spin Miniprep Kit (Qiagen Inc.). An overview of bacterial strains used in this study is given in Supplementary **Table S1**. The antibiotics used for bacterial cultures were: ampicillin (100 µg/ml), carbenicillin (50 µg/ml), chloramphenicol (20 µg/ml), kanamycin (50 µg/ml), and tetracycline (15 µg/ml). The *E. coli* K5-specific ΦK5B was kindly provided by Prof. Dr. Dieter Bitter-Suermann (Hannover Medical School) and sequenced at the Institute of Clinical Biochemistry. The sequence of KfIB is provided in Supplementary **Fig. S4**. All primers used for cloning are listed in Supplementary **Tables S2, S3**.

2.4.2. HPLC-based anion-exchange chromatography (HPLC-AEC)

HPLC-AEC was performed on a Prominence UFLC-XR system (Shimadzu) using a CarboPac PA-100 column (Dionex™; 2 x 250 mm). Separation of the analytes was performed at 50°C at a flow rate of 0.6 ml/min using ultrapure water and 1 M NaCl as the mobile phase. The details of the elution conditions are described for each relevant experiment. All the samples from bacterial lysates were filtrated using Ultrafree™-cetrifugal filter units (Millipore; 0.22 µm; 16,000 x g, 5 min) prior to the HPLC-AEC analysis.

2.4.3. Cloning

*Engineered heparosan synthases*__To generate the pBAD33-based plasmid expressing untagged KfiCA under one promoter, first *kfiC* from pEX-A258-*kfiCs* (Eurofins Genomics, see Supplementary **Fig. S2**), and then *kfiA* from pBS-*kfiA* (Priem et al., 2017) were inserted into pBAD33 via KpnI/HindIII and XbaI/PstI, respectively, yielding pBAD33-*kfiCA* (6225, see Supplementary **Table S2**). To obtain pBAD33-*kfiCAB* (5744), *kfiB* from pBBR1-MCS3-*kfiB* (unpublished) was cloned via XhoI, PvuI into pBAD33-*kfiCA*. The pBBR1-MCS3-*kfiB* itself was obtained from the laboratory; it was constructed from SpeI, SacI digestion of pBBR-*glcAT-kfiDB* (Priem et al., 2017) and cloning of the *kfiB* fragment into SpeI/SacI sites of pBBR-MCS3 (Kovachs et al., 1995). A plasmid encoding KfiD (pBBR1-MCS3-*kfiD*) was obtained

from (Leroux & Priem, 2016) and labeled as 5691-tetR. The tetracycline resistance of pBBR1-MCS3-kfiD was changed into a kanamycin (5691-kanR) resistance using restriction-free cloning (Bond & Naus, 2012) and the primers MS18 and MS19 (Supplementary **Table S3**), to allow expression in *E. coli* BL21-Gold (DE3).

Tagged KfiA and KfiB proteins intended for purifications were generated as follows: *kfiA* and *kfiB* were amplified by PCR from pBS-*kfiA* (Priem et al., 2017) and pBBR1-MCS3-kfiB (Barreteau et al., 2012) using primers MS05, MS02 and MS10, MS11 (Supplementary **Table S2** and **S3**), respectively. Via restriction-free cloning (Bond & Naus, 2012), the obtained PCR products coding for KfiA and KfiB were cloned into modified versions of pMBP-S3N10-*csxA*-His₆ (tac) (Fiebig et al., 2014). In the resulting construct pMal-c-His₆-*kfiA* (5644) and pMal-c-MBP-S3N10-preScission-*kfiB*-His₆ (5668), KfiA is N-terminally fused to a His₆ tag, and KfiB is N-terminally linked via a PreScission cleavage site and an S3N10 linker to a maltose binding protein (MBP), and C-terminally fused to a His₆ tag. For improved solubility, KfiC was fused with a trigger factor (TF) to create a construct as previously described (Leroux & Priem, 2016; Sugiura et al., 2010). TF was cloned via NdeI/BamHI from pBS-His₆-TF (Leroux & Priem, 2016) into pET28a. Subsequently, KfiC from pEX-A258-*kfiCs* (Eurofins Genomics) was cloned via BamHI/HindIII into pET28a-His₆-TF vector, yielding pET28a-His₆-TF-kfiC (5822).

*Engineered K5 lyase B and its mutants*__The gene encoding KflB was amplified using the primers MM315 and MM317 (Supplementary **Table S2**). The amplicon was inserted via BamHI/XhoI into pStrep-D245*EndoNF*-His₆ (Schwarzer et al., 2009), replacing D245*EndoNF*. The start codon (ATG) at nucleotide position 1-3 of *kflB* was removed via restriction-free cloning (Bond & Naus, 2012) with the primers MS49 and MS50 (Supplementary **Table S2**) generating pStrep-*kflB*-His₆ (6118). pStrep-*kflB*-Y229A-His₆ (6119) and pStrep-*kflB*-S505A-His₆ (6122) were generated using site-directed mutagenesis with primers DS210, DS211 and MR01, MR02, respectively, and 6118 as a template (Supplementary **Table S2** and **S3**). Single alanine substitutions in 6112, 6113, and 6114 (F202, E206, K208A) were introduced according to Liu *et al.* (Liu & Naismith, 2008) using the primers indicated in Supplementary **Table S2** and **S3** and Strep-*kflB*-His₆ (6118) as DNA template. Additional mutations were generated using the single mutants as a DNA template. The sequence encoding eGFP-GS24 was amplified from pET22b-Strep II-*eGFP*-GS24- Δ N-*endoNF* (Kiermaier et al., 2016)

with primers MS73 and MS74 (Supplementary Table S3) and cloned into 6168 using the RF-cloning protocol (Bond & Naus, 2012) generating 6338. The ALFA-tag (Götzke et al., 2019) was introduced by restriction-free cloning into Strep-*kflB*-E206A, K208A-His₆ using the primers indicated in Supplementary **Table S2**. All plasmids were propagated in *E. coli* XL1 Blue, and the sequence identity of the coding sequences was confirmed by Sanger sequencing (Eurofins Genomics).

2.4.4. Heparosan preparation

Heparosan production in high cell density cultures__Heparosan was expressed using *E. coli* BJ transformed with a plasmid expressing KfiCAB (5744) and KfiD (5691-tetR, see Supplementary **Table S2**), respectively, as previously described by Barreteau *et al.* (Barreteau et al., 2012) with the following changes: The fed-batch phase started with the addition of the continuous feeding solution (50% glycerol and 0.5% L-arabinose. After 30 min, isopropyl 1-thio- β -D-galactopyranoside (IPTG) was added at a final concentration of 0.2 mM.

*Heparosan fractionation*__Heparosan (100 mg/ml) was dissolved in water and loaded on a HiTrap Q HP column (Cytiva) precalibrated with water. Heparosan was eluted with 1 M NaCl using a linear gradient from 0-480 mM NaCl over 40 column volumes. Fractions were collected and analyzed on 15% PAGE stained with alcian blue/silver. Fractions, as indicated in Supplementary **Fig. S6**, were pooled, lyophilized, dialyzed (Zellu Trans, Roth, 10 kDa MWCO) against water, and lyophilized again.

2.4.5. Preparation of the fluorescently labeled K5 oligosaccharide acceptor

*Production of intracellular K5-derived oligosaccharides*__Generation, extraction, and purification of heparosan oligosaccharides were performed according to a previously published protocol (Barreteau et al., 2012). The acidic supernatant was neutralized with the Amberlite FPA66 anion-exchange free base (Sigma-Aldrich), and the centrifugation step (30 min, 7000 rpm) was repeated to remove the resin. Generated oligosaccharides were detected by a thin layer chromatography-based protocol developed by Z. Zhang *et al.* (Z. Zhang et al., 2007). In the final step, oligomers were separated by size exclusion chromatography with a Hiload Superdex S30x3 column (26 x 600 mm, GE Healthcare) with 0.1 M carbonate ammonium at a flow rate of 1.2 ml/min as a mobile phase. The appropriate fractions of oligomers were pooled,

concentrated *in vacuo*, and lyophilized, yielding heparosan oligomers with a degree of polymerization (DP) from 2 to 12.

*Fluorescent labeling of the K5 oligosaccharide acceptor*__Heparosan decamer **1** underwent treatment with mercury (II) acetate as described by Ludwigs *et al.* (Ludwigs *et al.*, 1987). The solution was passed through a column filled with Dowex 50W-X8 (H⁺ form; Sigma-Aldrich). The filtrate was lyophilized to give a white, solid product **2**. According to the literature (France *et al.*, 2000), nonamer **2** underwent reductive amination, with the following changes: the crude product was precipitated with cold acetone and centrifuged (4 °C, 20 min, 4000 g). The supernatant was discarded, and residues were neutralized with 1 M HCl for 1 hour at room temperature under stirring. The reaction mixture was concentrated *in vacuo*. The crude product was purified by silica chromatography (acetonitrile/water = 9:1 → 0:1) to give a yellowish solid **3**.

2.4.6. Protein Expression and Purification

*Expression and purification*__Overnight bacterial cultures (*E. coli* M15[pREP4] for 5644, 5668, BL21(DE3) for 5822, and KflB wt and all mutants) were grown in PowerBroth (Athena Enzyme System) supplemented with carbenicillin (pKfiA, pKfiB, and all KflB constructs) or kanamycin (pKfiC) at 37 °C with shaking until an optical density (OD₆₀₀) of ~1.0 was reached. After cooling at 4°C for app. 30 min, expression of pKfiC, A, B and KflB constructs was induced by 0.1 mM IPTG and 0.5 mM IPTG, respectively. Bacteria were harvested after overnight incubation at 15 °C. Cells expressing pKfiCAB were lysed by sonication in Buffer A (50 mM Tris, pH 8.0, 500 mM NaCl), while KflB proteins were sonicated in Buffer W (100 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA). Both buffers contained EDTA-free protease inhibitors (complete EDTA-free, Roche). The resulting lysates were centrifuged for 3 min at 16,000 x g. Soluble fractions containing pKfiCAB were loaded on His Trap HP columns (1 ml; GE Healthcare) and gradually eluted with Buffer B (50 mM Tris, pH 8.0, 500 mM NaCl, 500 mM imidazole; linear elution with 50-500 mM imidazole over 20 min). Recombinant KflB proteins were purified using StrepTactin columns (IBA) and eluted according to the instruction provided by the manufacturer.

Size-Exclusion Chromatography (SEC)—Proteins were concentrated by centrifugation using centrifugal devices (Amicone, Millipore). The concentrated sample was loaded onto a Superdex 200 HR 10/30 column (Amersham Biosciences). Column equilibration and elution were performed with 10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1 mM DTT at 0.5 ml/min. In addition, SEC of Strep-KfIB-E206A, K208A intended for fluorescent labeling was performed in phosphate-buffered saline (PBS).

2.4.7. Elongation of the K5 oligosaccharide acceptor – fluorescent heparosan

Non-tagged KfiCA were expressed from 6225 in BL21(DE3) grown in PowerBroth with chloramphenicol selection (**Fig. 5B**). KfiCAB, D and KfiCA, D were expressed from 5744, 5691-kanR, and 6225, 5691-kanR, respectively, in BL21-Gold (DE3) grown in PowerBroth with chloramphenicol and kanamycin selection. The procedure was similar to the expression of tagged proteins, with the following changes: expression was induced at $OD_{600} = 0.5$ with L-arabinose (Roth) at the final concentration of 0.1%. A 1 ml aliquot of bacterial culture with $OD_{600} = 2.0$ was centrifuged and the obtained pellet was washed four times with buffer A [100 mM Bis-Tris (Sigma), pH 6.5, 5 mM NaCl]. Next, the pellet was resuspended in buffer A containing 0.1 mg/mL lysozyme (Serva) and EDTA-free protease inhibitors (complete EDTA-free, Roche). After 10 minutes of incubation in the lysis mixture, cells were lysed by sonication. The soluble fraction was used in *in vitro* assays. To elongate 3 μ M (**Fig. 5B**) or 10 μ M (**Fig. 5C**) of the K5 oligosaccharide acceptor, 90 μ L (**Fig. 5B**) or 10 μ L (**Fig. 5C**) of expressed proteins were used in a total volume of 1 mL (**Fig. 5B**) or 60 μ L (**Fig. 5C**) in the presence of 3 mM (**Fig. 5B**) or 1 mM (**Fig. 5C**) uridine diphosphate *N*-acetylglucosamine [UDP-GlcNAc (Sigma-Aldrich)] and uridine diphosphate D-glucuronic acid [UDP-GlcA (Sigma-Aldrich)], 1 mM dithiothreitol [DTT (Sigma)] and 5% glycerol (Sigma) in a buffer comprising 200 mM Bis-Tris pH 6.5, 10 mM NaCl and 10 mM $MnCl_2$ (**Fig. 5C**). In agreement with previous studies (Sugiura et al., 2010), adding 10 mM $MgCl_2$ was tested but did not considerably alter enzyme activity (**Fig. 5B**). The reaction was incubated at 37 °C for the indicated time points. Enzymes were inactivated by incubation at 70 °C for 20 min, and the precipitate was removed by centrifugation (16,000 x g, 5 min). The supernatants samples (25 μ L, **Fig. 5B** or 20 μ L, **Fig. 5C**) were injected into the HPLC-AEC injector coupled to a fluorescence detector (Exc. $\lambda = 330$ nm, Emi. $\lambda = 420$ nm) and analyzed in a 20% -3 curved gradient from 3 to 8 min, followed by a linear gradient

of 52% over the next 39 min (**Fig. 5B**) or in a 60% linear elution gradient from 3 to 42 min (**Fig. 5C**).

In vitro reactions with purified, tagged proteins (pKfiC, pKfiA and if indicated pKfiB) were performed as described for non-tagged proteins with the following changes: 2 μM of pKfiCA(B) (5-7 μL in total) were used in a total volume of 60 μL (**Fig. 5C**). The K5 oligosaccharide acceptor was used in a concentration of 10 μM , the donor substrates in a concentration of 1 mM each. The reaction buffer contained 50 mM Bis-Tris pH 6.5, 10% glycerol (if indicated, see Supplementary **Fig. S3C**), and 10 mM MnCl_2 . The resulting products (35 μL) were analyzed using HPLC-AEC in a 60% linear gradient from 3 to 42 min.

2.4.8. K5 lyase B activity assays

Activity__KfIB wt or KfIB mutants (0.2-2 μM) were incubated with 10 $\mu\text{g}/\mu\text{L}$ of heparosan in assay buffer (25 mM Tris-HCl, pH 8.0, 50 mM NaCl) with DTT at the final concentration of 1 mM at 37 °C. Samples (18 μL) were taken as indicated (**Fig. 7B, D**) and immersed in liquid nitrogen to inhibit reaction. Analysis was carried out using HPLC-AEC coupled to UV detection ($\lambda = 250 \text{ nm}$) in a 16% linear gradient from 3 to 10 min, followed by a -1 curved gradient of 30% over the next 30 min and a +1 curved gradient of 50% over the next 6 min. Moreover, samples were analyzed via alcian blue/silver staining after PAGE (**Fig. 7A, C**).

Fluorescently labeled heparosan was produced by KfiCA, and the enzymes were heat-inactivated as described above. KfIB wt (1 nM, 3.9 μL of the stock solution) was added to the inactivated heparosan synthesis reaction mixture (295 μL), and water was supplemented to a total volume of 300 μL . The reaction mixture was incubated at 37 °C, samples (25 μL) were taken at the indicated time points and analyzed via HPLC-AEC coupled to a fluorescence detector (Exc. $\lambda = 330 \text{ nm}$, Emi. $\lambda = 420 \text{ nm}$) and analyzed in a 20% -3 curved gradient from 3 to 8 min, followed by a linear gradient of 52% over the next 39 min.

The minimal length of a KfIB substrate__Heparosan oligosaccharides (DP = 4, 6, 8, 10) with a concentration of 10 $\mu\text{g}/\mu\text{L}$ were incubated at 37 °C overnight in assay buffer with 1 μM KfIB wt in a total reaction volume of 80 μL . Samples (4 - 8 μL) analyzed via HPLC-AEC coupled to UV detection ($\lambda = 232 \text{ nm}$) in a 30% linear gradient from 3 to 14 min.

KflB specificity determination—The reaction mixture contained 1 mg/ml of a substrate as indicated [hyaluronic acid (Dextra), heparan sulfate (Iduron), heparosan, chondroitin sulfate (Sigma Aldrich), dermatan sulfate (Iduron)] and was started upon adding 4.5 μ M KflB wt in a total volume of 25 μ L at 37 °C. Samples were taken after half an hour, two hours, and overnight incubation. The reactions were stopped by immersion in liquid nitrogen, and polymers were visualized by alcian blue/silver staining after PAGE.

2.4.9. Labeling of an inactive KflB

Purified Strep-KflB-E206A, K208A (6168) (73.3 μ l, 3.13 nmol, 1 equiv.) in PBS was mixed with IR Dye 800CW NHS Ester (1.1 μ l, 9.4 nmol, 3 equiv.; LI-COR Biosciences) in DMF and incubated at 4 °C for 15 hours. The non-reacted dye was removed with Zeba™ Dye and Biotin Removal Spin Columns (ThermoFisher Scientific) according to the manufacturer's guidelines. The obtained fluorescently labeled protein is referred to as Strep-iKflB IR800.

2.4.10. Dot blots

A sample of the engineered heparosan-producing strain was prepared as described in 2.4.4. *E. coli* K1, *E. coli* K5, and *E. coli* BJ (Supplementary **Table S1**) were cultivated in PowerBroth overnight at 37 °C without antibiotic selection and with tetracycline, respectively. Liquid bacterial cultures ($OD_{600} = 2.6$) were transferred (80 μ L, 40 μ L and 20 μ L) onto a nitrocellulose membrane and air-dried. Subsequently, membranes were incubated in Intercept® (PBS) Blocking Buffer (BB, LI-COR Biosciences) and PBS (1:2) at 4 °C overnight. Heparosan was directly detected using Strep-iKflB-IR800 [1:3000 in BB:PBS, (1:2)]. If Strep-eGFP-GS24-iKflB (6338) [20 μ g/ml in BB:PBS, (1:2)] was used for detection, additional incubation steps with Avidin [2 μ g/ml in BB:PBS, (1:2) for 10 min; (Sigma-Aldrich) followed by addition of Strep-Tactin IR 800CW (1:2000) was performed. Each incubation was performed for 1 h unless stated otherwise at room temperature with orbital shaking at 13 rpm. The membrane was washed four times with 0.1% Tween® 20 in PBS after each incubation with detection agents. For visualization, membranes were scanned using an Odyssey® infrared imaging system (LI-COR Biosciences).

2.5. Footnotes

The abbreviations: Φ , bacteriophage; Δ , unsaturated uronic acid at the non-reducing end; λ , wavelength; A, alanine; 2AB, 2-aminobenzamide; AEC, anion-exchange chromatography; Bis-Tris, Bis-Tris methane; CPS, capsular polysaccharide; CTD, C-terminal chaperone domain; DNA, deoxyribonucleic acid; DP, degree of polymerization; DTT, dithiothreitol; *E. coli*, *Escherichia coli*; F, phenylalanine; GAG, glycosaminoglycan; GlcA, glucuronic acid; GlcNAc, *N*-acetylglucosamine; HPLC, high-performance liquid chromatography; IPTG, isopropyl β -D-1-thiogalactopyranoside; kanR, kanamycin resistance; kDa, kilo Dalton; KfiA, *N*-acetylglucosaminyltransferase; KfiC, D-glucuronyltransferase; KfiD, UDP-Glc-dehydrogenase; KfiA, K5 lyase A; KfiB, K5 lyase B; M, mol per liter; MBP, maltose-binding protein; MS, mass spectrometry; MW, molecular weight; MWCO, molecular weight cut-off; NA, highly sulfated domain of heparan sulfate; NA/NS, intermediately sulfated domain of heparan sulfate; NS, nonsulfated domain of heparan sulfate; OD, optical density; PAGE, polyacrylamide gel electrophoresis; PBS, phosphate-buffered saline; S, serine; SDS, sodium dodecyl sulfate; SEC, Size-Exclusion Chromatography; tetR, tetracycline resistance; TF, trigger factor; TSP, tailspike protein; Tris, tris(hydroxymethyl)aminomethane; Tween[®] 20, polyethylene glycol sorbitan monolaurate; UDP, uridine diphosphate; UDP-GlcA, uridine diphosphate glucuronic acid; UDP-GlcNAc, uridine diphosphate *N*-acetylglucosamine; UV, ultraviolet; wt, wildtype, Y, tyrosine

SUPPLEMENTARY INFORMATION

Extending the enzymatic toolbox for heparosan polymerization, depolymerization, and detection

Małgorzata Sulewska^{ab}, Monika Berger^a, Manuela Damerow^a, David Schwarzer^a,
Martina Mühlenhoff^a, Rita Gerardy-Schahn^a, Bernard Priem^b, Timm Fiebig^a

^aInstitute of Clinical Biochemistry, Hannover Medical School, Hannover, Germany

*^bCentre de Recherche sur les Macromolécules Végétales, Groupe Chimie et
Biotechnologie des Oligosaccharides, 601 rue de la Chimie, BP 53X, 38041
Grenoble, Cedex 09, France*

TABLE OF CONTENTS

SUPPLEMENTARY TABLES

Supplementary Table S1. Overview of bacterial strains used in this study.

Supplementary Table S2. Overview of plasmids and recombinant proteins used in this study.

Supplementary Table S3. Primers used in this study.

Supplementary Table S4. Protein expression, purification and obtained yield.

SUPPLEMENTARY FIGURES

Supplementary Figure S1. Analysis of the K5 oligosaccharides by mass spectrometry.

Supplementary Figure S2. The coding sequence of KfiC used in this study.

Supplementary Figure S3. Characterization of heparosan synthases.

Supplementary Figure S4. The K5 lyase B sequence.

Supplementary Figure S5. Coomassie-stained polyacrylamide gels of KflB and its mutants (final pools).

Supplementary Figure S6. Heparosan fractionation.

Supplementary Figure S7. Heparosan degradation catalyzed by KflB wt and KflB S505A.

Supplementary Table S1. Overview of bacterial strains used in this study.

Strain	Relevant characteristics	Purpose	Source or references
<i>Escherichia coli</i> K1	neuDBACES kpsFEDUCSTM	Polysialic acid production	Frosch <i>et al.</i> , 1985; Daines <i>et al.</i> , 2000
<i>Escherichia coli</i> K5	kfiACBD kpsFEDUCSTM	Heparosan production	Schwarzer <i>et al.</i> , 2007; Silver <i>et al.</i> , 2001
<i>Escherichia coli</i> HB13 (K-12 derivative)	Δ wcaJ, pBBR1-kfiAB, pSU2718-kfiCD, pWSK29-elmA	K5 oligomers production	Barreteau <i>et al.</i> , 2012
<i>Escherichia coli</i> BJ (BL21 derivative)	Δ wcaJ kpsFEDUCSTM	Heparosan production	Leroux & Priem, 2016
<i>Escherichia coli</i> BL21(DE3)	lon ompT (λ DE3) dcm	Protein expression	Novagen
<i>Escherichia coli</i> XL1-Blue	recA1 endA1 tetR	Cloning	Invitrogen
<i>Escherichia coli</i> M15[pREP4]	recA+ uvr+ kanR	Protein expression	Diagen

Supplementary Table S2. Overview of plasmids and recombinant proteins used in this study.

Protein	Accession number	Identifier	Recombinant construct	Cloning primers	Vector
KfiA	Q47332	5644	His ₆ -KfiA (pKfiA)	MS03, MS04	pMal-c
KfiB	TEZ99055.1	5668	MBP-S3N10-Pre-scission-KfiB-His ₆	MS10, MS11	pMal-c
KfiD	CAE55820.1	5691-tetR	KfiD-tetR	-	pBBR1-MCS3
KfiD	CAE55820.1	5691-kanR	KfiD-kanR	MS18, MS19	pBBR1-MCS3
KfiC KfiA		6225	KfiCA	-	
KfiC KfiA KfiB		5744	KfiCAB	-	pBAD33
KfiC		5822	His ₆ -TF-KfiC (pKfiC)	-	pET28a
KfiB	This study	6112	Strep-KfiB-F202A-His ₆	MS43, MS44	Modified pET22b (Schwarzer <i>et al.</i> , 2007)
		6113	Strep-KfiB-E206A-His ₆	MS45, MS46	
		6114	Strep-KfiB-K208A-His ₆	MS47, MS48	
		6118	Strep-KfiB-His ₆	MM315, MM317; MS49, MS50	

	6119	Strep-KfIB- Y229A-His ₆	DS210, DS211; MS49, MS50
	6122	Strep-KfIB- S505A-His ₆	MR01, MR02; MS49, MS50
	6167	Strep-kfIB- F202A,E206A -His ₆	MS57, MS58
	6168	Strep-KfIB- E206A,K208 A-His ₆	MS59, MS60
	6169	Strep-KfIB- E206A,Y229 A-His ₆	MS45, MS46
	6170	Strep-KfIB- E206A,S505A -His ₆	MS45, MS46
	6223	Strep-ALFA- KfIB- E206A,K208 A-His ₆	MS65, MS66
	6224	Strep-KfIB- E206A,K208 A-ALFA- His ₆	MS61, MS62
	6336	Strep-ALFA- GS24-KfIB- E206A,K208 A-His ₆	MS69, MS70
eGFP- KfIB	6338	Strep-GFP- GS24-KfIB- E206A,K208 A-His ₆	MS73, MS74

Supplementary Table S3. Primers used in this study.

Primer	Sequence (5' to 3')	Source
Cloning primers		
DS210	gcagattgccagcatgtaactgccaacggcactgaaggg	This study
DS211	cccttcagtgccgttgccagttacatgctggcaatctgc	
MS03	ccaccaccaccaccacggatccattgttgcaaatatgcatcataaccacc	
MS04	cggccagtgccaaagctttacctaggtattccacattatacactaattcgag gttaagttttg	
MS10	ctggaggtgctgtttcaaggtccgaataaattagtgctagtcggacatcctg	
MS11	gtggtggtggtggtgctcgaggcccttgattttagctctcctatattataaac	
MS18	gtagtttatacacgttaaattgctaacgcagtcaggcacatgagccatattc aacgggaaacg	
MS19	ggtgcccccggcttccatttagaaaaactcatcgagcatcaaatgaaactg	
MS43	ctcagggcggagcagtagagttgaaaggtacagcca	
MS44	ctccgcctgaggtagttcttagcaatgacgttatc	
MS45	gcagtagcgttgaaaggtacagccagttacaacatagtcagtaatg	
MS46	ctttcaacgctactgctccgaactgaggtagttcttag	
MS47	gtggcaggtacagccagttacaacatagtcagtaatgttataggg	
MS48	gctgtacctgccaaactctactgctccgaactgag	
MS49	cctggtccgcgtggatccgctaaattaaccaaactaataactgaag	
MS50	gctgctttattagtagcatatccggaatagctacctcgaatgccttfaatcata agtcc	
MS57	aactaccctcaggccggagcagtagcgttgaag	
MS58	ggcctgaggtagttcttagcaatgacgttatctatgag	
MS59	cagtagcgttggcaggtacagccagttacaacatagtc	
MS60	tgccaacgctactgctccgaactgagggtagttc	

MS61	ggttttactcaatcagcattcactgtgtgccctcccagtcgaagaggaatt gcgccg	
MS62	aggctctgtttatccctagcatctgaaggctcggcagcctacggcgcaat tcctct	
MS65	ctggttccgcgtggatcccctcccagtcgaagaggaattgcgccgtag	
MS66	aagattccttcagtattaggtttggttaatttagcaggctcggcagcctacg gcgcaat	
MS69	gtaggttcgaatggcggattaagcccctcccagtcgaagaggaattgc gccgtag	
MS70	agaacgcctgaaggagtgacataggctcggcagcctacggcgcaatt cctct	
MS73	cctggttccgcgtggatccggtaaaggagaagaactttcactggag	
MS74	gcaagattccttcagtattaggtttggttaatttagcggccgctgaagacc g	
MR01	gcattcactgtgtcgcagatgctagggataaaacag	
MR02	ctgtttatccctagcatctgcgacaacagtgaatgc	
MM315	cgggatccatggctaaattaaccaaacc	
MM317	gtccgccgagcttaggcaggaagctag	
Sequencing primers		
pQE_FP	cggataacaatttcacacag	Eurofins Genomics
M13_FP	tgtaaacgacggccagt	Eurofins Genomics
MS51	ccatcggatgatgctg	This study
DS212	gctacagtgctacagg	This study
DS213	cctgtagcactgtagc	This study

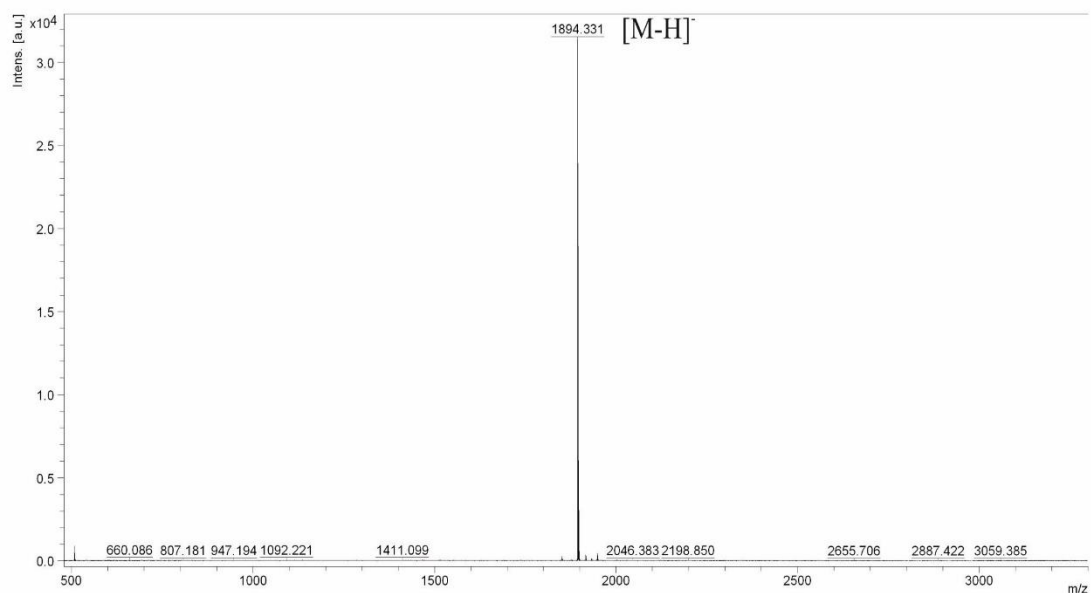
Supplementary Table S4. Protein expression, purification and obtained yield.

Identifier	Expression strain	MW of mature protein, kDa	Affinity chromatography system	Yield, mg/L
5644	M15[pREP4]	28.3	His-tag	1.2
5668		109.2		5.2
5822	BL21(DE3)	110.6		12.8
6112		54.0	Strep-tag	10.8
6113		54.1		13.5
6114		54.1		15.8
6118		54.1		20.0
6119		54.0		22.0
6122		69.8		9.0
6167		54.0		22.0
6168		54.0		18.6
6169		54.0		11.4
6170		69.8		18.7
6223		55.8		5.8
6224		55.8*		Not obtained
6336		57.7		8.9
6338		83.0		4.9

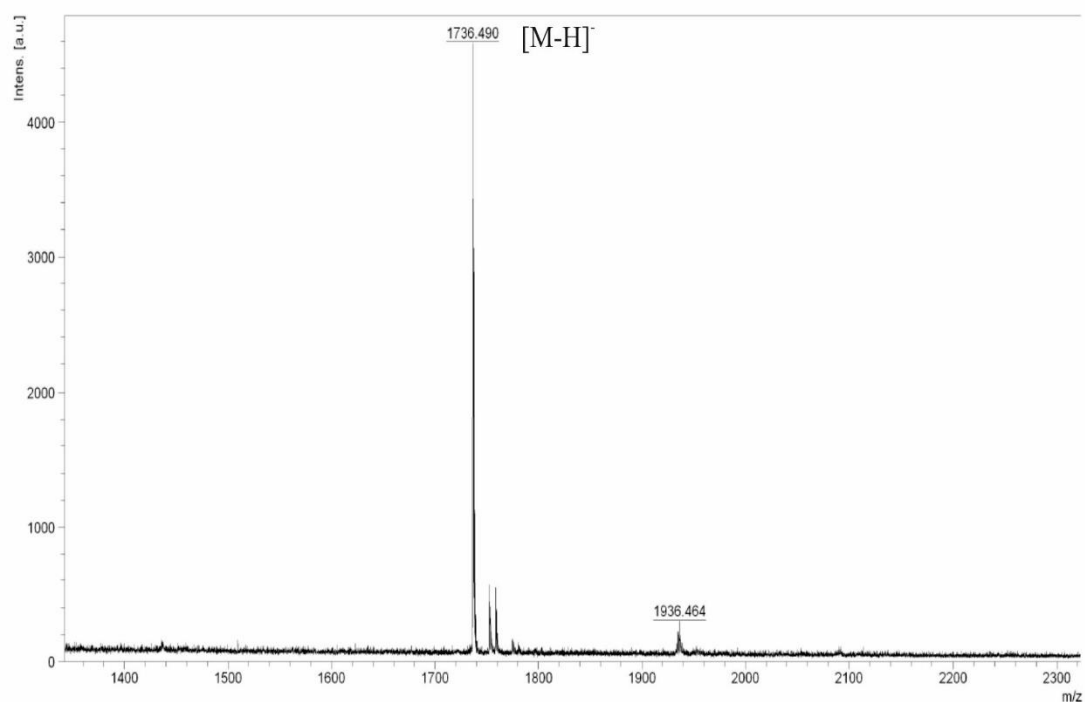
* MW expected after proteolytic cleavage. However, the obtained data (not shown) suggest it did not occur.

Supplementary Figure S1. Mass spectrometry analysis of the K5 oligosaccharides. The major molecular ions associated with the isolated products: decamer (**1**), unlabeled nonamer (**2**), octamer, hexamer, tetramer, 2-aminobenzamid labeled nonamer (**3**) had mass to charge ratios (m/z) of **A.** 1894.331 ($[M-H]^-$), **B.** 1736.49 ($[M-H]^-$), **C.** 1516.218 ($[M]^+$), **D.** 1136.851 ($[M-H]^-$), **E.** 758.164 ($[M]^+$), **F.** 374.906 ($[M+NH_3]^{5+}$), respectively.

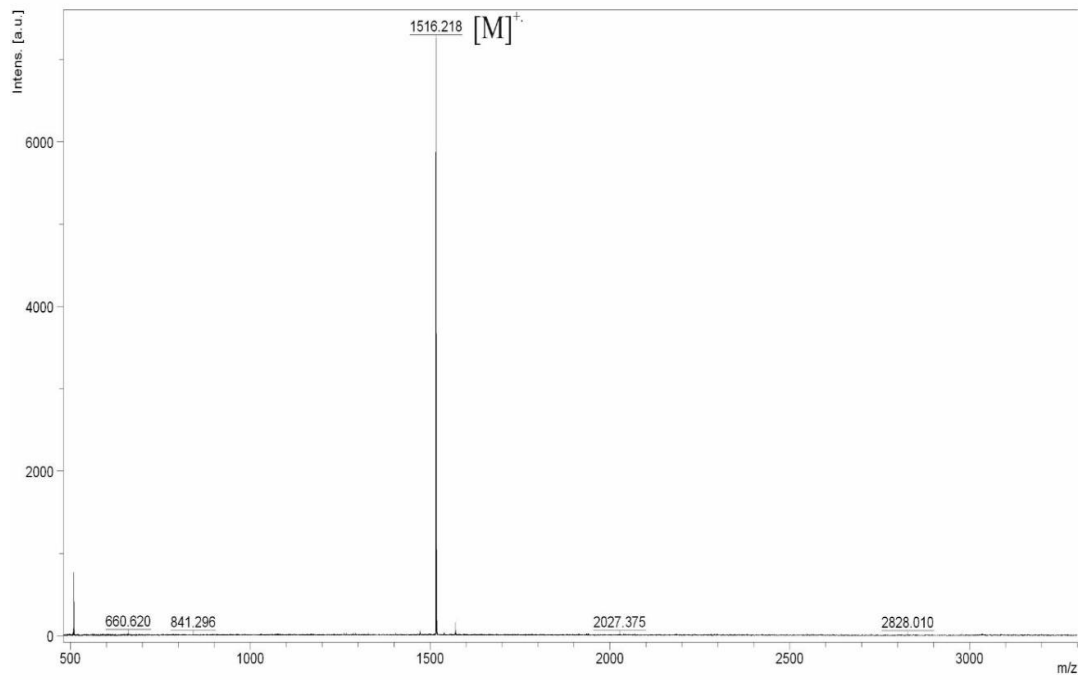
A.



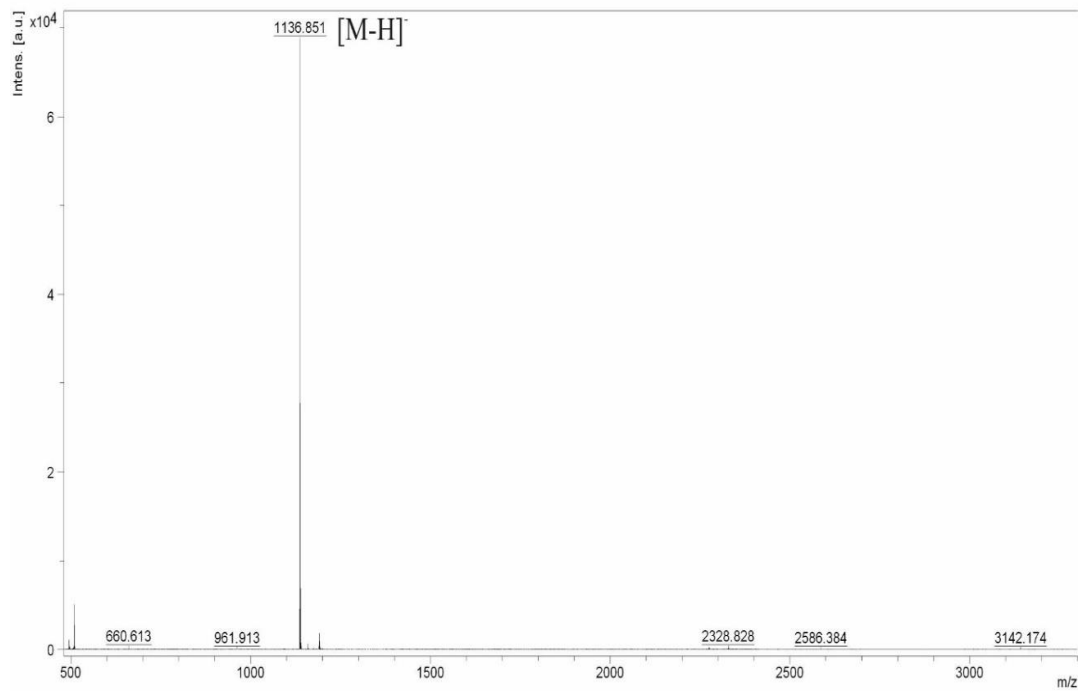
B.



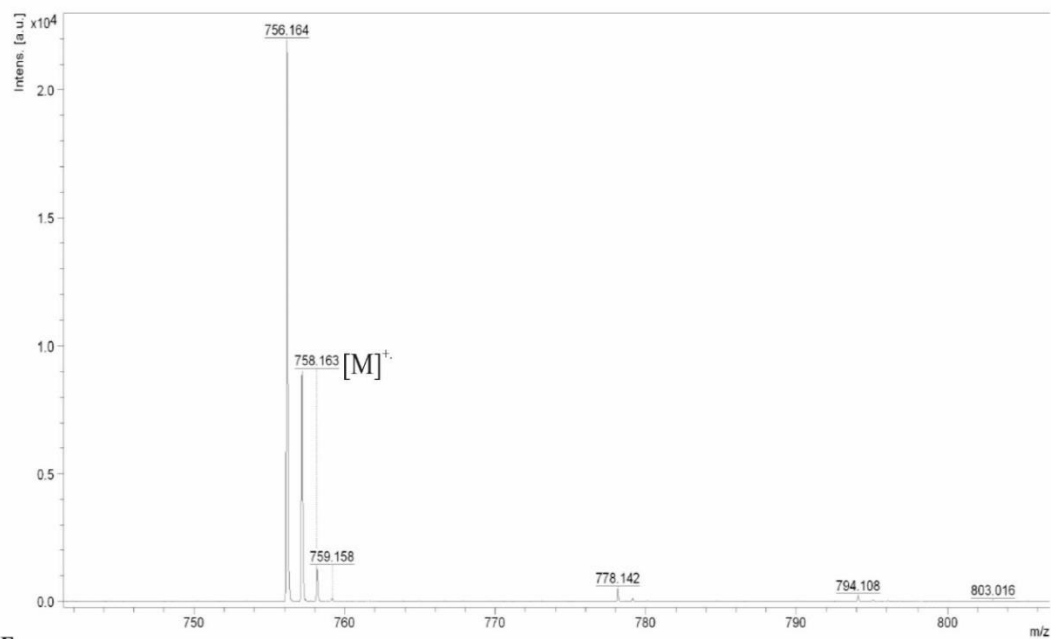
C.



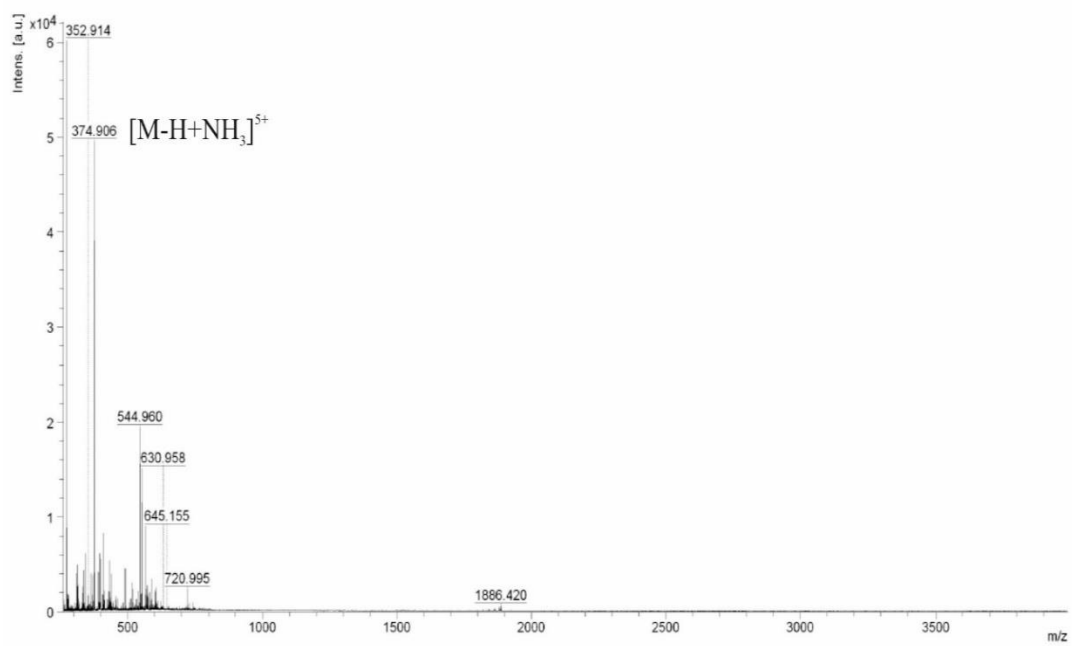
D.



E.



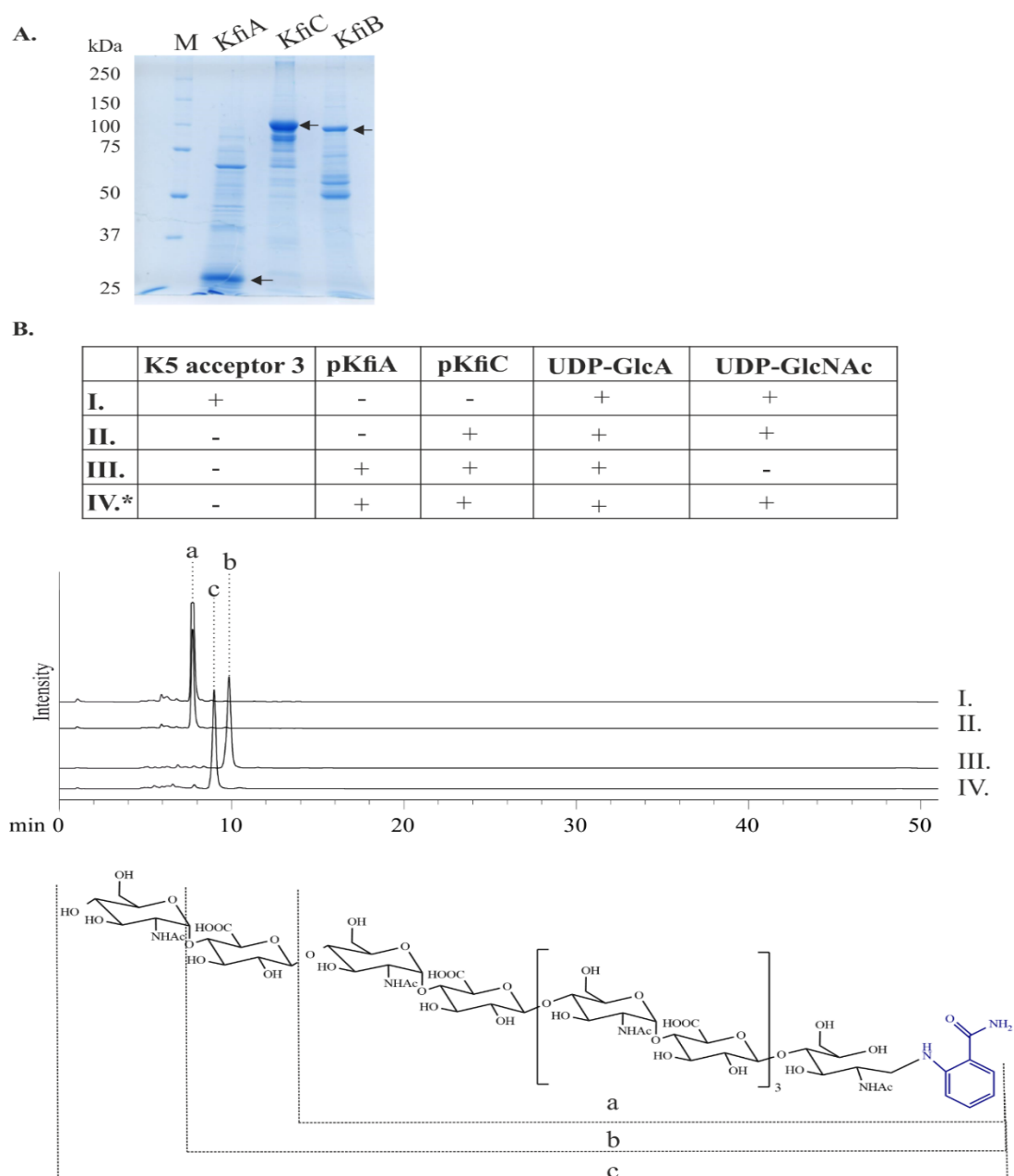
F.



Supplementary Figure S2. The coding sequence of KfiC used in this study. The synthetic version of *kfiC* was designed to remove unwanted restriction sites, and add cloning sites (underlined). *KfiC* start and stop codon are in bold.

GGTACCATAAGAAGGAGATATACCATGGATCCTATGAACGCAGAATATATAAAT
TTAGTTGAACGTAAAAAGAAATTAGGGACAAATATTGGTGCTCTTGATTTTTTAT
TATCAATTCATAAGGAGAAAGTTGATCTTCAACATAAAAACTCGCCTTTAAAAG
GTAACGATAACCTTATTCACAAAAGAATAAACGAATACGACAATGTA^{ATG}ACTTGAAC
TATCTAAGAATGTATCAGCTCAGAACTCTGGCAATGAGTTTTCTATTTATTGGG
ATATGCAGATTCTCTTAGAAAAGTTGGTATGTTGGATACTTATATTA^{AAA}ATTGTT
TGTTATCTAACAATTCAATCTCGTTATTTTAAAAATGGCGAACGAGTTAACTTT
TTGAACATATAAGTAACGCTCTACGGTATTCAAGGAGTGATTTTCTCATTAACTCT
TATTTTTGAACGATATATCGAATATATAAACCATCTAAAATTGTCGCCCAAACAA
AAAGATTTTTATTTTTGTACGAAGTTTTCAAATTTTCATGATTATACTAAAAATG
GATATAAATATTTAGCATTTGATAATCAAGCCGATGCAGGGTATGGCCTGACTTT
ATTATTA^{AAATG}CAAACGATGATATGCAAGATAGTTATAATCTACTCCCTGAGCA
AGA^{ACTTTTT}ATTTGTAATGCTGTAATAGATAATATGAATATTTATAGGAGTCAA
TTAACAAATGTCTACGAAAATACGATTTATCAGAAATAACTGATATATACCCA
AATAAAATTATATTGCAAGGAATTAAGTTTCGATAAGAAAAAAAAATGTTTATGGA
AAAGATCTTGTTAGTATAATAATGTCAGTATTCAATTCAGAAGATACTATTGCAT
ACTCATTACATTCATTGTTGAATCAAACCTATGAAAATATTGAAATTCTCGTGTG
CGATGATTGTTTCATCGGACAAAAGCCTTGAAATAATTAAGAGCATAGCTTATTCT
AGTTCAAGAGTGAAAGTATATAGCTCACGAAAAACCAAGGCCCTTATAATATA
AGAAATGAGCTAATAAAAAAAGCACACGGTAATTTTCATCACCTTTCAAGATGCA
GATGATCTTTCTCATCCGGAGAGAATACAAAGACAAGTTGAGGTTCTTCGCAAT
AATAAGGCTGTAATCTGTATGGCTAACTGGATACGTGTTGCGTCAAATGGAAAA
ATTCAATTCCTTCTATGATGATAAAGCCACAAGAATGTCTGTTGTATCGTCAATGA
TAAAAAAGATATTTTTGCGACAGTTGGTGGCTATAGACAATCTTTAATTGGTGC
AGATACGGAGTTTTATGAAACAGTAATAATGCGTTATGGGCGAGAAAGTATTGT
AAGATTACTGCAACCATTGATATTGGGGTTATGGGGAGACTCCGGACTTACCAG
GAATAAAGGAACAGAAGCTCTACCTGATGGATATATATCACAATCTCGAAGAGA
ATATAGTGATATCGCGGCAAGACAACGAGTGTTAGGGAAAAGTATCGTAAGTGA
TAAAGATGTACGTGGTTTTATTATCTCGCTATGGTTTGTTTAAAGATGTATCAGGA
ATAATTGAACAATAGGTCGACTCTAGAATAACTGCAGTACTCGAGAACGATCGC
GGCCGCAAGCTT

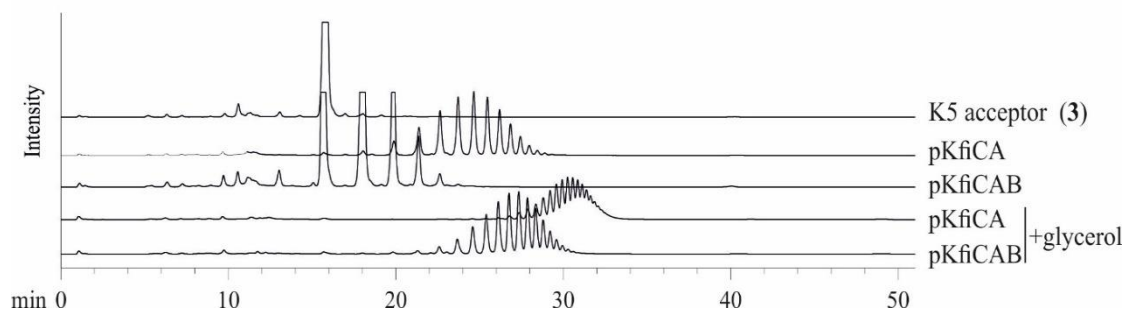
Supplementary Figure S3. Characterization of purified heparosan synthases. A. Coomassie-stained polyacrylamide gels of KfiA (5664), KfiC (5822) and KfiB (5668). **B.** Single transfers of UDP-donors to the K5 oligosaccharide acceptor. As previously reported (Sugiura *et al.* 2010), KfiC exhibits catalytic activity only in the presence of KfiA (II, III). **C.** Heparosan of higher MW is synthesized when both enzymes and both donor substrates are present. However, non-tagged version of KfiA and KfiC tested under reproducible conditions produced longer chains (see. **Fig. 5C**).



*The reaction mixture (without UDP-GlcNAc) was supplemented with KfiCA and incubated overnight at 37 °C. The reaction was stopped by incubation at 65 °C for 10

minutes. Subsequently, KfiA and UDP-GlcNAc was added and the ON incubation at 37 °C repeated.

C.



Supplementary Figure S4. The Kf1B sequence. **A.** The coding sequence of the K5 lyase B from the *Escherichia coli* K5B bacteriophage; **B.** The translation product – the primary sequence of the K5 lyase B.

A.

ATGGCTAAATTAACCAAACCTAATACTGAAGGAATCTTGCATAAAGGACAATCT
TTGTATGAGTACCTTGATGCGAGAGTTTTAACATCAAAGCCGTTTGGTGCTGTAG
GTGACGCCACTACTGATGATACGGAGGTTATAGCTGCTTCATTAAACTCTCAGA
AAGCTGTCACAATCTCAGACGGTGTATTCTCTAGCTCTGGTATTAATAGTAATTA
CTGTAACCTTAGACGGTAGAGGTAGTGGGGTGCTAAGTCACCGTTCCAGTACAGG
TAACTACTTAGTATTTAACAAATCCACGTACAGGTTCGCTTAAGTAATATTACGGTA
GAAAGTAATAAGGCAACTGATACTCAGGGACAGCAGGTATCTCTTGCGGGT
GGAAGTGATGTTACTGTAAGTGACGTTAACTTCTCAAACGTTAAAGGTAAGTGGT
TCAGTTTAATCGCATACCCTAATGATGCGCCACCTGATGGACTTATGATTAAGG
CATTTCGAGGTAGCTATCCGGATATGCTACTAATAAAGCAGCCGGATGCGTACT
TGCTGACTCCTCAGTTAACTCCCTCATAGATAACGTCATTGCTAAGAACTACCCT
CAGTTCGGAGCAGTAGAGTTGAAAGGTACAGCCAGTTACAACATAGTCAGTAAT
GTTATAGGGGCAGATTGCCAGCATGTAACCTTACAACGGCACTGAAGGGCCAATA
GCCCCTTCTAATAACCTTATCAAGGGGGTGATGGCTAATAACCCTAAGTATGCA
GCGGTTGTTGCAGGCAAAGGAAGTACGAACCTAATCTCAGACGTGCTCGTAGAT
TACTCAACTTCTGATGCTAGGCAGGCTCATGGTGTACCGTAGAAGGGTCTGATA
ACGTCATAAATAATGTGCTTATGTCAGGATGTGATGGTACTAACTCTTTAGGACA
AGGGCAGACTGCTACAATTGCACGCTTTATAGGTACAGCTAATAACAACCTATGC
GTCTGTATTTCTAGCTACAGTGCTACAGGTGTTATTACTTTTGAATCCGGCTCT
ACCCGTAACCTTCGTAGAGGTAAGCACCCTGGCAGGAGAAACGACCTTCTCAGT
TCTGCTAGTACTATTGACGGTGCAGCTACTATTGACGGCACTAATAATAGTAACG
TAGTGCACGCACCTGCCTTAGGGCAGTACGTAGGTAGTATGTCAGGTAGGTTTCG
AATGGCGGATTAAGTCCATGTCACCTCCCTTCAGGCGTTCTTACTTCTGCTGATAA
GTACAGAATGCTTGGGGATGGTGCTGTGTCATTAGCTGTAGGTGGGGTACTTCT
TCTCAAGTTCGCCTATTTACTTCTGATGGTACTTCTCGGACAGTGTCCCTCACCA
ACGGTAACGTGCGTCTTTCTACCAGTAGCACAGGCTATTTGCAGTTAGGTGCTGA
TGCAATGACCCAGATAGTACTGGTACATACGCATTAGGTTCCGCCAGTCGAGC
ATGGTCTGGCGGTTTTACTCAATCAGCATTCACTGTTGTCTCAGATGCTAGGGAT
AAAACAGAGCCTCTTATTATCTCAGATGCTTTACTGGATGCTTGGTCTGAAGTTG
ACTTTGTGCAGTTTCAGTACTTGGACCGAATTGAGGAGAAGGGTGCAGACTCAG
CTAGATGGCACTTCGGTATCATCGCTCAGAGAGCTAAGGAGGCTTTCGAACGTC
ACGGTATAGATGCACATCGCTATGGCTTCTTGTGCTTCGACAGTTGGGATGATGT
ATACGAGGAAGATGCCAATGGCTCTCGTAAACTGATTACACCAGCAGGTTCCCG
CTACGGTATTCGTTACGAGGAAGTACTGATATTAGAGGCTGCGTTGATGCGGAG
GACTATTAAGCGTATGCAGGAAGCACTAGCTTCCCTGCCTAAGTAA

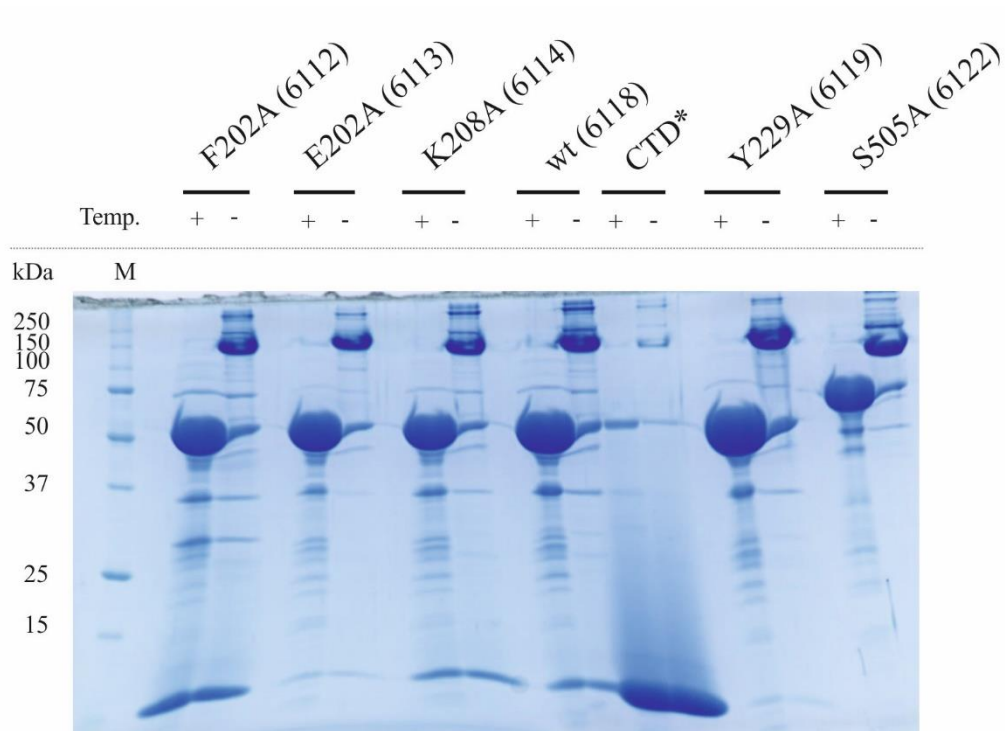
B.

5'3' Frame 1

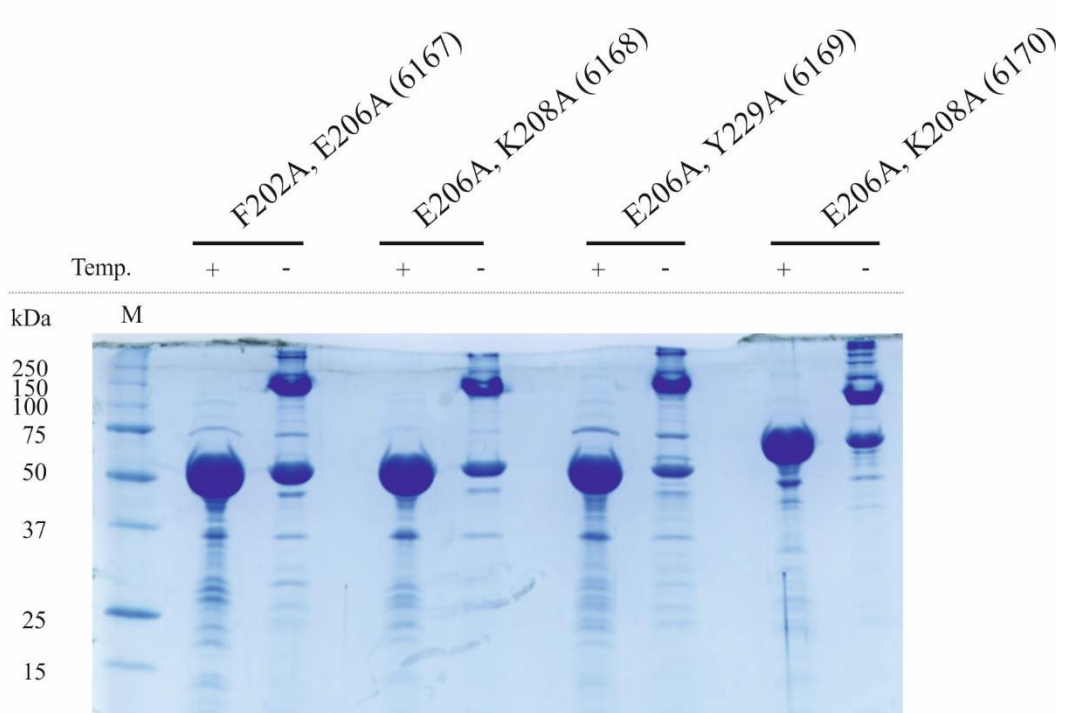
MAKLTKPNTREGILHKGQSLYEYLDARVLTSKPFGAVGDATDDTEVIAASLNSQKA
VTISDGVFSSSGINSNYCNLDGRGSGVLSHRSSSTGNLYLFNPNRTGRLSNITVESNKA
TDTTQGGQVSLAGGSDVTVSDVNFSNVKGTGFSLIAYPNDAPPDGLMIKGIRGSYSG
YATNKAAGCVLADSSVNSLIDNVIKNYPQFGAVELKGTASYNIVSNVIGADCQHV
TYNGTEGPIAPSNLIKGVMANPKYAAVVAGKGSTNLISDVLVDYSTSDARQAHG
VTVEGSDNVINNVLMMSGCDGTNSLGQGQTATIARFIGTANNNYASVFPSYSATGVIT
FESGSTRNFVEVKHPGRRNDLLSSASTIDGAATIDGTNNSNVVHAPALGQYVGSMS
GRFEWRIKSMSLPSGVLTSAKRYRMLGDGAVSLAVGGGTSSQVRLFTSDGTSRTVS
LTNGNVRLSTSSTGYLQLGADAMTPDSTGTALGSASRAWGGFTQSAFTVVSDAR
DKTEPLIISDALLDAWSEVDFVQFQYLDRIEEKGADSARWHFGIIAQRKEAFERHGI
DAHRYGFLCFDSWDDVYEEDANGSRKLITPAGSRYGIRYEEVLILEAALMRRTIKR
MQEALASLPK

Supplementary Figure S5. Coomassie-stained polyacrylamide gels of KflB and its mutants (final pools) (A, B). M – marker, T - temperature „+” – denatured samples. The lanes showing 6118 and 6122 in Fig. 6 were excised from this gel. * Sample is not relevant for this study.

A.



B.

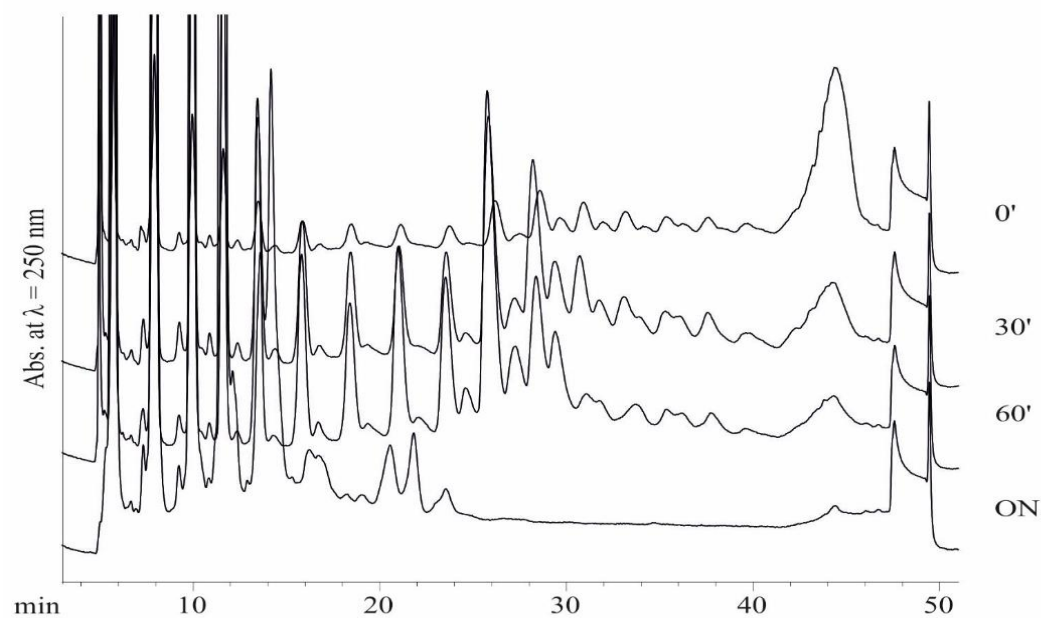


Supplementary Figure S6. Heparosan fractionation. Highly disperse heparosan was fractionated by anion-exchange chromatography and visualized by AlcianBlue/Silver stained 15% PA gel. The fractions 27-30 were pooled for performing *in vitro* assays with Kf1B.

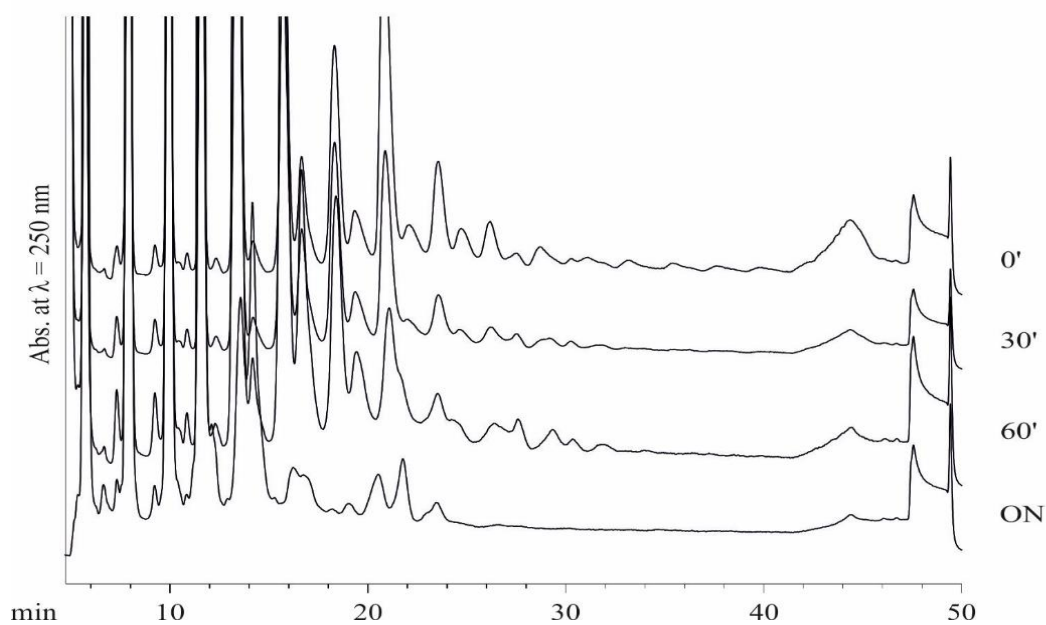


Supplementary Figure S7. Heparosan degradation catalyzed by KfIB wt (6118) and KfIB S505A (6122). Zoomed-in sections of the chromatograms shown in Fig. 7 to emphasize differences in the depolymerization mode utilized by KfIB wt and KfIB-S505A.

A. KfIB wt



B. KfIB S505A



Chapter 3 – Analysis of glycosaminoglycan-derivatives from multipotent pancreatic progenitor cells by xCGE-LIF

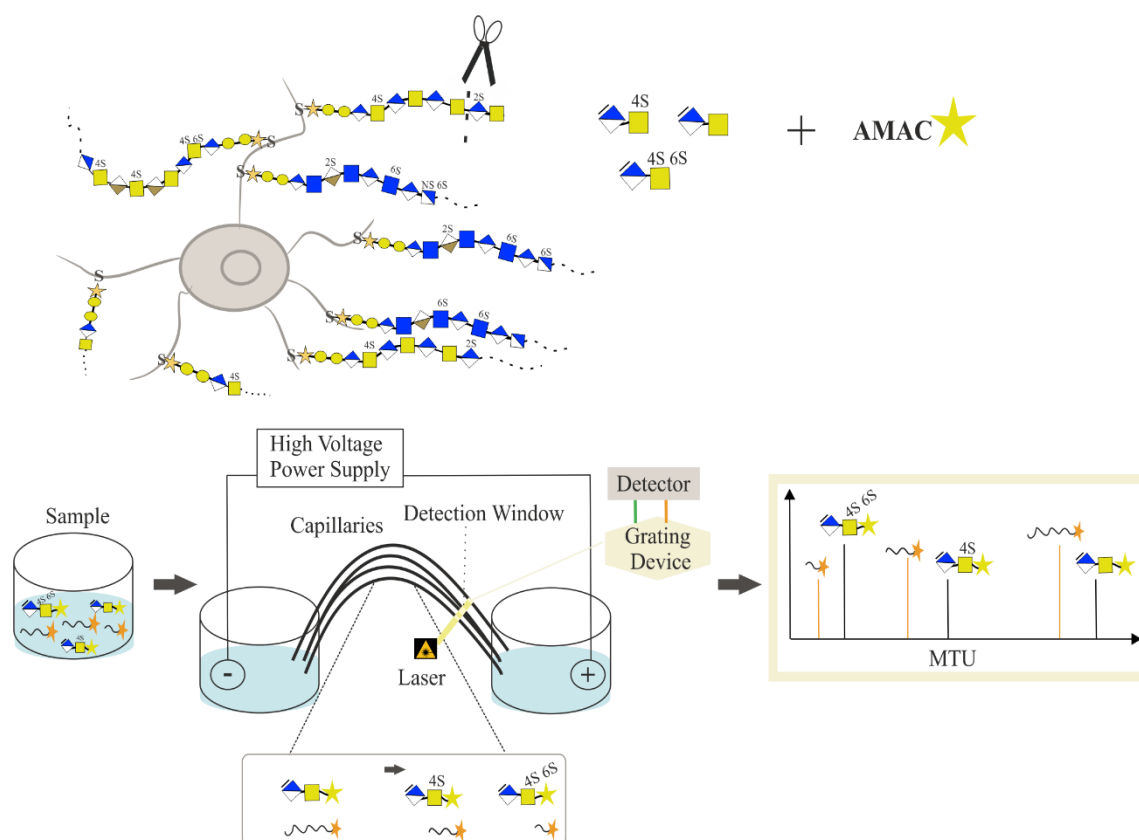
Małgorzata Sulewska^{ab§}, Charlotte Rossdam^{a§}, Julia Beimdiek^a,
Jonas Scholz^a, Ortwin Naujok^a, Astrid Oberbeck^a, Timm Fiebig^a
Bernard Priem^b, Rita Gerardy-Schahn^a, Falk F. R. Buettner^a

^a*Institute of Clinical Biochemistry, Hannover Medical School, Hannover, Germany*

^b*Centre de Recherche sur les Macromolécules Végétales, Groupe Chimie et
Biotechnologie des Oligosaccharides, 601 rue de la Chimie, BP 53X, 38041
Grenoble, Cedex 09, France*

§These authors contributed equally to this study

Manuscript in preparation



Preface

The successful development of a multiplexed capillary gel electrophoresis with laser-induced fluorescence detection (xCGE-LIF) protocol for the analysis of glycosphingolipids laid the foundation for the conceptualization of a study to analyze glycosaminoglycans (GAGs) by Dr. Charlotte Rossdam and Prof. Falk F. R. Buettner. The current study aimed towards the establishment of a standard library containing fluorescently labeled GAG-derived disaccharides. Based on the standards we aimed to investigate changes in the expression of GAGs that may occur during cell differentiation. Pilot experiments interrogating the basic question whether fluorescently labeled GAGs can be separated and visualized by xCGE-LIF were carried out by the Master student Jonas Scholz. I further optimized initially established protocols and refined data presentation. To enlarge the library of standards, I prepared heparosan and heparosan-derived oligomers under the supervision of Dr. Bernard Priem. The experiments on the cultured cells were made possible through collaboration with PD Dr. Ortwin Naujok, an expert in stem cell biology. Together with Dr. Charlotte Rossdam, Julia Beimdiek, Prof. Rita Gerardy-Schahn and Prof. Falk F. R. Buettner, I analyzed and interpreted the results. This manuscript was conceptualized and written by myself with help from Dr. Charlotte Rossdam. Prof. Falk F. R. Buettner edited the manuscript.

Chapitre 3 – Analyse des dérivés de glycosaminoglycanes de cellules progénitrices pancréatiques multipotentes par la technique de xCGE-LIF

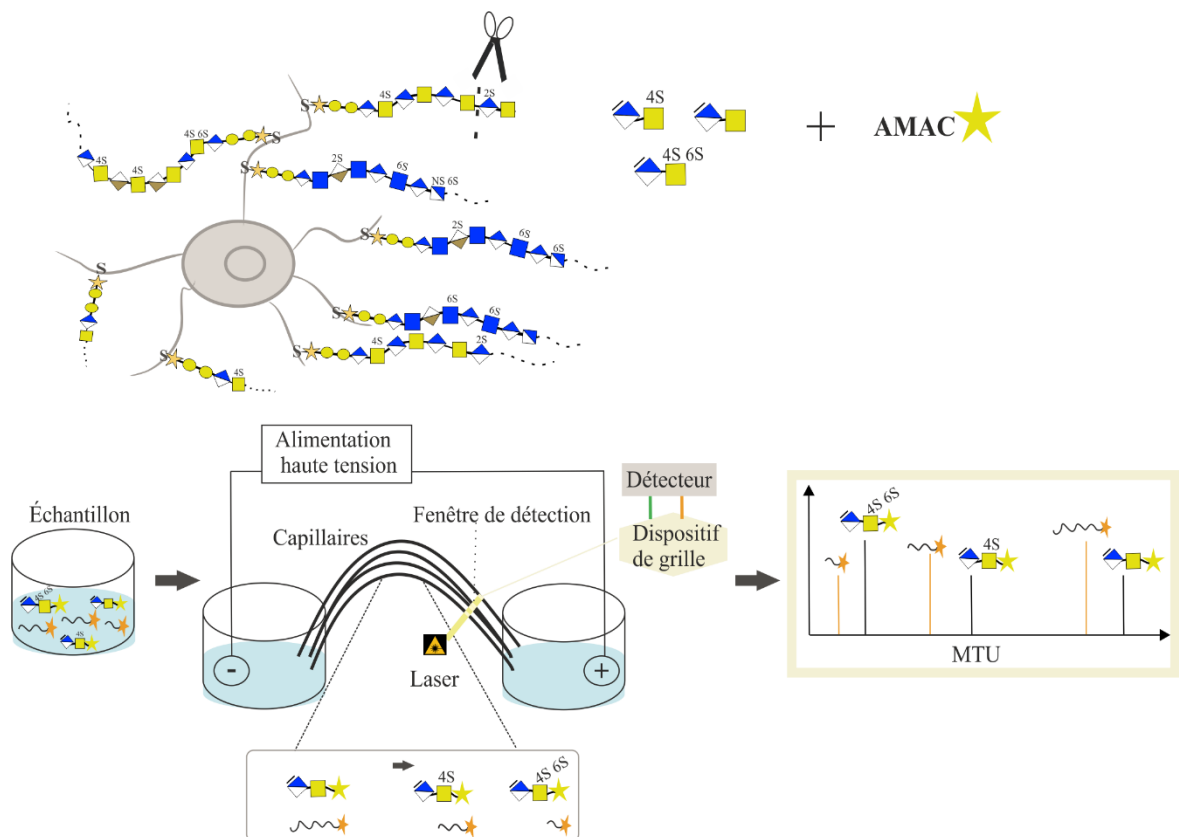
Małgorzata Sulewska^{ab§}, Charlotte Rossdam^{a§}, Julia Beimdiek^a,
Jonas Scholz^a, Ortwin Naujok^a, Astrid Oberbeck^a, Timm Fiebig^a
Bernard Priem^b, Rita Gerardy-Schahn^a, Falk F. R. Buettner^a

^a*Institute of Clinical Biochemistry, Hannover Medical School, Hannover, Germany*

^b*Centre de Recherche sur les Macromolécules Végétales, Groupe Chimie et Biotechnologie des Oligosaccharides, 601 rue de la Chimie, BP 53X, 38041 Grenoble, Cedex 09, France*

§ *Ces auteurs ont contribué à parts égales à cette étude*

Manuscrit en preparation



Préface

Le succès obtenu de la méthode d'analyse en gel d'électrophorèse capillaire multiplexe à détection en fluorescence laser (xCGE-LIF) pour l'analyse des glycosphingolipides a conduit la fondation pour la conceptualisation à l'évaluer pour l'analyse des glycosaminoglycanes (GAGs) à l'initiative des Dr Charlotte Rossdam et Prof Falk F. R. Buettner. Le présent travail a pour objectif l'établissement d'une bibliothèque de standards de disaccharides de GAGs marqués à la fluorescence. Dans un deuxième temps, celle-ci est exploitée pour l'étude des changements potentiels de l'expression des GAGs durant la différenciation cellulaire. Les expériences premières traitant de la capacité de l'approche xCGE-LIF à séparer et visualiser les GAGs marqués à la fluorescence ont été menées par Jonas Scholz dans le cadre d'un stage de Master. J'ai ensuite optimisé le protocole ainsi que le traitement de données. Afin d'élargir la banque de standards, j'ai préparé de l'héparosan et des oligosaccharides d'héparosan sous la direction du Dr Bernard Priem. Les expériences sur les cultures cellulaires ont été rendues possibles grâce à une collaboration avec le Dr Ortwin Naujok, expert en biologie cellulaire des stèmes. Avec les Dr Charlotte Rossdam, Julia Beimdiek, Prof Rita Gerardy-Schahn et le Prof FR Buettner, j'ai analysé et interprété les résultats. Le Prof Falk Buettner a édité le manuscrit.

.

3.1. Introduction

Human pluripotent stem cells (hPSCs), including both human embryonic stem cells (hESCs) and human induced pluripotent stem cells (hiPSCs), offer remarkable potential in tissue engineering and regenerative medicine (Okano et al., 2014; Pera et al., 2004). In the future, hPSCs might be used to regenerate damaged tissue in various diseases, such as diabetes (Soria et al., 2000). To use cells for such purpose, it is critical to thoroughly understand, successfully guide, and control stem cell differentiation into the desired lineage and tissue. However, while proteomic, transcriptomic, and genomic processes during stem cell differentiation are well studied, knowledge of the significance of the glycome is still rare.

The entirety of glycans in a cell, the glycome, has been shown to play essential roles in stem cell biology. For example, the characterization of hPSCs is based on the expression of numerous cell surface markers, exclusively expressed on the cell surface of human stem cells. The majority of these are, in fact, glycoconjugates, such as the tumor rejection antigens TRA-1-60 and TRA-1-81 (Wright & Andrews, 2009), or the stage-specific embryonic antigens 3, 4, and 5 (Kannagi et al., 1983, Breimer et al., 2017; Tang et al., 2011), or the somewhat recently proposed sialyl lactotetra (Barone et al., 2014). Moreover, proteoglycans are majorly involved in determining stem cell fate by interaction with chemokines, growth factors, and morphogens, thereby controlling signal transduction (Johnson et al., 2007; “Essentials of Glycobiology,” 2022; Gasimli et al., 2012; Gasimli et al., 2014; Smith et al., 2011). Proteoglycans are ubiquitous constituents of the extracellular matrix. They consist of a transmembrane core protein attached to a long glycosaminoglycan (GAG) residue facing the cell surface. The presence of GAGs at the cell surface is vital for their various roles in the cell (“Essentials of Glycobiology,” 2022). Alteration of their repeating disaccharide structures can significantly impact cell adhesion and interaction with signaling molecules which then can affect signal transduction. For example, the glycosaminoglycan heparan sulfate has been shown to structurally transition from a low-sulfated variant in stem cells to a highly-sulfated variant in neural, endodermal, or mesodermal progenitor cells (Johnson et al., 2007; Smith et al., 2011, Gasimli et al., 2014; Mikami & Kitagawa, 2017). In other studies, GAGs and particularly HS have been shown to play significant roles in pancreatic β -cells survival in diabetes by

protecting the function of the β -cells from autoimmune destruction (Ziolkowski et al., 2012).

These and many other studies underline the diverse cellular functions of GAGs and their potential in regenerative medicine, drug development, or therapeutics. However, an essential requirement in successful GAGs research is having access to appropriate and high-quality GAGs analytics. However, structural GAGs analysis still is challenging due to their high molecular mass and their structural diversity regarding charge, sulfation, or size (Solakyildirim, 2019; Pepi et al., 2021). Therefore, structural analysis of GAGs is frequently performed somewhat passively by immunological experiments or by analyzing the expression of responsible biosynthetic or degrading enzymes (Gasimli et al., 2014; Johnson et al., 2007). Direct structural analysis is mainly performed by GAG-derived disaccharides (diGAGs) analysis using mass spectrometry (MS) (Gasimli et al., 2014). While there have been remarkable improvements in structural GAGs analysis using MS in recent years (Pepi et al., 2021), there are also significant disadvantages. For instance, conventional MS analytics is usually not applicable to differentiate structural or linkage isomers due to the same underlying mass (Liang et al., 2010; Liang et al. 2011; Lin et al., 2020). Recently, Poyer *et al.* reported a protocol for differentiating chondroitin sulfate disaccharides using chromatographic methods coupled with MS (Poyer et al., 2021).

Multiplexed capillary gel electrophoresis with laser-induced fluorescence detection (xCGE-LIF) offers another approach to the structural analysis of GAGs. xCGE-LIF has been applied for the analysis of N-glycans in our laboratory (Konze et al., 2017; Beimdiek et al., 2022; Thiesler et al., 2016; Janosz et al., 2021; Taubert et al., 2022) and by others as well (Callewaert et al., 2004; Huffman et al., 2014; Hennig et al. 2015) for several years now. Furthermore, this technique has been expanded for analyzing glycosphingolipid (GSL)-derived glycans (Rossdam et al., 2019). This method offers significant advantages over other glyco-analytical technologies, such as MS, regarding isomer separation capacity, sensitivity, and the potential for high throughput.

In this study, we have used our expertise in xCGE-LIF to analyze N- and GSL-derived glycans to establish a protocol for GAG-derived disaccharides. We describe the development of a novel analytical approach to diGAGs analysis, including enzymatic digestion, fluorescent labeling, purification, and xCGE-LIF analysis. Based

on an established migration time database of diGAGs, we performed diGAGs profiling of hPSCs. Moreover, as a proof-of-concept study, we monitored changes in diGAGs composition upon hESCs differentiation into multipotent pancreatic progenitor cells (MPPCs). We were able to capture structural alterations upon differentiation.

3.2. Results and discussion

3.2.1. Standards library development

Establishing the GAG-derived disaccharides (diGAGs) based standards library was our approach's first and fundamental step. Commercially available diGAGs were fluorescently labeled with 2-aminoacridone (AMAC) in a reductive amination reaction with 2-picoline borane complex (pic-BH₃) used as a reducing agent (**Fig. 14**).

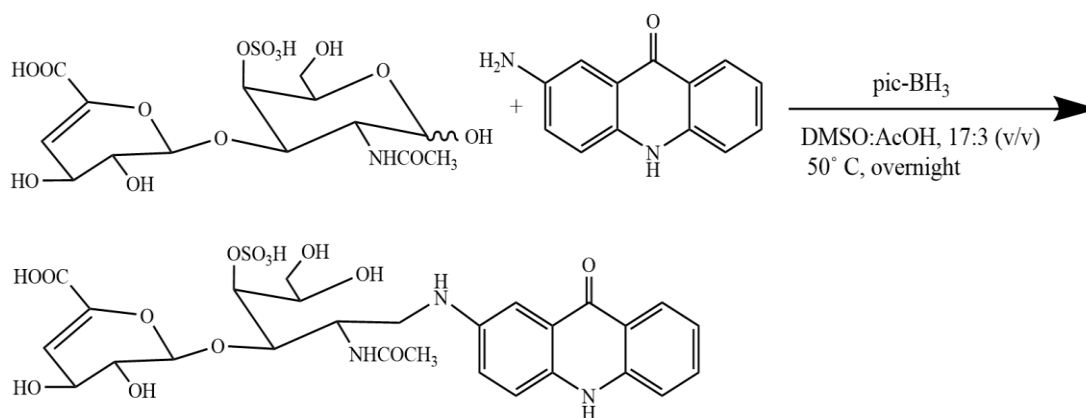


Figure 14. Labeling of diGAGs with AMAC and pic-BH₃ complex on the example of the CS-derived disaccharide.

The advantage of pic-BH₃ over the traditional, widely used sodium cyanoborohydride is based on features, such as environmental friendliness and the ability to be used in the presence of water (Sato et al., 2004). The combination of the chemical and physical properties, such as high fluorescence, the lack of charge, and low polarity made AMAC the ideal reagent for the fluorescent labeling of highly negatively charged molecules as diGAGs and their electrophoretic separation (Li et al., 2015). The excessive amount of AMAC was removed by hydrophilic interaction liquid chromatography (HILIC) using cellulose as the stationary phase and 96% acetonitrile as the mobile phase.

The established library constitutes 15 fluorescently labeled diGAGs (diGAG-AMAC) from purchased building blocks of chondroitin sulfate (CS)/dermatan sulfate

(DS), heparan sulfate (HS)/heparin, hyaluronic acid (HA), and heparosan with a net charge of -1 to -3. GAGs, especially heparin, are highly negatively charged polymers consisting of building blocks with a total net charge of even -4 per disaccharide. However, they are undetectable in our protocol due to the high electrophoretic mobility, which results in their elution before the first marker oligomer (data not included). Therefore, diGAGs with a -4 net charge are not included in the standards library.

The standards library was measured by xCGE-LIF (**Fig. 15, Table S6**), making it possible to determine the electrophoretic mobility of each diGAG-AMAC which is characterized by its migration time units (MTUs). The dominant signal present in the respective electropherograms was then assigned to the known diGAG-AMAC. Multi-replicate analysis of the entire protocol showed that the MTU is characteristic to each diGAG and highly reproducible (**Fig. 16**), with an average deviation of 0.3 MTU in technical replicates performed on the same day. The obtained electropherograms facilitate rapid determination of the net charge of an unknown diGAG-AMAC. The molecules with a -3 net charge elute in circa 20-30 MTU range, disaccharides with a -2 and -1 net charge elute at 150-165 MTU and 534-544 MTU, respectively. Our protocol enables separation and discrimination between isomers differing in a glycosidic bond (**Fig. 15D**) or a sulfate group position.

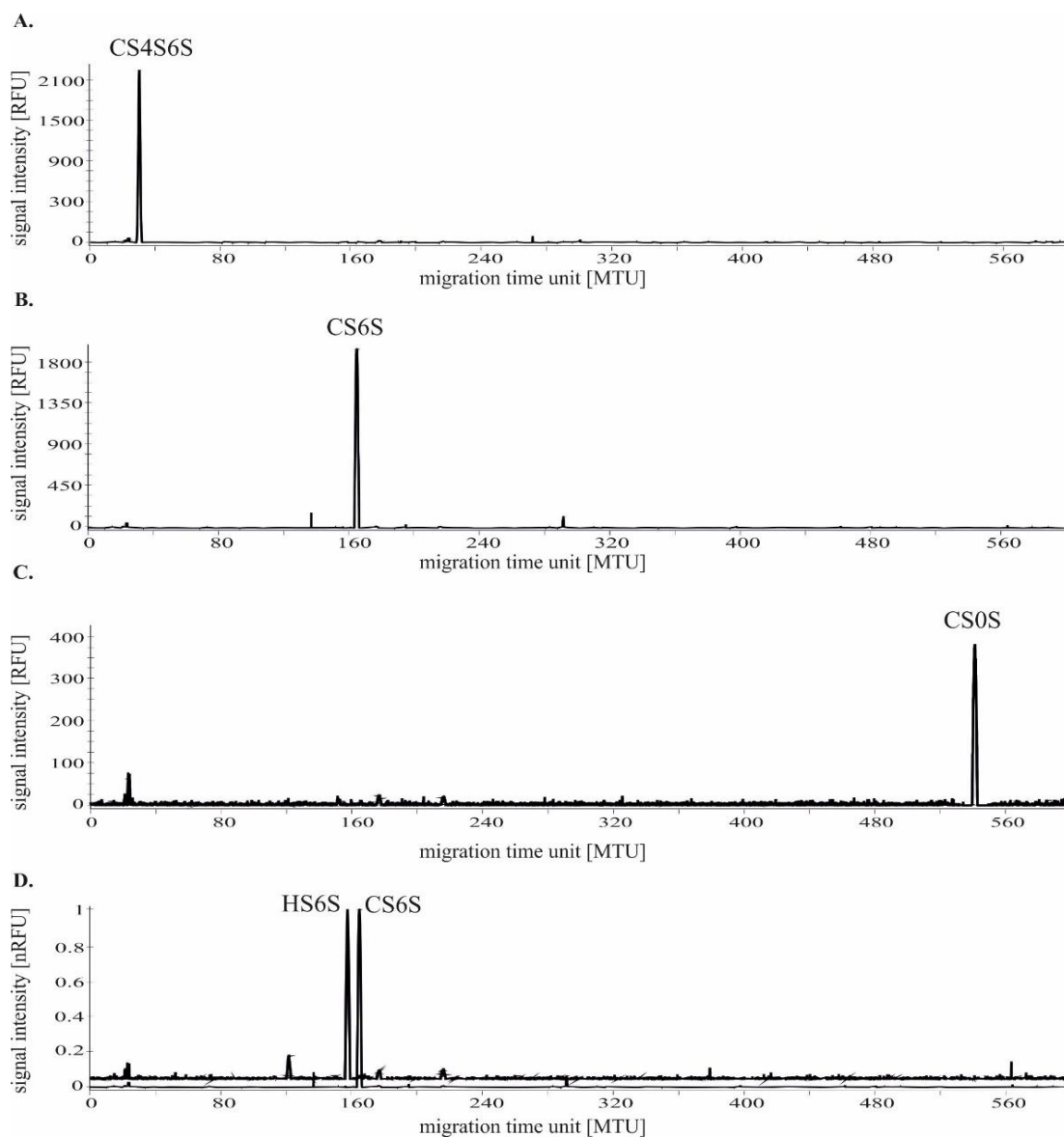


Figure 15. xCGE-LIF analysis of diGAGs-AMAC. In each electropherogram, the dominant peak was assigned to the purchased labeled GAG disaccharide. **A.** CS4S6S. **B.** CS6S. **C.** CS0S. **D.** Overlays of electropherograms displaying CS6S and HS6S. RFU = Relative fluorescence units; nRFU = normalized RFU (normalized to the most intense peak).

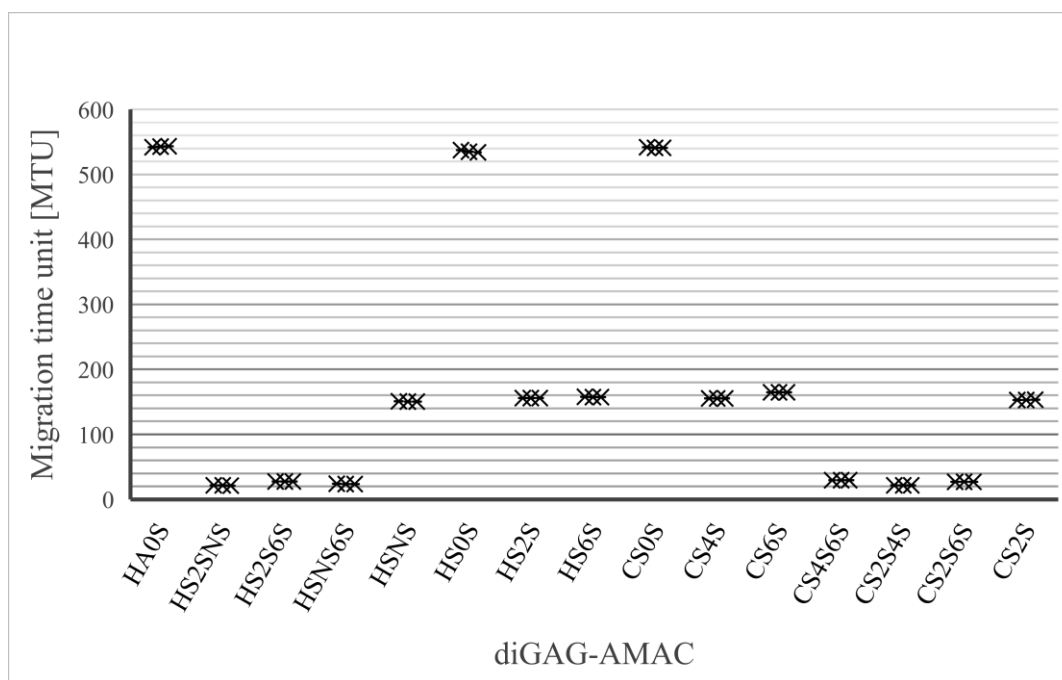
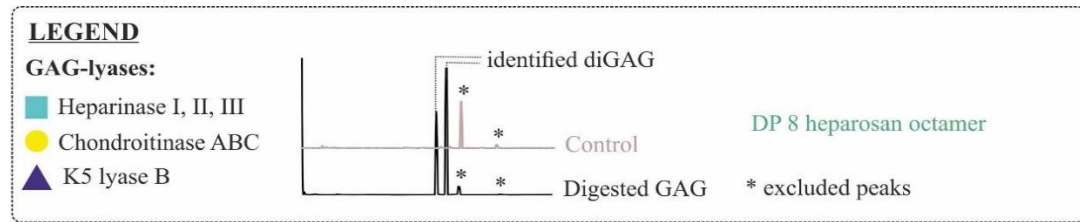


Figure 16. Reproducibility of xCGE-LIF migration times. Three technical replicates determining MTU of diGAGs-AMAC by xCGE-LIF are depicted. The exact information regarding mean and standard deviation is displayed in **Table S6**.

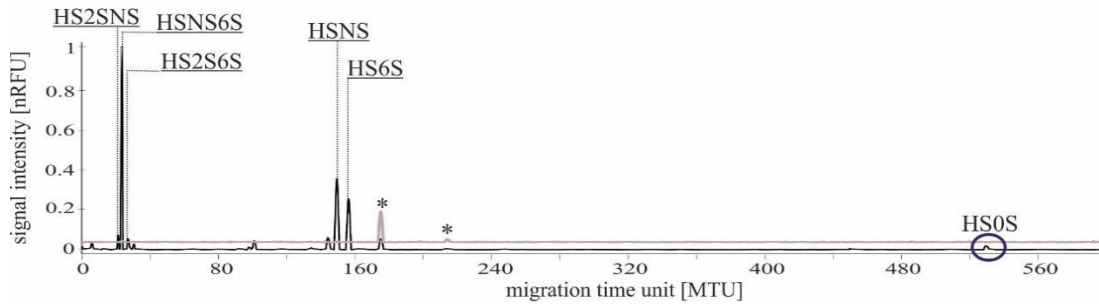
3.2.2. Analysis of diGAGs from purified GAGs

Our next aim was to identify diGAGs from purified GAGs using our database. Therefore, HA, HS, heparosan, heparin, CS, and DS were digested overnight with the respective GAG-specific lyases (**Table S7**). In a β -elimination reaction, these lyases cleave specific glycosidic bonds resulting in a sugar chain degradation and concomitant formation of an unsaturated uronic acid (Linhardt, 1999/2003). Degradation products, diGAGs, were isolated by centrifugal filtration, labeled with AMAC, and analyzed using xCGE-LIF. To confirm that peaks in electropherograms arise from diGAGs-AMAC, we included negative controls of undigested GAGs. Using our database, we assigned several peaks to specific diGAGs (**Fig. 17**). Although the repeating sequences of HA and heparosan are perfectly homogenous, additional peaks are present in the electropherograms suggesting mixtures of oligomers. This observation corresponds to our study on Kf1B (**Chapter 2**). By labeling heparosan octamer and decamer, we determined that both mixtures constitute oligomers with DP2-8, and their migration velocity is increasing along with their charge and chain length (**Fig. 17E**). This result also suggests that peaks with MTU of 40-120 present in

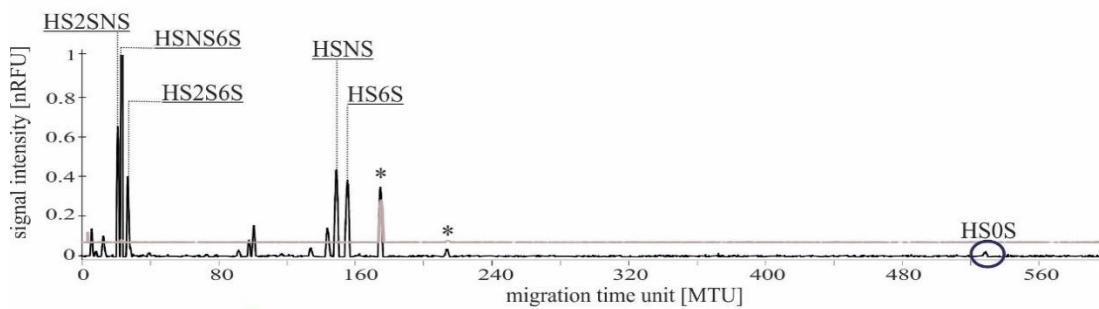
electropherograms from digested samples of HS/CS/DE/heparin could represent products of incomplete digestion, e.g., a tetramer.



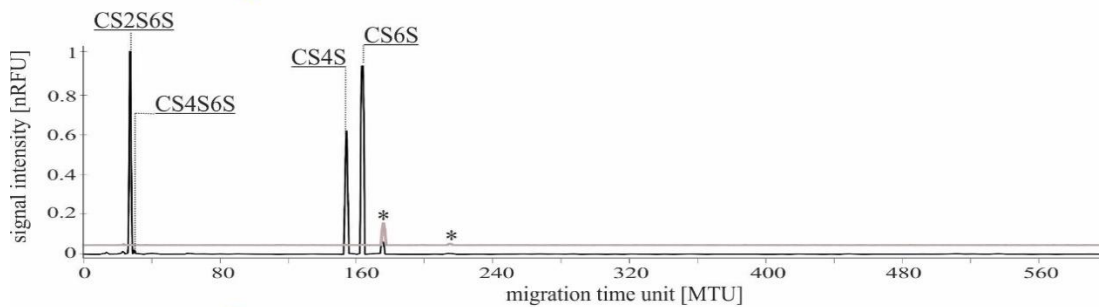
A. Heparan sulfate ■



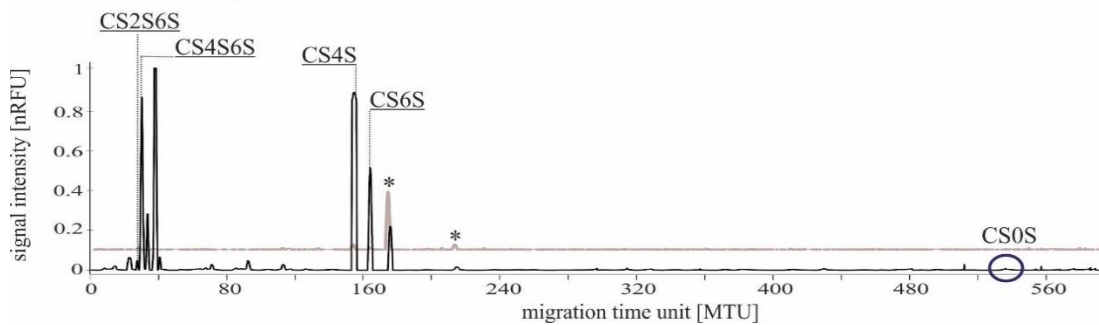
B. Heparin ■



C. Chondroitin sulfate ●



D. Dermatan sulfate ●



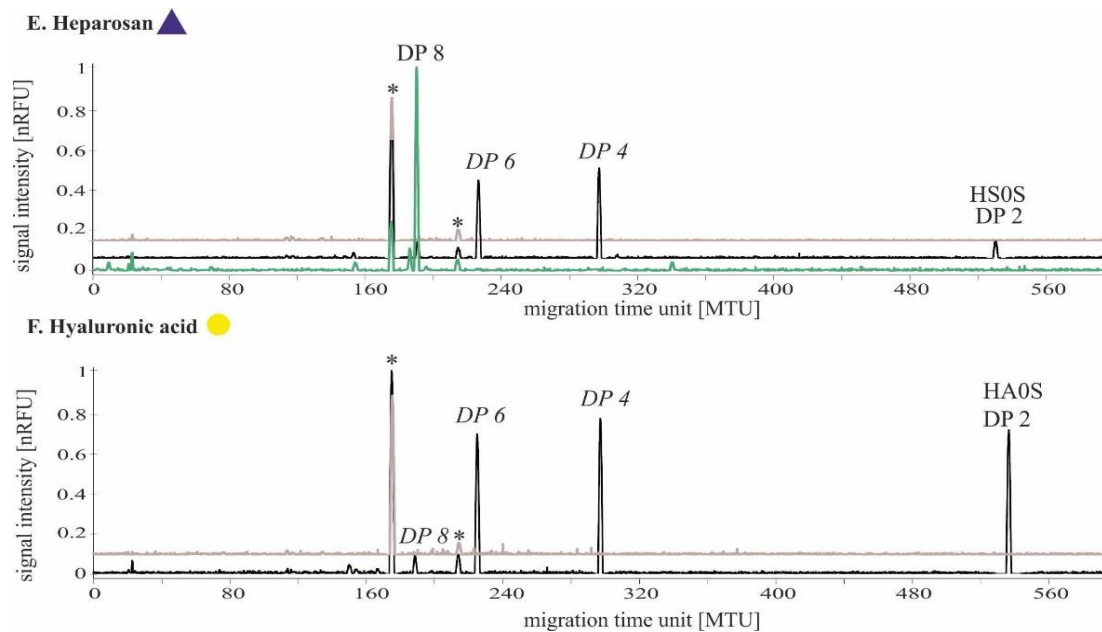


Figure 17. xCGE-LIF analysis of AMAC-labeled products of GAG digestions and peaks' assignment using our established database. Overlays of electropherograms display digestion products (black) and controls (without digestion; grey). The blue, empty circle represents the detection of a diGAG at a very low concentration. Peaks assigned without an experimental confirmation are depicted in italics. GAGs were enzymatically digested with Heparinase I, II, III (■); Chondroitinase ABC (●); and K5 lyase B (▲). GAGs and used lyases: **A.** Heparan sulfate + ■. **B.** Heparin + ■. **C.** Chondroitin sulfate + ●; **D.** Dermatan sulfate + ●. **E.** Heparosan + ▲; labeled heparosan octamer (DP 8) is depicted in green; **F.** Hyaluronic acid + ●. DP = degree of polymerization; nRFU = normalized RFU (normalized to the most intense peak). * The peaks with MTU ~ 177, 216 are excluded as related to a labeling procedure (data not included).

3.2.3. Characterization of diGAGs derived from hiPSCs

The use of the developed protocol, which successfully identified several building blocks of GAGs, provided the basis for the structural analysis of GAG derivatives in stem cells. The crucial step in broadening the application of the xCGE-LIF protocol was the extraction of diGAGs from stem cells. One of the reported strategies started from the extraction of GAGs followed by enzymatic digestion (Volpi et al., 2014; Ucakurk et al., 2014). The second approach relies on direct GAG digestion from a cell lysate with subsequent diGAGs extraction (Li et al., 2015; Carvalho et al., 2020; Silva et al., 2020). We used a straightforward and time-effective protocol developed by Li *et al.* (**Fig. 18**) with minor modifications to analyze a GAG profile in human induced pluripotent stem cells (hiPSC). Due to the high stability of GAGs (van der Meer et al., 2017; Lowry & Beavers, 1994), diGAGs generation could be pursued in a

thermal cell disruption, followed by direct digestion with GAG-lyases, isolation, and labeling.

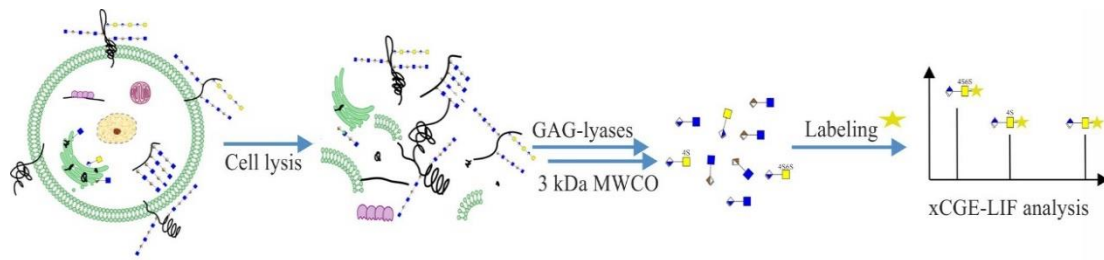
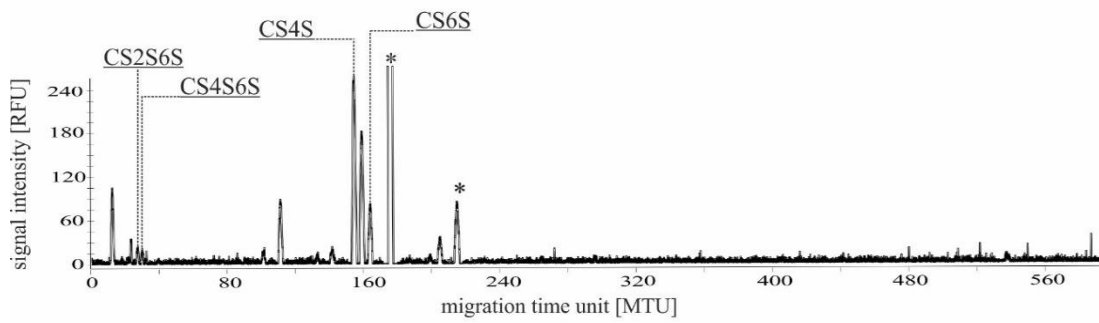


Figure 18. Workflow of AMAC-labeled diGAGs preparation from the cell lysates (adapted from Li *et al.*, 2015).

At the time of writing this chapter, only one biological replicate representing the characterization of GAGs from hiPSCs was performed. Therefore, the results should be treated with reserve. Using our database, we could identify six diGAGs - four CS/DS-derived (**Fig. 19A**) and two HS-derived (**Fig. 19B**) - from the obtained electropherograms. Several peaks could not be assigned due to limitations of the currently available database. The mixture, upon digestion, likely contained oligomers of $DP > 2$ represented by peaks with MTUs of 66-120 and ~ 206 . We can observe that some of the peaks in both electropherograms are characterized by similar MTU. Unspecific digestion was one of the possible causes. However, obtained data argued against it (data not included). There is a high likelihood that unassigned peaks are GAG-derived digestion products characterized by similar MTU and indistinguishable due to limitations in the resolution of the applied analytical method and/or they are not present in the database, yet.

A. Digestion with Chondroitinase ABC



B. Digestion with Heparinase I, II, III

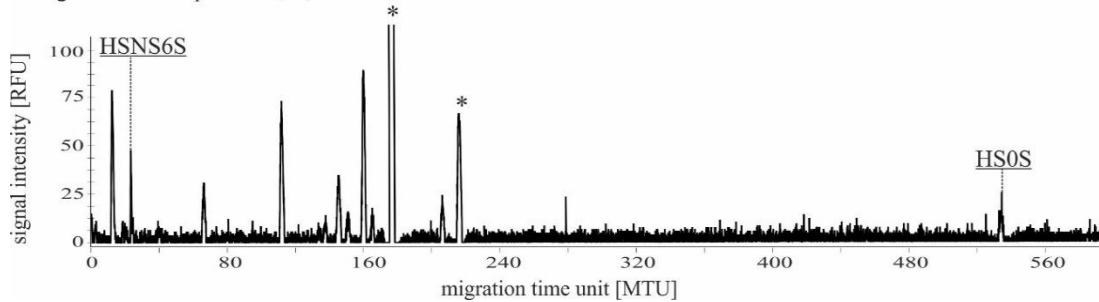


Figure 19. xCGE-LIF analysis of diGAG-AMAC from 2.5×10^6 hiPSCs. Peaks were assigned based on the database of GAG standards and their migration time units (see **Fig. 15**, **Table S6**). **A.** GAGs digestion with Chondroitinase ABC lyase generating CS, DS, and HA-derived diGAGs. **B.** GAG digestion with Heparinase I, II, III generating HS-derived diGAGs. * The peaks with MTU ~ 177, 216 are excluded as related to a labeling procedure (data not included).

3.2.4. Determining changes in diGAGs during cell differentiation

The structure and concentration changes of GAGs are an inherent part of physiological processes such as cell signaling (Lepedda et al., 2021) or differentiation (Smith et al., 2011). Interestingly, it has been speculated that strictly guided cell differentiation could be achieved by adding defined HS oligomers to the culture medium (Smith et al., 2011) or by controlling the expression of enzymes involved in GAG biosynthesis or degradation (Mikael et al., 2019). We applied our xCGE-LIF-based protocol to investigate compositional changes of GAGs during hESC differentiation into multipotent pancreatic progenitor cells (MPPCs). The electropherogram for CS on d12 of differentiation, from 2.5×10^6 of cells, revealed the appearance of 0S, along with 4S, and 6S, present on d0 (**Fig. 20A**), while the disaccharide profile for HS displayed the disappearance of NS in favor of 2SNS and NS6S (**Fig. 20B**). Assuming that all peaks correspond to GAG oligomers (excluding peaks of MTU ~177 and 2216 MTU) and their intensity to their concentration, we

could conclude that during hESCs differentiation into MPPCs, the concentration of HS increases by 18%, while CS/DS decreases by 8%.

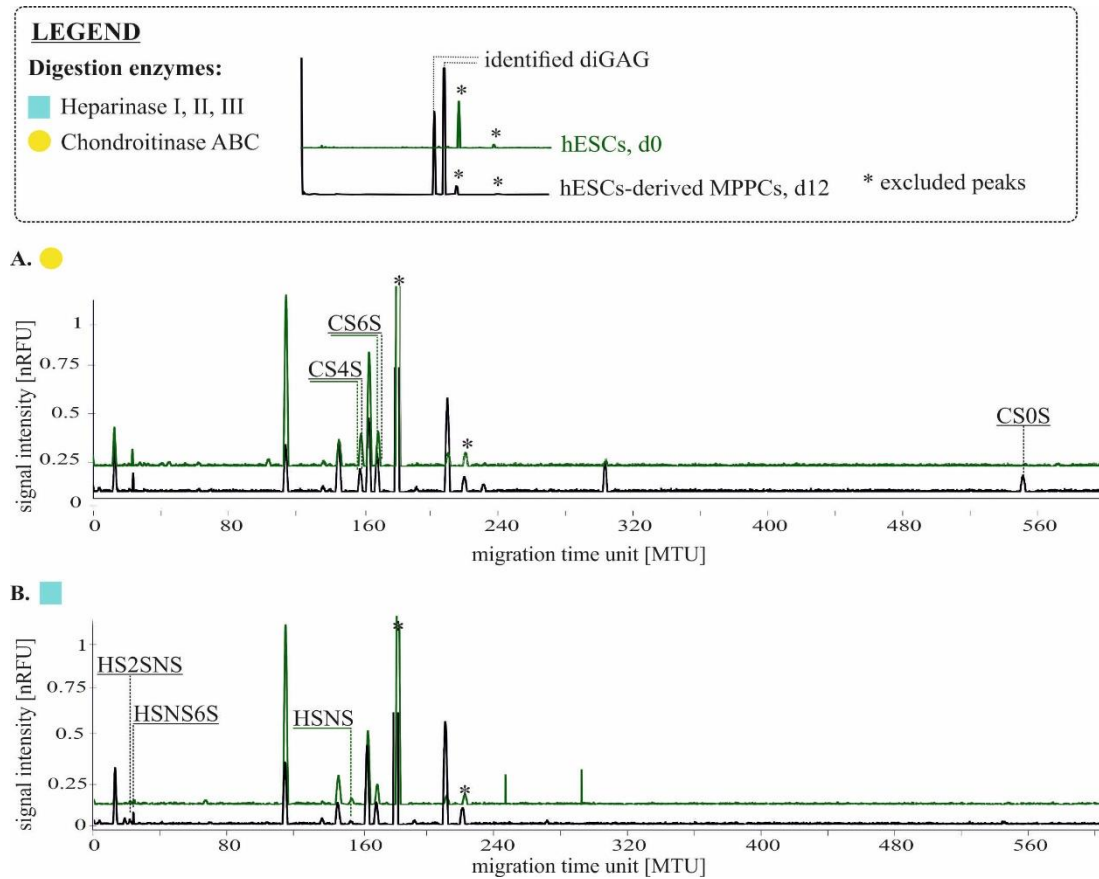


Figure 20. Comparison of GAG composition of hESCs and hESC-derived MPPCs by xCGE-LIF. Comparison of GAG composition of hESCs and hESC-derived MPPCs by xCGE-LIF. Peaks were assigned based on the database of GAG standards and their migration time units (see Fig. 2, Table S6). Overlays of electropherograms display hESCs (green) and hESCs-derived MPPCs (black). GAGs were enzymatically digested with Heparinase I, II, III (■); Chondroitinase ABC (●). **A.** GAGs were digested with Chondroitinase ABC lyase generating CS, DS, and HA-derived diGAGs. **B.** GAGs were digested with Heparinase I, II, III generating HS-derived diGAGs. nRFU = normalized RFU (normalized to the most intense peak). * The peaks with MTU ~ 177, 216 are excluded as related to a labeling procedure (data not included).

3.3. Conclusions

Methods enabling fast, cost-effective, and high-throughput analysis of GAGs are crucial for their comprehensive understanding. To our knowledge, we present the first xCGE-LIF-based protocol for compositional analysis of GAG-derived disaccharides from stem cells and their differentiated derivatives. The protocol is efficient and straightforward. Ultimately, xCGE-LIF has the potential to become a powerful tool for GAG analysis, building the basis for the identification of novel biomarkers. However,

for the successful closure of this project, the GAG disaccharide database needs to be expanded as well as additional biological replicates with stem cell-derived GAGs need to be performed.

3.4. Materials and methods

3.4.1. hPSC culture and differentiation

The hPSC lines Hes-3 SC30 (ESC, ES Cell International) and Phoenix NSC20 (iPSC, MHHi001-A) (Haase et al., 2017) were cultured on cell culture plastic coated with Matrigel (MG, Corning, Amsterdam, Netherlands) using StemMACS™ iPS-Brew XF medium (Miltenyi Biotec, Bergisch Gladbach, Germany). Cell passaging was performed once a week by seeding 6.5×10^4 single cells in iPS-Brew plus 5 μ M Y-27632 (Selleck Chemicals, Munich, Germany). For differentiation, hPSC colonies were dissociated into single cells by Trypsin/EDTA (T/E) (Biochrom, Berlin, Germany) and centrifuged. The cell pellet was then re-suspended in iPS-Brew plus 5 μ M Y-27632. 1.4×10^6 cells were subsequently seeded on MG-coated 10 cm cell culture dishes. Differentiation into MPPCs (multipotent pancreatic progenitor cells) was initiated the next day and performed according to a previously published protocol until day 12 of differentiation (Dettmer et al., 2021; Dettmer et al., 2022). On day 12, the cells were dissociated, filtered and GFP2-positive MPPCs were purified by cell sorting at the central cell sorting of the Hannover Medical School. Cell pellets were stored at 80 °C for further use.

3.4.2. Extraction and enzymatic digestion of GAGs

Cell pellets (2.5×10^6) were resuspended in 90 μ L of distilled water and incubated at 70 °C for 20 min. Thermally released GAGs from cell pellets or 10 μ g of purified GAGs (HS, CS, heparin, DS, HA and heparosan, see Table S1) were digested overnight with 0.5 μ L Chondroitinase ABC (37 °C), 0.5 μ L Heparinase I, II, III (30 °C), or 0.5 μ L KfIB (37 °C) in 1x Bacteroides Heparinase Reaction Buffer (New England BioLabs) in a total volume of 100 μ L (see **Table S7**). Enzymatic degradation products were isolated using Amicon centrifugal devices (3 kDa MWCO; Sigma Aldrich). Centrifugation (4000 rpm) lasted 90 minutes with pauses, every 15 min, for adding 100 μ L of prewarmed water. The eluate was dried in a SpeedVac concentrator (Thermo Fisher Scientific) in the dark and stored at -20 °C. Digestion of purified GAGs

was confirmed by polyacrylamide (PA) gel stained with AlcianBlue/Silver (**Fig. S8**). Detailed information about an enzymatic cleavage pattern is available in Linhardt, 1998/2003 (Linhardt, 1998/2003).

3.4.3. *Fluorescent labeling and HILIC-solid phase extraction (SPE)*

Purchased diGAGs (**Table S5**) or diGAGs obtained by enzymatic degradation of GAGs (**Table S7**) were fluorescently labeled with 2-aminoacridone (2-AMAC, Sigma Aldrich). The GAG disaccharides in distilled water (2 μ L) were mixed with 5 μ L 0.1 M 2-AMAC (in DMSO:AcOH, 17:3 (v/v)), 5 μ L 0.2 M 2-picoline borane complex (2-PB; in DMSO, Merck) and incubated overnight at 50 °C in the dark. The reaction was quenched with 100 μ L acetonitrile. The crude product was purified according to the protocol of Hitchcock *et al.* with the following modifications (Hitchcock *et al.*, 2008). The cellulose column was rinsed with four 200 μ L volumes of 96% acetonitrile. Excessive amount of amination reagents were removed by rinsing with five 100 μ L volumes of 96% acetonitrile upon incubation of the column in Eppendorf® Thermomixed Compact (Merck) with shaking at 500 rpm in the dark and subsequent centrifugation at 900 rpm for 1 min. Fluorescently labeled glycans were eluted with three 100 μ L volumes of water by centrifugation at 900 rpm for 1 min upon incubation with shaking for 5 min at 500 rpm. The eluate was dried in a SpeedVac concentrator (Thermo Fisher Scientific) in the dark and stored at -20 °C.

3.4.4. *xCGE-LIF*

All the equipment and chemicals were purchased from Applied Biosystem™ (Thermo Fisher Scientific, Foster City, CA) unless stated otherwise. AMAC-labeled glycans were resuspended in 30 μ L water (LC-MS grade, Merck). Analyzed samples consisted of 0.5-3 μ L of glycan solution, 1 μ L GeneScan™ 500 LIZ™ dye Size Standard (1:20-1:50 in HiDi™ Formamide), and HiDi™ Formamide up to a total volume of 10 μ L in a MicoAmp™ Optical 96-Well Reaction Plate. The migration times' calibration was obtained by applying the size standard consisting of 16 base pair DNA fragments of defined lengths (20-500 bps). The analyzed mixture was separated by multiplexed capillary gel electrophoresis. Labeled molecules were detected by laser-induced fluorescence using ABI PRISM® 3100-*Avant* Genetic Analyzer (remodeled by Axel Künzler, advanced biolab service GmbH, Munich, Germany) and Run 3100-*Avant* Data Collection Software v.2.0. Simultaneous analysis of four samples was

possible with a 4-Capillary Array of 50 cm length filled with POP-7™ Polymer. Measurements were performed according to the following program. Upon heating the system to 60 °C, the samples were injected into the capillaries for 15 s at 1.6 kV. Next, a pause lasting 50 s was followed by data acquisition for 50 min at 12 kV. The obtained data were analyzed with Genemapper™ Software v.3.7. Generated electropherograms displayed peaks representing the size standard in the orange channel and the glycans in the blue channel at specific migration time units (MTUs) and intensities (relative fluorescence units, RFUs). A signal having a peak amplitude threshold of ≥ 20 RFU was considered a peak. We built a database of glycan MTUs based on purchased disaccharides and used it to analyze GAG composition in a sample.

3.4.5. Data processing and statistics

Generated electropherograms in the range of 10 – 600 MTU were processed in CorelDraw 2017. In electropherograms with the y-axis labeled as nRFU (normalized relative fluorescence unit), the dominant peak intensity was set to 1. In quantitative analysis, the intensity of all peaks in the 10 – 600 MTU range was summed up except for GAG-unrelated peaks with MTU of 177 and 216.

3.5. Footnotes

The abbreviations: AcOH, acetic acid; AMAC, 2-aminoacridone; CS, chondroitin sulfate; diGAGs, glycosaminoglycan-derived disaccharides; diGAG-AMAC, AMAC-labeled diGAGs; DNA, deoxyribonucleic acid; DP, degree of polymerization; DS, dermatan sulfate; DMSO, dimethyl sulfoxide; GAG, glycosaminoglycan; GSL, glycosphingolipid; HA, hyaluronic acid; hESCs, human embryonic stem cells; HILIC, hydrophilic interaction liquid chromatography; hiPSCs, human induced pluripotent stem cells; hPSCs, human pluripotent stem cells; HS, heparan sulfate; IdoA, iduronic acid; kDa, kilo dalton; KS, keratan sulfate; MPPCs, multipotent pancreatic progenitor cells; MS, mass spectrometry; MTUs, migration time units; MWCO, molecular weight cut-off; nRFU, normalized RFU; pic-BH₃, 2-picoline borane complex; RFU, relative fluorescence units; xCGE-LIF, multiplexed capillary gel electrophoresis with laser-induced fluorescence detection.

All diGAGs and GAGs used in this study are stated in **Table S5**.

SUPPLEMENTARY INFORMATION

Chapter 3 – Analysis of glycosaminoglycan-derivatives from multipotent pancreatic progenitor cells by xCGE-LIF

Małgorzata Sulewska^{ab§}, Charlotte Rossdam^{a§}, Julia Beimdiek^a,
Jonas Scholz^a, Ortwin Naujok^a, Astrid Oberbeck^a, Timm Fiebig^a
Bernard Priem^b, Rita Gerardy-Schahn^a, Falk F. R. Buettner^a

^aInstitute of Clinical Biochemistry, Hannover Medical School, Hannover, Germany

*^bCentre de Recherche sur les Macromolécules Végétales, Groupe Chimie et
Biotechnologie des Oligosaccharides, 601 rue de la Chimie, BP 53X, 38041
Grenoble, Cedex 09, France*

[§]These authors contributed equally to this study

SUPPLEMENTARY TABLES

Supplementary Table S5. Glycosaminoglycans, GAG-derived oligomers used in this study.

Supplementary Table S6. Migration time units (MTUs) of defined AMAC-labeled GAG-derived disaccharides.

Supplementary Table S7. Enzymes used for the digestion of glycosaminoglycans.

SUPPLEMENTARY FIGURES

Supplementary Figure S8. Alcian blue/silver-stained polyacrylamide gel of glycosaminoglycan degradation.

Table S5. Glycosaminoglycans, GAG-derived oligomers used in this study.

Symbol	Structure	Charge	Supplier
Hyaluronic acid HA			
HA			Dextra
HA0S	β - Δ UA-(1 \rightarrow 3)-GlcNAc	-1	Iduron
Heparan sulfate HS, heparin, heparosan			
HS			Iduron
Heparin			Sigma-Aldrich
Heparosan			Chapter 2
HS2SNS	β - Δ UA-2S-(1 \rightarrow 4)-GlcN-NS	-3	Dextra
HS2S6S	β - Δ UA-2S-(1 \rightarrow 4)-GlcNAc-6S	-3	Iduron
HSNS6S	β - Δ UA-(1 \rightarrow 4)-GlcN-NS, 6S	-3	Iduron
HSNS	β - Δ UA-(1 \rightarrow 4)-GlcN-NS	-2	Iduron
HS0S	β - Δ UA-(1 \rightarrow 4)-GlcNAc	-1	Iduron
HS2S	β - Δ UA-2S-(1 \rightarrow 4)-GlcNAc	-2	Iduron
HS6S	β - Δ UA-(1 \rightarrow 4)-GlcNAc-6S	-2	Iduron
Hep DP8	β - Δ UA-(1 \rightarrow 4)-GlcNAc-(GlcA-GlcNAc) ₃	-4	Chapter 2
Hep DP10	β - Δ UA-(1 \rightarrow 4)-GlcNAc-(GlcA-GlcNAc) ₄	-5	Chapter 2
Chondroitin sulfate CS, Dermatan sulfate DS			
CS			Sigma-Aldrich
DS			Iduron
CS0S	β - Δ UA-(1 \rightarrow 3)-GalNAc	-1	Carbosynth
CS4S	β - Δ UA-(1 \rightarrow 3)-GalNAc-4S	-2	Carbosynth
CS6S	β - Δ UA-(1 \rightarrow 3)-GalNAc-6S	-2	Carbosynth
CS4S6S	β - Δ UA-(1 \rightarrow 3)-GalNAc-4S, 6S	-3	Carbosynth
CS2S4S	β - Δ UA-2S-(1 \rightarrow 3)-GalNAc-4S	-3	Iduron
CS2S6S	β - Δ UA-2S-(1 \rightarrow 3)-GalNAc-6S	-3	Iduron
CS2S	β - Δ UA-2S-(1 \rightarrow 3)-GalNAc	-2	Iduron

Supplementary Table S6. Migration time units (MTUs) of defined AMAC-labeled GAG-derived disaccharides. The MTUs of three replicates for each glycan were determined using xCGE-LIF. Corresponding electropherograms of selected diGAGs* and scatter plots are depicted in **Fig. 15** and **Fig. 16**, respectively.

Glycan symbol	Migration Time Unit; MTU			Mean MTU	Standard deviation
HA0S	541.59	542.8	543.38	542.59	0.9133
HS2SNS	21.34	21.42	21.00	21.25	0.2230
HS2S6S	27.59	27.53	27.53	27.55	0.0346
HSNS6S	23.87	23.61	23.47	23.65	0.2030
HSNS	150.79	150.46	150.46	150.57	0.1905
HS0S	537.13	534.94	534.06	535.38	1.5809
HS2S	156.03	156.00	155.87	155.97	0.0850
HS6S*	157.55	157.49	157.28	157.44	0.1418
CS0S*	541.59	540.97	540.91	541.16	0.3765
CS4S	155.52	155.23	155.13	155.29	0.2026
CS6S*	164.79	164.58	164.64	164.67	0.1082
CS4S6S*	29.64	29.65	29.55	29.61	0.0551
CS2S4S	21.75	21.68	21.54	21.66	0.1069
CS2S6S	26.86	26.93	26.74	26.84	0.0961
CS2S	153.00	152.94	153.18	153.04	0.1249

Supplementary Table S7. Enzymes used for glycosaminoglycan digestion.

Enzyme	Supplier	Catalog number	GAG
Chondroitinase ABC (from <i>Proteus vulgaris</i>)	Sigma-Aldrich	C3667	CS, DS, HA
Heparin Lyase I	New England Biolabs®	P0735S	HS, Heparin
Heparin Lyase II	New England Biolabs®	P0736S	
Heparin Lyase III	New England Biolabs®	P0737S	
K5 lyase B	Chapter 2		Heparosan

Supplementary Figure S8. Alcian blue/silver-stained polyacrylamide gel of glycosaminoglycan degradation. GAG degradation was catalyzed by adequate lyases and visualized with alcian blue/silver in 15% PA gel. ChABC = Chondroitinase ABC. Degradation products were AMAC-labeled and analyzed by xCGE-LIF. The resulting electropherograms are depicted in **Fig. 17**.



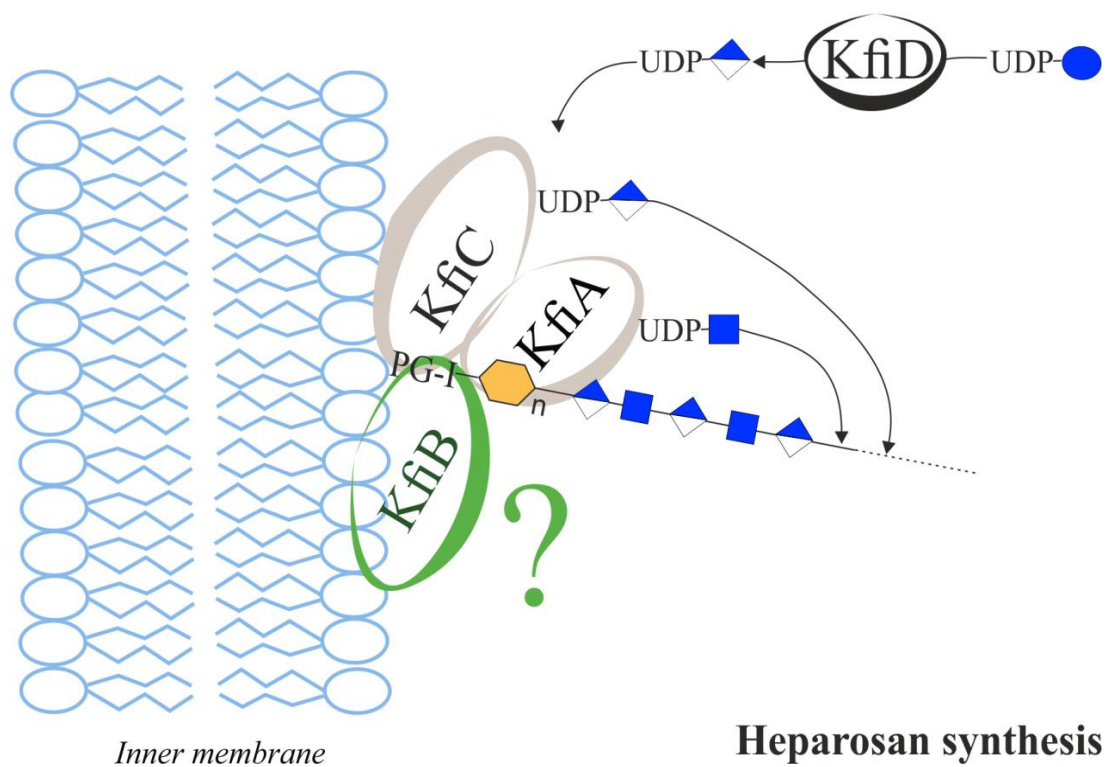
Chapter 4 – Effect of KfiB on heparosan synthesis in recombinant *Escherichia coli* BL-21

Małgorzata Sulewska^{ab}, Marta Ferrer Xalabarder^b, Juliette Tridon^b,
Timm Fiebig^a, Rita Gerardy-Schahn^a, Bernard Priem^b

^a*Institute of Clinical Biochemistry, Hannover Medical School, Hannover, Germany*

^b*Centre de Recherche sur les Macromolécules Végétales, Groupe Chimie et
Biotechnologie des Oligosaccharides, 601 rue de la Chimie, BP 53X, 38041
Grenoble, Cedex 09, France*

Manuscript in preparation



Preface

Bacteria producing group 2 capsular polysaccharides (CPS) assemble all genes involved in this process in the so-called capsular gene complex. The capsular gene complex comprises three regions, whereby region 2 harbours enzymes that are responsible for the synthesis of the serotype specific CPS. *E. coli* K5 which produces heparosan as CPS the region 2 comprises four genes called *kfi*ABCD. While the catalytic functions of enzymes encoded in the genes *kfi*ACD are well described and known to be heparosan synthases (KfiA and KfiC) and a UDP-Glc dehydrogenase (KfiD), the function of KfiB is still under debate, although, literature data suggest a role of KfiB in the location and/or sterical organization of the heparosan synthases KfiA and KfiC. When we started our experiments on the role of KfiB we realized that KfiB exists in different isoforms, differing by the number of tandem repeats (TRs). As TRs have been reported to regulate enzymatic activity and/or exhibit a function in the control of the length of a growing polymer, we asked if the number of TRs identified in KfiB impacts the length of heparosan chains synthesized in recombinant bacteria. Supervised by Dr. Bernard Priem and Dr. Timm Fiebig, I developed the concept for this study and designed KfiB constructs with varying numbers of the TR. Data acquisition was carried out in fed-batch bacterial cultures. Heparosan quantifications were supported by Marta Ferrer Xalabarder and Juliette Tridon. All authors participated in the analysis and interpretation of the results. This manuscript was conceptualized and written by myself with help of Dr. Bernard Priem. Prof. Rita Gerardy-Schahn helped with editing the manuscript.

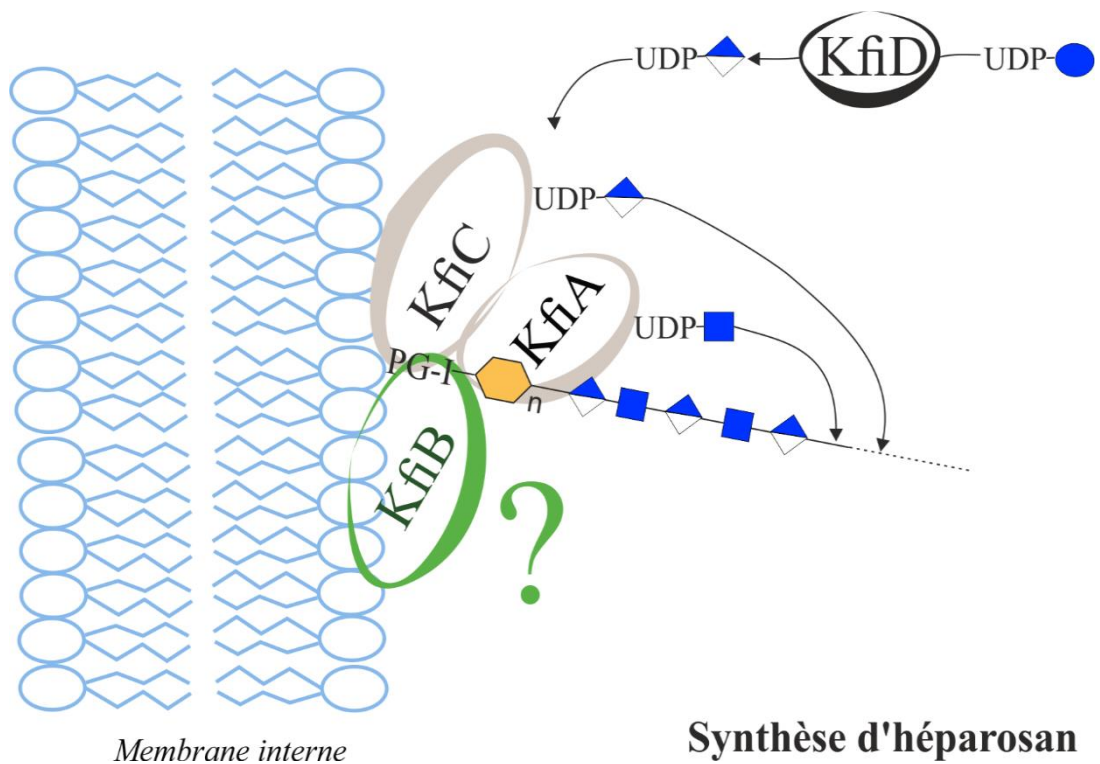
Chapitre 4 – Effet de KfiB sur la synthèse d'héparosan recombinant dans *Escherichia coli* BL-21

Małgorzata Sulewska^{ab}, Marta Ferrer Xalabarder^b, Juliette Tridon^b,
Timm Fiebig^a, Rita Gerardy-Schahn^a, Bernard Priem^b

^a*Institute of Clinical Biochemistry, Hannover Medical School, Hannover, Germany*

^b*Centre de Recherche sur les Macromolécules Végétales, Groupe Chimie et
Biotechnologie des Oligosaccharides, 601 rue de la Chimie, BP 53X, 38041
Grenoble, Cedex 09, France*

Manuscrit en préparation



Préface

Les bactéries produisant des polysaccharides capsulaires du groupe 2 (CPS) possèdent les gènes associés au sein du complexe capsulaire génétique. Celui-ci est composé de trois régions, la région 2 contenant les gènes sérotype-dépendant. La région 2 de la souche *E. coli* K5 qui produit le CPS héparosan est composée des 4 gènes *kfi*ABCD. Bien que les propriétés catalytiques des enzymes codées par les gènes *kfi*ACD soient bien décrites et définissent le complexe héparosan synthase (KfiA et KfiC) et l'UDP-Glc deshydrogénase (KfiD), la fonction de KfiB n'est pas clairement définie, à l'exception d'un rôle dans la localisation et/ou dans l'organisation stérique du complexe héparosan synthase KfiA et KfiC. Quand nous avons commencé cette étude, nous avons réalisé que KfiB existait sous différentes isoformes se distinguant par la présence d'un nombre variable d'une séquence peptidique répétée (tandem repeat ou TR). Dans la mesure où les TRs ont été décrites comme pouvant réguler une activité enzymatique et/ou contrôler la taille du polymère en formation, nous nous sommes interrogés sur la possible implication du nombre de TRs sur la taille de l'héparosan synthétisé dans des bactéries recombinantes, sous la direction de Dr Bernard Priem et Dr Timm Fiebig. J'ai réalisé la conception de ce travail et construit les vecteurs d'expression exprimant les différents variants TRs. Les résultats ont été obtenus par réalisation de cultures en fed-batch. Les quantifications d'héparosan ont été réalisées par Marta Ferrer Xalabarder and Juliette Tridon. Tous les auteurs ont participé à l'analyse et l'interprétation des résultats. J'ai écrit cet article avec l'aide du Dr Bernard Priem. Prof Rita Gerardy-Schahn a également contribué à son écriture.

4.1. Introduction

Heparosan [α -1,4-GlcNAc- β -1,4-GlcA-] is a capsular polysaccharide (CPS; also known as K-antigen from the German word Kapsel) of Gram-negative bacteria such as *Escherichia coli* K5 and *Pasteurella multocida*. Most heparosan-producing strains are pathogenic, whereby the capsule is involved in attachment to the host cell and protection from the host immune response. The mechanism of heparosan synthesis (also known as K5 polysaccharide) in *E. coli* K5 has become a model for studying K antigens of group 2 capsule-producing bacteria (Whitfield, 2006; Jann & Jann, 1983). The capsule polysaccharide locus within group 2 capsule-producing bacteria comprises three regions; encoding conserved Kps proteins involved in the CPS export (Regions 1 and 3) and proteins (Region 2) responsible for synthesizing the serotype-specific CPS. Region 2 of the *E. coli* K5 capsule gene cluster contains four genes *kfi*ABCD that encode proteins forming the "heparosan synthase complex". The functions of KfiA, KfiC, and KfiD have been identified and shown to be α -GlcNAc-transferase, β -GlcA-transferase, and UDP-Glc dehydrogenase, respectively (Hodson et al., 2000; Sugiura et al., 2010; Sieberth et al., 1995). Notably, *kfi* genes could be expressed as recombinant proteins in different bacterial strains, including Gram-positive bacteria, to produce heparosan (Jin et al., 2016).

In contrast to KfiA, C, and D, the function of KfiB remains puzzling, and some authors still qualify KfiB as an "unknown protein" (Nzakizwanayo et al., 2015). However, the available data suggest the involvement of the protein in heparosan synthesis and capsule formation. For instance, based on the study by Hodson (Hodson et al., 2000), the hypothesis has been formulated that KfiB, along with KfiA, may play a role in the stable association of KfiC with the membrane. Later studies showed that *kfiB* knockout in *E. coli* Nissle 1917 modified the bacterial attachment properties. Based on the increased attachment of the mutant to intestinal epithelium cells, the authors suggested a modification in the capsule composition (Nzakizwanayo et al., 2015). Other studies analyzing *in vivo* production of heparosan showed differences in the yield and the degree of polymerization depending on whether *kfiB* was co-expressed with *kfiCAD* in recombinant *E. coli* BL21 (C. Zhang et al., 2012; Leroux & Priem, 2016).

KfiB homologous sequences are found in many pathogenic *E. coli* isolates, defining a family of KfiB isoforms. A starting point for the present work was our discovery that

variation in *kfiB* is limited to a small, spatially defined sequence domain of 84 nucleotides, which appeared twice in the "normal" sequence of *E. coli* K5 *kfiB* (i.e., at position 1114 and 1198). To facilitate further reading, we designated the repeat of 28 amino acids KfiB-repeat. In bacteria, DNA tandem repeats (TRs) are frequently found within genes involved in the synthesis of products present on the cell surface (Zhou et al., 2013). TRs are reported to regulate enzymatic activity (Bergfeld et al., 2007) or the length of growing polymers (Hagelueken et al., 2015). The KfiB-repeat identified in KfiB represents such a TR.

The protein prediction tool Phyre2 (Kelley et al., 2015) suggests KfiB to be an alpha-helix-rich two-domain protein that in part attains a coiled-coil architecture (**Fig. 21A**). The translation products of TRs (highlighted in green and magenta) separate these two domains. Notably, 12/28 amino acids in the KfiB-repeat are acidic, and a systematic search of gene sequences demonstrated that the KfiB-repeat could be present at different numbers (1 – 4). Furthermore, a zwitterionic motif formed by glutamic acid and lysine (EK-motif) was found to occur at two positions per KfiB-repeat (**Fig. 21A**). An analysis carried out by Nowinski et al. showed that E and K are the most prominent amino acids on the proteins surfaces. The EK motif, like zwitterions, interacts electrostatically with water molecules leading to the formation of a hydration layer (Nowinski et al., 2012; Yuan et al., 2020). In addition, the Kyte-Doolittle-Hydrophathy plot of KfiB revealed KfiB-repeat to be the most hydrophilic segment in the sequence (**Fig. 21B**) (Kyte & Doolittle, 1982; Duvaud et al., 2021).

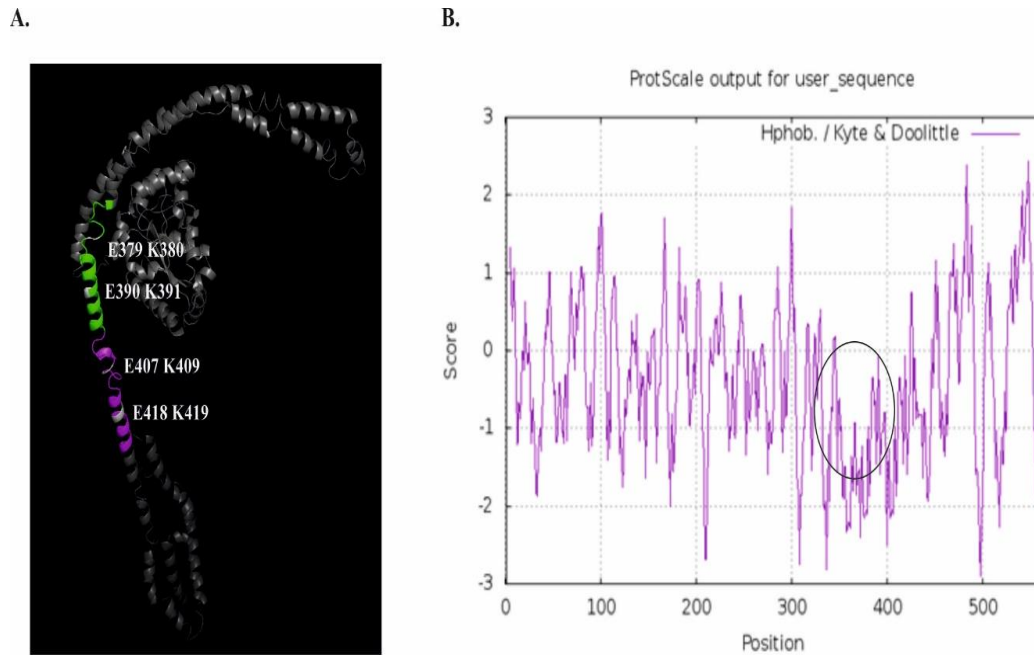


Figure 21. Analysis of the predicted KfiB structure. **A.** Structure of KfiB predicted in the Phyre2 web portal and visualized in Pymol (Kelley et al., 2015). The KfiB-repeats are highlighted in green and magenta, the EK motif in white. **B.** The Kyte-Doolittle-Hydrophobicity plot of KfiB. The KfiB-repeats are marked with a circle. The plot is based on the work of Kyte and Doolittle, 1982 and generated by expasy.org (Kyte & Doolittle, 1982; Duvaud et al., 2021).

To analyze the function of KfiB and interrogate the relevance of the number of expressed KfiB-repeats, we expressed defined KfiB isoforms carrying from one to four repeats in the nonpathogenic *E. coli* BL21 derivative, and we investigated their impact on the synthesis (production yield and molecular weight) of heparosan.

4.2. Results and discussion

4.2.1. Design of recombinant strains carrying different isoforms of KfiB

Available data concerning the synthesis and transport of heparosan exploit a KfiB variant containing two KfiB-repeats in its sequence (Leroux & Priem, 2016). However, the GenBank data display *E. coli* strains with varying numbers of KfiB-repeats (**Table 1**).

Strain	KfiB-repeat	GenBank
<i>Escherichia coli</i> ABU 83972	1	CP001671.1
<i>Escherichia coli</i> AR_0104	1	CP020116.1
<i>Escherichia coli</i> Nissle 1917	2	CP022686.1
<i>Escherichia coli</i> EH01-18-04-A	3	CP063515.1
<i>Escherichia coli</i> YD736 YD736_45	4	NZ_PGHK01000045.1

Table 1. *E. coli* strains with a varying number of KfiB-repeats in *kfiB* and its accession codes.

Zhang *et al.* demonstrated *in vivo* production of heparosan in the absence of KfiB, KfiD, or both concomitantly (Zhang *et al.*, 2012). Based on that, to investigate the significance of KfiB-repeat on heparosan production and export, we designed vectors that harbor only the *kfiCA* genes (Δ KfiB) and vectors containing the *kfiCAB* genes. As shown in **Fig. 22**, three variants of the latter vector containing *kfiB* with variant numbers of KfiB-repeats were cloned: KfiB_{1r} contains one, KfiB_{2r} two, and KfiB_{4r} four repeats (**Fig. 22**). The translation product of one KfiB-repeat corresponds to the following peptide sequence: KYEEEEISEKESKLTQAISEKEQILKQLH.

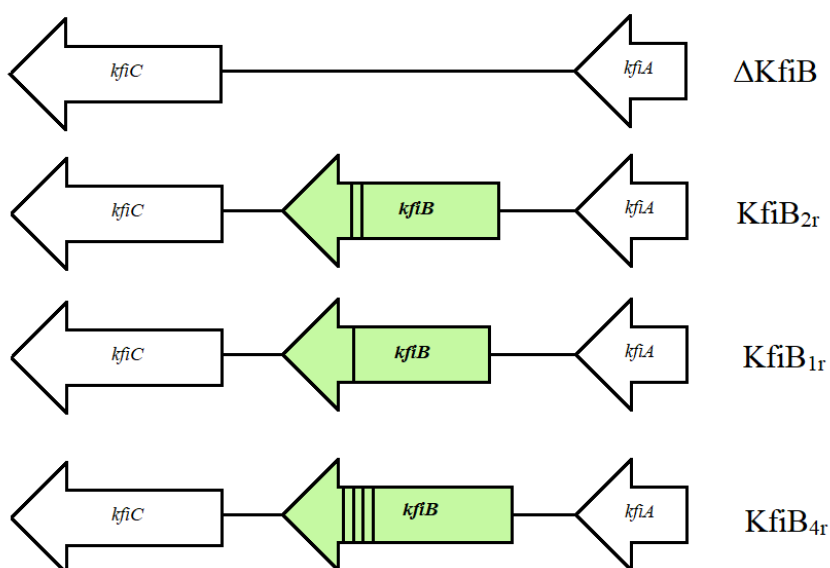


Figure 22. Schematic representation of the generated vectors. Schematic representation of generated vectors, without *kfiB* or with its variants, harboring different numbers of KfiB-repeats represented by vertical lines.

4.2.2. *In vitro* assays with heparosan synthases and KfiB isoforms

As a starter molecule (acceptor) for the *in vitro* synthesis of heparosan we used the heparosan K5 nonamer-2AB prepared in **Chapter 2**. The use of this homogenous acceptor allowed the comparative testing of heparosan synthesis complexes in the absence of KfiB (Δ KfiB) and the presence KfiB variants as indicated in **Fig. 23**. All tested constructs were active and elongated K5 nonamer-2AB. Overall, the product profiles were similar, however, with bacterial lysates containing KfiB_{1r} and KfiB_{4r} we detected longer chains. The effect was most prominent for KfiB_{4r}. Although this result hints towards a function of the KfiB-motif as a molecular ruler, the effect remained mild. Additionally, more work is needed to understand why bacterial lysates expressing KfiB_{2r} behaved similarly to lysates lacking KfiB (Δ KfiB).

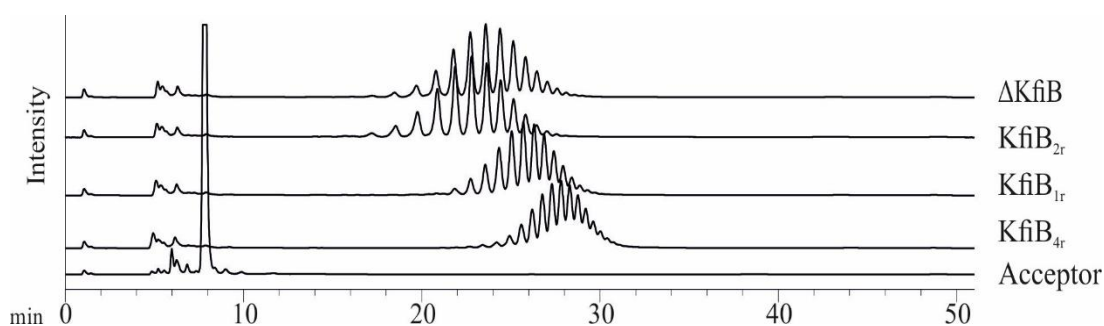


Figure 23. Elongation of K5 nonamer-2AB catalyzed by KfiCA/and KfiB and its variants. HPLC-AEC coupled to fluorescence detection.

4.2.3. *Influence of KfiB isoforms in vivo*

In a second step we asked, if the presence of KfiB may impact the yield of heparosan production *in vivo*. Therefore, the vectors as described (**Fig. 22**) were expressed in *E. coli* BL21-Gold (DE3) in a minimal medium (Priem et al., 2002). Since Hickey *et al.* have shown that *E. coli* Nissle 1917 and *E. coli* K5 shed heparosan into the medium (Hickey et al., 2013), we separated bacterial cultures by centrifugation and used a calorimetric assay to determine the amount of heparosan in the culture medium (sHep) and the pellet (caHep for cell-associated heparosan). Cultures were carried out using a fed-batch system.

Since no heparosan was found in the culture medium after 24 hours, we used a PCR analysis to test BL21-Gold (DE3) for the presence of the *kps* genes and also searched the KEGG database (Kanehisa et al., 2000; 2019; 2021). With the exception of KpsF, we confirmed the absence of *kps* genes (Kps proteins), thus explaining why

heparosan is accumulated intracellularly. In the pilot study (**Fig. 24**), highest yields of ceHep were found in bacteria expressing KfiB_{2r}, while KfiB_{1r} transformants produced the lowest yields.

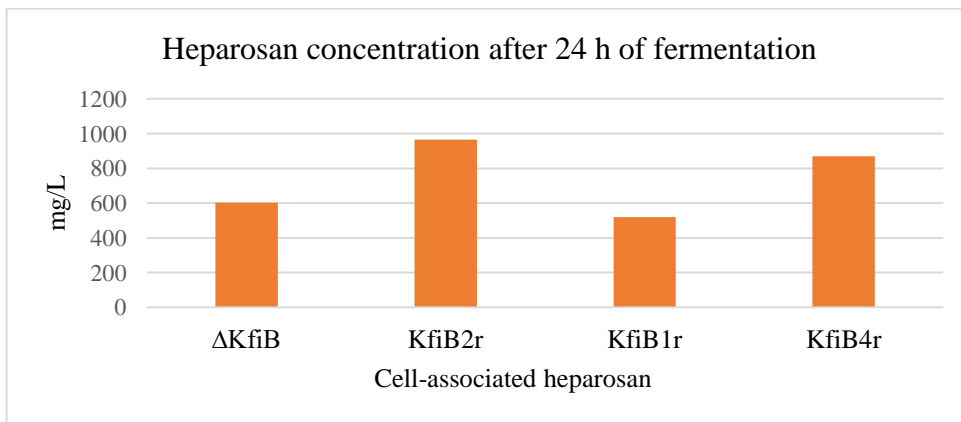


Figure 24. Heparosan production in *E. coli* BL21-Gold (DE3). The concentration of cell-associated heparosan produced by BL21-Gold (DE3) expressing recombinant plasmids in fed-batch cultures after 24 h.

Since isolation of sHep has many advantages over the isolation of ceHep, a second production system was tested. Here, we used the *E. coli* BJ strain carrying all *kps* genes (Leroux & Priem, 2016). As before, bacteria were cultured in fed-batch systems and the presence of heparosan tested in supernatant and pellet. Estimation of the total Hep concentration showed us that ΔKfiB produced less Hep than all variants expressing KfiB (**Fig. 25**). This result agrees with the reported data showing that KfiB helps with the attachment of KfiC to the inner membrane, thus participating in the interaction of the heparosan synthase complex with the Kps proteins export system (Hodson et al., 2010; Sugiura et al., 2010). A positive effect on heparosan synthesis was observed with all KfiB variants, at a first glance not suggesting an impact of the number of repeats on productivity in *E. coli* BJ.

After 24 h of feeding, produced heparosan splits equally between sHep and ceHep in ΔKfiB, while in KfiB variants sHep prevailed. In KfiB_{2r}, sHep represented 2/3 of total heparosan. After 48 h, the total heparosan concentration increased in each strain by ca. 22-40%. However, heparosan produced in the late feeding stage remained primarily caHep, leading to a roughly proportional ratio of sHep/caHep. Leroux & Priem observed a similar phenomenon in the study dealing with chaperone-fused KfiC (Leroux & Priem, 2016). This indicates that the heparosan shedding is highly

controlled, apparently dominating in the first stage of heparosan synthesis in the presence of KfiB and decreasing in the course of heparosan production.

The yield of heparosan produced in the engineered *E. coli* BL21-Gold (DE3) and BJ strains is comparable (Fig. 24 and 25), except for KfiB_{1r}. The expression of intracellularly trapped heparosan in BL21-Gold (DE3) KfiB_{1r} gave the yield ca. 60% of that of BJ KfiB_{1r}. Overall, our data confirm that KfiB promotes heparosan synthesis.

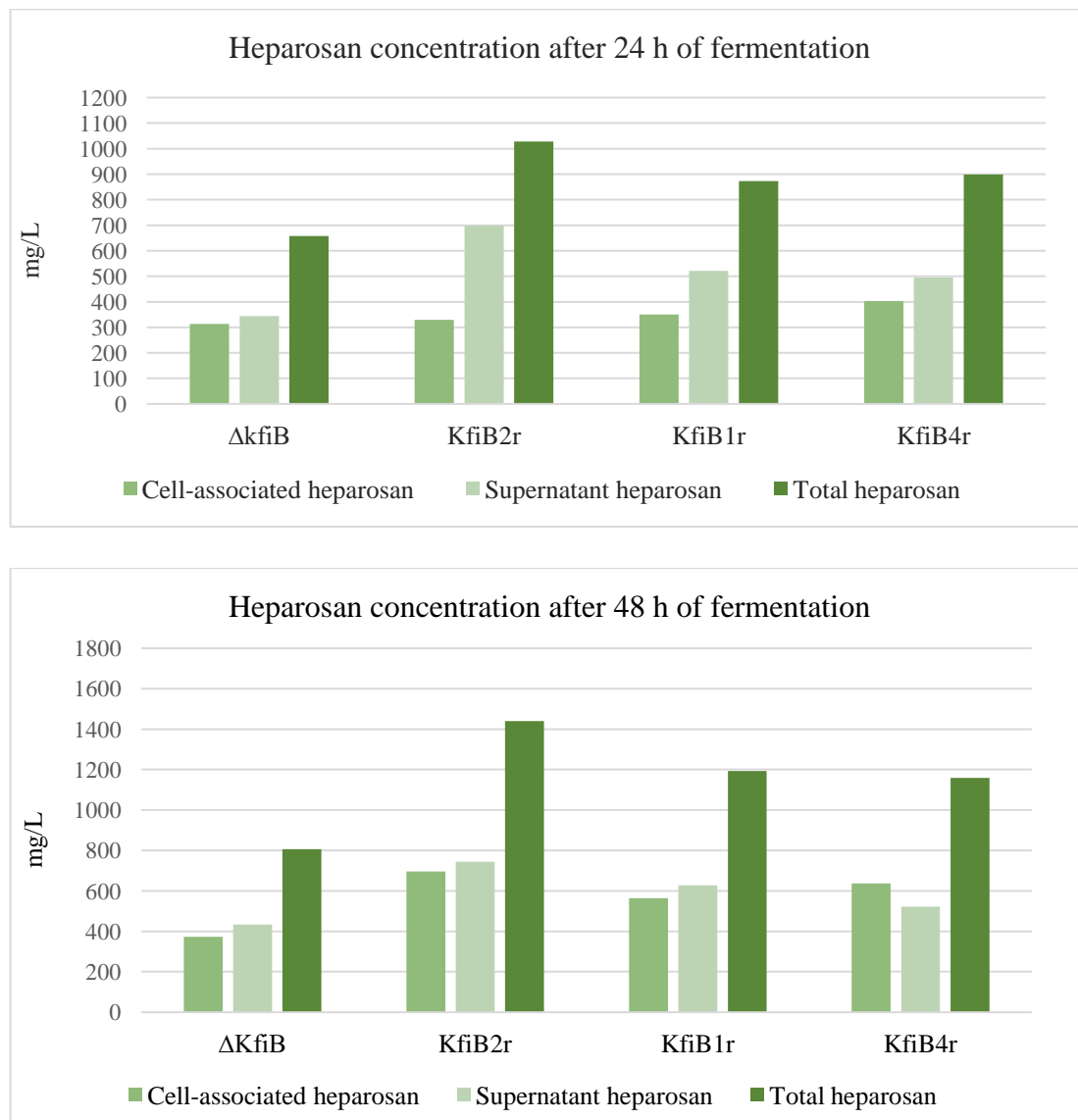


Figure 25. Heparosan production in recombinant *E. coli* BJ. Heparosan was produced in fed-batch cultures in recombinant *E. coli* BJ, after 24 h of feeding (upper panel) and 48 h of feeding (lower panel). Data represent the results of three biological replicates.

We used the method described by Blumenkrantz *et al.* to determine heparosan concentrations (Blumenkrantz & Asboe-Hansen, 1973). The disadvantage of this system is that the browning of neutral sugars when heated in concentrated sulfuric acid

interferes with the colorimetric determination of uronic acids (Filisetti-Cozzi & Carpita, 1991). Although the measurement results are overestimated by this method, it is still suitable for polymers containing many uronic acid residues (Filisetti-Cozzi & Carpita, 1991). As a second method we used our newly developed heparosan detection agent, Strep-eGFP-iKfIB (see **Chapter 2**), in the dot blot assay. *E. coli* BJ and *E. coli* K5 were used as negative and positive control, respectively. Overnight liquid bacterial cultures were adjusted to the optical density (OD_{600}) = 1.6. Being aware that caHep may be intracellular or may form a bacterial capsule, we prepared three samples: (1) the supernatant sample (sHep) was obtained upon liquid bacterial culture (LBC) centrifugation. For the second (2) sample the bacterial cultures were boiled before spotting to lyse the cells. This sample contains sHep and caHep. For the third (3) sample the LBC was directly spotted onto nitrocellulose, allowing detection of sHep and the fraction of caHep that is presented on the cell surface. The dot blot analysis confirmed the uronic acid quantification results all engineered strains produced heparosan. However, heparosan production in Δ KfIB is significantly compromised, resulting in very low signal intensity in samples (1) and (2) and no signal in sample (3). Different from the engineered *E. coli* strains, *E. coli* K5 expresses the enzyme eliminase (ElmA), previously shown to contribute to heparosan shedding (Hickey et al., 2013). Importantly, our dot blot results suggest that heparosan is shed also from engineered strains, if KfIB variants are present. Thereby, the dot blot evaluation did not show significant differences in sHep levels between *E. coli* K5 and engineered strains (**Fig. 26**).

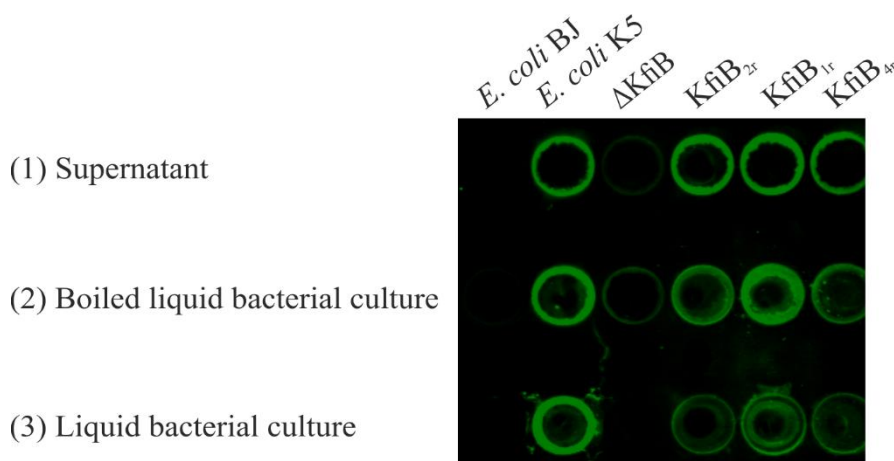


Figure 26. Detection of heparosan produced by engineered strains (green channel). Indirect detection with 1) Strep-eGFP-iKfIB (18.4 μ g/ml); 2) Strep-Tactin IR800 (1:2000).

4.2.4. Influence of Influence of KfiB isoforms on the size of Heparosan

The negative charge of heparosan allows its electrophoretic separation, followed by staining to compare MW and dispersity (\mathcal{D}) of polymers. In **Fig. 27** this type of experiment is shown. Total heparosan isolated from bacteria engineered as shown was run on a 15 % polyacrylamide gel and stained with alcian Blue/Silver. Particularly the higher contrast image demonstrates that heparosan isolated from Δ KfiB and KfiB_{2r} are similar with respect to MW and \mathcal{D} . However, in contrast to KfiB_{2r} fraction, heparosan isolated from the Δ KfiB culture contains a pool of very high MW heparosan, which is similarly present in the KfiB_{1r} and KfiB_{4r} samples. Overall, this analysis indicates that Δ KfiB and KfiB_{2r} produce polymers of higher \mathcal{D} and lower MW.

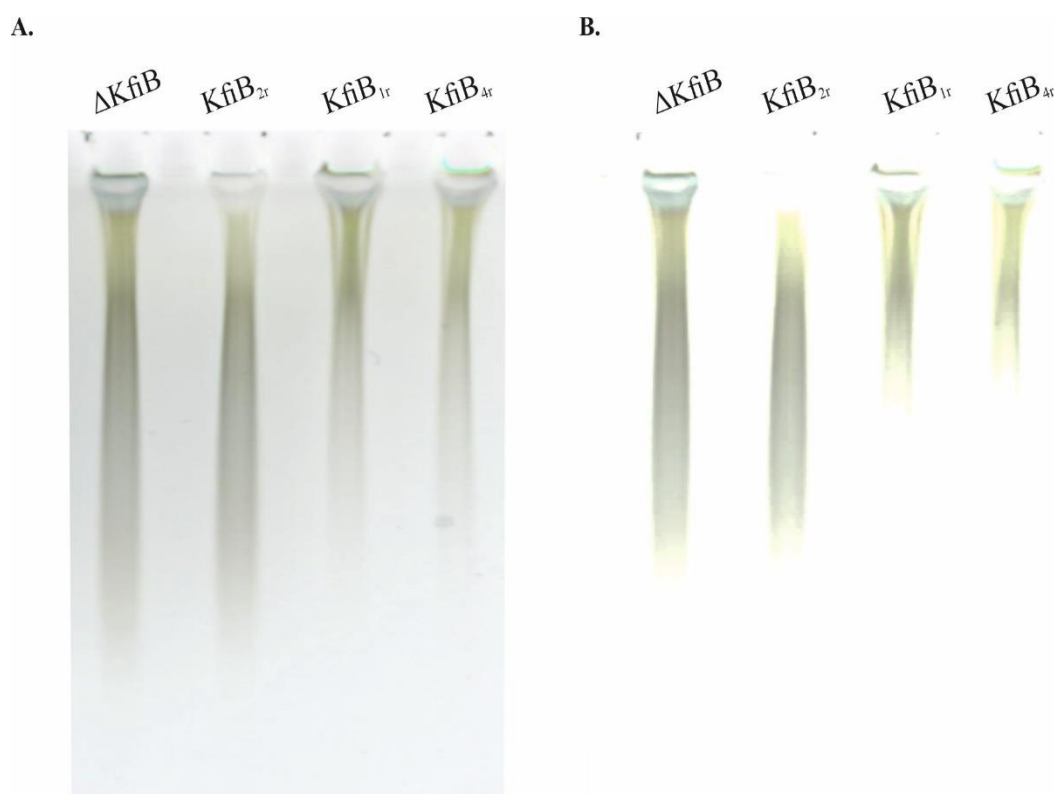


Figure 27. Comparison of heparosan molecular weight produced by engineered strains. A. Lower contrast image. **B.** Higher contrast image. Polymers (4.5 μ g) were run a 15% PA gel and visualized with alcian blue/silver.

To confirm the above data, we measured the MW also by size exclusion chromatography with multi-angle light scattering (SEC-MALS). This powerful technique can be used to determine the molar mass, conformation, or conjugation rate of biomolecules. However, despite many attempts and additional purification steps, produced heparosan samples contained aggregates. Aggregates were revealed by a

strong light scattering response accompanied by a low refractive index. The overlapping elution of aggregates with heparosan led to the overestimation of MW, and the lack of reproducibility between culture replicates.

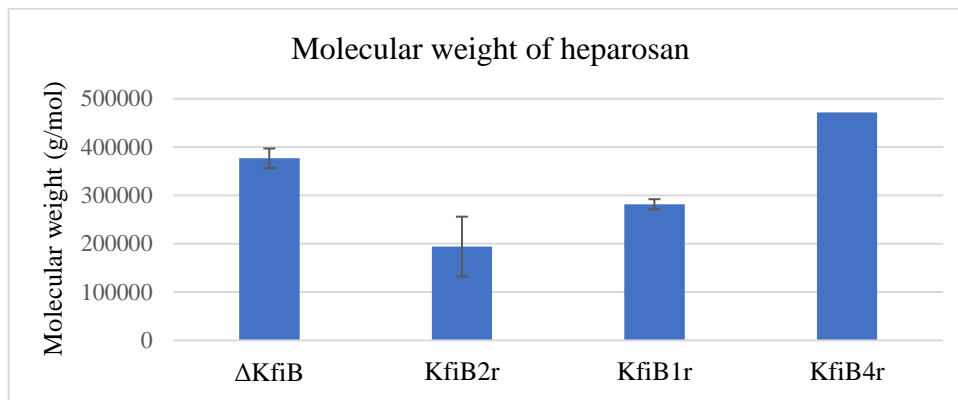


Figure 28. SEC-MALS analysis of heparosan. SEC-MALS was performed to determine the molecular weight of sHep expressed by *E. coli* BJ, 24 hours post-induction. Data represent the results of two technical replicates, except for KfiB_{4r}.

4.3. Conclusions

In this study, we wanted to elucidate the role of KfiB and of the so called KfiB-repeat on the quality and quantity of heparosan produced in engineered bacterial systems. Although the exact function remains unknown, KfiB is highly important in the heparosan synthesis machinery *in vivo*, having a beneficiary effect on the yield. However, the KfiB-repeat does not seem to affect this aspect, as the overall yields were comparable for each KfiB-containing strain. The formation of aggregates hindered the SEC-MALS analysis; therefore, the precise determination of MWs was not possible. Nevertheless, the polyacrylamide gel analysis showed modest differences between the KfiB variants, with KfiB_{1r} and KfiB_{4r} producing chains of slightly higher MW. This result corresponds to the *in vitro* assay. The obtained data suggest that heparosan shedding is a controlled process that decreases over time. The higher concentration of sHep 24 hours post-induction and its variability between KfiB strains suggest its role in shedding, which is influenced by the number of KfiB-repeats. In conclusion, although the role of KfiB still has to be fully elucidated, we have made a step forward to understand the mechanism of heparosan synthesis. Additionally, we determined that BL21-Gold (DE3) lacks some of the *kps* genes. This is vital information in the studies of capsular polymers.

4.4. Materials and methods

All repeating procedures were performed as in **Chapter 2** unless stated otherwise.

4.4.1. Bacterial strains, plasmids, application

*KfiB*_{2r} containing two KfiB-repeats of a length of 84 each was ligated downstream of *kfiCA* into engineered XhoI/PvuI restriction sites of the expression vector pBAD33-*kfiCA* (**Chapter 2**), labeled in this study as ΔKfiB. The part of the *kfiB*_{2r} gene, located between BbvCI/SacI sites, from pBAD33-*kfiCAB*_{2r} was substituted with synthesized (GenScript) gene fragments containing the following modifications, one KfiB-repeat deletion, insertion of two additional KfiB-repeats for KfiB_{1r} and KfiB_{4r}, respectively.

E. coli BJ (Leroux & Priem, 2016) and *E. coli* BL21-Gold(DE3) (Novagene, Darmstadt, Germany) strains were co-transformed with pBAD33 (*CmR*, *ara* induction) plasmid variants carrying different numbers of KfiB-repeats, and pBBR-*kfiD*-tetR, pBBR-*kfiD*-kanR, respectively (**Table 2**).

Strain	Plasmids	Relevant characteristics	Experiments
<i>E. coli</i> BL21-Gold (DE3)	pBAD33- <i>kfiCA</i> + pBBR- <i>kfiD</i>	<i>cmR kanR araC lac dcm ompT</i> (λ DE3)	<i>in vitro</i> dot blot fed-batch culture
	pBAD33- <i>kfiCAB</i> _{2r} + pBBR- <i>kfiD</i>		
	pBAD33- <i>kfiCAB</i> _{1r} + pBBR- <i>kfiD</i>		
	pBAD33- <i>kfiCAB</i> _{4r} + pBBR- <i>kfiD</i>		
<i>E. coli</i> BJ	pBAD33- <i>kfiCA</i> + pBBR- <i>kfiD</i>	<i>kpsFEDUCSTM cmR tetR araC lac dcm ompT wcaJ</i>	fed-batch culture PAGE
	pBAD33- <i>kfiCAB</i> _{2r} + pBBR- <i>kfiD</i>		
	pBAD33- <i>kfiCAB</i> _{1r} + pBBR- <i>kfiD</i>		
	pBAD33- <i>kfiCAB</i> _{4r} + pBBR- <i>kfiD</i>		

Table 2. Escherichia coli strains developed for this study and relevant information.

4.4.2. In vitro assay

The procedure was analogous to the elongation of the K5 oligosaccharide acceptor (**Chapter 2**). However, with the following changes: 50 mM Bis-Tris, the mixture did not contain Mg²⁺, and 5 μL of bacterial lysate was used in the total volume of 100 μl. Samples were analyzed in a 20% -3 curved gradient from 3 to 8 min,

followed by a linear gradient of 52% over the next 39 min using HPLC-AEC coupled to a fluorescent detector (Exc. $\lambda = 330$ nm, Emi. $\lambda = 420$ nm).

4.4.3. *Fed-batch culture*

Bacterial cultures were carried out at 28 °C, pH 6.8 in 0.5 L fed-batch bioreactors containing 0.2 L of the minimal medium described by Priem *et al.* (Priem *et al.*, 2002). With the following changes, upon the total exhaustion of the initially added glucose and glycerol, the fed-batch was started with 60 ml of 50% glycerol. Next, gene expressions were induced with 1 ml of 1 g/L L-arabinose and, with a half an hour delay, isopropyl β -D-1-thiogalactopyranoside (IPTG) at the final concentration of 0.2 mM. The cultures were carried out for 24 or 48 hours.

4.4.4. *Heparosan quantification*

Culture samples (1 ml) were centrifuged (3 min, 13,000 rpm) 24 or 48 hours post-induction. The supernatant was labeled as the "medium fraction", and the pellet was boiled for 20 min to release intracellular heparosan. Subsequently, the pellet was resuspended in 1 ml of water and centrifuged again (3 min, 13,000 rpm). The obtained supernatant was labeled as the "cell-associated fraction". To 250 μ l of each fraction, 750 μ l of 95% ethanol was added and centrifuged (3 min, 13,000 rpm) after 15 minutes. The obtained pellets were resuspended in 250 μ l of distilled water. The heparosan quantification was based on the sulfuric acid/tetraborate/m-hydroxydiphenyl method to determine the uronic acid amount (Blumenkrantz & Asboe-Hansen, 1973) with glucuronic acid used as a standard.

4.4.5. *Heparosan extraction*

The cells were harvested after 24/48 hours of the fed-batch culture by centrifugation (20 min, 9000 rpm), and the supernatant heparosan (medium) was separated from the cell-associated heparosan (pellet). The pellet was resuspended in distilled water, and cells were lysed by autoclaving (105 °C, 20 min) and subsequently centrifuged (20 min, 9000 rpm). The obtained pellet was discarded, and the supernatant was acidified to pH 4 with IR120 H⁺ (Sigma-Aldrich) to remove proteins and centrifuged (30 min, 7000 rpm). The supernatant pH was neutralized. Afterward, to the supernatant and the cell-associated heparosan mixture were added three volumes

of 95% ethanol and stored at 4 °C for 24 hours for heparosan precipitation. The precipitated heparosan was separated by centrifugation (20 min, 7000 rpm).

4.4.6. Heparosan purification

Samples containing heparosan (10-20 mg) were resuspended in water (1 ml), filtrated (0.22 µm), and loaded on an ion exchanger column, Mono Q 10/100 ("Mono Q[®] 10/100 GL" n.d.). The colorimetric assay detected heparosan-containing fractions (3 ml), desalted with HiPrep 26/10 Desalting column ("HiPrepTM 26/10 Desalting Cytiva, 17-5087-01, Column L × ID 10 Cm × 26 Mm, 17-132 Mm Particle Size (Dry)" n.d.) and lyophilized.

4.4.7. High-Performance Size-Exclusion Chromatography with Multi-Angle Light Scattering (HPSEC-MALS) analysis

Samples and solvent were filtrated (0.2 µm and 0.1 µm, respectively, Millipore). To determine weight average molecular weight and molecular weight distribution high-performance size-exclusion chromatography (HPSEC) (LC-20AD, Shimadzu, Japan) with online multi-angle light scattering (MALS) (miniDAWN TREOS, Wyatt Technology Corp., Santa Barbara, CA, USA) fitted with a K5 cell, a laser ($\lambda = 660$ nm), a refractive index detector, and a viscometer was used. Columns (OHPAK SB-G guard column and OHPAK SB 806MHQ column (Shodex)) were eluted with 0.1 M NaNO₃ containing 0.03 % NaN₃ at 0.5 mL min⁻¹.

4.5. Footnotes

The abbreviations: λ , wavelength; \mathcal{D} , dispersity; AEC, anion-exchange chromatography; CPS, capsular polysaccharide; caHep, cell-associated heparosan; cmR, chloramphenicol resistance; DNA, deoxyribonucleic acid; E, glutamic acid; *E. coli*, *Escherichia coli*; GlcA, glucuronic acid; GlcNAc, *N*-acetylglucosamine; HPSEC, high-performance size-exclusion chromatography; HPLC, high-performance liquid chromatography; IPTG, isopropyl β -D-1-thiogalactopyranoside; K, lysine; kanR, kanamycin resistance; KfiA, *N*-acetylglucosaminyltransferase; KfiC, D-glucuronyltransferase; KfiD, UDP-glucose-dehydrogenase; MALS, multi-angle light scattering; MW, molecular weight; OD, optical density; sHept, supernatant heparosan; tetR, tetracycline resistance; TRs, tandem repeats; UDP, uridine diphosphate.

Chapter 5 – General Discussion

During my PhD studies, experimental work was undertaken to establish improved conditions for the synthesis and detection of heparosan, an important material used for biomedical applications and as precursor for the synthesis of heparin and HS. Moreover, my goal was to expand the toolbox with respect to reagents for the detection of heparosan and other glycosaminoglycans. As heparosan is the capsular polysaccharide of *E. coli* Nissle and *E. coli* K5, investigations were focused at the use of bacterial enzymes to advance heparosan synthesis *in vitro* and *in vivo*. In this context, we investigated the functional mechanisms of heparosan synthases (KfiA/C) and also interrogated how KfiB, a protein of yet unknown function, contributes to the quality of synthesized heparosan. To gain access to a ready to use heparosan detection reagent, we took advantage of the Φ K5B borne lyase KfiB and in a structure guided approach transformed this protein into an artificial heparosan-specific lectin. Moreover, we characterized the mode of action of KfiB. Finally, we established a xCGE-LIF-based analytical method to profile cellular GAGs. In a first step, the migration times were determined for fluorescently labeled diGAG standards. Using this basis, we were able to analyze diGAGs isolated from hiPSC and hESC and to capture changes in diGAGs composition during hESC differentiation.

5.1. Extending the enzymatic toolbox for heparosan polymerization, depolymerization, and detection

My scientific contribution begins with **Chapter 2**, devoted to heparosan in the context of its synthesis, degradation, and detection. In the first stage of this project, we investigated the action mode of heparosan synthases *in vitro* and obtained fluorescent heparosan of HMW and low dispersity. Our results indicate that heparosan synthesis proceeds according to the distributive mechanism. An enzyme-substrate complex is transient and enzymes undergo repeating cycles of association, elongation, and dissociation from the product before the end of the polymerization reaction. While several studies showed a positive effect of KfiB on the heparosan yield *in vivo* (Zhang et al., 2012; Leroux & Priem, 2016), it did not correspond to our *in vitro* assays.

The emergence of bacteria exhibiting antibiotic resistance provides an increasing problem and alternative strategies are requested to reduce the over-use of antibiotics.

On this background, the application of bacteriophages is under intensive research. Prerequisite towards the application of bacteriophages is a comprehensive understanding of the phage infection process. We dedicated our research towards understanding the heparosan lyase (KflB) forming the tailspike protein of *E. coli* K5 specific Φ K5B. The bioinformatics analysis demonstrated that KflB like other TSPs contains a chaperone-like domain located at the C-terminal domain (Schwarzer et al., 2007), which is essential for correct folding and released from the folded protein after a highly conserved serine residue. The combination of two KflB substrates, unlabeled and labeled heparosan, enabled the development of two methods to detect the products of heparosan degradation. Based on a double bond in uronic acid at the non-reducing end and a fluorescent label at the reducing end for unlabeled and labeled heparosan, respectively.

The degradation product profile defines KflB as a processive endo-lyase. The processive mode was also reported for other TSPs (Schwarzer et al., 2009) and agrees with the observations that during the infection process, phage-borne depolymerases create a narrow tunnel in encapsulated bacteria (Lindberg, 1977). Interestingly, the retention of CTD in KflB S505A appears to alter its mode of action, changing KflB into a distributive lyase without affecting the reaction rate if contacted with heparosan in solution. The same phenomenon was described for a homotrimer forming endoNF and its mutant with the CTD retention (Schwarzer et al., 2009). A possible explanation is that the retention of the CTD interferes with the trimeric complex formation and, consequently, the substrate-enzyme complex. The structural studies on the KflA provide a possible explanation of the altered mode in the mutant KflB S505A (Thompson et al., 2010). The crystal structure of KflA indicated that the C-terminal domain of a mature KflA monomer interacts with a neighboring C-terminal domain of the two other KflA monomers, ultimately contributing to the homotrimer complex. Retaining the chaperone CTD could limit the full range of interactions prohibiting the formation of a stable homotrimer. Alternatively, based on the crystal structure of endoNF in complex with its substrate (Stummeyer et al., 2005) one may suggest that the polymer interacts with the binding sites of two monomers while being cleaved by the third. This cooperation enables the simultaneous degradation of three polysialic acid chains. The SDS-PAGE analysis suggests that KflB S505A may form a dimer indicating the quaternary protein structure modification. The extrapolation of the

results reported by Stummeyer *et al.* on KfIB may lead to the presumption that the lack of participation of the two binding sites is accountable for the altered mode of action in KfIB S505A. However, an unequivocal answer to the question what exactly alters the mode of action in KfIB S505A is still awaited and studies to deeper investigate the question were beyond the scope of my thesis.

Inspired by the artificial polysialic acid-specific lectin (Stummeyer *et al.*, 2005), we addressed the gap in heparosan-specific detection agents by developing a homologous reagent using a recombinant KfIB. The two prerequisites in KfIB engineering were to eliminate catalytic activity while maintaining specific heparosan binding. Multiple sequence alignment of heparosan lyases and the KfIA active site studies defined as E206 and K208 (Thompson *et al.* 2010) indicated the potential active sites of KfIB. Subsequent site-directed mutagenesis led to the generation of an inactive KfIB carrying alanine substitutions of catalytic residues. Interestingly, the collected data show significant catalytic inhibition for constructs having F202 and E206 mutations suggesting these two residues as part of the active site of KfIB. The construct initially used to engineer KfIB into an artificial heparosan-specific lectin was a N-terminal Strep-tagged sequence. However, Western blots carried out to interrogate accessibility of the Step-tag demonstrated poor results, most probably because of an unfavorable steric positioning of the tag. Several new constructs were cloned to overcome this obstacle. As an efficient construct we identified KfIB with N-terminally fused eGFP. In parallel, we labeled inactive KfIB with a fluorescent dye carrying a N-hydroxy succinimide activated ester. Although, the succinimide ester is commonly described to only react with primary amines, side reactions have been reported (Madler *et al.*, 2009). Clearly, this presents a disadvantage because it results in the generation of a heterogeneously labeled mixture of proteins. Moreover, numerous bulky groups on the protein surface could interfere with binding sites. However, we successfully detected heparosan with the fluorescently labeled inactive KfIB, indicating that binding sites were not compromised. Ultimately, we developed an artificial lectin enabling direct and indirect heparosan detection. Both detection agents showed high specificity, but a protocol using direct detection is more time efficient. Furthermore, as KfIB-catalyzed limited degradation of heparan sulfate most probably occurs in non-sulfated regions of a chain, KfIB opens the possibility to recognize them and quantify,

for instance in tissues, using Kf1B-based lectin. Bacterial serotyping is another exciting area of heparosan-specific lectin application.

5.2. Analysis of glycosaminoglycan-derivatives from multipotent pancreatic progenitor cells by xCGE-LIF

The high level of heterogeneity combined with similarity makes glycosaminoglycans challenging molecules to analyze and distinguish (Solakyildirim et al. 2019; Pepi et al. 2021). Nevertheless, their involvement in various biological processes brings them attention as promising biomarkers, potential therapeutic targets, or biomaterials (G. Li, et al., 2017; M. Li, et al., 2021; Nikitovic et al., 2021). Developing high-throughput analytical tools can significantly accelerate the elucidation of structure-function relationships in GAGs, thus, providing fundamentals to extend their application.

Many research groups have shown xCGE-LIF to stand out from other glycoanalytical techniques as being very sensitive, cost-effective, and capable of high throughput (Rossdam et al., 2019; Huffman et al., 2014). This led to our efforts to establish an xCGE-LIF-based method suitable to GAGs analysis reported in **Chapter 3**. The critical component of glycoanalysis by xCGE-LIF is a database, which we established in this study. Currently the database comprises 15 disaccharides and two heparosan oligomers. Ideally, in the future, the database would be extended to include other diGAGs, especially KS-derived. However, because only a minority of diGAGs is commercially available, the expansion of the database needs chemoenzymatic approaches to deliver diGAGs of defined structure.

We can observe subtle deviations in the MTU of analyzed diGAGs; unsulfated diGAGs are primarily prone to shifted migration times. Every analyzed sample was mixed with a size standard to prevent MTU deviations caused by the changes in the polymerization of the separation matrix. The size standards used in this study consisted of fluorescently labeled single-stranded DNA fragments. Migration times of analyzed glycans were calculated on a linear regression of the migration times of DNA fragments of defined length. Another factor affecting reproducibility is different charges of the separation matrix that influences migration times of analytes, especially glycans. Thus, an additional glycan-based size standard would greatly improve reproducibility. For instance, Fischöder *et al.* reported that tailored human milk

oligosaccharides could serve as well-defined reference standards (Fischöder et al., 2018).

Ultimately, we plan additional analysis replicates using cell-derived GAGs before publishing our results.

Despite of several advantages, our protocol shares some limitations with other GAG analysis technologies, which are based on the enzymatic generation of diGAGs (Osago et al., 2014; G. Li, et al., 2017). For instance, after enzymatic cleavage information about the epimeric state of uronic acid residues is lost thus preventing the distinguishing between IdoA and GlcA residues. Chemical depolymerization can be used to retain the epimeric state of UA but at the cost of losing information about acetylation (Shaklee & Conrad, 1984; Guo et al., 1989). A chemical approach to generate GAG-derived disaccharides was used by Gill *et al.* to develop a protocol combining chromatographic methods and MS (Gill et al., 2012).

The obtained data showed that our protocol has potential for the analysis of stem and differentiated cells. Other notable areas of the possible applications of the xCGE-LIF-based protocol include analysis of GAGs composition of the glycocalyx and their shedding in pathological events such as sepsis (Colbert & Schmidt, 2016). Furthermore, modifications in GAGs composition and their sulfation patterns have been linked to cancer progression (Morla, 2019). Thus, xCGE-LIF-based analysis lays a path to biomarkers development for disease diagnosis and progression.

Overall, xCGE-LIF is a promising powerful tool for analyzing GAG-derived disaccharides. The collected data show the successful separation of linkage and positional isomers, allowing their identification within the established standards database. This provided the basis for analyzing GAG profiles in stem and differentiated cells and, in the future, for determining GAG-based biomarkers.

5.3. Effect of KfiB on heparosan synthesis in recombinant

***Escherichia coli* BL-21**

Chapter 4 is dedicated to elucidate the effect of KfiB on heparosan synthesis in a recombinant *E. coli* BL21 derivative. The function of KfiB is unknown, and even its expression and purification have not yet been described in the English language literature. Initially, KfiB was perceived as an essential protein for heparosan synthesis along with KfiC and KfiA (Petit et al., 1995). Further research showed that heparosan

production also occurs without KfiB but is significantly impaired considering the yield (Zhang et al., 2012). However, Leroux & Priem showed that the fusion of KfiC with the *E. coli* trigger factor (TF) might compensate for the lack of KfiB, leading to higher yields *in vivo* (Leroux & Priem, 2016). Interestingly, Nzakizwanayo *et al.* reported a *E. coli* Nissle 1917 *kfiB* knock-out mutant to be capsule deficient based on the loss of sensitivity to *E. coli* K5 bacteriophage (Nzakizwanayo et al., 2015). It is important to note that the lack of a capsule is not synonymous with the ceased production. Analysis of

E. coli Nissle 1917 *kfiB* and *kfiC* knock-out mutants revealed that although both are capsule deficient, they exhibit different phenotypes, with *E. coli* Nissle 1917 $\Delta kfiB$ having higher biofilm formation capacity, cytotoxicity, and adherence to epithelial cells compared to *E. coli* Nissle 1917 wt and $\Delta kfiC$ (Nzakizwanayo et al., 2015). However, the underlying cause for various phenotypes has yet to be elucidated.

Our study was based on the finding that the KfiB sequence varies in the number of KfiB-repeats. We used a fluorescent acceptor, described in **Chapter 2**, to perform *in vitro* assays and conducted several fed-batch cultures to systematically address the impact of variant numbers of KfiB-repeats as well as the absence of KfiB on heparosan synthesis and further processing in terms of yield, MW, and shedding to a growth medium. Heparosan was absent in the supernatant of the first bacterial cultures, pointing us to the discovery that BL21-Gold (DE3) lacks *kps* genes. As transport and shedding of heparosan were one of the study objectives, other cultures were carried out in *E. coli* BJ. This *E. coli* BL21 derivative was designed by Leroux *et al.* to promote heparosan synthesis in the absence of the competing pathway of colanic acid biosynthesis (Leroux & Priem, 2016).

Our data suggest that heparosan shedding is controlled. The purpose of this process is unclear; however, it could have some role in host invading, as has been reported for the meningococcal strains (Vianzon et al., 2017).

Significant efforts have been made to elucidate the heparosan machinery, but many questions remain. The function of KfiB, mechanism, and significance of shedding still need to be determined. Moreover, the mechanism controlling the size of a polymer is unclear. Some polymerization systems use a molecular ruler to control the length of growing chains. A well-characterized example is the assembly of LPS in *E. coli* O9a. The polymannose glycan part of LPS is synthesized by the WbdA protein,

while WbdD terminates chain elongation by adding a closing residue (Hagelueken et al., 2015). Next to the WbdA:WbdD stoichiometry, the final chain length correlates with the length of the coiled-coil of WbdD (Hagelueken et al., 2015). Thus, the WbdD protein, acts as a molecular ruler through the varying length of the coiled-coil domain. The finding by Hagelueken *et al.* motivated us to investigate whether an analogous mechanism is used in KfiB by varying the number of KfiB-repeats, which in part attain a coiled-coil structure. Although we observe subtle changes in the MW of heparosan, qualitative differences do not directly correlate with the number of KfiB-repeats. Additional studies are needed to comprehend how heparosan synthesis is terminated. The possibilities include the kinetics of the chain translocation, loss of the glycosyltransferases affinity towards the nascent glycan chain if a certain length is reached, or an allosteric effect induced by other components (Whitfield, 2006). The existence of a chain capping, homologous to WbdD, also cannot be excluded as an additional residue at the non-reducing end could be easily overlooked in structural analysis (Whitfield, 2006). Another aspect likely connected to chain termination is their readiness for exportation. In the case of the mentioned O9a antigen, phosphomethyl modification at the non-reducing end of the O9a antigen is recognized by the ABC transporter, ensuring that only completed glycans are exported for assembling into LPS (Cuthbertson et al., 2010).

References

- Abatangelo, G., Vindigni, V., Avruscio, G., Pandis, L., & Brun, P. (2020). Hyaluronic Acid: Redefining Its Role. *Cells*, *9*(7), 1743. <https://doi.org/10.3390/cells9071743>
- Ackermann, Hans-W. (1998). Tailed Bacteriophages: The Order Caudovirales. *Advances in Virus Research*, *51*, 135–201. [https://doi.org/10.1016/s0065-3527\(08\)60785-x](https://doi.org/10.1016/s0065-3527(08)60785-x)
- Ackermann, Hans-W. (2011). The first phage electron micrographs. *Bacteriophage*, *1*(4), 225–227. <https://doi.org/10.4161/bact.1.4.17280>
- Auray-Blais, C., Lavoie, P., Zhang, H., Gagnon, R., Clarke, J. T. R., Maranda, B., ... Millington, D. S. (2012). An improved method for glycosaminoglycan analysis by LC-MS/MS of urine samples collected on filter paper. *Clinica Chimica Acta*, *413*(7-8), 771–778. <https://doi.org/10.1016/j.cca.2012.01.012>
- Barone, A., Säljö, K., Benktander, J., Blomqvist, M., Månsson, J.-E., Johansson, B. R., ... Teneberg, S. (2014). Sialyl-lactotetra, a Novel Cell Surface Marker of Undifferentiated Human Pluripotent Stem Cells. *Journal of Biological Chemistry*, *289*(27), 18846–18859. <https://doi.org/10.1074/jbc.m114.568832>
- Barreteau, H., Richard, E., Drouillard, S., Samain, E., & Priem, B. (2012). Production of intracellular heparosan and derived oligosaccharides by lyase expression in metabolically engineered *E. coli* K-12. *Carbohydrate Research*, *360*, 19–24. <https://doi.org/10.1016/j.carres.2012.07.013>
- Beimdiek, J., Hennig, R., Burock, R., Puk, O., Biskup, S., Rapp, E., ... Das, A. M. (2022). Serum N-glycomics of a novel CDG-IIb patient reveals aberrant IgG glycosylation. *Glycobiology*, *32*(5), 380–390. <https://doi.org/10.1093/glycob/cwac003>
- Bergfeld, A. K., Claus, H., Vogel, U., & Mühlenhoff, M. (2007). Biochemical Characterization of The polysialic Acid-specific O-Acetyltransferase NeuO of *Escherichia coli* K1. *Journal of Biological Chemistry*, *282*(30), 22217–22227. <https://doi.org/10.1074/jbc.m703044200>
- Blumenkrantz, N., & Asboe-Hansen, G. (1973). New method for quantitative determination of uronic acids. *Analytical Biochemistry*, *54*(2), 484–489. [https://doi.org/10.1016/0003-2697\(73\)90377-1](https://doi.org/10.1016/0003-2697(73)90377-1)
- Bohlmann, L., Tredwell, G. D., Yu, X., Chang, C.-W., Haselhorst, T., Winger, M., ... von Itzstein, M. (2015). Functional and structural characterization of a heparanase. *Nature Chemical Biology*, *11*(12), 955–957. <https://doi.org/10.1038/nchembio.1956>

- Bond, S. R., & Naus, C. C. (2012). RF-Cloning.org: an online tool for the design of restriction-free cloning projects. *Nucleic Acids Research*, *40*(W1), W209–W213. <https://doi.org/10.1093/nar/gks396>
- Breimer, M. E., Säljö, K., Barone, A., & Teneberg, S. (2017). Glycosphingolipids of human embryonic stem cells. *Glycoconjugate Journal*, *34*(6), 713–723. <https://doi.org/10.1007/s10719-016-9706-y>
- Bucay, V., Gold, M. H., & Andriessen, A. (2020). Low molecular weight heparan sulfate containing facial skin care for reducing inflammation and restoring aged-skin homeostasis. *Journal of Cosmetic Dermatology*, *19*(8), 1851–1856. <https://doi.org/10.1111/jocd.13528>
- Budde, I., Litschko, C., Führung, J. I., Gerardy-Schahn, R., Schubert, M., & Fiebig, T. (2020). An enzyme-based protocol for cell-free synthesis of nature-identical capsular oligosaccharides from *Actinobacillus pleuropneumoniae* serotype 1. *Journal of Biological Chemistry*, *295*(17), 5771–5784. <https://doi.org/10.1074/jbc.ra120.012961>
- Buffet, A., Rocha, E. P. C., & Rendueles, O. (2020). *Selection for the bacterial capsule in the absence of biotic and abiotic aggressions depends on growth conditions*. <https://doi.org/10.1101/2020.04.27.059774>
- Busse, M., & Kusche-Gullberg, M. (2003). In Vitro Polymerization of Heparan Sulfate Backbone by the EXT Proteins. *Journal of Biological Chemistry*, *278*(42), 41333–41337. <https://doi.org/10.1074/jbc.m308314200>
- Callewaert, N., Vlierberghe, H. V., Hecke, A. V., Laroy, W., Delanghe, J., & Contreras, R. (2004). Noninvasive diagnosis of liver cirrhosis using DNA sequencer-based total serum protein glycomics. *Nature Medicine*, *10*(4), 429–434. <https://doi.org/10.1038/nm1006>
- Carlsson, P., & Kjellén, L. (2012). Heparin biosynthesis. *Handbook of Experimental Pharmacology*, (207), 23–41. https://doi.org/10.1007/978-3-642-23056-1_2
- Carvalho, M. S., Silva, J. C., Hoff, C. M., Cabral, J. M. S., Linhardt, R. J., Silva, C. L., & Vashishth, D. (2020). Loss and rescue of osteocalcin and osteopontin modulate osteogenic and angiogenic features of mesenchymal stem/stromal cells. *Journal of Cellular Physiology*, *235*(10), 7496–7515. <https://doi.org/10.1002/jcp.29653>
- Caterson, B., & Melrose, J. (2018). Keratan sulfate, a complex glycosaminoglycan with unique functional capability. *Glycobiology*, *28*(4), 182–206. <https://doi.org/10.1093/glycob/cwy003>

- Chavaroche, A. A. E., van den Broek, L. A. M., Boeriu, C., & Eggink, G. (2011). Synthesis of heparosan oligosaccharides by *Pasteurella multocida* PmHS2 single-action transferases. *Applied Microbiology and Biotechnology*, *95*(5), 1199–1210. <https://doi.org/10.1007/s00253-011-3813-2>
- Chen, M., Bridges, A., & Liu, J. (2006). Determination of the Substrate Specificities of *N*-Acetyl-D-glucosaminyltransferase. *Biochemistry*, *45*(40), 12358–12365. <https://doi.org/10.1021/bi060844g>
- Chen, X., Zaro, J. L., & Shen, W.-C. (2013). Fusion protein linkers: Property, design and functionality. *Advanced Drug Delivery Reviews*, *65*(10), 1357–1369. <https://doi.org/10.1016/j.addr.2012.09.039>
- Chen, Y., Li, Y., Yu, H., Sugiarto, G., Thon, V., Hwang, J., ... Chen, X. (2013). Tailored Design and Synthesis of Heparan Sulfate Oligosaccharide Analogues Using Sequential One-Pot Multienzyme Systems. *Angewandte Chemie International Edition*, *52*(45), 11852–11856. <https://doi.org/10.1002/anie.201305667>
- Clarke, B. R., Esumeh, F., & Roberts, I. S. (2000). Cloning, Expression, and Purification of the K5 Capsular Polysaccharide Lyase (KflA) from Coliphage K5A: Evidence for Two Distinct K5 Lyase Enzymes. *Journal of Bacteriology*, *182*(13), 3761–3766. <https://doi.org/10.1128/jb.182.13.3761-3766.2000>
- Colbert, J. F., & Schmidt, E. P. (2016). Endothelial and microcirculatory function and dysfunction in sepsis. *Clinics in Chest Medicine*, *37*(2), 263–275. <https://doi.org/10.1016/j.ccm.2016.01.009>
- Couchman, J. R., & Pataki, C. A. (2012). An Introduction to Proteoglycans and Their Localization. *Journal of Histochemistry and Cytochemistry*, *60*(12), 885–897. <https://doi.org/10.1369/0022155412464638>
- Cress, B. F., Englaender, J. A., He, W., Kasper, D., Linhardt, R. J., & Koffas, M. A. G. (2014). Masquerading microbial pathogens: Capsular polysaccharides mimic host-tissue molecules. *FEMS Microbiology Reviews*, *38*(4), 660–697. <https://doi.org/10.1111/1574-6976.12056>
- Cuthbertson, L., Kos, V., & Whitfield, C. (2010). ABC Transporters Involved in Export of Cell Surface Glycoconjugates. *Microbiology and Molecular Biology Reviews*, *74*(3), 341–362. <https://doi.org/10.1128/mubr.00009-10>
- d'Hérelle, F. H. (1922). The Bacteriophage: Its Role in Immunity.
- d'Hérelle, F. H. (2011). On an invisible microbe antagonistic to dysentery bacilli. Note by M. F. d'Herelle, presented by M. Roux. *Comptes Rendus Academie des Sciences*

- 1917; 165:373–5. *Bacteriophage*, 1(1), 3–5. <https://doi.org/10.4161/bact.1.1.14941>
- Daines, D. A., Wright, L. F., Chaffin, D. O., Rubens, C. E., & Silver, R. P. (2000). NeuD plays a role in the synthesis of sialic acid in *Escherichia coli* K1. *FEMS Microbiology Letters*, 189(2), 281–284. <https://doi.org/10.1111/j.1574-6968.2000.tb09244.x>
- Damerell, D., Ceroni, A., Maass, K., Ranzinger, R., Dell, A., & Haslam, S. M. (2012). The GlycanBuilder and GlycoWorkbench glycoinformatics tools: updates and new developments. *Bchm*, 393(11), 1357–1362. <https://doi.org/10.1515/hsz-2012-0135>
- De Risi, M., Tufano, M., Alvino, F. G., Ferraro, M. G., Torromino, G., Gigante, Y., ... De Leonibus, E. (2021). Altered heparan sulfate metabolism during development triggers dopamine-dependent autistic-behaviours in models of lysosomal storage disorders. *Nature Communications*, 12(1). <https://doi.org/10.1038/s41467-021-23903-5>
- DeAngelis, P. L. (2002). Microbial glycosaminoglycan glycosyltransferases. *Glycobiology*, 12(1), 9R16. <https://doi.org/10.1093/glycob/12.1.9r>
- DeAngelis, P. L. (2015). Heparosan, a promising “naturally good” polymeric conjugating vehicle for delivery of injectable therapeutics. *Expert Opinion on Drug Delivery*, 12(3), 349–352. <https://doi.org/10.1517/17425247.2015.978282>
- DeAngelis, P. L., Liu, J., & Linhardt, R. J. (2013). Chemoenzymatic synthesis of glycosaminoglycans: Re-creating, re-modeling and re-designing nature’s longest or most complex carbohydrate chains. *Glycobiology*, 23(7), 764–777. <https://doi.org/10.1093/glycob/cwt016>
- DeAngelis, P. L., & White, C. L. (2002). Identification and Molecular Cloning of a Heparosan Synthase from *Pasteurella multocida* Type D. *Journal of Biological Chemistry*, 277(9), 7209–7213. <https://doi.org/10.1074/jbc.m112130200>
- DeAngelis, P. L., & White, C. L. (2004). Identification of a Distinct, Cryptic Heparosan Synthase from *Pasteurella multocida* Types A, D, and F. *Journal of Bacteriology*, 186(24), 8529–8532. <https://doi.org/10.1128/jb.186.24.8529-8532.2004>
- Dettmer, R., Niwolik, I., Cirksena, K., Yoshimoto, T., Tang, Y., Mehmeti, I., ... Naujok, O. (2022). Proinflammatory cytokines induce rapid, NO-independent apoptosis, expression of chemotactic mediators and interleukin-32 secretion in human pluripotent stem cell-derived beta cells. *Diabetologia*, 65(5), 829–843. <https://doi.org/10.1007/s00125-022-05654-0>
- Dettmer, R., Niwolik, I., Mehmeti, I., Jörns, A., & Naujok, O. (2021). New hPSC SOX9 and INS Reporter Cell Lines Facilitate the Observation and Optimization of

- Differentiation into Insulin-Producing Cells. *Stem Cell Reviews and Reports*, 17(6), 2193–2209. <https://doi.org/10.1007/s12015-021-10232-9>
- Duvaud, S., Gabella, C., Lisacek, F., Stockinger, H., Ioannidis, V., & Durinx, C. (2021). Expaty, the Swiss Bioinformatics Resource Portal, as designed by its users. *Nucleic Acids Research*, 49(W1), W216–W227.
- Essentials of Glycobiology. (2022). In A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, ... P. H. Seeberge (Eds.), *Glycobiology* (4th ed., Vol. 18). Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press.
<https://doi.org/10.1093/glycob/cwn065>
- Fenzl, C., Teramoto, K., & Moshirfar, M. (2015). Ocular manifestations and management recommendations of lysosomal storage disorders I: mucopolysaccharidoses. *Clinical Ophthalmology*, 9, 1633–1644. <https://doi.org/10.2147/ophth.s78368>
- Fiebig, T., Berti, F., Freiberger, F., Pinto, V., Claus, H., Romano, M. R., ... Gerardy-Schahn, R. (2013). Functional expression of the capsule polymerase of *Neisseria meningitidis* serogroup X: A new perspective for vaccine development. *Glycobiology*, 24(2), 150–158. <https://doi.org/10.1093/glycob/cwt102>
- Fiebig, T., Litschko, C., Freiberger, F., Bethe, A., Berger, M., & Gerardy-Schahn, R. (2018). Efficient solid-phase synthesis of meningococcal capsular oligosaccharides enables simple and fast chemoenzymatic vaccine production. *Journal of Biological Chemistry*, 293(3), 953–962. <https://doi.org/10.1074/jbc.ra117.000488>
- Filisetti-Cozzi, T. M. C. C., & Carpita, N. C. (1991). Measurement of uronic acids without interference from neutral sugars. *Analytical Biochemistry*, 197(1), 157–162.
[https://doi.org/10.1016/0003-2697\(91\)90372-z](https://doi.org/10.1016/0003-2697(91)90372-z)
- Fischöder, T., Cajic, S., Reichl, U., Rapp, E., & Elling, L. (2018). Enzymatic Cascade Synthesis Provides Novel Linear Human Milk Oligosaccharides as Reference Standards for xCGE-LIF Based High-Throughput Analysis. *Biotechnology Journal*, 14(3), 1800305. <https://doi.org/10.1002/biot.201800305>
- France, R. R., Cumpstey, I., Butters, T. D., Fairbanks, A. J., & Wormald, M. R. (2000). Fluorescence labelling of carbohydrates with 2-aminobenzamide (2AB). *Tetrahedron: Asymmetry*, 11(24), 4985–4994. [https://doi.org/10.1016/s0957-4166\(00\)00477-8](https://doi.org/10.1016/s0957-4166(00)00477-8)
- Frosch, M., Görgen, I., Boulnois, G. J., Timmis, K. N., & Bitter-Suermann, D. (1985). NZB mouse system for production of monoclonal antibodies to weak bacterial antigens: isolation of an IgG antibody to the polysaccharide capsules of *Escherichia coli* K1

- and group B meningococci. *Proceedings of the National Academy of Sciences*, 82(4), 1194–1198. <https://doi.org/10.1073/pnas.82.4.1194>
- Funderburgh, J. L. (2002). Keratan Sulfate Biosynthesis. *IUBMB Life (International Union of Biochemistry and Molecular Biology: Life)*, 54(4), 187–194. <https://doi.org/10.1080/15216540214932>
- Gasimli, L., Hickey, A. M., Yang, B., Li, G., dela Rosa, M., Nairn, A. V., ... Linhardt, R. J. (2014). Changes in glycosaminoglycan structure on differentiation of human embryonic stem cells towards mesoderm and endoderm lineages. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1840(6), 1993–2003. <https://doi.org/10.1016/j.bbagen.2014.01.007>
- Gill, V. L., Wang, Q., Shi, X., & Zaia, J. (2012). Mass Spectrometric Method for Determining the Uronic Acid Epimerization in Heparan Sulfate Disaccharides Generated Using Nitrous Acid. *Analytical Chemistry*, 84(17), 7539–7546. <https://doi.org/10.1021/ac3016054>
- Götzke, H., Kilisch, M., Martínez-Carranza, M., Sograte-Idrissi, S., Rajavel, A., Schlichthaerle, T., ... Frey, S. (2019). The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications. *Nature Communications*, 10(1), 4403. <https://doi.org/10.1038/s41467-019-12301-7>
- Guo, Y., & Conrad, H. Edward. (1989). The disaccharide composition of heparins and heparan sulfates. *Analytical Biochemistry*, 176(1), 96–104. [https://doi.org/10.1016/0003-2697\(89\)90278-9](https://doi.org/10.1016/0003-2697(89)90278-9)
- Guzzo, T., Barile, F., Marras, C., Bellini, D., Mandaliti, W., Nepravishta, R., ... Topai, A. (2020). Stability Evaluation and Degradation Studies of DAC[®] Hyaluronic-Polylactide Based Hydrogel by DOSY NMR Spectroscopy. *Biomolecules*, 10(11), 1478. <https://doi.org/10.3390/biom10111478>
- Haase, A., Göhring, G., & Martin, U. (2017). Generation of non-transgenic iPS cells from human cord blood CD34 + cells under animal component-free conditions. *Stem Cell Research*, 21, 71–73. <https://doi.org/10.1016/j.scr.2017.03.022>
- Hagelueken, G., Clarke, B. R., Huang, H., Tuukkanen, A., Danciu, I., Svergun, D. I., ... Naismith, J. H. (2015). A coiled-coil domain acts as a molecular ruler to regulate O-antigen chain length in lipopolysaccharide. *Nature Structural & Molecular Biology*, 22(1), 50–56. <https://doi.org/10.1038/nsmb.2935>
- Hampton, H. G., Watson, B. N. J., & Fineran, P. C. (2020). The arms race between bacteria and their phage foes. *Nature*, 577(7790), 327–336. <https://doi.org/10.1038/s41586->

- Hänfling, P., Shashkov, A. S., Jann, B., & Jann, K. (1996). Analysis of the enzymatic cleavage (beta elimination) of the capsular K5 polysaccharide of *Escherichia coli* by the K5-specific coliphage: reexamination. *Journal of Bacteriology*, 178(15), 4747–4750. <https://doi.org/10.1128/jb.178.15.4747-4750.1996>
- Hennig, R., Rapp, E., Kottler, R., Cajic, S., Borowiak, M., & Reichl, U. (2015). N-Glycosylation Fingerprinting of Viral Glycoproteins by xCGE-LIF. *Methods in Molecular Biology*, 123–143. https://doi.org/10.1007/978-1-4939-2874-3_8
- Hickey, A. M., Bhaskar, U., Linhardt, R. J., & Dordick, J. S. (2013). Effect of eliminase gene (*elmA*) deletion on heparosan production and shedding in *Escherichia coli* K5. *Journal of Biotechnology*, 165(3-4), 175–177. <https://doi.org/10.1016/j.jbiotec.2013.03.018>
- Higashi, K., Ly, M., Wang, Z., Masuko, S., Bhaskar, U., Sterner, E., ... Linhardt, R. J. (2011). Controlled photochemical depolymerization of K5 heparosan, a bioengineered heparin precursor. *Carbohydrate Polymers*, 86(3), 1365–1370. <https://doi.org/10.1016/j.carbpol.2011.06.042>
- Hitchcock, A. M., Bowman, M. J., Staples, G. O., & Zaia, J. (2008). Improved workup for glycosaminoglycan disaccharide analysis using CE with LIF detection. *Electrophoresis*, 29(22), 4538–4548. <https://doi.org/10.1002/elps.200800335>
- Hlozek, J., Kuttel, M. M., & Ravenscroft, N. (2018). Conformations of *Neisseria meningitidis* serogroup A and X polysaccharides: The effects of chain length and O-acetylation. *Carbohydrate Research*, 465, 44–51. <https://doi.org/10.1016/j.carres.2018.06.007>
- Hodson, N., Griffiths, G., Cook, N., Pourhossein, M., Gottfridson, E., Lind, T., ... Roberts, I. S. (2000). Identification That KfiA, a Protein Essential for the Biosynthesis of the *Escherichia coli* K5 Capsular Polysaccharide, Is an α -UDP-GlcNAc Glycosyltransferase. *Journal of Biological Chemistry*, 275(35), 27311–27315. [https://doi.org/10.1016/s0021-9258\(19\)61512-7](https://doi.org/10.1016/s0021-9258(19)61512-7)
- Hsu, H.-P., Chen, Y.-T., Chen, Y.-Y., Lin, C.-Y., Chen, P.-Y., Liao, S.-Y., ... Dzhagalov, I. L. (2021). Heparan sulfate is essential for thymus growth. *Journal of Biological Chemistry*, 296, 100419. <https://doi.org/10.1016/j.jbc.2021.100419>
- Huffman, J. E., Pučić-Baković, M., Klarić, L., Hennig, R., Selman, M. H. J., Vučković, F., ... Deelder, A. M. (2014). Comparative Performance of Four Methods for High-throughput Glycosylation Analysis of Immunoglobulin G in Genetic and

- Epidemiological Research. *Molecular & Cellular Proteomics*, 13(6), 1598–1610.
<https://doi.org/10.1074/mcp.m113.037465>
- Hughes, K. A., Sutherland, I. W., & Jones, M. V. (1998). Biofilm susceptibility to bacteriophage attack: the role of phage-borne polysaccharide depolymerase. *Microbiology*, 144(11), 3039–3047. <https://doi.org/10.1099/00221287-144-11-3039>
- Hynnekleiv, L., Magno, M., Vernhardsdottir, R. R., Moschowits, E., Tønseth, K. A., Dartt, D. A., ... Utheim, T. P. (2022). Hyaluronic acid in the treatment of dry eye disease. *Acta Ophthalmologica*, 100, 844–860.
- Itano, N., Sawai, T., Yoshida, M., Lenas, P., Yamada, Y., Imagawa, M., ... Kimata, K. (1999). Three Isoforms of Mammalian Hyaluronan Synthases Have Distinct Enzymatic Properties. *Journal of Biological Chemistry*, 274(35), 25085–25092. <https://doi.org/10.1074/jbc.274.35.25085>
- Jann, K., & Jann, B. (1983). The K Antigens of Escherichia coli. *Host Parasite Relationships in Gram-Negative Infections*, 33, 53–79. <https://doi.org/10.1159/000407421>
- Janosz, E., Hetzel, M., Spielmann, H., Tumpara, S., Rossdam, C., Schwabbauer, M., ... Moritz, T. (2021). Pulmonary transplantation of alpha-1 antitrypsin (AAT)-transgenic macrophages provides a source of functional human AAT in vivo. *Gene Therapy*, 28(9), 477–493. <https://doi.org/10.1038/s41434-021-00269-3>
- Jin, L., Abrahams, J. P., Skinner, R., Petitou, M., Pike, R. N., & Carrell, R. W. (1997). The anticoagulant activation of antithrombin by heparin. *Proceedings of the National Academy of Sciences*, 94(26), 14683–14688. <https://doi.org/10.1073/pnas.94.26.14683>
- Jin, P., Zhang, L., Yuan, P., Kang, Z., Du, G., & Chen, J. (2016). Efficient biosynthesis of polysaccharides chondroitin and heparosan by metabolically engineered *Bacillus subtilis*. *Carbohydrate Polymers*, 140, 424–432. <https://doi.org/10.1016/j.carbpol.2015.12.065>
- Jing, W., & DeAngelis, P. L. (2004). Synchronized chemoenzymatic synthesis of monodisperse hyaluronan polymers. *The Journal of Biological Chemistry*, 279(40), 42345–42349. <https://doi.org/10.1074/jbc.M402744200>
- Jing, W., Roberts, J. W., Green, D. E., Almond, A., & DeAngelis, P. L. (2017). Synthesis and characterization of heparosan-granulocyte-colony stimulating factor conjugates: a natural sugar-based drug delivery system to treat neutropenia. *Glycobiology*, 27(11), 1052–1061. <https://doi.org/10.1093/glycob/cwx072>
- Johnson, C. E., Crawford, B. E., Stavridis, M., ten Dam, G., Wat, A. L., Rushton, G., ...

- Merry, C. L. R. (2007). Essential Alterations of Heparan Sulfate During the Differentiation of Embryonic Stem Cells to Sox1-Enhanced Green Fluorescent Protein-Expressing Neural Progenitor Cells. *STEM CELLS*, 25(8), 1913–1923. <https://doi.org/10.1634/stemcells.2006-0445>
- Kamimura, K., & Maeda, N. (2021). Glypicans and Heparan Sulfate in Synaptic Development, Neural Plasticity, and Neurological Disorders. *Frontiers in Neural Circuits*, 15. <https://doi.org/10.3389/fncir.2021.595596>
- Kanehisa, M. (2019). Toward understanding the origin and evolution of cellular organisms. *Protein Science*, 28(11). <https://doi.org/10.1002/pro.3715>
- Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., & Tanabe, M. (2020). KEGG: integrating viruses and cellular organisms. *Nucleic Acids Research*, 49(D1), D545–D551. <https://doi.org/10.1093/nar/gkaa970>
- Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, 28(1), 27–30. <https://doi.org/10.1093/nar/28.1.27>
- Kannagi, R., Cochran, N. A., Ishigami, F., Hakomori, S., Andrews, P. W., Knowles, B. B., & Solter, D. (1983). Stage-specific embryonic antigens (SSEA-3 and -4) are epitopes of a unique globo-series ganglioside isolated from human teratocarcinoma cells. *The EMBO Journal*, 2(12), 2355–2361. <https://doi.org/10.1002/j.1460-2075.1983.tb01746.x>
- Keire, D., Szajek, A., Gray, E., Mulloy, B., Al-Hakim, A., Chase, C., ... Morris, T. S. (2015). Diversifying the Global Heparin Supply Chain: Reintroduction of Bovine Heparin in the United States? *Pharmaceutical Technology*, 39(11).
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Pyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10(6), 845–858. <https://doi.org/10.1038/nprot.2015.053>
- Keys, T. G., Berger, M., & Gerardy-Schahn, R. (2012). A high-throughput screen for polysialyltransferase activity. *Analytical Biochemistry*, 427(1), 60–68. <https://doi.org/10.1016/j.ab.2012.04.033>
- Keys, T. G., Freiburger, F., Ehrit, J., Krueger, J., Eggers, K., Buettner, F. F. R., & Gerardy-Schahn, R. (2012). A universal fluorescent acceptor for high-performance liquid chromatography analysis of pro- and eukaryotic polysialyltransferases. *Analytical Biochemistry*, 427(2), 107–115. <https://doi.org/10.1016/j.ab.2012.05.011>
- Keys, T. G., Fuchs, H. L. S., Ehrit, J., Alves, J., Freiburger, F., & Gerardy-Schahn, R. (2014). Engineering the product profile of a polysialyltransferase. *Nature Chemical*

Biology, 10(6), 437–442. <https://doi.org/10.1038/nchembio.1501>

Khan, S. A., Mason, R. W., Kobayashi, H., Yamaguchi, S., & Tomatsu, S. (2020). Advances in glycosaminoglycan detection. *Molecular Genetics and Metabolism*, 130(2). <https://doi.org/10.1016/j.ymgme.2020.03.004>

Khurana. (2008). *Essentials of Medical Physiology*. Elsevier India.

Kiermaier, E., Moussion, C., Veldkamp, C. T., Gerardy-Schahn, R., de Vries, I., Williams, L. G., ... Sixt, M. (2016). Polysialylation controls dendritic cell trafficking by regulating chemokine recognition. *Science*, 351(6269), 186–190. <https://doi.org/10.1126/science.aad0512>

Kizer, M., Li, P., Cress, B. F., Lin, L., Jing, T. T., Zhang, X., ... Wang, X. (2018). RNA Aptamers with Specificity for Heparosan and Chondroitin Glycosaminoglycans. *ACS Omega*, 3(10), 13667–13675. <https://doi.org/10.1021/acsomega.8b01853>

Knecht, L. E., Veljkovic, M., & Fieseler, L. (2020). Diversity and Function of Phage Encoded Depolymerases. *Frontiers in Microbiology*, 10. <https://doi.org/10.3389/fmicb.2019.02949>

Konze, S. A., Cajic, S., Oberbeck, A., Hennig, R., Pich, A., Rapp, E., & Buettner, F. F. R. (2017). Quantitative Assessment of Sialo-Glycoproteins and N-Glycans during Cardiomyogenic Differentiation of Human Induced Pluripotent Stem Cells. *ChemBioChem*, 18(13), 1317–1331. <https://doi.org/10.1002/cbic.201700100>

Krusius, T., Finne, J., Margolis, R. K., & Margolis, R. U. (1986). Identification of an O-glycosidic mannose-linked sialylated tetrasaccharide and keratan sulfate oligosaccharides in the chondroitin sulfate proteoglycan of brain. *Journal of Biological Chemistry*, 261(18), 8237–8242. [https://doi.org/10.1016/s0021-9258\(19\)83901-7](https://doi.org/10.1016/s0021-9258(19)83901-7)

Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology*, 157(1), 105–132. [https://doi.org/10.1016/0022-2836\(82\)90515-0](https://doi.org/10.1016/0022-2836(82)90515-0)

Lane, R. S., Haller, F. M., Chavarroche, A. A. E., Almond, A., & DeAngelis, P. L. (2017). Heparosan-coated liposomes for drug delivery. *Glycobiology*, 27(11), 1062–1074. <https://doi.org/10.1093/glycob/cwx070>

Legoux, R., Lelong, P., Jourde, C., Feuillerat, C., Capdevielle, J., Sure, V., ... Salomé, M. (1996). N-acetyl-heparosan lyase of *Escherichia coli* K5: gene cloning and expression. *Journal of Bacteriology*, 178(24), 7260–7264. <https://doi.org/10.1128/jb.178.24.7260-7264.1996>

- Leiman, P. G., Battisti, A. J., Bowman, V. D., Stummeyer, K., Mühlhoff, M., Gerardy-Schahn, R., ... Molineux, I. J. (2007). The Structures of Bacteriophages K1E and K1-5 Explain Processive Degradation of Polysaccharide Capsules and Evolution of New Host Specificities. *Journal of Molecular Biology*, 371(3), 836–849. <https://doi.org/10.1016/j.jmb.2007.05.083>
- Leiphraipam, P. D., Patil, P. P., Remmers, N., Swanson, B., Grandgenett, P. M., Qiu, F., ... Radhakrishnan, P. (2019). Role of keratan sulfate expression in human pancreatic cancer malignancy. *Scientific Reports*, 9(1), 9665. <https://doi.org/10.1038/s41598-019-46046-6>
- Lepedda, A. J., Nieddu, G., Formato, M., Baker, M. B., Fernández-Pérez, J., & Moroni, L. (2021). Glycosaminoglycans: From Vascular Physiology to Tissue Engineering Applications. *Frontiers in Chemistry*, 9. <https://doi.org/10.3389/fchem.2021.680836>
- Leroux, M., & Priem, B. (2016). Chaperone-assisted expression of KfiC glucuronyltransferase from Escherichia coli K5 leads to heparosan production in Escherichia coli BL21 in absence of the stabilisator KfiB. *Applied Microbiology and Biotechnology*, 100(24), 10355–10361. <https://doi.org/10.1007/s00253-016-7745-8>
- Li, G., Li, L., Joo, E. J., Son, J. W., Kim, Y. J., Kang, J. K., ... Linhardt, R. J. (2017). Glycosaminoglycans and glycolipids as potential biomarkers in lung cancer. *Glycoconjugate Journal*, 34(5), 661–669. <https://doi.org/10.1007/s10719-017-9790-7>
- Li, G., Li, L., Tian, F., Zhang, L., Xue, C., & Linhardt, R. J. (2015). Glycosaminoglycanomics of Cultured Cells Using a Rapid and Sensitive LC-MS/MS Approach. *ACS Chemical Biology*, 10(5), 1303–1310. <https://doi.org/10.1021/acscchembio.5b00011>
- Li, M., Sun, J., Zhang, W., Zhao, Y., Zhang, S., & Zhang, S. (2021). Drug delivery systems based on CD44-targeted glycosaminoglycans for cancer therapy. *Carbohydrate Polymers*, 251, 117103. <https://doi.org/10.1016/j.carbpol.2020.117103>
- Li, Y., Yu, H., Thon, V., Chen, Y., Muthana, M. M., Qu, J., ... Chen, X. (2013). Donor substrate promiscuity of the N-acetylglucosaminyltransferase activities of Pasteurella multocida heparosan synthase 2 (PmHS2) and Escherichia coli K5 KfiA. *Applied Microbiology and Biotechnology*, 98(3), 1127–1134. <https://doi.org/10.1007/s00253-013-4947-1>
- Liang, Y.-J., Kuo, H.-H., Lin, C.-H., Chen, Y.-Y., Yang, B.-C., Cheng, Y.-Y., ... Yu, J. (2010). Switching of the core structures of glycosphingolipids from globo- and

- lacto- to ganglio-series upon human embryonic stem cell differentiation. *Proceedings of the National Academy of Sciences*, 107(52), 22564–22569. <https://doi.org/10.1073/pnas.1007290108>
- Liang, Y.-J., Yang, B.-C., Chen, J.-M., Lin, Y.-H., Huang, C.-L., Cheng, Y.-Y., ... Yu, J. (2011). Changes in Glycosphingolipid Composition During Differentiation of Human Embryonic Stem Cells to Ectodermal or Endodermal Lineages. *STEM CELLS*, 29(12), 1995–2004. <https://doi.org/10.1002/stem.750>
- Lin, R.-J., Kuo, M.-W., Yang, B.-C., Tsai, H.-H., Chen, K., Huang, J.-R., ... Yu, J. (2020). B3GALT5 knockout alters glycosphingolipid profile and facilitates transition to human naïve pluripotency. *Proceedings of the National Academy of Sciences*, 117(44), 27435–27444. <https://doi.org/10.1073/pnas.2003155117>
- Lindahl, B., Eriksson, L., Spillmann, D., Caterson, B., & Lindahl, U. (1996). Selective Loss of Cerebral Keratan Sulfate in Alzheimer's Disease. *Journal of Biological Chemistry*, 271(29), 16991–16994. <https://doi.org/10.1074/jbc.271.29.16991>
- Lindberg, AA. (1977). *Bacterial Surface Carbohydrates and Bacteriophage Adsorption* (pp. 289–356; I. W. Sutherland, Ed.). Academia Press, London; United Kingdom.
- Linhardt, R. J. (2003). *Current Protocols in Molecular Biology*. John Wiley & Sons Inc. (Original work published 1999)
- Liu, H., & Naismith, J. H. (2008). An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnology*, 8(1), 91. <https://doi.org/10.1186/1472-6750-8-91>
- Lowry, K. M., & Beavers, E. M. (1994). Thermal stability of sodium hyaluronate in aqueous solution. *Journal of Biomedical Materials Research*, 28(10), 1239–1244. <https://doi.org/10.1002/jbm.820281014>
- Ludwigs, U., Elgavish, A., Esko, J. D., Meezan, E., & Rodén, L. (1987). Reaction of unsaturated uronic acid residues with mercuric salts. Cleavage of the hyaluronic acid disaccharide 2-acetamido-2-deoxy-3-O-(β -d-gluco-4-enopyranosyluronic acid)-d-glucose. *Biochemical Journal*, 245(3), 795–804. <https://doi.org/10.1042/bj2450795>
- Ma, S.-N., Mao, Z.-X., Wu, Y., Liang, M.-X., Wang, D.-D., Chen, X., ... Tang, J.-H. (2020). The anti-cancer properties of heparin and its derivatives: a review and prospect. *Cell Adhesion & Migration*, 14(1), 118–128. <https://doi.org/10.1080/19336918.2020.1767489>
- Mädler, S., Bich, C., Touboul, D., & Zenobi, R. (2009). Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. *Journal of Mass Spectrometry*:

JMS, 44(5), 694–706. <https://doi.org/10.1002/jms.1544>

- Maloney, F. P., Kuklewicz, J., Corey, R. A., Bi, Y., Ho, R., Mateusiak, L., ... Zimmer, J. (2022). Structure, substrate recognition and initiation of hyaluronan synthase. *Nature*, 604(7904), 195–201. <https://doi.org/10.1038/s41586-022-04534-2>
- Mao, J., Qiu, L., Ge, L., Zhou, J., Ji, Q., Yang, Y., ... Chen, J. (2021). Overcoming multidrug resistance by intracellular drug release and inhibiting p-glycoprotein efflux in breast cancer. *Biomedicine & Pharmacotherapy*, 134, 111108. <https://doi.org/10.1016/j.biopha.2020.111108>
- Mikael, P. E., Willard, C., Koyee, A., Barlao, C.-G., Liu, X., Han, X., ... Dordick, J. S. (2019). Remodeling of Glycosaminoglycans During Differentiation of Adult Human Bone Mesenchymal Stromal Cells Toward Hepatocytes. *Stem Cells and Development*, 28(4), 278–289. <https://doi.org/10.1089/scd.2018.0197>
- Mikami, T., & Kitagawa, H. (2013). Biosynthesis and function of chondroitin sulfate. *Biochimica et Biophysica Acta*, 1830(10), 4719–4733. <https://doi.org/10.1016/j.bbagen.2013.06.006>
- Mikami, T., & Kitagawa, H. (2016). Sulfated glycosaminoglycans: their distinct roles in stem cell biology. *Glycoconjugate Journal*, 34(6), 725–735. <https://doi.org/10.1007/s10719-016-9732-9>
- Miller, T., Goude, M. C., McDevitt, T. C., & Temenoff, J. S. (2014). Molecular engineering of glycosaminoglycan chemistry for biomolecule delivery. *Acta Biomaterialia*, 10(4), 1705–1719. <https://doi.org/10.1016/j.actbio.2013.09.039>
- Millis, G. (2002). *System and Method for Selecting and Transmitting Images of Interest To a User*.
- Mizumoto, S., & Yamada, S. (2021). An Overview of in vivo Functions of Chondroitin Sulfate and Dermatan Sulfate Revealed by Their Deficient Mice. *Frontiers in Cell and Developmental Biology*, 9. <https://doi.org/10.3389/fcell.2021.764781>
- Morla, S. (2019). Glycosaminoglycans and Glycosaminoglycan Mimetics in Cancer and Inflammation. *International Journal of Molecular Sciences*, 20(8). <https://doi.org/10.3390/ijms20081963>
- Mühlenhoff, M., Stummeyer, K., Grove, M., Sauerborn, M., & Gerardy-Schahn, R. (2003). Proteolytic Processing and Oligomerization of Bacteriophage-derived Endosialidases. *Journal of Biological Chemistry*, 278(15), 12634–12644. <https://doi.org/10.1074/jbc.m212048200>

- Murphy, K. J., Merry, C. L. R., Lyon, M., Thompson, J. E., Roberts, I. S., & Gallagher, J. T. (2004). A New Model for the Domain Structure of Heparan Sulfate Based on the Novel Specificity of K5 Lyase. *Journal of Biological Chemistry*, 279(26), 27239–27245. <https://doi.org/10.1074/jbc.m401774200>
- Na, L., Yu, H., McArthur, J. B., Ghosh, T., Asbell, T., & Chen, X. (2020). Engineer *P. multocida* Heparosan Synthase 2 (PmHS2) for Size-Controlled Synthesis of Longer Heparosan Oligosaccharides. *ACS Catalysis*, 10(11), 6113–6118. <https://doi.org/10.1021/acscatal.0c01231>
- Nikitovic, D., & Pérez, S. (2021). Preface for the Special Issue on the Exploration of the Multifaceted Roles of Glycosaminoglycans: GAGs. *Biomolecules*, 11(11), 1630. <https://doi.org/10.3390/biom11111630>
- Nowinski, A. K., Sun, F., White, A. D., Keefe, A. J., & Jiang, S. (2012). Sequence, Structure, and Function of Peptide Self-Assembled Monolayers. *Journal of the American Chemical Society*, 134(13), 6000–6005. <https://doi.org/10.1021/ja3006868>
- Nzakizwanayo, J., Kumar, S., Ogilvie, L. A., Patel, B. A., Dedi, C., Macfarlane, W. M., & Jones, B. V. (2015). Disruption of Escherichia coli Nissle 1917 K5 Capsule Biosynthesis, through Loss of Distinct kfi genes, Modulates Interaction with Intestinal Epithelial Cells and Impact on Cell Health. *PLOS ONE*, 10(3), e0120430. <https://doi.org/10.1371/journal.pone.0120430>
- O’Leary, T. R., Xu, Y., & Liu, J. (2012). Investigation of the substrate specificity of K5 lyase A from K5A bacteriophage. *Glycobiology*, 23(1), 132–141. <https://doi.org/10.1093/glycob/cws136>
- Oguma, T., Toyoda, H., Toida, T., & Imanari, T. (2001). Analytical method of chondroitin/dermatan sulfates using high performance liquid chromatography/turbo ionspray ionization mass spectrometry: application to analyses of the tumor tissue sections on glass slides. *Biomedical Chromatography*, 15(5), 356–362. <https://doi.org/10.1002/bmc.74>
- Okano, H., & Yamanaka, S. (2014). iPS cell technologies: significance and applications to CNS regeneration and disease. *Molecular Brain*, 7(1), 22. <https://doi.org/10.1186/1756-6606-7-22>
- Osago, H., Shibata, T., Hara, N., Kuwata, S., Kono, M., Uchio, Y., & Tsuchiya, M. (2014). Quantitative analysis of glycosaminoglycans, chondroitin/dermatan sulfate, hyaluronic acid, heparan sulfate, and keratan sulfate by liquid chromatography–electrospray ionization–tandem mass spectrometry. *Analytical Biochemistry*, 467,

62–74. <https://doi.org/10.1016/j.ab.2014.08.005>

- Osawa, T., Sugiura, N., Shimada, H., Hirooka, R., Tsuji, A., Shirakawa, T., ... Kakuta, Y. (2009). Crystal structure of chondroitin polymerase from *Escherichia coli* K4. *Biochemical and Biophysical Research Communications*, 378(1), 10–14. <https://doi.org/10.1016/j.bbrc.2008.08.121>
- Otto, N. J., Green, D. E., Masuko, S., Mayer, A., Tanner, M. E., Linhardt, R. J., & DeAngelis, P. L. (2012). Structure/Function Analysis of *Pasteurella multocida* Heparosan Synthases. *Journal of Biological Chemistry*, 287(10), 7203–7212. <https://doi.org/10.1074/jbc.m111.311704>
- Pepi, L. E., & Amster, I. J. (2021). Electron-Activated Tandem Mass Spectrometry Analysis of Glycosaminoglycans. *Current Protocols*, 1(4). <https://doi.org/10.1002/cpz1.83>
- Pepi, L. E., Sanderson, P., Stickney, M., & Amster, I. J. (2021). Developments in Mass Spectrometry for Glycosaminoglycan Analysis: A Review. *Molecular & Cellular Proteomics*, 20, 100025. <https://doi.org/10.1074/mcp.r120.002267>
- Pera, M. F., & Trounson, A. O. (2004). Human embryonic stem cells: prospects for development. *Development*, 131(22), 5515–5525. <https://doi.org/10.1242/dev.01451>
- Pérez-López, N., Martín, C., García, B., Solís-Hernández, M. P., Rodríguez, D., Alcalde, I., ... Quirós, L. M. (2021). Alterations in the Expression of the Genes Responsible for the Synthesis of Heparan Sulfate in Brains With Alzheimer Disease. *Journal of Neuropathology & Experimental Neurology*, 80(5), 446–456. <https://doi.org/10.1093/jnen/nlab028>
- Peters, H., Jürs, M., Jann, B., Jann, K., Timmis, K. N., & Bitter-Suermann, D. (1985). Monoclonal antibodies to enterobacterial common antigen and to *Escherichia coli* lipopolysaccharide outer core: demonstration of an antigenic determinant shared by enterobacterial common antigen and *E. coli* K5 capsular polysaccharide. *Infection and Immunity*, 50(2), 459–466. <https://doi.org/10.1128/iai.50.2.459-466.1985>
- Petit, C., Rigg, G. P., Pazzani, C., Smith, A., Sieberth, V., Stevens, M., ... Roberts, I. S. (1995). Region 2 of the *Escherichia coli* K5 capsule gene cluster encoding proteins for the biosynthesis of the K5 polysaccharide. *Molecular Microbiology*, 17(4), 611–620. https://doi.org/10.1111/j.1365-2958.1995.mmi_17040611.x
- Poyer, S., Seffouh, I., Lopin-Bon, C., Jacquinet, J.-C., Neira, J. L., Salpin, J.-Y., & Daniel, R. (2021). Discrimination of sulfated isomers of chondroitin sulfate disaccharides by HILIC-MS. *Analytical and Bioanalytical Chemistry*, 413(28), 7107–7117. <https://doi.org/10.1007/s00216-021-03679-9>

- Prehm, P. (1984). Hyaluronate is synthesized at plasma membranes. *Biochemical Journal*, 220(2), 597–600. <https://doi.org/10.1042/bj2200597>
- Priem, B., Gilbert, M., Wakarchuk, W. W., Heyraud, A., & Samain, E. (2002). A new fermentation process allows large-scale production of human milk oligosaccharides by metabolically engineered bacteria. *Glycobiology*, 12(4), 235–240. <https://doi.org/10.1093/glycob/12.4.235>
- Priem, B., Peroux, J., Colin-Morel, P., Drouillard, S., & Fort, S. (2017). Chemo-bacterial synthesis of conjugatable glycosaminoglycans. *Carbohydrate Polymers*, 167, 123–128. <https://doi.org/10.1016/j.carbpol.2017.03.026>
- Radhakrishnamurthy, B., Dalperes, E. R., Ruiz, H., & Berenson, G. S. (1977). Determination of aorta glycosaminoglycans by automated ion-exchange chromatography. *Analytical Biochemistry*, 82(2), 445–454. [https://doi.org/10.1016/0003-2697\(77\)90182-8](https://doi.org/10.1016/0003-2697(77)90182-8)
- Rippe, M., Stefanello, T. F., Kaplum, V., Britta, E. A., Garcia, F. P., Poirot, R., ... Auzély-Velty, R. (2019). Heparosan as a potential alternative to hyaluronic acid for the design of biopolymer-based nanovectors for anticancer therapy. *Biomaterials Science*, 7(7), 2850–2860. <https://doi.org/10.1039/c9bm00443b>
- Rittershaus, Emily S. C., Baek, S.-H., & Sasseti, Christopher M. (2013). The Normalcy of Dormancy: Common Themes in Microbial Quiescence. *Cell Host & Microbe*, 13(6), 643–651. <https://doi.org/10.1016/j.chom.2013.05.012>
- Rodgers, K. D., San Antonio, J. D., & Jacenko, O. (2008). Heparan sulfate proteoglycans: A GAGgle of skeletal-hematopoietic regulators. *Developmental Dynamics*, 237(10), 2622–2642. <https://doi.org/10.1002/dvdy.21593>
- Rossdam, C., Konze, S. A., Oberbeck, A., Rapp, E., Gerardy-Schahn, R., von Itzstein, M., & Buettner, F. F. R. (2019). Approach for Profiling of Glycosphingolipid Glycosylation by Multiplexed Capillary Gel Electrophoresis Coupled to Laser-Induced Fluorescence Detection To Identify Cell-Surface Markers of Human Pluripotent Stem Cells and Derived Cardiomyocytes. *Analytical Chemistry*, 91(10), 6413–6418. <https://doi.org/10.1021/acs.analchem.9b01114>
- Samson, J. E., Magadán, A. H., Sabri, M., & Moineau, S. (2013). Revenge of the phages: defeating bacterial defences. *Nature Reviews Microbiology*, 11(10), 675–687. <https://doi.org/10.1038/nrmicro3096>
- Sande, C., & Whitfield, C. (2021). Capsules and Extracellular Polysaccharides in *Escherichia coli* and *Salmonella*. *EcoSal Plus*, 9(2).

<https://doi.org/10.1128/ecosalplus.esp-0033-2020>

- Sarrazin, S., Lamanna, W. C., & Esko, J. D. (2011). Heparan Sulfate Proteoglycans. *Cold Spring Harbor Perspectives in Biology*, 3(7), a004952–a004952.
<https://doi.org/10.1101/cshperspect.a004952>
- Sato, S., Sakamoto, T., Miyazawa, E., & Kikugawa, Y. (2004). One-pot reductive amination of aldehydes and ketones with α -picoline-borane in methanol, in water, and in neat conditions. *Tetrahedron*, 60(36), 7899–7906.
<https://doi.org/10.1016/j.tet.2004.06.045>
- Schaefer, L., & Schaefer, R. M. (2010). Proteoglycans: from structural compounds to signaling molecules. *Cell and Tissue Research*, 339(1), 237–246.
<https://doi.org/10.1007/s00441-009-0821-y>
- Scholl, D., Rogers, S., Adhya, S., & Merrill, C. R. (2001). Bacteriophage K1-5 Encodes Two Different Tail Fiber Proteins, Allowing It To Infect and Replicate on both K1 and K5 Strains of *Escherichia coli*. *Journal of Virology*, 75(6), 2509–2515.
<https://doi.org/10.1128/jvi.75.6.2509-2515.2001>
- Schwarzer, D., Browning, C., Stummeyer, K., Oberbeck, A., Mühlenhoff, M., Gerardy-Schahn, R., & Leiman, P. G. (2015). Structure and biochemical characterization of bacteriophage phi92 endosialidase. *Virology*, 477, 133–143.
<https://doi.org/10.1016/j.virol.2014.11.002>
- Schwarzer, D., Stummeyer, K., Gerardy-Schahn, R., & Mühlenhoff, M. (2007). Characterization of a novel intramolecular chaperone domain conserved in endosialidases and other bacteriophage tail spike and fiber proteins. *The Journal of Biological Chemistry*, 282(5), 2821–2831. <https://doi.org/10.1074/jbc.M609543200>
- Scott, J. E., Quintarelli, G., & Dellovo, M. C. (1964). The chemical and histochemical properties of Alcian Blue. *Histochemie*, 4(2), 73–85.
<https://doi.org/10.1007/bf00306149>
- Shaklee, P. N., & Conrad, H. E. (1984). Hydrazinolysis of heparin and other glycosaminoglycans. *Biochemical Journal*, 217(1), 187–197.
<https://doi.org/10.1042/bj2170187>
- Sieberth, V., Rigg, G. P., Roberts, I. S., & Jann, K. (1995). Expression and characterization of UDPGlc dehydrogenase (KfiD), which is encoded in the type-specific region 2 of the *Escherichia coli* K5 capsule genes. *Journal of Bacteriology*, 177(15), 4562–4565. <https://doi.org/10.1128/jb.177.15.4562-4565.1995>

- Silva, J. C., Han, X., Silva, T. P., Xia, K., Mikael, P. E., Cabral, J. M. S., ... Linhardt, R. J. (2020). Glycosaminoglycan remodeling during chondrogenic differentiation of human bone marrow–/synovial-derived mesenchymal stem/stromal cells under normoxia and hypoxia. *Glycoconjugate Journal*, *37*(3), 345–360. <https://doi.org/10.1007/s10719-020-09911-5>
- Silver, R. P., Prior, K., Nsahlai, C., & Wright, L. F. (2001). ABC transporters and the export of capsular polysaccharides from Gram-negative bacteria. *Research in Microbiology*, *152*(3-4), 357–364. [https://doi.org/10.1016/s0923-2508\(01\)01207-4](https://doi.org/10.1016/s0923-2508(01)01207-4)
- Sismey-Ragatz, A. E., Green, D. E., Otto, N. J., Rejzek, M., Field, R. A., & DeAngelis, P. L. (2007). Chemoenzymatic Synthesis with Distinct Pasteurella Heparosan Synthases. *Journal of Biological Chemistry*, *282*(39), 28321–28327. <https://doi.org/10.1074/jbc.m701599200>
- Smith, Raymond A. A., Meade, K., Pickford, Claire E., Holley, Rebecca J., & Merry, Catherine L. R. (2011). Glycosaminoglycans as regulators of stem cell differentiation. *Biochemical Society Transactions*, *39*(1), 383–387. <https://doi.org/10.1042/bst0390383>
- Sokaribo, A. S., Perera, S. R., Sereggela, Z., Krochak, R., Balezantis, L. R., Xing, X., ... White, A. P. (2021). A GMMA-CPS-Based Vaccine for Non-Typhoidal Salmonella. *Vaccines*, *9*(2), 165. <https://doi.org/10.3390/vaccines9020165>
- Solakyildirim, K. (2019). Recent advances in glycosaminoglycan analysis by various mass spectrometry techniques. *Analytical and Bioanalytical Chemistry*, *411*(17), 3731–3741. <https://doi.org/10.1007/s00216-019-01722-4>
- Soria, B., Roche, E., Berná, G., León-Quinto, T., Reig, J. A., & Martín, F. (2000). Insulin-secreting cells derived from embryonic stem cells normalize glycemia in streptozotocin-induced diabetic mice. *Diabetes*, *49*(2), 157–162. <https://doi.org/10.2337/diabetes.49.2.157>
- Soundararajan, M., von Büнау, R., & Oelschlaeger, T. A. (2019). K5 Capsule and Lipopolysaccharide Are Important in Resistance to T4 Phage Attack in Probiotic E. coli Strain Nissle 1917. *Frontiers in Microbiology*, *10*. <https://doi.org/10.3389/fmicb.2019.02783>
- Stachteá, X. N., Tykesson, E., van Kuppevelt, T. H., Feinstein, R., Malmström, A., Reijmers, R. M., & Maccarana, M. (2015). Dermatan Sulfate-Free Mice Display Embryological Defects and Are Neonatal Lethal Despite Normal Lymphoid and Non-Lymphoid Organogenesis. *PLOS ONE*, *10*(10), e0140279.

<https://doi.org/10.1371/journal.pone.0140279>

- Studelska, D. R., Giljum, K., McDowell, L. M., & Zhang, L. (2005). Quantification of glycosaminoglycans by reversed-phase HPLC separation of fluorescent isoindole derivatives. *Glycobiology*, *16*(1), 65–72. <https://doi.org/10.1093/glycob/cwj037>
- Stummeyer, K., Dickmanns, A., Mühlenhoff, M., Gerardy-Schahn, R., & Ficner, R. (2005). Crystal structure of endosialidase NF – the polysialic acid degrading tailspike of bacteriophage K1F. *GBM Annual Spring Meeting Mosbach 2005*, 2005(Spring). https://doi.org/10.1240/sav_gbm_2005_m_001110
- Sugiura, N., Baba, Y., Kawaguchi, Y., Iwatani, T., Suzuki, K., Kusakabe, T., ... Watanabe, H. (2010). Glucuronyltransferase Activity of KfiC from Escherichia coli Strain K5 Requires Association of KfiA. *Journal of Biological Chemistry*, *285*(3), 1597–1606. <https://doi.org/10.1074/jbc.m109.023002>
- Suzuki, M. (1939). Biochemical studies on carbohydrates. *The Journal of Biochemistry*, *30*(2), 185–191. <https://doi.org/10.1093/oxfordjournals.jbchem.a125896>
- Szekeres, G. P., Pagel, K., & Heiner, Z. (2021). Analytical challenges of glycosaminoglycans at biological interfaces. *Analytical and Bioanalytical Chemistry*, *414*(1), 85–93. <https://doi.org/10.1007/s00216-021-03705-w>
- Tang, C., Lee, A. S., Volkmer, J.-P., Sahoo, D., Nag, D., Mosley, A. R., ... Drukker, M. (2011). An antibody against SSEA-5 glycan on human pluripotent stem cells enables removal of teratoma-forming cells. *Nature Biotechnology*, *29*(9), 829–834. <https://doi.org/10.1038/nbt.1947>
- Taubert, R., Engel, B., Diestelhorst, J., Hupa-Breier, K. L., Behrendt, P., Baerlecken, N. T., ... Thevis, M. (2021). Quantification of polyreactive immunoglobulin G facilitates the diagnosis of autoimmune hepatitis. *Hepatology*, *75*(1), 13–27. <https://doi.org/10.1002/hep.32134>
- Thiesler, C. T., Cajic, S., Hoffmann, D., Thiel, C., van Diepen, L., Hennig, R., ... Buettner, F. F. R. (2016). Glycomic Characterization of Induced Pluripotent Stem Cells Derived from a Patient Suffering from Phosphomannomutase 2 Congenital Disorder of Glycosylation (PMM2-CDG). *Molecular & Cellular Proteomics*, *15*(4), 1435–1452. <https://doi.org/10.1074/mcp.m115.054122>
- Thompson, J. E., Pourhossein, M., Waterhouse, A., Hudson, T., Goldrick, M., Derrick, J. P., & Roberts, I. S. (2010). The K5 Lyase KfiA Combines a Viral Tail Spike Structure with a Bacterial Polysaccharide Lyase Mechanism. *Journal of Biological Chemistry*, *285*(31), 23963–23969. <https://doi.org/10.1074/jbc.m110.127571>

- Tomatsu, S., Shimada, T., Mason, R., Montaña, A., Kelly, J., LaMarr, W., ... Orii, T. (2014). Establishment of Glycosaminoglycan Assays for Mucopolysaccharidoses. *Metabolites*, 4(3), 655–679. <https://doi.org/10.3390/metabo4030655>
- Twort, F. W. (1915). An investigation on the nature of ultramicroscopic viruses. *The Lancet*, 186(4814), 1241–1243. [https://doi.org/10.1016/s0140-6736\(01\)20383-3](https://doi.org/10.1016/s0140-6736(01)20383-3)
- Ucakturk, E., Cai, C., Li, L., Li, G., Zhang, F., & Linhardt, R. J. (2014). Capillary electrophoresis for total glycosaminoglycan analysis. *Analytical and Bioanalytical Chemistry*, 406(19), 4617–4626. <https://doi.org/10.1007/s00216-014-7859-8>
- van der Meer, J.-Y., Kellenbach, E., & van den Bos, L. J. (2017). From Farm to Pharma: An Overview of Industrial Heparin Manufacturing Methods. *Molecules*, 22(6), 1025. <https://doi.org/10.3390/molecules22061025>
- Vianzon, V., Illek, B., & Moe, G. R. (2017). Effect of Vaccine-Elicited Antibodies on Colonization of *Neisseria meningitidis* Serogroup B and C Strains in a Human Bronchial Epithelial Cell Culture Model. *Clinical and Vaccine Immunology*, 24(10). <https://doi.org/10.1128/cvi.00188-17>
- Volpi, N. (2010). High-performance liquid chromatography and on-line mass spectrometry detection for the analysis of chondroitin sulfates/hyaluronan disaccharides derivatized with 2-aminoacridone. *Analytical Biochemistry*, 397(1), 12–23. <https://doi.org/10.1016/j.ab.2009.09.030>
- Volpi, N., Galeotti, F., Yang, B., & Linhardt, R. J. (2014). Analysis of glycosaminoglycan-derived, precolumn, 2-aminoacridone-labeled disaccharides with LC-fluorescence and LC-MS detection. *Nature Protocols*, 9(3), 541–558. <https://doi.org/10.1038/nprot.2014.026>
- Weinstein-Opppenheimer, C. R., Brown, D. I., Coloma, R., Morales, P., Reyna-Jeldes, M., Díaz, M. J., ... Acevedo, C. A. (2017). Design of a hybrid biomaterial for tissue engineering: Biopolymer-scaffold integrated with an autologous hydrogel carrying mesenchymal stem-cells. *Materials Science & Engineering. C, Materials for Biological Applications*, 79, 821–830. <https://doi.org/10.1016/j.msec.2017.05.116>
- Wessels, M. R., Goldberg, J. B., Moses, A. E., & DiCesare, T. J. (1994). Effects on virulence of mutations in a locus essential for hyaluronic acid capsule expression in group A streptococci. *Infection and Immunity*, 62(2), 433–441. <https://doi.org/10.1128/iai.62.2.433-441.1994>
- Whitfield, C. (2006). Biosynthesis and Assembly of Capsular Polysaccharides in *Escherichia coli*. *Annual Review of Biochemistry*, 75(1), 39–68.

<https://doi.org/10.1146/annurev.biochem.75.103004.142545>

- Williams, A., Gedeon, K. S., Vaidyanathan, D., Yu, Y., Collins, C. H., Dordick, J. S., ... Koffas, M. A. G. (2019). Metabolic engineering of *Bacillus megaterium* for heparosan biosynthesis using *Pasteurella multocida* heparosan synthase, PmHS2. *Microbial Cell Factories*, 18(1). <https://doi.org/10.1186/s12934-019-1187-9>
- Willis, L. M., Stupak, J., Richards, M. R., Lowary, T. L., Li, J., & Whitfield, C. (2013). Conserved glycolipid termini in capsular polysaccharides synthesized by ATP-binding cassette transporter-dependent pathways in Gram-negative pathogens. *Proceedings of the National Academy of Sciences*, 110(19), 7868–7873. <https://doi.org/10.1073/pnas.1222317110>
- Wright, A. J., & Andrews, P. W. (2009). Surface marker antigens in the characterization of human embryonic stem cells. *Stem Cell Research*, 3(1), 3–11. <https://doi.org/10.1016/j.scr.2009.04.001>
- Xu, D., & Esko, J. D. (2014). Demystifying Heparan Sulfate–Protein Interactions. *Annual Review of Biochemistry*, 83(1), 129–157. <https://doi.org/10.1146/annurev-biochem-060713-035314>
- Xu, Y., Masuko, S., Takeddin, M., Xu, H., Liu, R., Jing, J., ... Liu, J. (2011). Chemoenzymatic Synthesis of Homogeneous Ultralow Molecular Weight Heparins. *Science*, 334(6055), 498–501. <https://doi.org/10.1126/science.1207478>
- Yakovlieva, L., & Walvoort, M. T. C. (2019). Processivity in Bacterial Glycosyltransferases. *ACS Chemical Biology*, 15(1), 3–16. <https://doi.org/10.1021/acscchembio.9b00619>
- Yan, L., Fu, L., Xia, K., Chen, S., Zhang, F., Dordick, J. S., & Linhardt, R. J. (2020). A Revised Structure for the Glycolipid Terminus of *Escherichia coli* K5 Heparosan Capsular Polysaccharide. *Biomolecules*, 10(11), 1516. <https://doi.org/10.3390/biom10111516>
- Yang, X., Yang, X., Yu, H., Na, L., Ghosh, T., McArthur, J. B., ... Chen, X. (2021). A GH89 human α -N-acetylglucosaminidase (hNAGLU) homologue from gut microbe *Bacteroides thetaiotaomicron* capable of hydrolyzing heparosan oligosaccharides. *AMB Express*, 11(1). <https://doi.org/10.1186/s13568-021-01253-1>
- Yuan, Z., Li, B., Niu, L., Tang, C., McMullen, P., Jain, P., ... Jiang, S. (2020). Zwitterionic Peptide Cloak Mimics Protein Surfaces for Protein Protection. *Angewandte Chemie International Edition*, 59(50), 22378–22381. <https://doi.org/10.1002/anie.202004995>
- Zhang, C., Liu, L., Teng, L., Chen, J., Liu, J., Li, J., ... Chen, J. (2012). Metabolic engineering of *Escherichia coli* BL21 for biosynthesis of heparosan, a bioengineered

heparin precursor. *Metabolic Engineering*, 14(5), 521–527.

<https://doi.org/10.1016/j.ymben.2012.06.005>

Zhang, H., Uchimura, K., & Kadomatsu, K. (2006). Brain Keratan Sulfate and Glial Scar Formation. *Annals of the New York Academy of Sciences*, 1086(1), 81–90.

<https://doi.org/10.1196/annals.1377.014>

Zhang, X., Lin, L., Huang, H., & Linhardt, R. J. (2019). Chemoenzymatic Synthesis of Glycosaminoglycans. *Accounts of Chemical Research*, 53(2), 335–346.

<https://doi.org/10.1021/acs.accounts.9b00420>

Zhou, K., Aertsen, A., & Michiels, C. W. (2014). The role of variable DNA tandem repeats in bacterial adaptation. *FEMS Microbiology Reviews*, 38(1), 119–141.

<https://doi.org/10.1111/1574-6976.12036>

Ziolkowski, A. F., Popp, S. K., Freeman, C., Parish, C. R., & Simeonovic, C. J. (2012).

Heparan sulfate and heparanase play key roles in mouse β cell survival and autoimmune diabetes. *Journal of Clinical Investigation*, 122(1), 132–141.

<https://doi.org/10.1172/jci46177>

Curriculum Vitae

PERSONAL INFO

Name: Małgorzata Sulewska

Date and place of birth: May 18, 1991 in Grajewo, Poland

Nationality: Polish

E-mail: m.sulewska91@wp.pl

EDUCATION

Biochemistry, Doctoral studies Oct 2018 - currently
Université Grenoble Alpes, France and Hannover Medical School, Germany
Supervisors: Prof. Dr. Rita Gerardy-Schahn and Dr. Bernard Priem

Chemistry, Master's degree Oct 2015 – Nov 2017
University of Gdansk, Poland

Chemistry, Bachelor's degree Oct 2011 – Sep 2014
University of Gdansk, Poland

EXPERIENCE

Research assistant Feb – Aug 2018
University of Natural Resources and Life Science in Vienna, Austria
Project: Synthetic bacterial analogs of mammalian oligomannose for eliciting neutralizing antibodies to the high-mannose patch on HIV envelope. Supervisor: Prof. Dr. Paul Kosma

Intern Jun – Sep 2017
University of Stuttgart in Stuttgart, Germany
Department of Biochemistry, Supervisor: Dr. Tomasz Jurkowski

Intern Nov – Feb 2017
Free University of Berlin in Berlin, Germany
Department of Organic Chemistry, Supervisor: Prof. Dr. Beate Kokschi

Intern Jun 2016
ZIAJA Ltd. in Gdansk, Poland
Laboratory of Physico-Chemical Analysis

Intern Sep 2013
Avantor Performance Materials Poland S.A. in Gliwice, Poland
Quality Control Department and Production Department
