

**Netzwerkanalyse neuartiger Glycosyltransferasen aus dem
Phylum Bacteroidota**

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaft (Dr. rer. nat.)

in der Abteilung für Biologie

Unterabteilung an der Fakultät für Mathematik, Informatik und Naturwissenschaften
der Universität Hamburg

vorgelegt von

Tim Peuker

Hamburg 2025

Die folgenden Gutachter empfehlen die Zulassung zur Dissertation

Prof. Dr. Wolfgang Streit

Prof. Dr. Susanne Dobler

Tag der Disputation: 20.06.2025

Eidesstattliche Versicherung:

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben. Sofern im Zuge der Erstellung der vorliegenden Dissertationsschrift generative Künstliche Intelligenz (gKI) basierte elektronische Hilfsmittel verwendet wurden, versichere ich, dass meine eigene Leistung im Vordergrund stand und dass eine vollständige Dokumentation aller verwendeten Hilfsmittel gemäß der Guten wissenschaftlichen Praxis vorliegt. Ich trage die Verantwortung für eventuell durch die gKI generierte fehlerhafte oder verzerrte Inhalte, fehlerhafte Referenzen, Verstöße gegen das Datenschutz- und Urheberrecht oder Plagiate.

Affidavit:

I hereby declare and affirm that this doctoral dissertation is my own work and that I have not used any aids and sources other than those indicated. If electronic resources based on generative artificial intelligence (gAI) were used in the course of writing this dissertation, I confirm that my own work was the main and value-adding contribution and that complete documentation of all resources used is available in accordance with good scientific practice. I am responsible for any erroneous or distorted content, incorrect references, violations of data protection and copyright law or plagiarism that may have been generated by the gAI.

Hamburg



Tim Peuker

Zusammenfassung

Glycosyltransferasen sind die entscheidenden Biokatalysatoren der enzymatischen Glykosylierung. Dieser Prozess ist einer von zahlreichen biologischen Modifizierungsvorgängen an Zellen, genetischem Material oder wie im Falle der Glykosylierung an Molekülen, die überall auf der Welt zu finden sind.

Beim Modifizierungsprozess der Glykosylierung, werden Kohlenhydratstrukturen an Proteine, Lipide oder andere Moleküle übertragen. Die Glykosylierung kann sowohl durch Enzyme katalysiert werden als auch ohne den Einfluss von Enzymen erfolgen. Dies wird dann als Glykation bezeichnet. Die enzymatische Glykosylierung wird durch Enzyme, die zu der Klasse der Transferasen zählen, genauer den Glycosyltransferasen katalysiert. Diese Enzyme übertragen aktivierte Zuckermoleküle, sogenannte Nukleotidzucker, auf verschiedenste Akzeptormoleküle. Der Modifikationsprozess verändert die Struktur und Funktion des jeweiligen Moleküls, wodurch sich Eigenschaften wie Stabilität oder Bioverfügbarkeit ändern können. Beispielhaft können hier Polyphenole genannt werden, die mit einer Vielzahl von vorteilhaften Eigenschaften, wie entzündungshemmend oder krebsvorbeugend in Verbindung gebracht werden, aber auf Grund ihrer schlechten Wasserlöslichkeit eine geringe Bioverfügbarkeit aufweisen. Diese Verfügbarkeit kann durch den Modifizierungsprozess der Glykosylierung erhöht werden.

Diese Arbeit befasste sich gezielt mit der Identifizierung und Charakterisierung neuartiger Glycosyltransferasen. Besonderer Fokus lag dabei auf der Erschließung neuer, dem Phylum Bacteroidota zuzuordnenden Glycosyltransferasen. Als Ausgangspunkt diente ein auf Grundlage von Sequenzen bereits beschriebener Glycosyltransferasen moduliertes Hidden Markov Model (Tabelle 25). Mit diesem Model wurden Datenbanken durchsucht, woraus mehr als 100 potenzielle Treffer für Glycosyltransferasen codierenden Genen resultierten. Eine Auswahl an 15 vielversprechenden Datenbankeinträgen, basierend auf Phylums-Zugehörigkeit, Sequenzlänge und/oder Verwandtschaft zu bereits bekannten GTs, wurden für genauere Untersuchungen ausgewählt. Von diesen konnte schließlich bei 6 Glycosyltransferasen (GT-HHo1 – GT-HHo6) glykosidische Aktivität nachgewiesen

werden. Anschließend wurde auch das Substratspektrum, sowie die Expressionsfähigkeit für zukünftige *in vitro* Assays untersucht.

Alle neu identifizierten Transferasen waren in der Lage die Flavonoide Quercetin und Naringenin zu glykosylieren. Besonders auffällig ist, dass die Glycosyltransferasen GT-HH01, GT-HH02 und GT-HH03 auch das Stilbenoid Oxyresveratrol glykosylieren können. Neben den experimentellen Untersuchungen, erfolgten auch eine *in-silico* Analyse der neuen Enzyme. Dabei wurde unter anderem der Faltungstyp der neuen Glycosyltransferasen moduliert, als auch erstmalig ein Interaktionsnetzwerk aller charakterisierten GTs erstellt. Alle sechs neuen Glycosyltransferasen weisen den Faltungstyp B auf, welcher aus zwei beta-alpha-beta Faltblättern besteht. Des Weiteren geht aus dem Interaktionsnetzwerk hervor, dass sich fast alle beschriebenen GTs in drei große Cluster unterteilen lassen. Die meisten der bekannten Glycosyltransferasen stammen aus den Phyla Pseudomonadota und Actinobacteria. Daneben stellte sich heraus, dass sich durch die Fokussierung auf das Phylum Bacteroidota in dieser Arbeit, die Anzahl der aus diesem Phylum bekannten Glycosyltransferasen nahezu verdoppelt werden konnte.

Diese Ergebnisse zeigen das enorme Potenzial des Phylums Bacteroidota in Bezug auf die Identifizierung und Erschließung neuartiger Glycosyltransferasen, sowie der daraus resultierenden potenziellen zukünftigen Anwendung in der Biotechnologie.

Abstract

Glycosyltransferases are the key biocatalysts of enzymatic glycosylation. This process is one of numerous biological modification processes on cells, genetic material or, as in the case of glycosylation, molecules that can be found all over the world.

At the glycosylation modification process, carbohydrate structures are transferred to proteins, lipids or other molecules. Glycosylation can be catalyzed by enzymes as well as without the influence of enzymes. This is then referred to as glycation. Enzymatic glycosylation is catalyzed by enzymes belonging to the class of transferases, more precisely glycosyltransferases. These enzymes transfer activated sugar molecules, so-called nucleotide sugars, to various acceptor molecules. The modification process changes the structure and function of the respective molecule, which can alter properties such as stability or bioavailability. Examples include polyphenols, which are associated with a variety of beneficial properties such as anti-inflammatory or cancer-preventing but have low bioavailability due to their poor water solubility. This availability can be increased by the modification process of glycosylation.

This work was specifically concerned with the identification and characterization of novel glycosyltransferases. Special focus was placed on the identification of new glycosyltransferases belonging to the phylum Bacteroidota. A Hidden Markov Model (Table 25) was modulated on the basis of sequences of already described glycosyltransferases served as a starting point. This model was used to search in various databases and results in more than 100 potential hits for genes encoding glycosyltransferases. A selection of 15 promising database entries, based on phylum affiliation, sequence length and/or relationship to already known GTs, were selected for further investigation. Of these, 6 glycosyltransferases (GT-HH01 - GT-HH06) were found to have glycosidic activity. Subsequently, the substrate spectrum as well as the expression capacity for future *in vitro* assays were investigated.

All newly identified transferases were able to glycosylate the flavonoids quercetin and naringenin. It is particularly striking that the glycosyltransferases GT-HH01, GT-HH02 and GT-HH03 can also glycosylate the stilbenoid oxyresveratrol. In addition to the experimental studies, the new enzymes were also analyzed *in silico*. Among other

things, the folding type of the new glycosyltransferases was modulated, and an interaction network of all characterized GTs was created for the first time. All six new glycosyltransferases show the folding type B, which consists of two beta-alpha-beta leaflets. Furthermore, the interaction network shows that nearly all described GTs can be divided into three large clusters. Most of the known glycosyltransferases originate from the phyla *Pseudomonadota* and *Actinobacteria*. In addition, it turned out that by focusing on the phylum Bacteroidota in this work, the number of glycosyltransferases known from this phylum could be almost doubled.

These results show the enormous potential of the phylum Bacteroidota with regard to the identification and development of novel glycosyltransferases, as well as the resulting potential future application in biotechnology.

Inhaltsverzeichnis

ABBILDUNGSVERZEICHNIS.....	VIII
TABELLENVERZEICHNIS	XI
ABKÜRZUNGSVERZEICHNIS	XIII
1. EINLEITUNG	1
1.1 GLYCOSYLTRANSFERASEN (GTs).....	1
1.2 MODELSUBSTRAT QUERCETIN	5
1.3 <i>BACTEROIDOTA</i>	8
1.4 ZIELSETZUNG.....	10
2. MATERIALIEN UND METHODEN	12
2.1 CHEMIKALIEN UND AGLYCONE	12
2.2 BAKTERIENSTÄMME, VEKTOREN, PRIMER.....	13
2.2.1 Bakterienstämme	13
2.2.2 Vektoren und Primer.....	14
2.3 MEDIEN UND MEDIENZUSÄTZE.....	15
2.3.1 Medien	15
2.3.2 Medienzusätze	17
2.4 KULTIVIERUNG UND LAGERUNG VON BAKTERIENSTÄMMEN	17
2.4.1 Anzucht von Bakterien	17
2.4.2 Lagerung von Bakterien	18
2.5 MOLEKULARBIOLOGISCHE METHODEN.....	18
2.5.1 Zelldichten Bestimmung	18
2.5.2 Agarose-Gelelektrophorese.....	19
2.5.3 Bestimmung der Nukleinsäure-Konzentration	20
2.5.4 Isolierung von genomischer DNA.....	20

2.5.5	Präparation von Plasmid-DNA	20
2.5.6	Polymerase-Kettenreaktion (PCR)	21
2.5.6.1	Primer	22
2.5.6.2	PCR-Ansatz	22
2.5.6.3	Kolonie-PCR	23
2.5.6.4	Restriktionsfreie-Klonierung (rf-Klonierung)	23
2.5.7	Sequenzierung von DNA-Fragmenten	24
2.5.8	Transfer von Plasmid-DNA	24
2.5.8.1	Hitzeschocktransformation	24
2.6	PROTEINBIOCHEMISCHE METHODEN	25
2.6.1	Überexpression von Proteinen	25
2.6.2	Natriumdodecylsulfat-Polyacrylamidgelelektrophorese (SDS-Page)	25
2.6.3	Western Blot	28
2.7	BIOKONVERSION	30
2.7.1	Biokonversion im Schüttelkolben	Fehler! Textmarke nicht definiert.
2.8	CHROMATOGRAPHISCHE METHODEN	30
2.8.1	Probenvorbereitung für Dünnschichtchromatographie (DC)	31
2.8.2	Probenvorbereitung für Hochleistungsflüssigkeitschromatographie (HPLC)	31
2.8.3	Dünnschichtchromatographie (DC)	31
2.8.4	Hochleistungsflüssigkeitschromatographie (HPLC)	32
2.9	<i>IN SILICO</i> METHODEN	33
2.9.1	Sequenzbasierte Suche mittels Hidden Markov Model	34
3.	ERGEBNISSE	35
3.1	SEQUENZBASIERTE DATENBANK SUCHE	35
3.1.1	Analyse des Hidden Markov Models	36

3.2 <i>IN SILICO</i> ANALYSE DER GLYCOSYLTRANSFERASEN.....	37
3.2.1 Faltungstypen.....	37
3.3 GLYCOSYLTRANSFERASEN CHARAKTERISIERUNG	39
3.3.1 Aktivitätsassay.....	39
3.3.2 Substratspektrum Bestimmung	41
3.3.3 Nachweis der Expression.....	44
3.3.4 Zusammenfassung über die herausgefundenen Eigenschaften der neu identifizierten.....	46
Glycosyltransferasen.....	46
3.4 NETZWERKANALYSE BESCHRIEBENER GLYCOSYLTRANSFERASEN	50
4. DISKUSSION	53
4.1 SEQUENZBASIERTES SCREENING NACH NEUEN BACTERIODOTALEN GLYCOSYLTRANSFERASEN	54
4.2 <i>IN SILICO</i> ANALYSEN	55
4.2.1 Strukturanalyse neuartiger Glycosyltransferasen.....	55
4.3 CHARAKTERISIERUNG VON GLYCOSYLTRANSFERASEN IDENTIFIZIERT DURCH SEQUENZBASIERTER DATENBANKSUCHE	58
4.3.1 Aktivitätsbestimmung potenziell neuartiger Glycosyltransferasen.....	59
4.3.2 Substratspektrumsanalyse von Glycosyltransferasen	60
4.4 SCHLUSSFOLGERUNGEN UND AUSBLICK	62
LITERATURVERZEICHNIS.....	65
ANHANG.....	70
DANKSAGUNG	73

Abbildungsverzeichnis

ABBILDUNG 1: SCHEMATISCHE DARSTELLUNG DER ENZYMATISCHEN GLYKOSYLIERUNG BEI DER GLYCOSYLTRANSFERASEN ALS SCHLÜSSELENZYME AGIEREN.	1
ABBILDUNG 2: STRUKTUREN DER BEKANNTEN FALTUNGSTYPEN VON GLYCOSYLTRANSFERASEN. DIE FALTUNGEN GT-A, GT-B UND GT-D KONNTEN BEREITS BEI LELOIR-GLYCOSYLTRANSFERASEN NACHGEWIESEN WERDEN. HIERVON SIND DIE FALTUNGEN GT-A UND GT-B DIE BEIDEN AM BESTEN UNTERSUCHTEN FALTUNGSTYPEN. BEIDE BESTEHEN AUS ZWEI $\beta/\alpha/\beta$ ROSSMANN DOMÄNEN, DIE EINMAL ENG (GT-A FALTUNG) UND EINMAL LOCKER (GT-B FALTUNG) ASSOZIIERT SIND. DIE GT-E UND DIE MEMBRANSTÄNDIGE GT-C FALTUNG SIND AUSSCHLIEßLICH BEI NICHT-LELOIR GLYCOSYLTRANSFERASEN ZU FINDEN. DIE ABBILDUNG WURDE MIT USCF CHIMERA ERSTELLT [9].	3
ABBILDUNG 3: REAKTIONSMECHANISMUS VON INVERTIERENDEN (A) UND BEIBEHALTENDEN (B UND C) GLYCOSYLTRANSFERASEN. DER REAKTIONSWEG ÜBER DOPPELTE S_N2 -ÄHNLICHE NUKLEOPHILE SUBSTITUTION (B) KONNTE BISHER NICHT NACHGEWIESEN WERDEN. DAHER GILT DER S_N1 REAKTIONSWEG ALS WAHRSCHEINLICHERER REAKTIONSMECHANISMUS VON BEIBEHALTENDEN-GLYCOSYLTRANSFERASEN. DAS M^{2+} STEHT FÜR EIN ZWEIWERDIGES METALL WELCHES KEINE ZWANGSLÄUFIGE VORAUSSETZUNG FÜR DIE KATALYTISCHE AKTIVITÄT IST. DIESE ABBILDUNG WURDE NACH MESTROM <i>ET AL.</i> MODIFIZIERT. [8].	4
ABBILDUNG 4: REAKTIONSMECHANISMUS DER GLYKOSYLIERUNG VON QUERCETIN ZU QUERCITRIN, BEI DER AUCH WASSER ALS NEBENPRODUKT ENTSTEHT. MIT BLAU IST DIE NUMMERIERUNG DER BENZOLRINGE UND DIE DER EINZELNEN ATOME DARGESTELLT, NACH DER SICH AUCH DIE JEWEILIGE BEZEICHNUNG DES ENTSTEHENDEN GLYKOSIDS RICHTET. IN DIESEM HIER DARGESTELLTEN BEISPIEL QUERCETIN 3-RHAMNOSE. ROT MARKIERT IST HYDROXYGRUPPE AN POSITION 3 AN DER DIE GLYKOSYLIERUNG IN DIESER BEISPIELREAKTION ERFOLGT. DIE ABBILDUNG WURDE MIT CHEMDRAW 21.0.0 ERSTELLT.	6
ABBILDUNG 5: DAS IN DIESER ARBEIT IM RAHMEN DES AKTIVITÄTSASSAYS BENUTZTE MODELSUBSTRAT QUERCETIN MIT DEN MÖGLICHEN GLYKOSIDEN QUERCITRIN, ISOQUERCITRIN UND RUTIN. DIE GLYKOSIDE WERDEN ANHAND DER DURCH DIE GLYCOSYLTRANSFERASEN ÜBERTRAGENDEM ZUCKERMOLEKÜL UNTERSCHIEDEN. IN ROT IST DAS HYDROXYMOLEKÜL BZW. SAUERSTOFFMOLEKÜL MARKIERT, ÜBER DAS DIE GLYKOSYLIERUNG ERFOLGT. DAS DABEI ENTSTEHENDE WASSERMOLEKÜL IST AUFGRUND DER BESSEREN ÜBERSICHTLICHKEIT NICHT DARGESTELLT. DIE ABBILDUNG WURDE MIT CHEMDRAW 21.0.0 ERSTELLT.	7
ABBILDUNG 6: ÜBERSICHT ÜBER DIE KLASSEN, ORDNUNGEN UND FAMILIEN DIE DEM PHYLUM BACTEROIDOTA ZUGEORDNET WERDEN. MODIFIZIERT NACH HAHNKE ET AL. 2016 [22]	9
ABBILDUNG 7: ARBEITSSCHEMA DER VORLIEGENDEN ARBEIT. AUSGANGSPUNKT BILDET EINE SEQUENZBASIERTE SUCHE MITTELS HIDDEN MARKOV MODEL UND ANSCHLIEßENDER KLONIERUNG ODER SYNTHESE VON POTENZIELLEN INTERESSANTEN GLYCOSYLTRANSFERASEN. NACHFOLGEND WURDEN DIE POTENZIELLEN	

GT'S MITTELS AKTIVITÄTSASSAY UNTERSUCHT. NACHGEWIESEN AKTIVE GT'S WURDEN CHARAKTERISIERT UND ANALYSIERT. DIE GRAFIK WURDE MIT BIORENDER ERSTELLT.	11
ABBILDUNG 8: VISUALISIERTES HIDDEN MARKOV MODEL. IN DIESEM GESAMT HMM WURDEN DIE KONSERVIERTEN MOTIVE AUS ABBILDUNG 9 IDENTIFIZIERT.....	36
ABBILDUNG 9: IDENTIFIZIERTE KONSERVIERTE MOTIVE DES HIDDEN MARKOV MODEL. MARKIERT SIND DIE AMINOSÄUREN DIE SPEZIELL FÜR DIE SUBSTRAT- UND DONORBINDUNG VERANTWORTLICH SIND.	37
ABBILDUNG 10: MITTELS COLABFOLD VORHERGESAGTE PROTEINSTRUKTUR VON GT-HH05. DAS ENZYM WEIST DIE GT-B FALTUNG AUF. GT-HH05 IST BEISPIELHAFT FÜR DIE ANDEREN 5 ENZYME (GT_HH01, GT-HH02, GT-HH03, GT-HH04 UND GT-HH06) GEWÄHLT DIE EBENFALLS DIE GT-B FALTUNG AUFWEISEN.....	38
ABBILDUNG 11: STRUKTURÜBERLAGERUNG DER SECHS NEU IDENTIFIZIERTEN GLYCOSYLTRANSFERASEN. HERVORGEHOBEN SIND DIE AMINOSÄUREN, DIE DAS AKTIVE ZENTRUM BILDEN UND IN DER ANALYSE DES HIDDEN MARKOV MODEL IDENTIFIZIERT WERDEN KONNTEN. OCKER: GT-HH01, BLAU: GT-HH02, PINK: GT-HH03, GRÜN: GT-HH04, ROT: GT-HH05, GRAU: GT-HH06	39
ABBILDUNG 12: DÜNNSCHICHTPLATTE NACH DER AUFTRENNUNG DER PROBEN SOWIE FÄRBen UNTER UV-LICHT MIT DER WELLENLÄNGE 356 NM. AUFGETRENNT SIND IN DEN SPALTEN 1 BIS 4 DIE VERWENDETEN REFERENZEN: 1: QUERCETIN, 2: QUERCITRIN, 3: ISOQUERCITRIN, 4: RUTIN. IN DEN NACHFOLGENDEN SPALTEN 5 BIS 10 SIND DIE PROBEN DER KONSTRUKTE GT-HH01 BIS GT-HH06 AUFGETRAGEN. MIT ROTEN RAHMEN SIND DIE BANDEN IN DEN JEWEILIGEN SPALTEN DER PROBEN MARKIERT, BEI DENEN EINE GLYKOSYLIERUNG ERFOLGT IST. DIE BLAUEN BANDEN IN DEN SPALTEN 5 BIS 10 SIND ANGEFÄRBTE SUBSTANZEN AUS DEM JEWEILIGEN MEDIUM UND UNABHÄNGIG VOM NACHWEIS DER ERFOLGREICHEN GLYKOSYLIERUNG. DIE EINDEUTIGKEIT DER BANDEN FÜR GLYKOSYLIERUNGSAKTIVITÄT IST AUFGRUND DER GLEICHEN FARBE SOWIE DER GLEICHEN POSITION WIE DIE JEWEILIGE REFERENZ AUF DER PLATTE GEGEBEN.	41
ABBILDUNG 13: GEFÄRBTE DÜNNSCHICHTPLATTE (UNTER UV-LICHT 366 NM) MIT DEM SUBSTRAT HESPERETIN DAS IM RAHMEN DES SUBSTRATSPEKTREN ASSAYS VERWENDET UND ANALYSIERTE WURDE. AUFGETRAGEN SIND DIE JEWEILIGEN UNTERSUCHTEN KONSTRUKTE (GT-HH01 BIS GT-HH06) DIE BEIDEN EINGESETZTEN KONTROLLEN (<i>E. COLI</i> BL21 DE3, pET21 UND <i>E. COLI</i> BL21 DE3), SOWIE DAS ENTSPRECHENDE AGLYKON. SPALTE 1 BIS 6 ZEIGT DIE UNTERSUCHTEN KONSTRUKTE IN CHRONOLOGISCHER REIHENFOLGE VON GT-HH01 BIS GT-HH06. IN DEN SPALTEN 7 UND 8 SIND DIE BEIDEN EINGESETZTEN NEGATIVKONTROLLEN ZU SEHEN. SPALTE 9 ZEIGT DAS VERWENDETE AGLYKON HESPERETIN UND IN SPALTE 10 DAS GLYKOSIDE HESPERETIN 7-O-GLYCOSID.	42
ABBILDUNG 14: SDS-GEL ZUM NACHWEIS DER EXPRESSION DER GLYCOSYLTRANSFERASEN GT-HH01. MARKIERT (ROTER RAHMEN) SIND DIE BANDEN BEI DER DIE DEUTLICHSTE EXPRESSION ERKENNBAR WAR. UNTERSUCHT WURDE DIE EXPRIMIERBARKEIT BEI UNTERSCHIEDLICHEN MEDIEN (LB-MEDIUM UND TB-MEDIUM) UND TEMPERATUREN (17°C, 22°C, 28°C UND 37°C). DIE PROBEN SIND IN DER REIHENFOLGE:	

MARKER, NICHT-INDUZIERT, 17°C, 22°C, 28°C UND 37°C AUFGETRAGEN SOWIE ERST DIE PROBEN AUS DEM LB-MEDIUM UND ALS ZWEITES AUS TB-MEDIUM.....	44
ABBILDUNG 15: WESTERN-BLOT MEMBRANEN DES KONSTRUKTES GT-HH01. ERKENNBAR SIND DIE BANDEN DER EXPRESSION UNTER DEN BEIDEN VERSCHIEDENEN TEMPERATURBEDINGUNGEN 17 °C (LINKS) UND 22°C (RECHTS). DEUTLICH IST ZUERKENNEN, DASS FÜR GT-HH01 DIE OPTIMALE EXPRESSIONSTEMPERATUR 17°C BETRÄGT.....	46
ABBILDUNG 16: NETZWERKANALYSE ALLER BEKANNTEN UND CHARAKTERISIERTEN GLYCOSYLTRANSFERASEN AUS DER CAZY-DATENBANK. PUNKTE STEHEN FÜR BAKTERIELLE UND DREIECKE FÜR ARCHEEÄLLE ENZYME. DIE WEITERE EINTEILUNG ERFOLGTE NACH DEM ENTSPRECHENDEN PHYLUM. BLAU: PSEUDOMONADOTA, ROT: ACTINOBACTERIA, CYAN: CYANOBACTERIA, PINK: ARCHAEA, GELB: BACTEROIDOTA. A DAS GESAMTE NETZWERK MIT SEINEN DREI DOMINIERENDEN CLUSTERN. B CLUSTER IN DEM ALLE IN DIESER ARBEIT BESCHRIEBENEN GLYCOSYLTRANSFERASEN ZU FINDEN SIND. DIE VISUALISIERUNG DES NETZWERKES ERFOLGTE MITTELS DER CYTOSCAPE SOFTWARE [36].	51
ABBILDUNG 17: NETZWERKANALYSE ALLER BEKANNTEN UND CHARAKTERISIERTEN GLYCOSYLTRANSFERASEN AUS DER CAZY-DATENBANK. DIE EINZELNEN EINTRÄGE DER CLUSTER LASSEN SICH BESTIMMTEN FUNKTIONEN ZUZUORDNEN. DIESE BERUHEN AUF DER AUSWERTUNG DER JEWEILIGEN IM CLUSTER VERSAMMELTEN ENZYME. DIE VISUALISIERUNG DES NETZWERKES ERFOLGTE MITTELS DER CYTOSCAPE SOFTWARE [36].	52
ABBILDUNG 18: DARSTELLUNG DER INTERAKTION VON UGT708C1 MIT UDP-GLUCOSE.....	57
ABBILDUNG 19: DARSTELLUNG VON PHENYLALANIN UND THYROSIN, SOWIE HERVORHEBUNG (ROTER RAHMEN) DES UNTERSCHIEDES ZWISCHEN BEIDEN AMINOSÄUREN ZUR VERDEUTLICHUNG IHRES EFFEKTES AUF DAS AKTIVE ZENTRUM BEI GT-HH06. ERSTELLT MIT CHEMDrWA 21.0.0 [33].	58
ABBILDUNG 20: DARGESTELLT IST DIE GLYCOSYLTRANSFERASE GT-HH01 MIT QUERCETIN, WELCHES AN DAS AKTIVE ZENTRUM BINDET. MIT GRÜN IST DIE BINDUNG DER 3'-HYDROXYGRUPPE DES B-RINGES VON QUERCETIN AN DAS ENZYM HERVORGEHOBEN. ÄHNLICH BINDET AUCH NARINGENIN AN DIE NEUEN GTs. DER UNTERSCHIED DABEI IST, DASS DIE BINDUNG VON DER 4'-HYDROXYGRUPPE ERFOLGT. DIE STRUKTUR DER GLYCOSYLTRANSFERASE, SOWIE DIE BINDUNG VON QUERCETIN AN DAS ENZYM WURDEN MITTELS COLABFOLD [35] BERECHNET UND MIT UCSF CHIMERA 1.16 [9] DARGESTELLT.	61

Tabellenverzeichnis

TABELLE 1: ÜBERSICHT ÜBER DIE IN DIESER ARBEIT VERWENDETEN CHEMIKALIEN UND AGLYCONE.....	12
TABELLE 2: VERWENDETE BAKTERIENSTÄMME	13
TABELLE 3: DIE IN DIESER ARBEIT GENUTZTEN PRIMER.	14
TABELLE 4: DIE IN DIESER ARBEIT VERWENDETEN VEKTOREN	15
TABELLE 5: ZUSAMMENSETZUNG LURIA BERTANI MEDIUM	16
TABELLE 6: ZUSAMMENSETZUNG M9 MINIMAL MEDIUM.....	16
TABELLE 7: BESTANDTEILE DER M9 SALZLÖSUNG FÜR DAS M9 MINIMAL MEDIUM	16
TABELLE 8: BESTANTEILE DER SPURENELEMENTLÖSUNG FÜR DAS M9 MINIMAL MEDIUM	17
TABELLE 9: ÜBERSICHT ÜBER DIE VERWENDETEN MEDIENZUSÄTZE.....	17
TABELLE 10: ZUSAMMENSETZUNG VOM LOADING DYE (6X)	19
TABELLE 11: ZUSAMMENSETZUNG TAE-PUFFERS	20
TABELLE 12: ALLGEMEINES PCR-SCHEMA	21
TABELLE 13: SCHEMA FÜR PCR-ANSATZ	22
TABELLE 14: ZUSAMMENSETZUNG SDS-PROBENPUFFER.....	25
TABELLE 15: ZUSAMMENSETZUNG POLYACRYLAMIDGEL.....	26
TABELLE 16: ZUSAMMENSETZUNG DER TRENNGEL STAMMLÖSUNG	26
TABELLE 17: ZUSAMMENSETZUNG DER SAMMELGEL STAMMLÖSUNG.....	26
TABELLE 18: ZUSAMMENSETZUNG COOMASSIE-FÄRBE LÖSUNG	27
TABELLE 19: ZUSAMMENSETZUNG DES ELEKTROPHORESEPUFFERS	27
TABELLE 20: ZUSAMMENSETZUNG DES TRANSFER-PUFFERS	29
TABELLE 21: ZUSAMMENSETZUNG DES TBS[T]-PUFFERS.....	29
TABELLE 22: ZUSAMMENSETZUNG DES DETEKTIONS-PUFFER	29
TABELLE 23: ZUSAMMENSETZUNG DER DETEKTIONS-LÖSUNG	30

TABELLE 24: ÜBERSICHT ÜBER DIE IN DIESER ARBEIT VERWENDETEN BIOINFORMATISCHEN SOFTWARES.....	33
TABELLE 25: PROTEINSEQUENZEN DIE ZUR ERSTELLUNG DES HMM GENUTZT WURDEN.	34
TABELLE 26: AUFLISTUNG DER 6 TREFFER DIE DURCH DEN SEQUENZBASIERTE DATENBANK ABGLEICH ALS VIELVERSPRECHENDSTE EINTRÄGE IDENTIFIZIERT WURDEN UND IM RAHMEN DIESER ARBEIT WEITER ANALYSIERT WURDEN.	35
TABELLE 27: UNTERSUCHTE SUBSTRATE IM RAHMEN DER SUBSTRATSPEKTRUMANALYSE UND AUFLISTUNG DURCH WELCHE GLYCOSYLTRANSFERASEN DIESE GLYKOSYLIERT WERDEN KONNTEN.	43
TABELLE 28: ÜBERSICHT ÜBER DIE ERGEBNISSE DER EXPRESSIONSSTUDIEN DER NEU IDENTIFIZIERTEN GLYCOSYLTRANSFERASEN GT-HHo1 – GT-HHo6	45
TABELLE 29: ZUSAMMENFASSUNG DER ERMITTELTEN EIGENSCHAFTEN DER GLYCOSYLTRANSFERASE GT-HHo1	47
TABELLE 30: ZUSAMMENFASSUNG DER ERMITTELTEN EIGENSCHAFTEN DER GLYCOSYLTRANSFERASE GT- HHo2	47
TABELLE 31: ZUSAMMENFASSUNG DER ERMITTELTEN EIGENSCHAFTEN DER GLYCOSYLTRANSFERASE GT-HHo3	48
TABELLE 32: ZUSAMMENFASSUNG DER ERMITTELTEN EIGENSCHAFTEN DER GLYCOSYLTRANSFERASE GT- HHo4	48
TABELLE 33: ZUSAMMENFASSUNG DER ERMITTELTEN EIGENSCHAFTEN DER GLYCOSYLTRANSFERASE GT- HHo5	49
TABELLE 34: ZUSAMMENFASSUNG DER ERMITTELTEN EIGENSCHAFTEN DER GLYCOSYLTRANSFERASE GT- HHo6	49

Abkürzungsverzeichnis

°C	Grad Celsius
Ad	adjust (auf angegebenes Volumen auffüllen)
Amp	Ampicillin
bp	Basenpaare
bzw.	beziehungsweise
ca.	circa
DC	Dünnschichtchromatographie
Dest./bidest	destilliert
DNA	Desoxyribonukleinsäure
<i>E. coli</i>	<i>Escherichia coli</i>
g	Gramm
GT's	Glycosyltransferasen
H ₂ O bidest	Destilliertes Wasser
HMM	Hidden Markov Model
HPLC	Hochleistungsflüssigkeitschromatographie
IPTG	Isopropyl-β-Dthiogalactopyranosid
kDa	Kilodalton
L	Liter
m	Milli
mL	Milliliter
mM	Millimolar
ng	Nanogramm
nm	Nanometer
OD	optische Dichte
PCR	Polymerase-Kettenreaktion

RF-Klonierung	Restriktionsfreie Klonierung
rpm	Umdrehungen pro Minute
SDS-Page	Natriumdodecylsulfat-Polyacrylamidgelelektrophorese
sec	Sekunde
Ta	Annealingtemperatur
Tris	Tris(hydroxymethyl)aminomethan
V	Volt
μ	Mikro
μL	Mikroliter

1. Einleitung

1.1 Glycosyltransferasen (GTs)

Glycosyltransferasen (nachfolgend auch GT/ GTs genannt) sind die Schlüsselenzyme bei der enzymatischen Glykosylierung. Dabei katalysieren diese Enzyme den Transfer von verschiedensten Donormolekülen auf ein breites Spektrum an Substraten bzw. Aglykonen und werden der Klasse der Transferasen (EC 2.4.x.y) zugeordnet (Liang et al. 2015). In der nachfolgenden Abbildung ist die enzymatische Glykosylierung schematisch dargestellt.

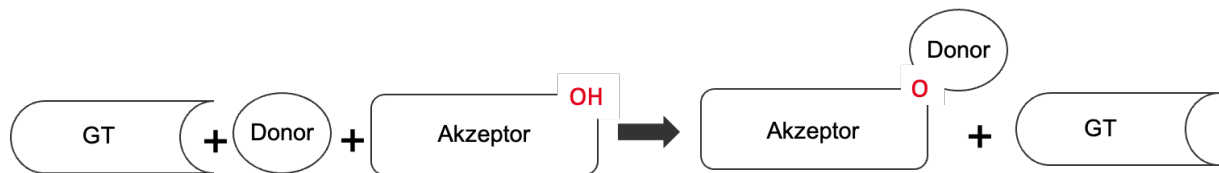


Abbildung 1: Schematische Darstellung der enzymatischen Glykosylierung, bei denen Glycosyltransferasen als Schlüsselenzyme agieren.

Durch diese enzymatische Modifikation können gezielt die biophysikalischen Eigenschaften der jeweiligen Aglykone verändert werden. So kann durch Glykosylierung die Wasserlöslichkeit erhöht werden oder auch die Substratstabilität beeinflusst werden (Chung et al. 2006; Woo et al. 2012; Mäkilä et al. 2016). Ebenfalls wichtig ist die Glykosylierung im Kontext der posttranslationalen Modifikation von Proteinen. So führt beispielsweise der gezielte enzymatische Transfer von Zuckermolekülen auf Proteine zu einem besseren Schutz dieser vor proteolytischen Abbau. Ebenfalls kann diese Modifikation dazu beitragen, dass Proteine korrekt gefaltet werden (Brooks 2006). Bei den dabei transferierten Donormolekülen handelt es sich zumeist um Nukleotidzucker. Diese setzen sich aus einem Monosaccharid und einem Nukleotid zusammen, die über ein Diphosphat miteinander verbunden sind. Beispielhaft zu nennen ist UDP-Glukose (Uridindiphosphat-Glukose) oder dTDP-Rhamnose (Deoxythymidindiphosphat-Rhamnose) (Mikkola 2020). Diese Nukleotidzucker-abhängigen Glycosyltransferasen werden auch Leloir-

Glycosyltransferasen genannt. Benannt nach Luis F. Leloir, der 1970 den Nobelpreis in Chemie für seine Erforschung der Glykosidbiosynthese und des Zuckerstoffwechsels erhalten hat (Lairson et al. 2008). Neben den Nukleotiddiphosphatzuckern gibt es auch Nukleotidmonophosphatzucker. Diese werden von Nicht-Leloir-Glycosyltransferasen genutzt die aufgrund dessen auch als Phosphorylasen bezeichnet werden können (Mestrom et al. 2019). Neben den Unterschieden in den genutzten Donormolekülen, lassen sich Glycosyltransferasen auch nach unterschiedlichen Faltungstypen einteilen. Nach aktuellem Stand der Forschung sind fünf verschiedene Faltungsarten von Glycosyltransferasen bekannt. Für Leloir-Glycosyltransferasen sind die Faltungstypen GT-A, GT-B und GT-D bekannt. Bei den Nicht-Leloir-Glycosyltransferasen sind die Faltungen GT-C und GT-E bekannt (Abbildung 2), wobei es sich bei den Glycosyltransferasen mit GT-C Faltung um Membran assoziierte Glycosyltransferasen handelt (Mestrom et al. 2019). Die am besten beschriebenen Faltungen sind die Faltungstypen GT-A und GT-B. Beide bestehen aus zwei $\beta/\alpha/\beta$ Rossmann Domänen, wobei an die N-terminale Domäne das Akzeptormolekül und an die C-terminale Domäne der Nukleotidzucker bindet (Lairson et al. 2008; Liang et al. 2015). Der Unterschied zwischen beiden Faltungen besteht in der Orientierung und Assoziation der beiden einzelnen Domänen zueinander. Die zwei $\beta/\alpha/\beta$ Rossmann Domänen der GT-A Faltung sind eng zueinander orientiert, wozu im Gegensatz die einzelnen Domänen der GT-B Faltung lockerer aneinandergebunden sind und dabei eine hufeisenartige Form bilden. In dem Freiraum zwischen den einzelnen $\beta/\alpha/\beta$ Rossmann Domänen der GT-B Faltung befindet sich das aktive Zentrum dieser Enzyme.

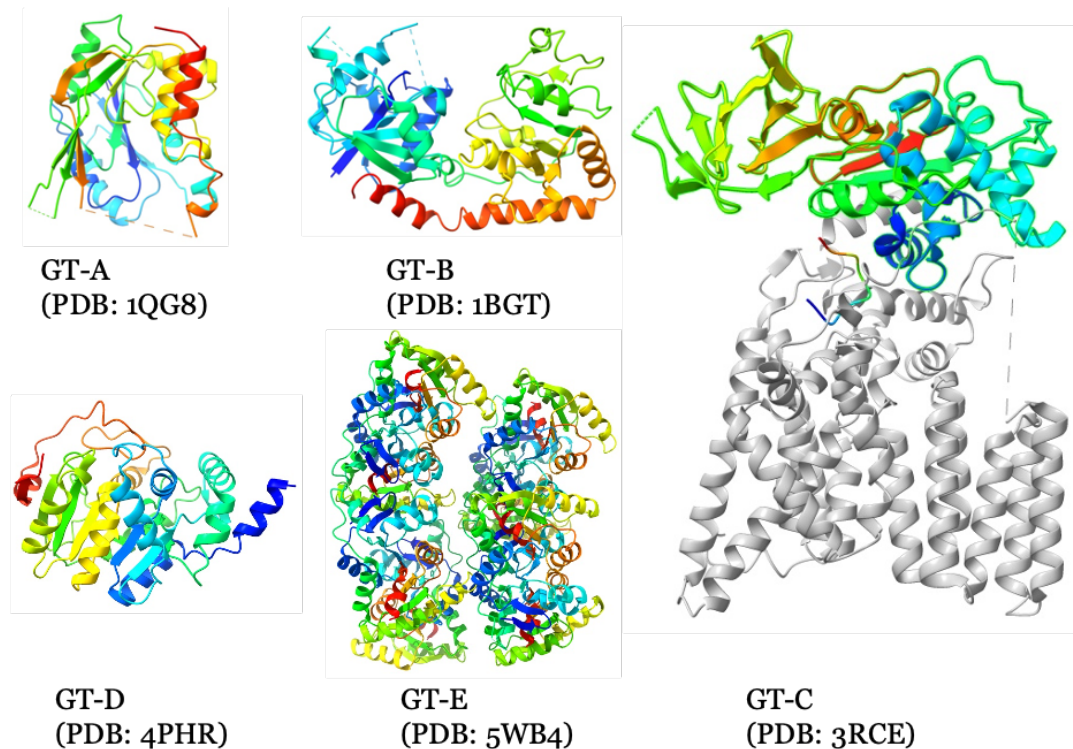


Abbildung 2: Strukturen der bekannten Faltungstypen von Glycosyltransferasen. Die Faltungen GT-A, GT-B und GT-D konnten bereits bei Leloir-Glycosyltransferasen nachgewiesen werden. Hiervon sind die Faltungen GT-A und GT-B die beiden am besten untersuchten Faltungstypen. Beide bestehen aus zwei $\beta/\alpha/\beta$ Rossmann Domänen, die einmal eng (GT-A Faltung) und einmal locker (GT-B Faltung) assoziiert sind. Die GT-E und die membranständige GT-C Faltung sind ausschließlich bei Nicht-Leloir Glycosyltransferasen zu finden. Die Abbildung wurde nach Pettersen et al. 2004 mit UCSF Chimera erstellt.

Neben unterschiedlichen Faltungstypen lassen sich auch verschiedene Reaktionsmechanismen bei der Glykosylierung durch GTs feststellen. Dabei unterscheidet man zwischen invertierenden und beibehaltenden Glycosyltransferasen. Die unterschiedlichen Faltungstypen haben dabei keinen Einfluss auf den jeweiligen Reaktionsmechanismus des Enzyms. Sowohl bei invertierenden als auch beibehaltenden Glycosyltransferasen konnten die Faltungen GT-A und GT-B nachgewiesen werden (Lairson et al. 2008). Beim invertierenden Reaktionsmechanismus handelt es sich um eine S_N2 -ähnliche nukleophile Substitution. Dabei agiert eine Seitenkette des aktiven Zentrums als basischerer Katalysator und deprotoniert das angreifende Nukleophil des Akzeptors. Dieser „Angriff“ des Akzeptors erfolgt von der Rückseite der Abgangsgruppe, da sich sonst die negativ geladene Abgangsgruppe und das negativ geladene angreifende Nukleophil im Weg stünden (Mestrom et al. 2019). Daraus resultiert, dass es beim invertierenden

Reaktionsmechanismus, wie der Name suggeriert, zur Invertierung der Bindung am Monosaccharid kommt (Lairson et al. 2008; Liang et al. 2015). Im Gegensatz zu invertierenden Glycosyltransferasen, ist der Reaktionsmechanismus von beibehaltenden Glycosyltransferasen nicht eindeutig aufgeklärt. So wäre eine doppelte S_N2 -ähnliche nukleophile Substitution denkbar. Dieser Reaktionsmechanismus wird aber als unwahrscheinlich angesehen, da bisher kein Glykosyl-Enzym Intermediat nachgewiesen werden konnte und einige GTs keine passenden Nukleophile nahe des Donorsubstrates aufweisen. Als realistischer wird ein S_Ni - Mechanismus angesehen. Dabei wird die Abgangsgruppe abgespalten und das Nukleophil greift von der gleichen Seite an (Ali et al. 2023). Der logischere Rückseitenangriff des Nukleophils, bei dem sich Abgangsgruppe und Nukleophil nicht blockieren würden, wird dabei durch das Enzym blockiert, womit nur der Angriff von der gleichen Seite erfolgen kann und die Konformation infolgedessen erhalten bleibt. Unklar ist dabei jedoch, ob es zur Bildung von ein oder zwei Intermediaten kommt. Experimentell konnte dieser Reaktionsmechanismus bereits bei der Glycosyltransferase OtsA nachgewiesen werden (Lee et al. 2011). In der sich anschließenden Abbildung 3 sind die drei potenziellen Reaktionsmechanismen grafisch dargestellt.

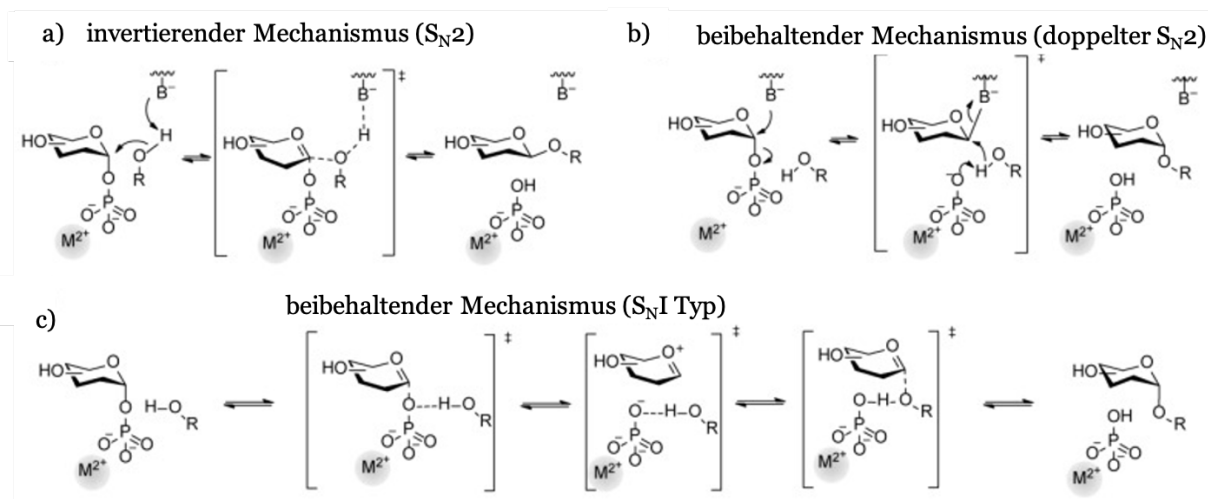


Abbildung 3: Reaktionsmechanismus von invertierenden (a) und beibehaltenden (b und c) Glycosyltransferasen. Der Reaktionsweg über doppelte S_N2 -ähnliche nukleophile Substitution (b) konnte bisher nicht nachgewiesen werden. Daher gilt der S_Ni Reaktionsweg als wahrscheinlicherer Reaktionsmechanismus von beibehaltenden-Glycosyltransferasen. Das M^{2+} steht für ein zweiwertiges Metall welches keine zwangsläufige Voraussetzung für die katalytische Aktivität ist. Diese Abbildung wurde von Mestrom *et al.* übernommen und an das Layout dieser Arbeit angepasst (Mestrom et al. 2019).

Nach aktuellem Stand der Forschung (Mai 2023) sind 1081422 Glycosyltransferasen bekannt (CAZY-Datenbank; [CAZy - GT](#)). Klassifizieren lassen sich all diese nach verschiedensten Gesichtspunkten. Dazu gehören, wie bereits aufgeführt, die Unterteilungen nach den jeweilig präferierten Donor- und/oder Akzeptormolekülen oder dem Reaktionsmechanismus. Durch die Bandbreite an bevorzugten Donor- und Akzeptormolekülen entstehen infolgedessen vielfach unterschiedlichste Glykoside. Demnach können Akzeptormoleküle, wie in Abbildung 4 dargestellt, sowohl über ein Sauerstoffatom, als auch über ein Stickstoff-, Schwefel-, oder Kohlenstoffatom glykosyliert werden. Ein weiteres Klassifizierungssystem von GTs ist die Unterscheidung basierend auf Sequenzähnlichkeiten. Nach diesem System lassen sich alle 1081422 Glycosyltransferasen in 113 GT-Familien unterteilen. Aufgelistet werden alle in der Carbonhydrate-Active enZymes Database (CAZY-Datenbank; [CAZy - GT](#)).

1.2 Modelsubstrat Quercetin

Bei Quercetin handelt es sich um ein Flavonoid, das zu der Gruppe der Polyphenole gehört. Diese zeichnen sich, wie der Namen suggeriert, durch eine Mehrzahl an phenolischen Ringen aus. Ein Ring besteht dabei aus sechs Kohlenstoffatomen, die aneinander gebunden sind und einen Ring formen. Die Grundstruktur eines Flavonoids, einer Untergruppe der Polyphenole, besteht aus zwei phenolischen Ringen (A und B), die über eine Kohlenstoffbrücke verbunden sind. Zusätzlich beinhaltet das Grundgerüst der Flavonoide noch einen sauerstoffhaltigen heterocyclischen Ring (C) (Braune and Blaut 2016). Beginnend am Sauerstoffatom erfolgt die Nummerierung der Einzelatome in Richtung des nächstliegenden Substituenten. Die Atome des B-Rings werden mit der entsprechenden Zahl und einem Strich nummeriert, damit keine Verwechslung erfolgen kann. Dargestellt ist dies in Abbildung 4. Blau beschriftet sind die einzelnen Ringe und die jeweiligen Nummern der Atome. In der entsprechenden Abbildung wurde das Ausgangssubstrat (Aglykon) Quercetin an der Position 3 glykosyliert. Daraus leitet sich dann der Name des Produktes (Glykosid) ab. Es handelt sich hierbei um Quercetin 3-Rhamnose.

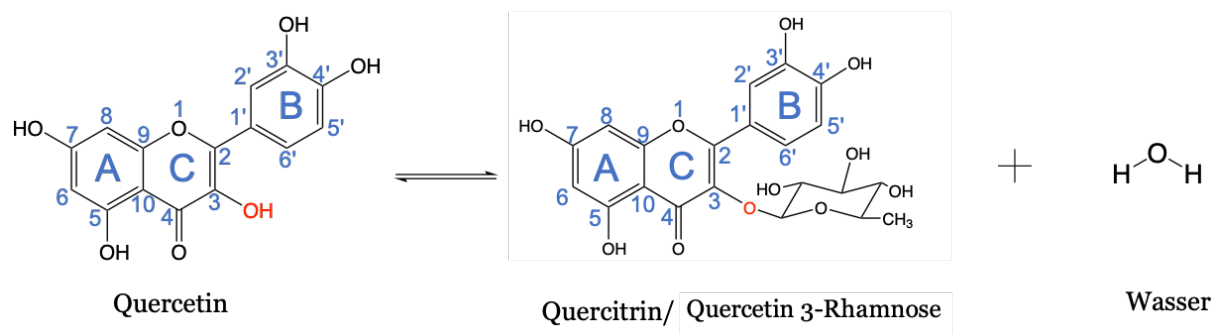


Abbildung 4: Reaktionsmechanismus der Glykosylierung von Quercetin zu Quercitrin, bei der auch Wasser als Nebenprodukt entsteht. Mit blau ist die Nummerierung der Benzolringe und die der einzelnen Atome dargestellt, nach der sich auch die jeweilige Bezeichnung des entstehenden Glykosids richtet. In diesem hier dargestellten Beispiel erfolgt die Glykosylierung über die Hydroxygruppe (rot markiert) an Position 3 aufgrund dessen das Glykosid als Quercetin 3-Rhamnose bezeichnet wird. Die Abbildung wurde mit ChemDraw 21.0.0 erstellt.

Neben dem Quercetin 3-Rhamnose, hier in Abbildung 4 mit dem Trivialnamen Quercitrin bezeichnet, spielten in dieser Arbeit auch die Glykoside Isoquercitrin und Rutin eine wichtige Rolle (Abbildung 5). Bei allen drei aufgezeigten Glykosiden erfolgte die Glykosylierung über die Hydroxygruppe an Position drei, womit es sich um eine O-glykosidische Bindung handelt, bei deren Entstehung ein Molekül Wasser freigesetzt wird. Neben der sogenannten O-Glykosylierung gibt es auch Glykosylierungen über Aminogruppen (N-Glykosylierung), Schwefel (S-Glykosylierung) und Kohlenstoff (C-Glykosylierung). Weiterhin können Glykoside auch anhand ihrer Anomerie eingeteilt werden. Hierbei unterscheidet man zwischen alpha- und beta-glykosidischen Bindungen (Lairson et al. 2008).

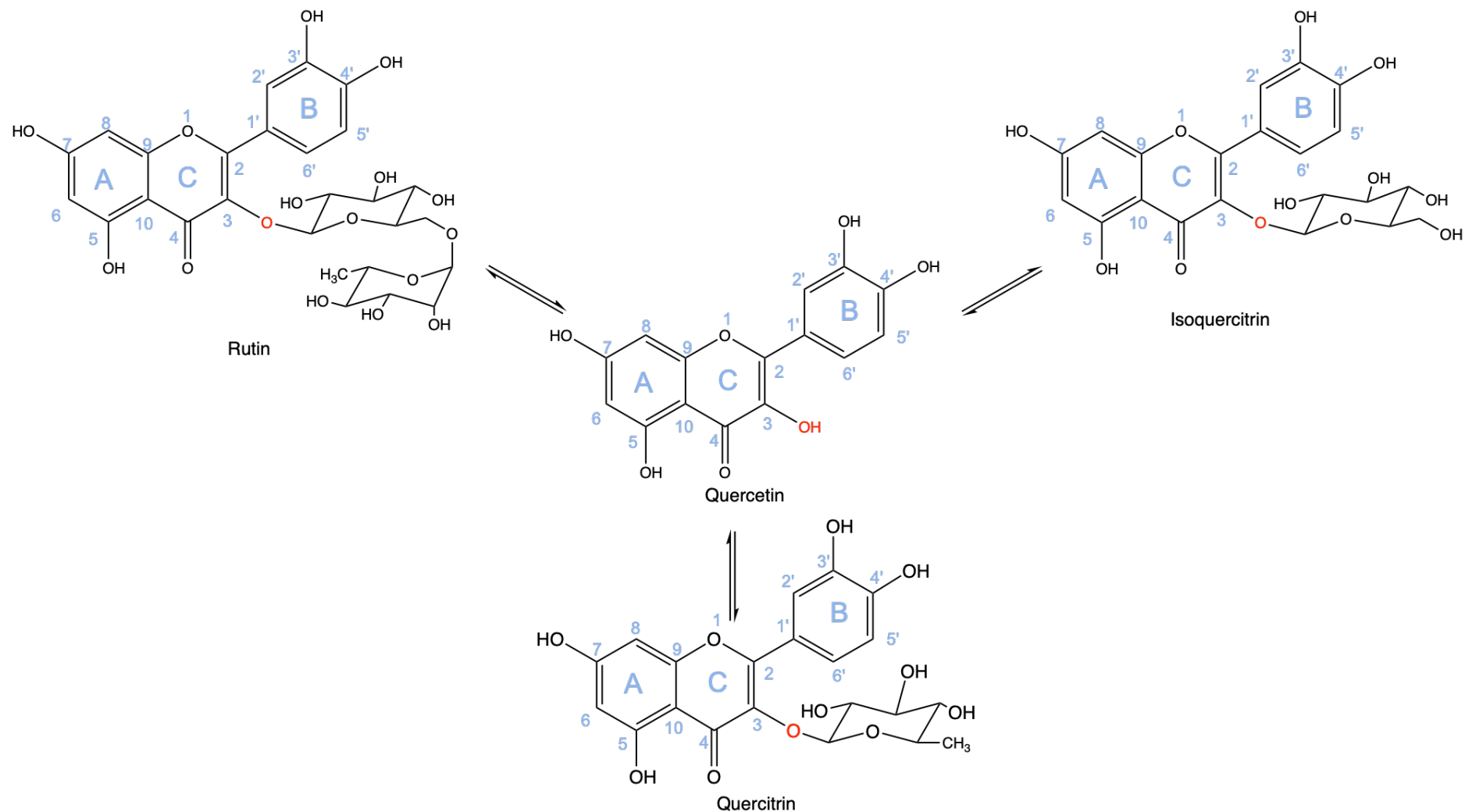


Abbildung 5: Das in dieser Arbeit im Rahmen des Aktivitätsassays benutzte Modellsubstrat Quercetin mit den möglichen Glykosiden Quercitrin, Isoquercitrin und Rutin. Die Glykoside werden anhand der durch die Glycosyltransferasen übertragendem Zuckermolekül unterschieden. In Rot ist das Hydroxymolekül bzw. Sauerstoffmolekül markiert, über das die Glykosylierung erfolgt. Das dabei entstehende Wassermolekül ist aufgrund der besseren Übersichtlichkeit nicht dargestellt. Die Abbildung wurde mit ChemDraw 21.0.0 erstellt.

1.3 Bacteroidota

Der Fokus zur Erschließung neuartiger Glycosyltransferasen aus dem Phylum Bacteroidota begründet sich aus der weltweiten Verbreitung dieses Phylums und dem damit einhergehenden riesigen Potenzial an möglichen neuartigen Enzymen. So beschreiben bereits Sun C-C *et al.* (Sun *et al.* 2023) die Vielzahl an Kohlenhydrat aktiven Enzymen aus Bacteroidota und deren entscheidenden Rolle bei der Zersetzung komplexer Kohlenhydratverbindungen. Diese ermöglichen die Zersetzung unterschiedlichster Substrate wie Cellulose, Hemicellulose, Chitin oder Agar. Die Verwertung dieser unterschiedlichsten Polysaccharide erfolgt unter der Kontrolle verschiedenster Polysaccharid-Verwertungsloki, auch PUL (polysaccharide utilization loci) genannt. Dabei besitzt jeder Bakterienstamm eine Vielzahl an unterschiedlichen PULs zur Verwertung verschiedenster Polysaccharide (Koropatkin *et al.* 2008; Mackenzie *et al.* 2012; Zheng *et al.* 2021). Erstmals wurde dieser aus 8 verschiedenen Genen (*susR*, *susA*, *susB*, *susC*, *susD*, *susE*, *susF* und *susG*) bestehende Lokus bei *Bacteroides thetaiotaomicron* entdeckt und beschrieben (Thomas *et al.* 2011). Als regulatorische Einheit fungiert dabei das *susR* Gen, das die Transkription der anderen Gene unter Anwesenheit von Stärke oder Maltose auslöst. Weiterführende, detaillierte Übersichten und Informationen zum Abbau von Kohlenhydratverbindungen unter der Kontrolle von PULs können bei Thomas *et al.* (Thomas *et al.* 2011) und Feng *et al.* (Feng *et al.* 2022) nachgelesen werden. Parallel dazu zeigen sowohl Veith *et al.* (Veith *et al.* 2023) als auch B. Fernández-Gómez *et al.* (Fernández-Gómez *et al.* 2013), dass Vertreter des Phylums Bacteroidota zahlreiche Gene besitzen, die für Glycosyltransferasen codieren. Unter anderem dienen diese GTs zur Bildung einer riesigen Bandbreite an Glykoproteinen.

Alle zugehörigen Vertreter des Phylum Bacteroidota sind Gram-negativ, bilden keine Sporen, sind unbeweglich, tragen Flagellen oder bewegen sich gleitend fort (Thomas *et al.* 2011). Sie besiedeln den Boden, Meer- und Süßwasser, und auch im Intestinaltrakt von Menschen und Tieren sind sie zu finden (Thomas *et al.* 2011). Ihre jeweiligen physiologischen Eigenschaften variieren daher stark und reichen von strikt anaeroben Vertretern bis hin zu strikt aeroben Vertretern. Unterteilt wird das Phylum

Bacteroidota in 8 Klassen (Bacteroidia, Chitinophagia, Cytophagia, Flavobacteriia, Saprospiria, Sphingobacteria, Bacteroidetes incertae sedis und Bacteroidetes Order IV Incertae sedis), denen wiederum 19 Familien zugeordnet werden (Abbildung 6). Vertreter der Klasse Bacteroidia sind hauptsächlich im Intestinaltrakt von Tieren oder uns Menschen zu finden. Dagegen sind Arten der Klassen Flavobacteriia, Cytophagia und Sphingobacteria vorwiegend in der Umwelt anzutreffen (Thomas et al. 2011). Insgesamt werden derzeit ca. 24.000 Arten dem Phylum Bacteroidota zugerechnet.

Klasse:	Ordnung:	Familie:		
 <i>Cytophagia</i>	 <i>Cytophagales</i>	 <i>Flammeovirgaceae</i>	 <i>Rhodothermaceae</i>	 <i>Salinibacteraceae</i>
 <i>Rhodothermia</i>	 <i>Rhodothermales</i>	 <i>Chitinophagaceae</i>	 <i>Balneolaceae</i>	 <i>Rikenellaceae</i>
 <i>Sphingobacteriia</i>	 <i>Sphingobacteriales</i>	 <i>Marinilabiliaceae</i>	 <i>Cyclobacteriaceae</i>	 <i>Cryomorphaceae</i>
 <i>Flavobacteriia</i>	 <i>Flavobacteriales</i>	 <i>Thermonemataceae</i>	 <i>Prevotellaceae</i>	 <i>Sphingobacteriaceae</i>
 <i>Bacteroidia</i>	 <i>Bacteroidales</i>	 <i>Saprospiraceae</i>	 <i>Prolixibacteraceae</i>	 <i>Hymenobacteriaceae</i>
 <i>Chitinophagia</i>	 <i>Chitinophagales</i>	 <i>Odoribacteraceae</i>	 <i>Porphyromonadaceae</i>	 <i>Flavobacteriaceae</i>
 <i>Balneolia</i>	 <i>Balneolales</i>	 <i>Cytophagaceae</i>	 <i>Crocinitomicaceae</i>	 <i>Bacteroidaceae</i>

Abbildung 6: Übersicht über die Klassen, Ordnungen und Familien, die dem Phylum Bacteroidota zugeordnet werden. Modifiziert nach Hahnke et al. 2016 (Hahnke et al. 2016)

1.4 Zielsetzung

Ziel der vorliegenden Arbeit war die Identifizierung bisher unbekannter Glycosyltransferasen aus dem Phylum Bacteroidota. Dieses bietet auf Grund seiner ubiquitär vorkommenden Vertreter ein enormes Potenzial an neuartigen Enzymen. Bisher werden die Vertreter des Phylums hauptsächlich im Zusammenhang mit dem menschlichen Mikrobiom betrachtet. Dem gegenüber legt jetzt diese Arbeit erstmalig den Fokus auf die gezielte Erschließung neuartiger Enzyme aus Vertretern dieses Phylums.

Basierend auf den Arbeiten von Rabausch *et al.* (Rabausch et al. 2013), Danso *et al.* (Danso et al. 2018) und Ruprecht *et al.* (Ruprecht et al. 2019) wurde ein Arbeitsschema entwickelt, das mittels sequenzbasierten Datenbanksuchen und hochspezifischer Dünnschichtchromatographie dazu beitragen sollte, neuartige Glycosyltransferasen zu identifizieren (Abbildung 7).

Die Basis der Datenbanksuche bildete ein auf Grundlage von Sequenzdaten charakterisierter Glycosyltransferasen erstelltes Hidden Markov Model. Damit identifizierte vielversprechende Treffer potenzieller Glycosyltransferasen wurden kloniert, synthetisiert und auf Aktivität getestet. Dabei diente das Polyphenol Quercetin als Modelsubstrat. Dieses stellte ein optimales Modelsubstrat dar, da Glycosyltransferasen von Mikroorganismen größtenteils dazu verwendet werden, um natürlich vorkommende phenolische und polyphenolische Verbindungen zu modifizieren. Diese haben wiederum eine wichtige Rolle in Stoffwechselprozessen, dienen als Antioxidantien, Stimulatoren des Immunsystems, sind an der Regulierung des Blutdrucks beteiligt und fungieren als Antikrebsmoleküle (Williamson 2017). Neue, als aktiv nachgewiesene Glycosyltransferasen sollten dann weiter charakterisiert werden und mittels *in-silico* Analysen phylogenetisch und interaktionsspezifisch untersucht werden.

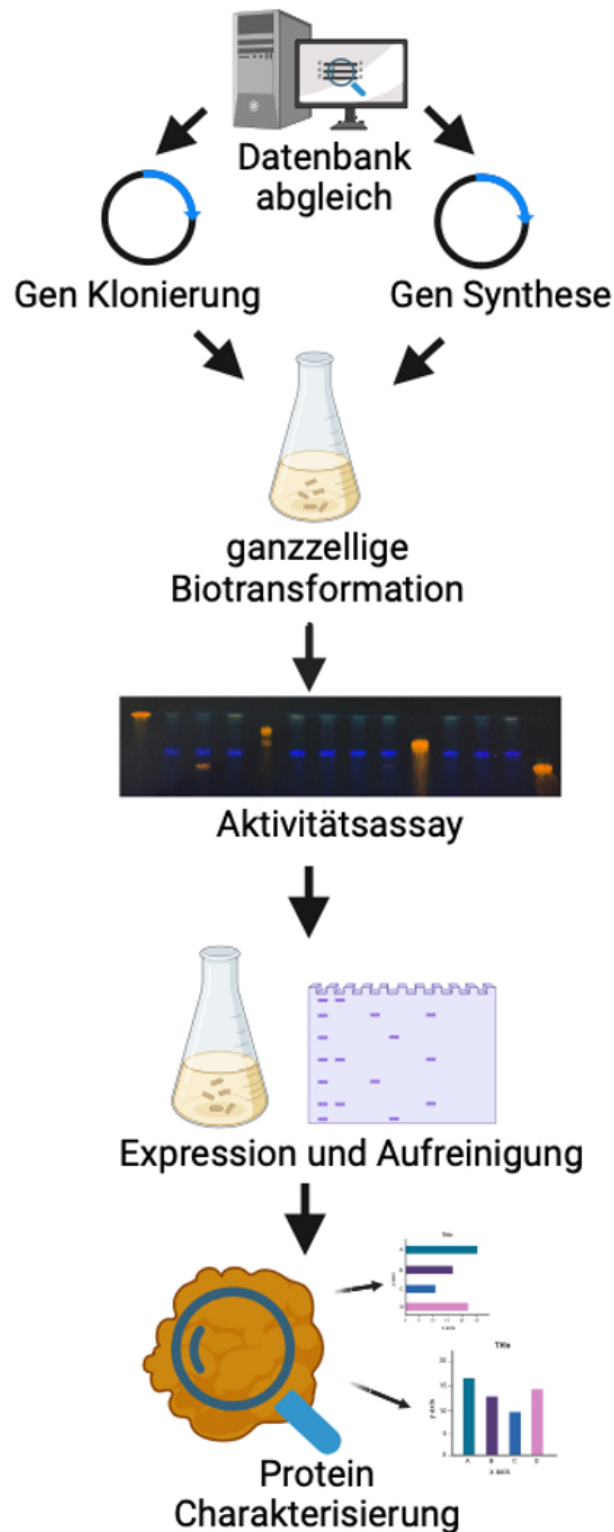


Abbildung 7: Arbeitsschema der vorliegenden Arbeit. Ausgangspunkt bildet eine sequenzbasierte Suche mittels Hidden Markov Model und anschließender Klonierung oder Synthese von potenziellen interessanten Glycosyltransferasen. Nachfolgend wurden die potenziellen GT's mittels Aktivitätsassay untersucht. Nachgewiesenen aktive GT's wurden charakterisiert und analysiert. Die Grafik wurde mit BioRender erstellt.

2. Materialien und Methoden

2.1 Chemikalien und Aglycone

Alle in dieser Arbeit verwendeten Chemikalien und Aglycone erfüllten Laborstandard. Nachfolgend sind diese tabellarisch aufgelistet.

Tabelle 1: Übersicht über die in dieser Arbeit verwendeten Chemikalien und Aglycone

Substanz	Molekulargewicht [g/mol]	Quelle
Quercetin	302.24	Alfa Aesar (Massachusetts, USA)
Quercitrin	448.38	Extrasynthes (Genay, Frankreich)
Isoquercetin	464.10	Carl Roth GmbH (Karlsruhe, Deutschland)
Rutin	610.51	Carl Roth GmbH (Karlsruhe, Deutschland)
Kaempferol	286.26	Molekula (München, Deutschland)
Naringenin	272.25	Carl Roth GmbH (Karlsruhe, Deutschland)
Hesperetin	302.27	Carl Roth GmbH (Karlsruhe, Deutschland)
Qxyresveratrol	244.42	Chengdu Kanghui Biotechnology (Chengdu, China)
Acetylsalicylsäure	180.16	Sigma Aldrich Biochemie GmbH (Hamburg, Deutschland)
Terephthalsäure	166.13	Sigma Aldrich Biochemie GmbH (Hamburg, Deutschland)
Naturstoffreagenz A	225.10	Carl Roth GmbH (Karlsruhe, Deutschland)

2.2 Bakterienstämme, Vektoren, Primer

2.2.1 Bakterienstämme

Nachfolgend sind alle verwendeten Bakterienstämme sowie deren Charakteristika und Quelle aufgeführt.

Tabelle 2: Verwendete Bakterienstämme

Bakterienstamm	Genotyp	Quelle
<i>E. coli</i> BL21	<i>F– ompT gal dcm lon hsdSB(rB–mB–)</i>	Erhalten von Thermo
DE3	<i>λ((DE3) [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS)</i>	Fischer (Eco114)
<i>E. coli</i> DH5α	<i>F– Φ 80lacZΔ M15 Δ (lacZYA -argF) U169 recA1 endA1 hsdR17 (rK– ,mK+) phoA supE44 thi-1 gyrA96 relA1 λ–</i>	DSMZ (Braunschweig, Deutschland)
<i>E. coli</i> T7	<i>huA2 lacZ::T7 gene1 [lon] ompT gal sulA11</i>	
Express lysY/Iq	<i>R(mcr-73::miniTn10--TetS)2 [dcm] R(zgb-210::Tn10--TetS) endA1 Δ(mcrC-mrr)114::IS10 [miniF-lysY, lacIq(CamR)]</i>	Erhalten von New England Biolabs
<i>Arenibacter algicola</i>	Wildtyp	Labor von Prof. Streit

2.2.2 Vektoren und Primer

In den nachfolgenden Tabellen sind die verwendeten Primer und Vektoren aufgelistet. Die für das RF-Klonieren benötigten Target (T) und Plasmid (P) Annealingtemperaturen sind hinter dem jeweiligen Anfangsbuchstaben angegeben.

Tabelle 3: Die in dieser Arbeit genutzten Primer.

Bezeichnung	DNA-Sequenz 5'→3'	Länge [bp]	Ta [°C]	Quelle
T7_Promotor	TAATACGACTCACTATAG	18	44	Aus dieser Studie
pET-rev	TCCGGATATAGTTCC	15	44	Aus dieser Studie
For_pEt21_03495	GTGGTGGTGGTGGTGGT GGTTGCCCAACACAAAA GAATATAG	42	T:54 P:59	Aus dieser Studie
Rev_pET21_03495	CCCTCTAGAAATAATTTT GTTTAACTTTAAGAAGG AGATATACATATGAAAA AATACTTGTTATTGCTGC T	71	T:54 P:59	Aus dieser Studie
For_pEt21_01957	GGTGGTGGTGGTGGTGG TTGTTAAGGAAGTCCAA ACT	37	T:52 P:56	Aus dieser Studie
Rev_pEt21_01957	AGAAATAATTTTGTTTAA CTTTAAGAAGGAGATAT ACATATGAAGTTATATTT AATTCTCCCAATG	66	T:52 P:56	Aus dieser Studie

Tabelle 4: Die in dieser Arbeit verwendeten Vektoren

Vektor	Charakteristika	Größe (bp)	Quelle
pET21a(+)	T7 Promotor/Terminator, His-Tag, Amp ^r , lacI	5443	Erhalten von Addgene (Massachusetts, USA)
pET21a(+)::GT-HH01	GT-HH01 als Insert durch RF-Klonierung	6726	Aus dieser Studie
pEt21a(+)::GT-HH02	GT-HH02 als Insert durch RF-Klonierung	6360	Aus dieser Studie
pET21a(+)::GT-HH03	GT-HH03 als Insert durch RF-Klonierung	6726	Aus dieser Studie
pET21a(+)::GT-HH04	GT-HH04 als Insert durch RF-Klonierung	6684	Aus dieser Studie
pEt21a(+)::GT-HH05	GT-HH05 als Insert durch RF-Klonierung	6684	Aus dieser Studie
pEt21a(+)::GT-HH06	GT-HH06 als Insert durch RF-Klonierung	6663	Aus dieser Studie

2.3 Medien und Medienzusätze

2.3.1 Medien

Die verwendeten Medien wurden vor der Verwendung mittels Dampfdrucksterilisation autoklaviert (121 °C, 20 Minuten, 1 bar Überdruck). Zur Kultivierung auf Festmedium wurden zu einem 1 L Medium 14 g Agar vor dem Autoklavieren hinzugegeben. Zum Gießen der Festmedien wurden 20 mL in sterile Petrischalen gegossen und nach dem Aushärten bei 4 °C gelagert.

Tabelle 5: Zusammensetzung Luria Bertani Medium

LB-Medium (Luria-Bertani)	
Trypton	10.0 g
Hefeextrakt	5.0 g
Natriumchlorid	5.0 g
H ₂ O bidest	ad 1000 mL

Tabelle 6: Zusammensetzung M9 Minimal Medium

M9 Minimal Medium	
M9 Salzlösung (10x)	100 mL
20 % Glukose	20 mL
1M MgSO ₄	1 mL
1M CaCl ₂	0.3 mL
Biotin (1 mg/mL)	1 mL
Thiamin (1 mg/mL)	1 mL
Spurenelementlösung (100x)	10 mL
H ₂ O bidest	867 mL

Tabelle 7: Bestandteile der M9 Salzlösung für das M9 Minimal Medium

M9 Salzlösung (10x)	
Na ₂ HPO ₄ -2H ₂ O	75.2 g/L
KH ₂ PO ₄	30 g/L
NaCl	5 g/L
NH ₄ Cl	5 g/L

Tabelle 8: Bestandteile der Spurenelementlösung für das M9 Minimal Medium

Spurenelementlösung (100x)	
EDTA	13.4 mM
FeCl ₃ -6H ₂ O	3.1 mM
ZnCl ₂	0.62 mM
CuCl ₂ -2H ₂ O	76 µM
CoCl ₂ -2H ₂ O	42 µM
H ₃ BO ₃	162 µM
MnCl ₂ -4H ₂ O	8.1 µM

2.3.2 Medienzusätze

Die thermolabil Medienzusätze wurden mit einer 0.2 µm Membran sterilfiltriert und nach dem Autoklavieren und Abkühlen des Mediums auf 60°C hinzugegeben. Nachfolgend sind die verwendeten Substanzen aufgelistet.

Tabelle 9: Übersicht über die verwendeten Medienzusätze

Substanz	Lösungsmittel	Stammlösung	Endkonzentration Medium
IPTG	H ₂ O _{bidest}	100 mM	100 µg/mL
Ampicillin	H ₂ O _{bidest}	100 mM	100 µg/mL

2.4 Kultivierung und Lagerung von Bakterienstämmen

2.4.1 Anzucht von Bakterien

Die Kultivierung von Bakterien erfolgte in 5 mL Flüssigmedium. Hierfür wurde mittels einer sterilen Pipettenspitze die jeweilige Bakterienkolonie aus einer Glycerinkultur oder Agar-Platte entnommen und ins Medium überführt. Bei notwendiger Selektion wurde der entsprechende antibiotische Zusatz beigelegt. Die Kultivierung erfolgte über Nacht bei entsprechender Temperatur (*E. coli* BL21 DE3, *E. coli* DH5α, *E. coli* T7

Express lysY/Iq Competent bei 37°C und *Arenibacter algicola* bei 28°C) auf einem Schüttelinkubator (GFL 3015, Gesellschaft für Labortechnik, Burgwedel, Germany) bei 150 rpm.

2.4.2 Lagerung von Bakterien

Die Lagerung der Bakterien erfolgte entweder für Kurzzeitlagerung auf Agar-Platten mit entsprechendem Antibiotikum bei 4°C oder für langfristige Lagerung wurden Glycerinkulturen angelegt. Dafür wurden Zellen aus einer Vorkultur mit 86 % Glycerin im Verhältnis 2:3 gemixt und in entsprechenden Kryoröhrchen bei - 70°C gelagert.

2.5 Molekularbiologische Methoden

Um für steriles Arbeiten zu sorgen, wurden die Medien autoklaviert und nicht thermostabile Zusätze sterilfiltriert. Impfösen und Drigalskispatel wurden mit 70 % Ethanol abgeflammt und die Arbeitsflächen vor und nach dem Arbeiten desinfiziert.

Standardisierte molekularbiologische Methoden wurden von Sambrook und Russell (Sambrook and Russell 2001) übernommen.

2.5.1 Photometrische Trübungsmessung

Durch die photometrische Trübungsmessung einer Bakteriensuspension konnte die Zelldichte ermittelt werden. Dafür wurde die optische Dichte (OD) in einer Halb-Mikroküvette (BRAND GMBH + CO KG, Wertheim, Deutschland) bei einer Wellenlänge von $\lambda=600$ nm gemessen. Als Messgerät wurde ein Absorptions-Einstrahlphotometer (BioPhotometer, Eppendorf AG, Hamburg, Deutschland) verwendet. Als Referenz wurde das jeweilige unbewachsene Medium verwendet. Ab einer gemessener OD₆₀₀ größer als 0,8 wurde die Probe verdünnt, um etwaige Messungenauigkeiten zu verhindern.

2.5.2 Agarose-Gelelektrophorese

Die Auftrennung von DNA-Fragmenten zur Überprüfung ihrer Größe erfolgte auf Grund von Polarität in einer Agarose-Gelelektrophosekammer. Verwendet wurden dazu 0.8 % TAE-Agarosegele und Elektrophoresekammern (Biometra Compact XS/S, Analytik Jena, Jena, Deutschland), in denen 1x TAE-Puffer als Laufmittel verwendet wurde.

Die zu untersuchende Probe wurde mit Loading Dye (TriTrack DNA Loading Dye (6x), Thermo Fisher Scientific, Massachusetts, USA) versetzt und in die Taschen des Gels aufgetragen. Die Gelelektrophorese erfolgte bei 100 V für eine Dauer von 35 Minuten. Für die richtige Abschätzung der Größe der untersuchten DNA-Fragmente wurde ein 1 kb+ Marker (GeneRuler 1kb plus DNA Ladder, Thermo Fisher Scientific, Massachusetts, USA) verwendet. Nach der elektrophoretischen Auftrennung wurde das Gel mit in Wasser gelösten Ethidiumbromid (0,5 µg/ml) für 15 min angefärbt und anschließend kurz in einem Wasserbad geschwenkt, um überschüssiges Ethidiumbromid zu entfernen. Zur Analyse und Dokumentation des Agarosegels wurde eine Gel-Dokumentationsanlage (ChemiDoc™ MP Imaging System, Bio-Rad Laboratores, Inc., Hercules, USA) genutzt.

Tabelle 10: Zusammensetzung vom Loading Dye (6x)

Loading Dye
Tris-HCL 10 mM (pH 7.6)
Bromphenolblau 0.03 %
Xylencyanol FF 0.03 %
Orange G 0.015 %
Glycerol 60 %
EDTA 60 mM

Tabelle 11: Zusammensetzung TAE-Puffers

50x TAE-Puffer
Tris-HCl 2 mM
Eisessig
EDTA 100 mM

2.5.3 Bestimmung der Nukleinsäure-Konzentration

Zur Bestimmung der Nukleinsäure-Konzentration von DNA-Proben wurde ein Spektral-Photometer (NaoPhotometer® NP80, Implen, Westlake Village, USA) verwendet. Hierfür wurde die Absorption bei einer Wellenlänge von $\lambda=260$ nm gemessen. Bei einem Extinktionswert von 1 entspricht das bei doppelsträngiger DNA einer Konzentration von 50 $\mu\text{g/mL}$. Für den Aufschluss über die Reinheit der Probe wurde zusätzlich die Absorption bei einer Wellenlänge von $\lambda=280$ nm gemessen und aus beiden Werten der Quotient gebildet. Ab einem Wert von 1.8 war ein guter Reinheitsgrad erreicht.

2.5.4 Isolierung von genomischer DNA

Die Isolierung genomischer DNA (gDNA) erfolgte mit dem NucleoSpin Microbial DNA Kit (Machery Nagel, Düren, Deutschland) unter Anwendung der Angaben des Herstellerprotokolls. Für die im Protokoll aufgeführten Zentrifugationsschritte wurde eine Tischzentrifuge (Mini Spin Plus, Eppendorf SE, Hamburg, Deutschland) verwendet. Für die Elution der gDNA wurden 30 μL H_2O bidest. genutzt.

2.5.5 Präparation von Plasmid-DNA

Zur Isolierung von Plasmid-DNA wurde das Presto™ Mini Plasmid Kit (Geneaid Biotech Ltd, New Taipei City, Taiwan) unter Verwendung des Herstellerprotokolls genutzt. Abweichend vom Protokoll des Herstellers wurde für die Elution der Plasmid-

DNA 30 µL H₂O bidest. verwendete und die Elution ein zweites Mal mit dem ersten Eluat wiederholt. Dadurch konnte eine höhere Ausbeute an Plasmid-DNA erzielt werden. Die Konzentration der isolierten Plasmid-DNA wurde, wie in Kapitel 2.5.3 beschrieben, überprüft.

2.5.6 Polymerase-Kettenreaktion (PCR)

Die Polymerase-Kettenreaktion (Polymerase Chain Reaction) ist eine Methode zur Amplifikation und Vervielfältigung von spezifischen Bereichen eines DNA-Stranges *in vitro* (Mullis et al. 1986). Diese Methode gleicht dem natürlichen Prinzip der Replikation und lässt sich in drei Schritte unterteilen. Zuerst wird der DNA-Doppelstrang in zwei komplementäre Einzelstränge aufgetrennt. Danach erfolgt die Anlagerung (Annealing) der Primer an die Einzelstränge und anschließend die Synthetisierung dieser durch die DNA-Polymerase in 5'→3'-Richtung, was als Elongation bezeichnet wird. Durch die mehrfache Wiederholung dieser einzelnen Schritte lässt sich die DNA bei jedem Zyklus verdoppeln. Ein allgemeines Schema über den Aufbau der PCR, wie sie für diese Arbeit genutzt wurde, ist in der folgenden Tabelle dargestellt.

Tabelle 12: Allgemeines PCR-Schema

Ta steht für die Annealingtemperatur der verwendeten Primer (siehe Tabelle 3 und 2.5.6.1), Y steht für die Zeit, die die Polymerase für die Elongation benötigt und hängt von der Anzahl zu amplifizierenden Basenpaaren ab.

Reaktionsschritt	Temperatur [°C]	Zeit [min:sec]	Zyklen
Initiale Denaturierung	95	5:00	1
Denaturierung	95	1:00	
Annealing	Ta	0:30	34
Elongation	72	Y	
Finale Elongation	72	3:00	1

2.5.6.1 Primer

Die individuelle Schmelztemperatur (T_m) jedes einzelnen Primers, der in dieser Arbeit verwendeten wurde, ist abhängig von der Anzahl der Basenpaare und des Anteils an Guanin und Cytosin. Aus dieser Schmelztemperatur (T_m) lässt sich die jeweilige Annealing Temperatur (T_a) der Primer bestimmen. Diese liegt etwa um 5°C niedriger und lässt sich mit folgender Formel berechnen („Dimethyl-Sulfoxide-Mediated-Primer- T_m -Reduction—A-Method-for_1993_Analytica.pdf“)

$$T_a = T_m - 5^\circ\text{C}$$

Für die in dieser Arbeit verwendeten Primerpaare wurde immer die niedrigere Annealingtemperatur der beiden Primer verwendet.

2.5.6.2 PCR-Ansatz

Die einzelnen Komponenten eines PCR-Ansatzes im Rahmen dieser Arbeit sind nachfolgend aufgeführt. Der Ansatz wurde immer auf Eis zusammen pipettiert, um einen vorzeitigen Start der PCR zu vermeiden. Verwendet wurde im Laufe dieser Arbeit eine DCS-Polymerase mit einer Elongationsgeschwindigkeit von 1000 bp/min und eine Phusion-Polymerase mit einer Elongationsgeschwindigkeit von 1000 bp/30 sec. Bei jedem PCR-Ansatz wurde eine Negativkontrolle mitgeführt.

Tabelle 13: Schema für PCR-Ansatz

	Volumen
DNA-Template (100 ng/ μL)	1 μL
Puffer (mit MgCl_2) 10x	5 μL
dNTP-Mix (5 mM)	1 μL
Primer (forward, 10 μM)	1 μL
Primer (reverse, 10 μM)	1 μL
DNA-Polymerase	0.5 μL
H_2O bidest	ad 25 μL

2.5.6.3 Kolonie-PCR

Die Kolonie-PCR wurde durchgeführt, um zu kontrollieren, dass das gewünschte Insert in den gewünschten Vektor erfolgreich aufgenommen wurde. Hierfür wurde von der entsprechenden Bakterien-Kolonie direkt Zellmaterial als DNA-Template eingesetzt. Dazu wurde dies vorher in 10 µL H₂O bidest. suspendiert und bei 95°C für 10 min aufgeköcht. Dies diente dem Zellaufschluss und nachfolgend wurde davon 1 µL als Template in der Kolonie-PCR eingesetzt. Der PCR-Ansatz entsprach dabei dem in Kapitel 2.5.6.2 aufgeführten Standard PCR-Ansatz.

2.5.6.4 Restriktionsfreie-Klonierung (rf-Klonierung)

Restriktionsfreie-Klonierung ist eine Methode, die als erstes von Van Den & Löwe im Jahr 2006 beschrieben wurde, die man auch als Ligations-unabhängiges Klonieren bezeichnen kann (van den Ent and Löwe 2006). Dafür nutzt man spezifische Primer, die zum einen die spezifische Sequenz zur Amplifikation des gewünschten DNA-Fragmentes besitzen, aber auch gleichzeitig einen Überhang haben der spezifisch für das gewünschte Plasmid (in dieser Arbeit pET21a(+)) ist. Zum Design dieser besonderen Primer wurde die „rf-cloning“ Webseite genutzt (Bond and Naus 2012). Die Methode des Restriktionsfreien-Klonieren beruht auf zwei aufeinanderfolgenden PCRs. Die erste PCR dient der Amplifikation des gewünschten DNA-Fragmentes, an das durch die speziellen Primer zusätzlich Überhänge für das gewünschte Plasmid gehängt werden. Dieses Amplifikat diente in der folgenden zweiten PCR als Primer und wird daher auch Mega-Primer genannt. In der zweiten PCR wird das gewünschte Ziel Plasmid als Target-DNA verwendet und mit Hilfe der Mega-Primer amplifiziert. Dabei wird das zuvor in PCR-Eins amplifizierte Insert aufgrund der Plasmid-Überhänge mit in das Plasmid eingebaut. Nachfolgend wird das Produkt aus der zweiten PCR für 1 h bei 37°C mit DpnI inkubiert. Dies diente dem Verdau von demethyliertem Plasmid, wodurch nur noch unmethyliertes, durch die zweite PCR entstandenes Plasmid übrigbleibt. Anschließend wird die Reaktion für 20 min bei 80°C hitzeinaktiviert.

2.5.7 Sequenzierung von DNA-Fragmenten

Um die durch Restriktionsfreie-Klonierung entstandenen Plasmide zu überprüfen, wurden die Plasmide sequenziert. Dafür wurde eine Übernacht-Kultur des plasmidtragenden Bakteriums angezogen und eine Plasmidpräparation durchgeführt. Das isolierte Plasmid wurde zu Microsynth Seqlab (Göttingen, Deutschland) geschickt, wo die Sequenzierung durchgeführt wurde. Die Ergebnisse wurden mittels Snapgene ausgewertet.

2.5.8 Transformatio von Plasmid-DNA

Transformation bezeichnet in der Mikrobiologie das Einschleusen von rekombinanten Plasmiden in Bakterienzellen. Die im Verlauf dieser Arbeit genutzte Transformationsmethode ist im folgendem aufgeführt.

2.5.8.1 Hitzeschocktransformation

Die für die Hitzeschocktransformation benötigten chemisch kompetenten Zellen von *E. coli* wurden von der Arbeitsgruppe Streit bereitgestellt. Die Zellen wurden für 5 min auf Eis aufgetaut und anschließend 1 µL das zu transformierenden Plasmids hinzu pipettiert. Der Ansatz wurde dann für 15 min auf Eis inkubiert. Danach erfolgte der Hitzeschock für 60 sec bei 42°C und eine erneute Inkubation für 5 min auf Eis. Anschließend wurde 1 mL LB-Medium zu den Zellen gegeben und diese für 45 min bei 37°C schüttelnd (180rpm) inkubiert. Nachfolgend wurden jeweils 50, 100 und 200 µL des Ansatzes auf LB-Agar Platten mit Selektionsantibiotika ausplattiert. Die Platten wurden über Nacht bei 37°C inkubiert.

2.6 Proteinbiochemische Methoden

Nachfolgend sind alle Proteinbiochemischen Methoden, die in der Arbeit genutzt wurden, aufgeführt.

2.6.1 Überexpression von Proteinen

Die Überexpression der Glycosyltransferasen erfolgte in 100 mL LB-Medium mit entsprechendem Selektionsantibiotikum. Die Expressionsstämme wurde 1 % in den mit Medium befüllten Schüttelkolben angeimpft und bis zu einer OD₆₀₀ von 0.6 bei 37°C schüttelnd (180 rpm) inkubiert. Danach wurden die Expressionskulturen mit IPTG (100 µM) induziert und bei 22°C über Nacht weiter inkubiert. Zur Überprüfung der Expression wurden zu den Zeitpunkten vor der Induktion und nach der Expression Proben genommen. Dazu wurden die genommenen Proben mit 1x SDS-Probenpuffer versetzt, aufgekocht und auf einer SDS-Page aufgetragen.

Tabelle 14: Zusammensetzung SDS-Probenpuffer

5x SDS-Probenpuffer	
Tris-HCl	0.25 M
Bromphenolblau	0.25 %
DTT	0.5 M
Glycerin	50 %
SDS	10 %

2.6.2 Natriumdodecylsulfat-Polyacrylamidgelelektrophorese (SDS-Page)

Natriumdodecylsulfat-Polyacrylamidgelelektrophorese dient der Auftrennung von Proteinen nach ihrer Größe. Dazu werden die zu untersuchenden Proteine mit Natriumdodecylsulfat umlagert und mittels eines elektrischen Feldes in einem Polyacrylamidgel aufgetrennt. Dieses besteht aus einem Trenn- und einem Sammelgel. Nachfolgend ist die Zusammensetzung der Gele aufgeführt.

Tabelle 15: Zusammensetzung Polyacrylamidgel

Substanz	10 % Trenngel	4 % Sammelgel
40 % Acrylamidmix Roth Geldoc	1.5 mL	240 µL
4x Trenngel-Stammlösung	1.5 mL	-
4x Sammelgel Stammlösung	-	360 µL
H ₂ O bidest.	2.5 mL	1.5 mL
TEMED	2.5 mL	2 µL
10 % Ammoniumpersulfat	50 µL	30 µL

Tabelle 16: Zusammensetzung der Trenngel Stammlösung

4x Trenngel Stammlösung	
Tris	18.2 g
SDS	0.4 g
H ₂ O bidest.	ad 100 mL
	pH 8,8

Tabelle 17: Zusammensetzung der Sammelgel Stammlösung

4x Sammelgel Stammlösung	
Tris	6.1 g
SDS	0.4 g
H ₂ O bidest.	ad 100 mL
	pH 8,8

Nach der Polymerisation der Polyacrylamidgele werden diese in eine Laufkammer (Mini PROTEAN® Tetra Cell, Bio-Rad Laboratoreis, Inc., Hercules, USA) eingesetzt und die Kammer an eine Spannungsquelle (PowerPac Basic, Bio-Rad Laboratoreis, Inc., Hercules, USA) angeschlossen. Zusätzlich zu den Proben wird auf das Gel ein Marker (Page Ruler Unstained Protein Ladder, Thermo Fisher Scientific, Massachusetts, USA) aufgetragen und das Gel für 1 h bei 120 V laufen gelassen. Der

Marker dient dem späteren Auswerten des Gels. Nach der Auftrennung der Proben im Gel wird dieses für 1 h mit einer Coomassie-Färbelösung gefärbt und abschließend über Nacht in einem Wasserbad schüttelnt entfärbt. Danach erfolgte die Auswertung und Dokumentierung des Gels.

Tabelle 18: Zusammensetzung Coomassie-Färbelösung

Coomassie-Färbelösung	
Coomassie Brilliant Blue R 250	0.3 g
Methanol	91 mL
Eisessig	16 mL
H ₂ O bidest	200 mL

Tabelle 19: Zusammensetzung des Elektrophoresepuffers

10x Elektrophoresepuffer	
Tris	30.3 g
Glycin	144.1 g
SDS	10.0 g
H ₂ O bidest	ad 1000 mL
	pH 8.4

2.6.3 Western Blot

Zur spezifischen Detektion und zum Nachweis von exprimierten Glycosyltransferasen wurde ein Western Blot durchgeführt. Für alle nachfolgenden beschriebenen Wasch- und Inkubationsschritte im Rahmen des Western Blots wurde ein Plattformschüttler (Polymax 1040, Heidolph Instruments Schwabach, Deutschland) genutzt, der mit 15-20 rpm betrieben wurde.

Als Grundlage für den Western Blot wurde wie unter 2.6.3 beschrieben ein SDS-Gel laufen gelassen, bei dem als einziger Unterschied ein anderer Marker (PageRuler™ Prestained Protein Ladder, Thermo Fisher Scientific, Massachusetts, USA) verwendet wurde. Nach der Polyacrylamidgelelektrophorese wurde das Gel auf eine Nitrocellulose-Membran (Roti®-NC Transfermembran, Roth, Karlsruhe, Deutschland) geblottet. Dazu wurde das Gel in einer Art Sandwich aus sechs Schichten Whatman-Papier (Gel-Blotting-Papier GBO03, Schleicher & Schuell BioScience GmbH, Dassel, Deutschland) und der Nitrocellulose-Membran, die alle jeweils in Transferpuffer getränkt waren, eingebracht. Zum Blotten wurde ein TransBlot Turbo Transfer System (Bio-Rad, Labororeis, Inc., Hercules, USA) genutzt. Danach wurde die Membran jeweils zweimal für 10 min mit TBS-Puffer gewaschen und anschließend, um unspezifische Bindungen der Antikörper zu verhindern, folgte die Inkubation der Membran für 1 h mit 5 % (w/v) Milchpulver in TBST-Puffer. Danach wurde die Membran wieder jeweils zweimal für 10 min mit TBST-Puffer gewaschen. Über Nacht erfolgte dann die Inkubation bei 4 °C mit dem primären Antikörper (anti-6-His IgG1, Maus, PROGEN Biotechnik GmbH, Heidelberg, Deutschland), der 1:5000 in TBST-Puffer verdünnt wurde. Am nächsten Tag wurde die Membran wieder jeweils zweimal für 10 min mit TBST-Puffer gewaschen und anschließend für 1 h bei Raumtemperatur mit dem sekundären Antikörper (Anti-Maus IgG Alkalische Phosphatase, Sigma-Aldrich® Biochemie GmbH, Hamburg, Deutschland) inkubiert. Dieser wurde 1:10000 verdünnt in TBST-Puffer mit 5 % (w/v) Milchpulver eingesetzt. Danach erfolgte erneutes Waschen für jeweils zweimal 10 min mit TBST-Puffer. Abschließend wurde die Membran für 3 min mit Detektions-Puffer äquilibriert und anschließend im Dunklen mit Detektions-Lösung inkubiert. Dies erfolgte so lange (ca. 15 min), bis die Banden deutlich zu erkennen waren. Zur Dokumentation wurde eine Gel-Dokumentationsanlage (ChemiDoc™ MP Imaging System, Bio-Rad Labororeis, Inc., Hercules, USA) genutzt.

Tabelle 20: Zusammensetzung des Transfer-Puffers

Das Methanol wurde nach dem Autoklavieren des Puffers hinzugegeben.

Transfer-Puffer	
Tris	15,14 g
Glycin	14,4 g
H ₂ O bidest	800 mL
Methanol	200 mL
	pH 8,6

Tabelle 21: Zusammensetzung des TBS[T]-Puffers

TBST [T]-Puffer	
Tris	15,14 g
NaCl	9 g
H ₂ O bidest	ad 1000 mL
Tween-20 (0.1 %w/v) [T]	1 mL
	pH 7,5

Tabelle 22: Zusammensetzung des Detektions-Puffer

Detektions-Puffer	
Tris	15,14 g
NaCl	9 g
H ₂ O bidest	ad 1000 mL
	pH 9,5

Tabelle 23: Zusammensetzung der Detektions-Lösung

Detektions-Lösung	
NBT	66 μ L
BCIP	33 μ L
Detektions-Puffer	ad 10 mL

Von NBT (Nitroblautetrazoliumchlorid) wurde eine Stammlösung von 75 mg/mL in 70 % (v/v) Dimethylformamid verwendet.

Von BCIP (5-Brom-4-chloro-3-indolylphosphat) wurde eine Stammlösung von 50 mg/mL in Dimethylformamid genutzt.

2.7 Biokonversion

Zur Identifizierung potenziell neuer Glycosyltransferasen und Substratscreening der neu identifizierten Glycosyltransferasen wurde eine ganzzellige Biokonversion genutzt.

Dazu wurden Schüttelkolben mit 20 mL LB-Medium und dem entsprechenden Antibiotikum (Ampicillin 100 μ M) befüllt und aus einer Vorkultur mit dem jeweiligen Bakterium 1 %ig angeimpft. Die Kolben wurde anschließend bei 37°C inkubiert, bis eine OD₆₀₀ von 0.6 erreicht wurde. Zu diesem Zeitpunkt erfolgte die Zugabe von IPTG (100 μ M) und des entsprechenden Substrates (100 μ M). Die Biokonversion wurde anschließend für 24 h bei 37°C schüttelnd inkubiert.

2.8 Chromatographische Methoden

Im Rahmen dieser Arbeit wurde Dünnschichtchromatographie für die Identifizierung potenziell neuer Glycosyltransferasen genutzt. Die Kombination der Systeme

Dünnschichtchromatographie und Hochleistungsflüssigkeitschromatographie wurden für Substratscreenings der neu identifizierten Glycosyltransferasen genutzt

2.8.1 Probenvorbereitung für die Dünnschichtchromatographie (DC)

Für die DC (Dünnschichtchromatographie) wurden 4 mL Probe aus den Kolben in den die Biokonversion durchgeführt wurde entnommen und mit 100 µL 1 M Phosphorsäure und 2 mL Essigsäureethylester versetzt. Dies wurde 10-mal invertiert und anschließend 10 Minuten lang bei 2000 g und 4°C zentrifugiert. Nach der Zentrifugation wurde der Überstand in Flachbodenvial überführt und für die DC-Analyse verwendet.

2.8.2 Probenvorbereitung für Hochleistungsflüssigkeitschromatographie (HPLC)

Proben für die Hochleistungsflüssigkeitschromatographie wurden in 10 % Acetonitril im Verhältnis 1:10 verdünnt und anschließend für 3 Minuten bei 10000 g zentrifugiert (Tischzentrifuge, MiniSpin, Eppendorf AG, Hamburg, Deutschland). Nach der Zentrifugation wurde der Überstand entnommen und in Flachbodenvials überführt und am Hitachi Elite LaChrome HPLC-System analysiert.

2.8.3 Dünnschichtchromatographie (DC)

Die DC-Analyse wurden mit einem automatischen DC-Sampler 4 (ATS 4; Camag, Muttenz, Schweiz) und 20 x 10 cm großen Kieselgelplatten auf Aluminiumträger 60 F254 (Merck KGaA, Darmstadt, Deutschland) durchgeführt. Die Platten wurden mit 20 µL der zu untersuchenden Probe durch den DC-Sampler 4 besprüht. Die eingesetzten Referenzen hatten eine Konzentration von 100µM. Als Laufmittel diente ein Gemisch aus Ethylacetat-Essigsäure-Ameisensäure-Wasser (100:11:11:27) (Wagner et al. 1983). Die entwickelte Platte wurde getrocknet und die Absorption der Banden mit einem DC-Scanner 3 (Camag, Muttenz, Schweiz) bestimmt. Danach wurde die Platte gefärbt, um die einzelnen Banden sichtbar zu machen. Dazu wurde in einem ersten Schritt die Platte in eine 1%ige (wt/vol) methanolische Lösung des

Naturstoffreagenz A eingetaucht und sofort anschließend getrocknet. Als zweiter Schritt wurde die Platte in eine 5 %ige (wt/vol) Lösung von Polyethylenglykol 4000 in Ethanol (70%, vol/vol) eingetaucht und ebenfalls sofort anschließend wieder getrocknet (Rabausch et al. 2013). Zur Visualisierung der Banden wurde eine UV-Lampe (365 nm) verwendet und die Platte fotografiert.

2.8.4 Hochleistungsflüssigkeitschromatographie (HPLC)

Die Hochleistungsflüssigkeitschromatographie wurde für die Analyse entstandener Glykoside verwendet. Dazu wurde ein Hitachi Elite LaChrome HPLC-System mit einer verschlossenen YMC-Triat C18-Säule mit einer Größe von 4,6 x 250 mm Länge und 5 µm Porengröße verwendet. Die Analysen wurden mit einem Reservephasenchromatographie-Gradienten aus Acetonitril und Wasser mit 0.05 % Trifluoressigsäure bei einer konstanten Flussrate von 1 ml/min durchgeführt. Alle Analysen bezogen sich auf die Absorption bei 254 nm und zur Quantifizierung der analysierten Proben wurden entsprechende Referenzsubstrate verwendet.

2.9 *In silico* Methoden

Nachfolgend sind alle im Rahmen dieser Arbeit genutzten und verwendeten bioinformatischen Softwares aufgelistet. Die Verwendung erfolgte nach den entsprechenden Anweisungen der Handbücher.

Tabelle 24: Übersicht über die in dieser Arbeit verwendeten bioinformatischen Softwares

Name	Anwendung/Verwendung	Quelle
BLAST	Abgleich von Protein- und Nukleotidsequenzen gegen die NCBI-Datenbanken	Boratyn et al, 2013; (Boratyn et al. 2013)
ChemDraw 21.0.0	Zeichnung von chemischen Strukturen	("ChemDraw." doi:10.1093/oi/authority.20110803095605254)
CD-search	Zur Identifizierung von konservierten Bereichen in Proteinsequenzen	Marchler-Bauer et al, 2017; (Marchler-Bauer et al. 2017)
ColabFold	Berechnung von möglichen Proteinstrukturen	Mirdita et al. 2022; (Mirdita et al. 2022)
Cytoscape 3.9.1	Für Netzwerkanalysen von Proteinsequenzen	Shannon et al, 2003; (Shannon et al. 2003)
UCSF Chimera 1.16	Darstellung von dreidimensionalen Proteinstrukturen	Pettersen et al, 2004; (Pettersen et al. 2004)
HMMER 3.3.2	Zum Datenbankabgleich von Sequenzhomologien	Mistry et al, 2013; (Mistry et al. 2013)
SnapGene 6.2.1	Zum Arbeit mit Protein- und Nukleotidsequenzen	("SnapGene software (www.snapgene.com)
Skyling	Zur Visualisierung des Hidden Markov Models	Wheeler et al, 2014; (Wheeler et al. 2014)
T-Coffee	Für die Erstellung von Alignements nach der „T-coffee Methoden“	Notredame et al, 2000, Wallace et al, 2006; (Notredame et al. 2000) (Wallace 2006)

2.9.1 Sequenzbasierte Suche mittels Hidden Markov Model

Zur Identifizierung neuer bisher unbekannter Glycosyltransferasen durch sequenzbasierte Suche in verschiedenen Datenbanken wurde ein Hidden Markov Model (HMM) erstellt. Dieses beruhte auf den Sequenzen bereits bekannter und funktionell getesteter Glycosyltransferasen (Tabelle 25). Die Erstellung des HMM erfolgte unter Verwendung des HMMer-Softwarepakets. Dafür wurde von den bekannten und funktionell aktiven Glycosyltransferasen ein Sequenz-Alignment auf Proteinebene mittels T-Coffee gemacht. Anschließend wurde das Programm HMMbuild genutzt, um aus dem Alignment ein Hidden Markov Model zu erstellen. Abschließend wurde unter Beibehaltung der Standardeinstellungen das HMM mittels HMM-Logo-Generator visualisiert (Wheeler et al. 2014).

Die Ergebnisse der sequenzbasierten Suche wurden weiter ausgewertet und die vielversprechendsten Sequenzen wurden synthetisiert oder amplifiziert. Diese wurden dann im Labor auf ihre funktionelle Aktivität überprüft. Die Sequenzen der funktionell aktiven neuen Glycosyltransferasen wurden genutzt, um das HMM zu verfeinern.

Tabelle 25: Proteinsequenzen, die zur Erstellung des HMM genutzt wurden.

UniProt Eintrag	Organismus	Quelle
C6VW33	<i>Dyadobacter fermentans</i>	Labor von Prof. Streit
I2GDW4	<i>Fibrisoma limi</i> BUZ 3	Labor von Prof. Streit
D2QIN4	<i>Spirosoma linguale</i>	Labor von Prof. Streit
C8X7J3	<i>Nakamurella multipartita</i> DSM 44233	Labor von Prof. Streit
AoAoC1IWM6	<i>Flaviumibacter solisilvae</i>	Labor von Prof. Streit
M7NFZ7	<i>Cesiribacter andamanensis</i> AMV16	Labor von Prof. Streit
AoAoE3V6S1	<i>Spirosoma radiotolerans</i>	Labor von Prof. Streit
IoKF42	<i>Fibrella aestuarina</i> BUZ 2	Labor von Prof. Streit
R4J7W2	uncultured bacterium	Rabausch et al. (Rabausch et al. 2013)

3. Ergebnisse

3.1 Sequenzbasierte Datenbanksuche

Das sequenzbasierte Screening nach potenziellen, neuartigen Glycosyltransferasen erfolgte auf Grundlage eines selbsterstellten Hidden Markov Models (2.9.1). Für das Screening wurden frei zugängliche Datenbanken und Genome aus der Arbeitsgruppe genutzt. Dazu gehörten unter anderem die UniProt- und nicht-redundante Datenbank des NCBI (National Center for Biotechnology Information). Die Suche wurde mit einem Schwellenwert >180 durchgeführt und resultierte in 100 Treffern von Genen, die für potenzielle, neue Glycosyltransferasen kodieren. Die Ergebnisse wurden anschließend weiter gefiltert, unter anderem nach der Zugehörigkeit zum Phylum Bacteroidota. Daraus ergaben sich 6 Hits (siehe Tabelle 26), die für weitere Analysen ausgewählt wurden.

Tabelle 26: Auflistung der 6 Treffer, die durch den sequenzbasierte Datenbankabgleich als vielversprechendste Einträge identifiziert wurden und im Rahmen dieser Arbeit weiter analysiert wurden.

GenBank Nummer	Name	Sequenz
OJY95357	GT-HH01	siehe Anhang
OJW59638	GT-HH02	siehe Anhang
GAO44619	GT-HH03	siehe Anhang
Sph21_4852	GT-HH04	siehe Anhang
WP_035084609.1	GT-HH05	siehe Anhang
WP_149627120.1	GT-HH06	siehe Anhang

3.1.1 Analyse des Hidden Markov Models

Für die Erstellung des beim sequenzbasierten Screening verwendeten Hidden Markov Models wurden 9 Sequenzen bereits beschriebener und charakterisierter Glycosyltransferasen verwendet (Tabelle 25). Die Konstruktion erfolgte wie im Abschnitt 2.9.1 beschrieben. Nach der Visualisierung des HMM (Abbildung 8) konnten vier konservierte Motive festgestellt werden (Abbildung 9). Diese wurden nachfolgend analysiert, wobei Funktionen im Rahmen der Donor- und Akzeptormolekülbindung zugeordnet werden konnten, die damit das aktive Zentrum des Enzyms bilden.



Abbildung 8: Visualisiertes Hidden Markov Model. In dieser gesamtheitlichen Darstellung des HMM wurden die konservierten Motive (hervorgehoben durch die roten Rahmen) aus Abbildung 9 identifiziert.

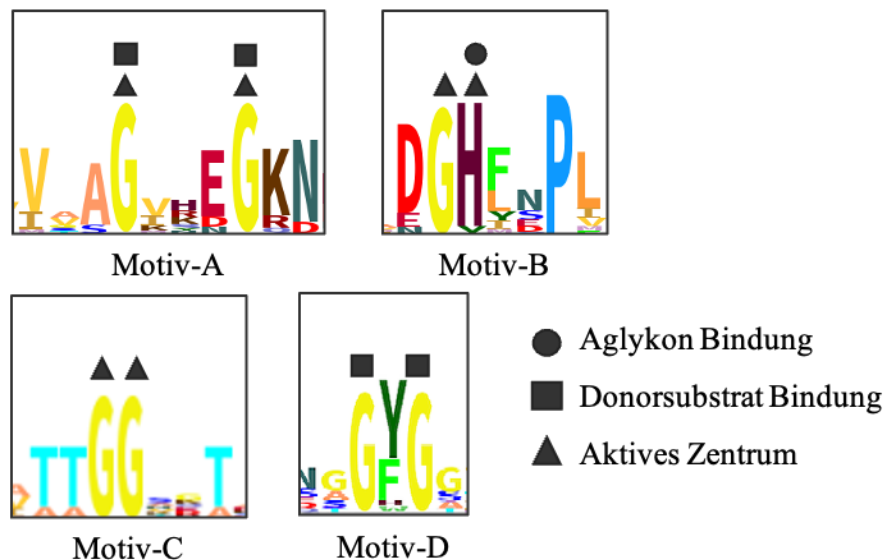


Abbildung 9: Identifizierte konservierte Motive des Hidden Markov Models. Markiert sind die Aminosäuren, die speziell für die Substrat- und Donorbindung verantwortlich sind.

3.2 *In silico* Analyse der Glycosyltransferasen

Neben der Charakterisierung der neu identifizierten Glycosyltransferasen erfolgten im Rahmen dieser Arbeit auch Analysen der neuen Enzyme mittels *in silico* Methoden. Dabei wurden die neuen Glycosyltransferasen auf ihren Faltungstyp untersucht und in Kontext mit bereits charakterisierten und beschriebenen Glycosyltransferasen gesetzt.

3.2.1 Faltungstypen

Zur Analyse und Auswertung der Faltungstypen der neu identifizierten Glycosyltransferasen wurde die Software ColabFold v1.5.2 (Mirdita et al. 2022) und für die Visualisierung UCSF Chimera 1.16 (Pettersen et al. 2004) genutzt.

Alle sechs neu entdeckten Glycosyltransferasen weisen die GT-B Faltung auf (Abbildung 10). Deutlich zu erkennen sind die beiden $\beta/\alpha/\beta$ Rossmann Domänen, die für diesen Faltungstyp typisch sind. Des Weiteren lässt sich durch die nur leichte Assoziation der beiden Domänen zueinander die durch diese Faltung entstehende typische Hufeisenform erkennen.

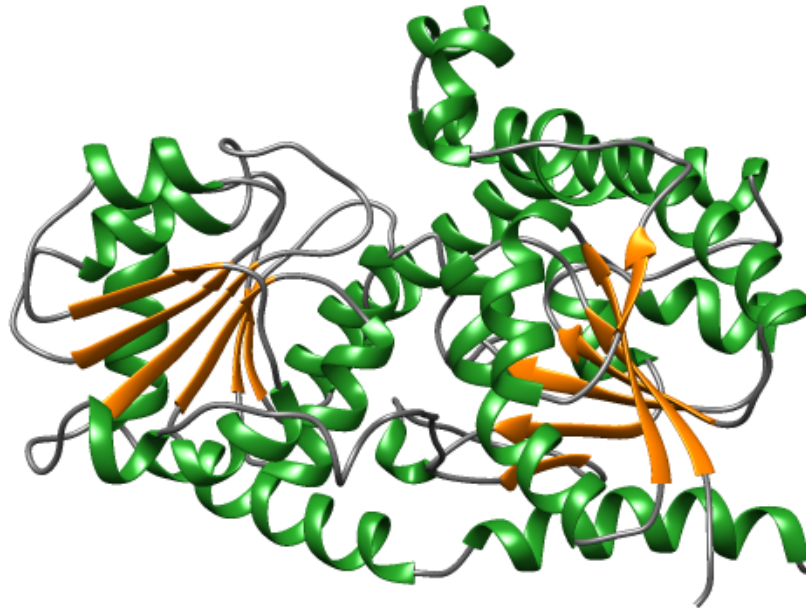


Abbildung 10: Mittels ColabFold vorhergesagte Proteinstruktur von GT-HHO5. Das Enzym weist die GT-B Faltung auf. GT-HHO5 ist beispielhaft für die anderen 5 Enzyme (GT_HHO1, GT-HHO2, GT-HHO3, GT-HHO4 und GT-HHO6) gewählt, die ebenfalls die GT-B Faltung aufweisen.

Die in Kapitel 3.1.1 beschriebenen konservierten Motive des Hidden Markov Models, welche bei den in dieser Arbeit neu beschriebenen Glycosyltransferasen gefunden wurden, lassen sich auch in den Proteinstrukturen der einzelnen Enzyme wiedererkennen. Gemeinsam bilden diese das aktive Zentrum der jeweiligen Transferasen. Dabei ist das Tyrosin aus dem Motiv D, sowie die beiden Glycin aus Motiv A (jeweils aus Abbildung 9) für die Bindung der Donormoleküle zuständig (Abbildung 11). Dieses Motiv konnte auch in fast allen Proteinsequenzen der Glycosyltransferasen gefunden werden. Einzig bei GT-HHO6 ist statt einem Tyrosin ein Phenylalanin zu finden und aus Motiv A fehlt das zweite Glycin. Hier geht die Proteinstruktur bereits in die α -Helix über. Neben diesen Abweichungen fehlt bei GT-HHO2 das Motiv A vollständig. Außerdem ist nicht nur die Bindung des Donormoleküle wichtig, auch die Bindung des jeweiligen Aglykons spielt eine entscheidende Rolle, damit die jeweilige Glycosyltransferase dieses glykosylieren kann. Hierfür ist das Histidin hauptverantwortlich. Dieses lässt sich in allen Sequenzen der neuen Glycosyltransferasen wiederfinden.

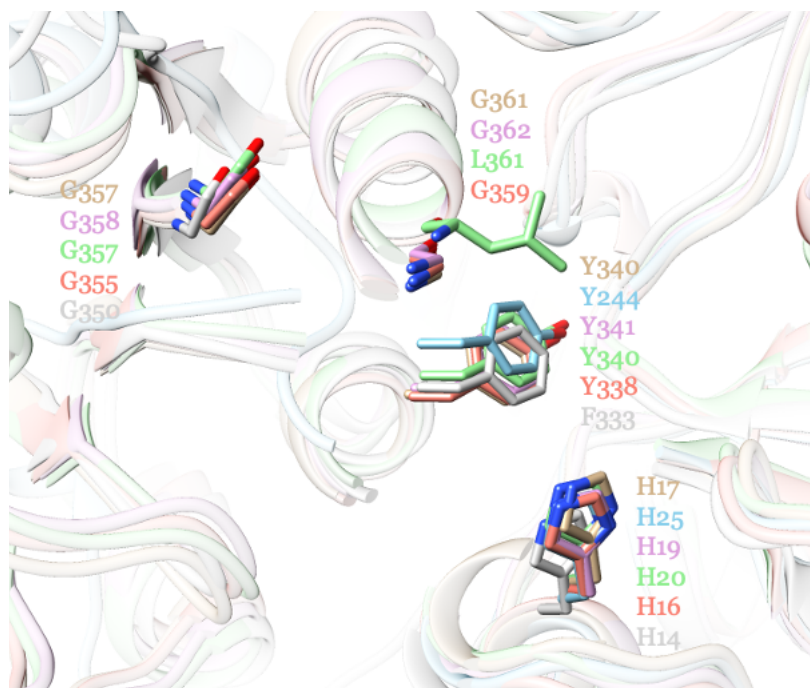


Abbildung 11: Strukturüberlagerung der sechs neu identifizierten Glycosyltransferasen. Hervorgehoben sind die Aminosäuren, die das aktive Zentrum bilden und in der Analyse des Hidden Markov Model identifiziert werden konnten. Ocker: GT-HH01, Blau: GT-HH02, Pink: GT-HH03, Grün: GT-HH04, Rot: GT-HH05, Grau: GT-HH06

3.3 Glycosyltransferasen Charakterisierung

Zur Glycosyltransferasen Charakterisierung wurde wie unter 2.7 beschrieben die Methode der Biokonversion verwendet. Dabei wurde Quercetin als Modells substrat im Rahmen des Aktivitätsnachweises der potenziellen neuen Glycosyltransferasen verwendet. Für die weiterführenden Analysen wie das Substratspektrum wurden die in Tabelle 1 aufgeführten Aglykone genutzt.

3.3.1 Aktivitätsassay

Die durch das sequenzbasierte Screening identifizierten potenziellen neuen Glycosyltransferasen wurden auf ihre Aktivität untersucht. Dafür wurden die das jeweilige zu untersuchende Konstrukte tragenden *E.coli* Stämme in LB-Medium

angeimpft und bei Erreichen der OD₆₀₀ 0.6 mit IPTG induziert und mit dem Aglykon Quercetin über Nacht inkubiert. Anschließend erfolgte die Analyse der potenziell entstandenen Glykoside mittels Dünnschichtchromatographie.

Die Auswertung der Dünnschichtplatten zeigt, dass die Konstrukte GT-HH01 bis GT-HH06 aktiv sind. Damit konnte nachgewiesen werden, dass die jeweiligen Gene eine aktive Glycosyltransferase codieren, die das Substrat Quercetin glykosylieren kann. Die Abbildung 12 zeigt die Dünnschichtplatte mit den jeweiligen verwendeten Referenzen (Spalte 1-4) und den untersuchten Konstrukten (Spalte 5-10). Als Referenzen dienten in Spalte 1 das Ausgangssubstrat (Aglykone) Quercetin und in den Spalten 2 bis 4 sind die möglichen Glykoside von Quercetin aufgetragen. Dabei handelt es sich um Quercitrin (Spalte 2), Isoquercitrin (Spalte 3) und Rutin (Spalte 4). In der in Spalte 5 aufgetrennten Probe der Biokonversion des *E. coli* BL21 DE3, pET21a(+):GT-HH01 sind zwei deutliche Banden auf der gleichen Höhe wie von den Referenzen Quercitrin und Rutin erkennbar. Daraus lässt sich schlussfolgern, dass hier eine Glykosylierung des Ausgangsstoffes Quercetin zu Quercitrin und Rutin stattgefunden hat. In Spalte 6, 7 und 8, in denen die Proben von *E. coli* BL21 DE3, pET21a(+):GT-HH02, *E. coli* BL21 DE3, pET21a(+):GT-HH03 und *E. coli* BL21 DE3, pET21a(+):GT-HH04 aufgetrennt wurden, sind leichte Banden für eine Umsetzung zu Quercitrin erkennbar. Die Probe vom *E. coli* BL21 DE3, pET21a(+):GT-HH05 in Spalte 9 weist eine Bande auf der gleichen Höhe wie die Referenz Isoquercitrin auf, was für eine Umsetzung des Aglykons Quercetin zum Glykosid Isoquercitrin steht. Die Probe von *E. coli* BL21 DE3, pET21a(+):GT-HH06 ist in Spalte 10 aufgeführt und zeigt eindeutig eine Bande für die Glykosylierung von Quercetin zu Rutin. Als Negativkontrolle fungierte der zur Klonierung genutzte Stamm *E. coli* BL21 DE3 mit nicht rekombinantem Vektor pET21.

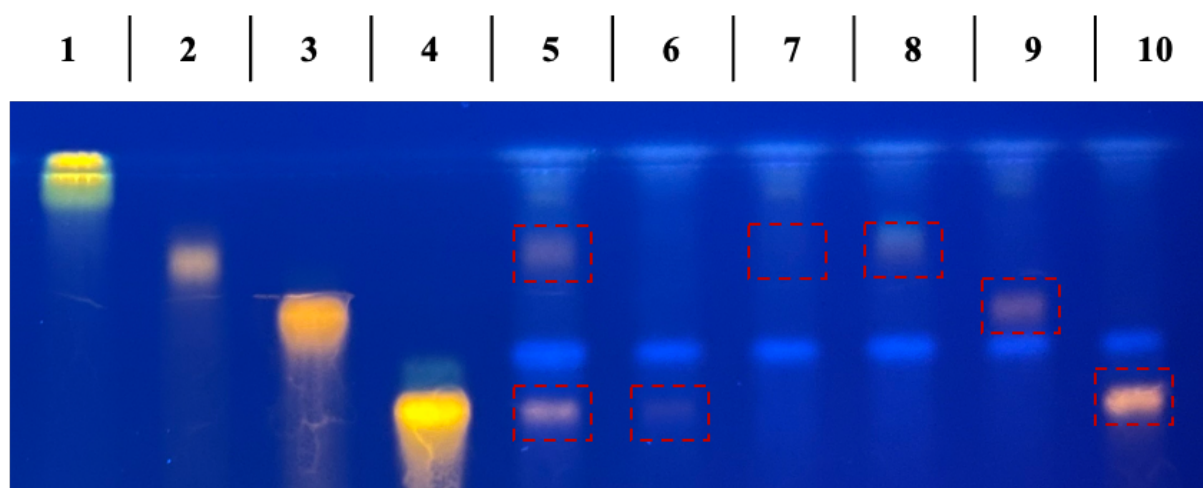


Abbildung 12: Dünnschichtplatte nach der Auftrennung der Proben sowie Färben unter UV-Licht mit der Wellenlänge 356 nm. Aufgetrennt sind in den Spalten 1 bis 4 die verwendeten Referenzen: 1: Quercetin, 2: Quercitrin, 3: Isoquercitrin, 4: Rutin. In den nachfolgenden Spalten 5 bis 10 sind die Proben der Konstrukte GT-HH01 bis GT-HH06 aufgetragen. Mit roten Rahmen sind die Banden in den jeweiligen Spalten der Proben markiert, bei denen eine Glykosylierung erfolgt ist. Die blauen Banden in den Spalten 5 bis 10 sind angefärbte Substanzen aus dem jeweiligen Medium und unabhängig vom Nachweis der erfolgreichen Glykosylierung. Die Eindeutigkeit der Banden für Glykosylierungsaktivität ist aufgrund der gleichen Farbe sowie der gleichen Position wie die jeweilige Referenz auf der Platte gegeben.

3.3.2 Substratspektrum Bestimmung

Nachdem die Aktivität der neuen Glycosyltransferasen nachgewiesen werden konnte, erfolgten weitere Untersuchungen dieser neuen Enzyme. Dazu gehört auch die Analyse des Substratspektrums der einzelnen GTs. Dazu wurden wieder Biokonversionen mit unterschiedlichen Substraten durchgeführt. Dafür wurden Substrate verwendet, die sich in ihrer Struktur immer mehr von dem in dieser Arbeit verwendeten Modellsupstrat Quercetin unterscheiden. Die jeweils genutzten Aglykone sind in Tabelle 1 einsehbar.

Die Untersuchungen der Substratspektren der neuen Glycosyltransferasen GT-HH01 bis GT-HH06 zeigten, dass all diese Enzyme in der Lage sind, Naringenin zu glykosylieren. Neben der Glykosylierung von Naringenin sind die im Rahmen dieser Arbeit identifizierten neuen GTs GT-HH01, GT-HH02 und GT-HH03 auch in der Lage, die Aglykone Kaempferol, Oxyresveratrol und Hesperetin zu verzuckern (Auflistung in Tabelle 27). In Abbildung 13 ist die Dünnschichtplatte der Substratspektrumsanalyse

für das Hesperetin gezeigt. Hier ist eindeutig zu erkennen, dass die Glycosyltransferasen GT-HHo1, GT-HHo2 und GT-HHo3 in der Lage sind, das Hesperetin zu glykosylieren. Dies ist erkennbar an den in den jeweiligen Spalten auftretenden grünlichen Banden (roter Rahmen). Die oberste Bande ist das eingesetzte Hesperetin (Aglykon) und die darunter befindliche, ebenfalls grünliche Bande, das durch die Aktivität der entsprechenden Enzyme neuentstandene Glykosid.

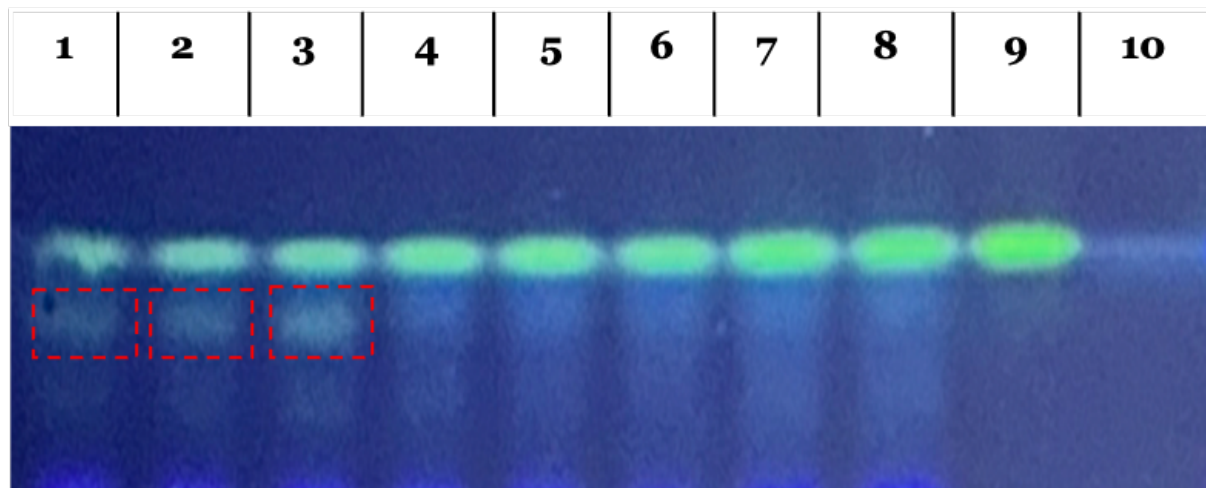
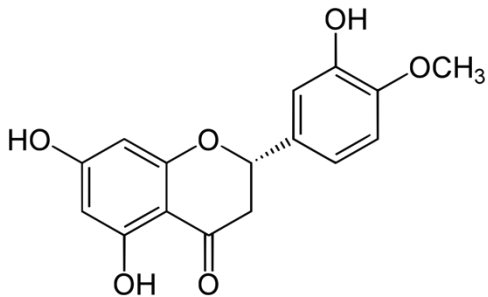
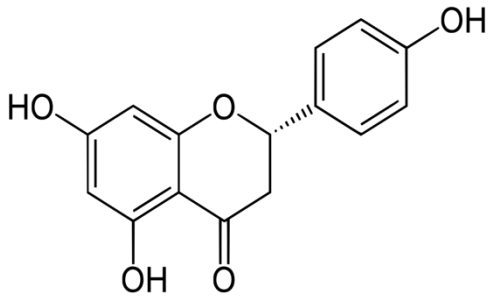
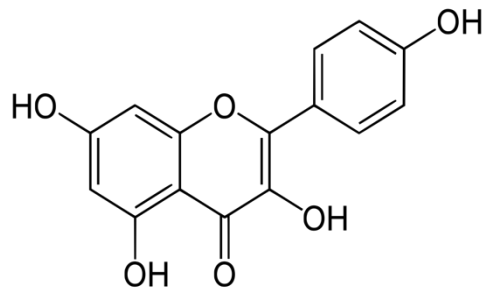
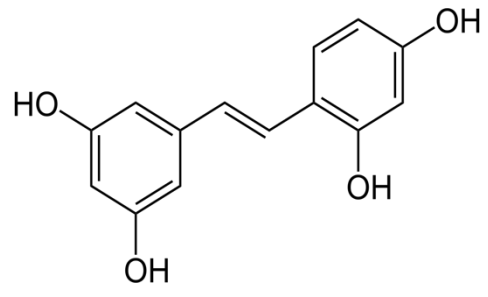


Abbildung 13: Gefärbte Dünnschichtplatte (unter UV-Licht 366 nm) mit dem Substrat Hesperetin, das im Rahmen des Substratspektrenassays verwendet und analysierte wurde. Aufgetragen sind die jeweiligen untersuchten Konstrukte (GT-HHo1 bis GT-HHo6), die beiden eingesetzten Kontrollen (*E. coli* BL21 DE3, pET21 und *E. coli* BL21 DE3), sowie das entsprechende Aglykon. Spalte 1 bis 6 zeigt die untersuchten Konstrukte in numerischer Reihenfolge von GT-HHo1 bis GT-HHo6. In den Spalten 7 und 8 sind die beiden eingesetzten Negativkontrollen zu sehen. Spalte 9 zeigt das verwendete Aglykon Hesperetin und in Spalte 10 das Glykosid Hesperetin 7-O-glycosid.

Tabelle 27: Untersuchte Substrate im Rahmen der Substratspektrumsanalyse und Auflistung, durch welche Glycosyltransferasen diese glykosyliert werden konnten.

<u>Substrat</u>	<u>Struktur</u>	<u>glykosylierbar durch</u>
Hesperetin		GT-HH01
		GT-HH02
		GT-HH03
Naringenin		GT-HH01
		GT-HH02
		GT-HH03
		GT-HH04
		GT-HH05
		GT-HH06
Kaempferol		GT-HH01
		GT-HH02
		GT-HH03
Oxyresveratrol		GT-HH01
		GT-HH02
		GT-HH03

3.3.3 Nachweis der Expression

Die als aktiv ermittelten Glykosyltransferasen (GT-HH01 bis GT-HH06) wurden neben ihren Substratspektren auch auf ihre Exprimierbarkeit untersucht. Dies erfolgte wie in Abschnitt 2.6 Proteinbiochemische Methoden beschrieben. Für den Western Blot wurden dabei aus Mäusen stammende Antikörper verwendet.

Auf den als Grundlage für die nachfolgende Western Blot Untersuchung dienenden SDS-Gelen sind die jeweiligen Proteine eindeutig zu erkennen (Abbildung 14). Hier lassen sich deutliche Unterschiede zwischen den nicht-induzierten Proben und den Proben nach der Induktion mit IPTG erkennen. Die erwarteten Größen der Proteine stimmen dabei mit den kalkulierten Größen überein. Diese liegen zwischen 36.6 kDa für GT-HH02, 47.9 kDa (GT-HH06), 48.2 kDa (GT-HH04), 49.6 kDa (GT-HH05) und 50.4 kDa für die Glykosyltransferase GT-HH01 und 03.

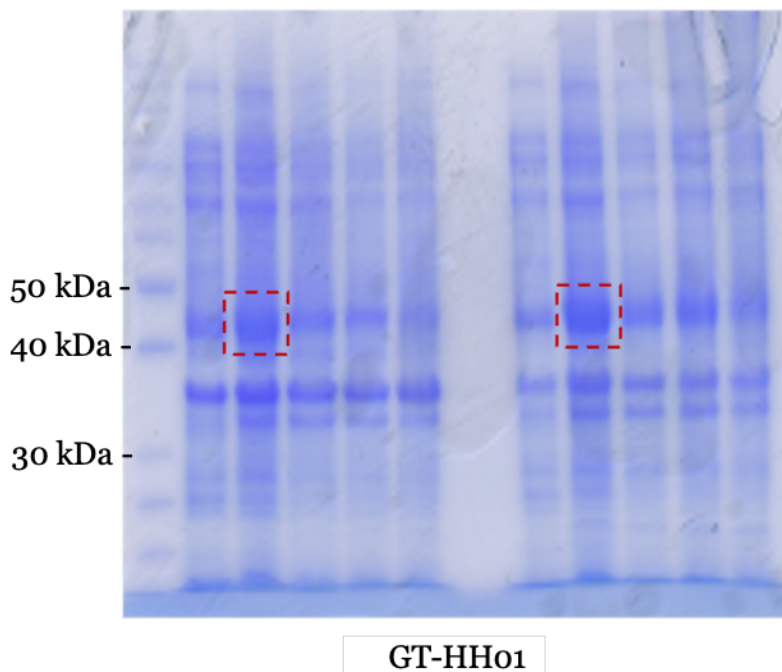


Abbildung 14: SDS-Gel zum Nachweis der Expression der Glycosyltransferase GT-HH01. Markiert (roter Rahmen) sind die Banden, bei denen deutlichste Expression erkennbar war. Untersucht wurde die Exprimierbarkeit in unterschiedlichen Medien (LB-Medium und TB-Medium) und bei unterschiedlichen Temperaturen (17°C, 22°C, 28°C und 37°C). Die Proben sind in der Reihenfolge: Marker, nicht-induziert, 17°C, 22°C, 28°C und 37°C aufgetragen, sowie erst die Proben aus dem LB-Medium und als zweites aus TB-Medium.

Tabelle 28: Übersicht über die Ergebnisse der Expressionsstudien der neu identifizierten Glycosyltransferasen GT-HH01 – GT-HH06

Glycosyltransferase	Optimale Expressionstemperatur		Proteingröße laut SDS-Gel	Kalkulierte Proteingröße
	LB-Medium	TB-Medium		
GT-HH01	17°C	17°C	45 kDa	50,4 kDa
GT-HH02	28°C	28°C	33 kDa	36,6 kDa
GT-HH03	22°C	22°C	46 kDa	50,4 kDa
GT-HH04	22°C	22°C	44 kDa	48,2 kDa
GT-HH05	28°C	28°C	48 kDa	49,6 kDa
GT-HH06	28°C	28°C	47 kDa	47,9 kDa

Aus den Ergebnissen geht deutlich hervor, dass für die neu identifizierten Glycosyltransferasen die optimale Expressionstemperaturen zwischen 17°C und 28°C liegen (Abbildung 14). Ebenfalls erkenntlich wird, dass die kalkulierten Proteingrößen größtenteils mit den erhaltenen bzw. beobachteten Größen übereinstimmen. Hier liegen die Unterschiede bei maximal 5.4 kDa und minimal bei ca. 0,9 kDa. Außerdem wurde das TB-Medium als das vielversprechendste Expressionsmedium identifiziert, bei dem die größten Proteinkonzentrationen erzielt werden konnten.

Mittels Western Blot ließen sich die Proteine GT-HH01, GT-HH03 und GT-HH04 nachweisen. Dabei stimmen auch wieder die Größen der exprimierten Proteine mit den vorher kalkulierten Größen überein. Als beste Temperatur für die Expression der Gene zeigte sich im Gegensatz zu den Ergebnissen von den SDS-Gelen, dass diese bei 17°C bzw. 22°C liegen. Die Konstrukte GT-HH02, GT-HH05 und GT-HH06 ließen sich dagegen mittels Western-Blot nicht nachweisen. Exemplarisch ist nachfolgend die Expression der Glykosyltransferase GT-HH01 dargestellt (Abbildung 15). Expressionsnachweise der Konstrukte GT-HH03 und GT-HH04 sind im Anhang einzusehen.

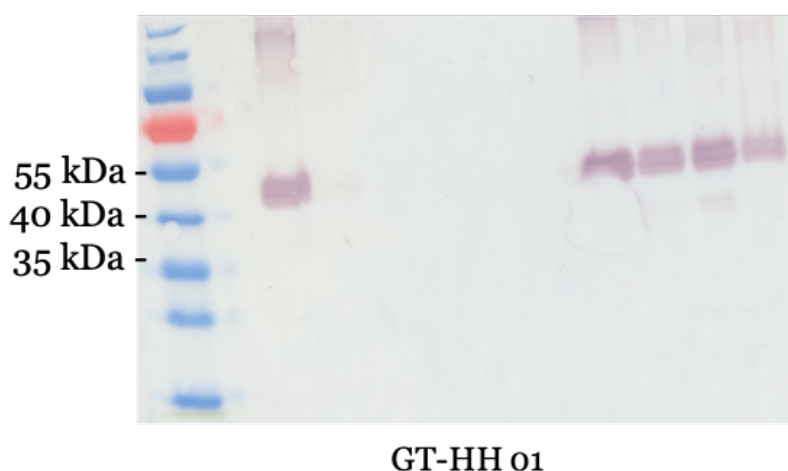


Abbildung 15: Western-Blot Membran des Proteins GT-HH01. Erkennbar sind die Banden der Expression unter den beiden verschiedenen Temperaturbedingungen 17°C (links) und 22°C (rechts). Deutlich ist zuerkennen, dass für GT-HH01 die optimale Expressionstemperatur 17°C beträgt.

3.3.4 Zusammenfassung über die herausgefundenen Eigenschaften der neu identifizierten Glycosyltransferasen

Nachfolgend sind die experimentell und *in silico* ermittelten Eigenschaften der neu identifizierten Glycosyltransferasen GT-HH01 bis GT-HH06 tabellarisch zusammengefasst (Tabelle 29-34). Auffällig ist, dass alle neuen GTs eine GT-B Faltung aufweisen. Zusätzlich zu der Eigenschaft, dass alle Enzyme Quercetin und Naringenin glycosylieren können, zeigten die Enzyme GT-HH01, GT-HH02 und GT-HH03 glycosidische Aktivität bei den Polyphenolen Hesperetin, Kaempferol und Oxyresveratrol.

Tabelle 29: Zusammenfassung der ermittelten Eigenschaften der Glykosyltransferase GT-HH01

GT-HH01	
Phylum	Bacteroidota
Organismus	<i>Sphingobacterium</i> sp.
GenBank Nummer	OJY95357
glykosyliert	Quercetin, Hesperetin, Naringenin, Kaempferol, Oxyresveratrol
Faltungstyp	GT-B
<i>In vitro</i> exprimierbar	ja, optimal bei 17°C

Tabelle 30: Zusammenfassung der ermittelten Eigenschaften der Glykosyltransferase GT-HH02

GT-HH02	
Phylum	Bacteroidota
Organismus	<i>Sphingobacterium</i> sp.
GenBank Nummer	OJW59638
glykosyliert	Quercetin, Hesperetin, Naringenin, Kaempferol, Oxyresveratrol
Faltungstyp	GT-B
<i>In vitro</i> exprimierbar	ja, optimal bei 28°C

Tabelle 31: Zusammenfassung der ermittelten Eigenschaften der Glykosyltransferase GT-HH03

GT-HH03	
Phylum	Bacteroidota
Organismus	<i>Flaviumibacter petaseus</i>
GenBank Nummer	GAO44619
glykosyliert	Quercetin, Hesperetin, Naringenin, Kaempferol, Oxyresveratrol
Faltungstyp	GT-B
<i>In vitro</i> exprimierbar	ja, optimal bei 22°C

Tabelle 32: Zusammenfassung der ermittelten Eigenschaften der Glykosyltransferase GT-HH04

GT-HH04	
Phylum	Bacteroidota
Organismus	<i>Sphingobacterium</i> sp.
GenBank Nummer	Sph21_4852
glykosyliert	Quercetin, Naringenin
Faltungstyp	GT-B
<i>In vitro</i> exprimierbar	ja, optimal bei 22°C

Tabelle 33: Zusammenfassung der ermittelten Eigenschaften der Glykosyltransferase GT-HH05

GT-HH05	
Phylum	Bacteroidota
Organismus	<i>Aquimarina latercula</i>
GenBank Nummer	WP_035084609.1
glykosyliert	Quercetin, Naringenin
Faltungstyp	GT-B
<i>In vitro</i> exprimierbar	ja, optimal bei 28°C

Tabelle 34: Zusammenfassung der ermittelten Eigenschaften der Glykosyltransferase GT-HH06

GT-HH06	
Phylum	Bacteroidota
Organismus	<i>Aquimarina</i> sp.
GenBank Nummer	WP_149627120.1
glykosyliert	Quercetin, Naringenin
Faltungstyp	GT-B
<i>In vitro</i> exprimierbar	ja, optimal bei 28°C

3.4 Netzwerkanalyse beschriebener Glycosyltransferasen

Für die Netzwerkanalyse der neuen Glycosyltransferasen wurden die Datensätze aller zum Zeitpunkt der Analyse als charakterisiert eingestuften bakteriellen und archeeälle Glycosyltransferasen der CAZY-Datenbank genutzt. Die Analyse der Datensätze erfolgte auf Proteinsequenzebene und wurde mittels Cytoscape (Shannon et al. 2003) visualisiert.

Die Auswertung der Daten ergab, dass im in Abbildung 16 dargestellten Netzwerk drei größere Cluster identifiziert werden konnten. Hierbei steht jeder Punkt für eine bakterielle und jedes Dreieck für eine archaeelle Glycosyltransferase. Weiterhin erfolgte die Einteilung der Enzyme entsprechend ihres Phylums. So stehen blaue Punkte für das Phylum Pseudomonadota, rote Punkte für Actinobacteria, cyane für Cyanobacteria und gelbe Punkte für Bacteroidota. Die Striche zwischen den Punkten und Dreiecken stehen jeweils für eine mindestens 60 %ige Übereinstimmung auf Proteinsequenzebene der einzelnen Enzyme.

Auffällig bei der Netzwerkanalyse ist, dass fast alle Glycosyltransferasen aus dem Phylum Actinobacteria sich in einem Cluster sammeln, wohingegen die GTs aus dem Phylum Pseudomonadota weiträumiger verteilt und vernetzter sind (Abbildung 16). Außerdem geht aus der Analyse hervor, dass es bisher nur eine sehr begrenzte Anzahl an charakterisierten und beschriebenen Glycosyltransferasen aus dem Phylum Bacteroidota gibt. Alle die in dieser Arbeit neu identifizierten Glycosyltransferasen sind in dem Cluster zu finden, das von den Actinobacteria dominiert wird. Besonders auffällig ist die Glycosyltransferase GT-HHo2, die deutlich weiter entfernt von den anderen fünf in dieser Arbeit identifizierten GTs lokalisiert ist.

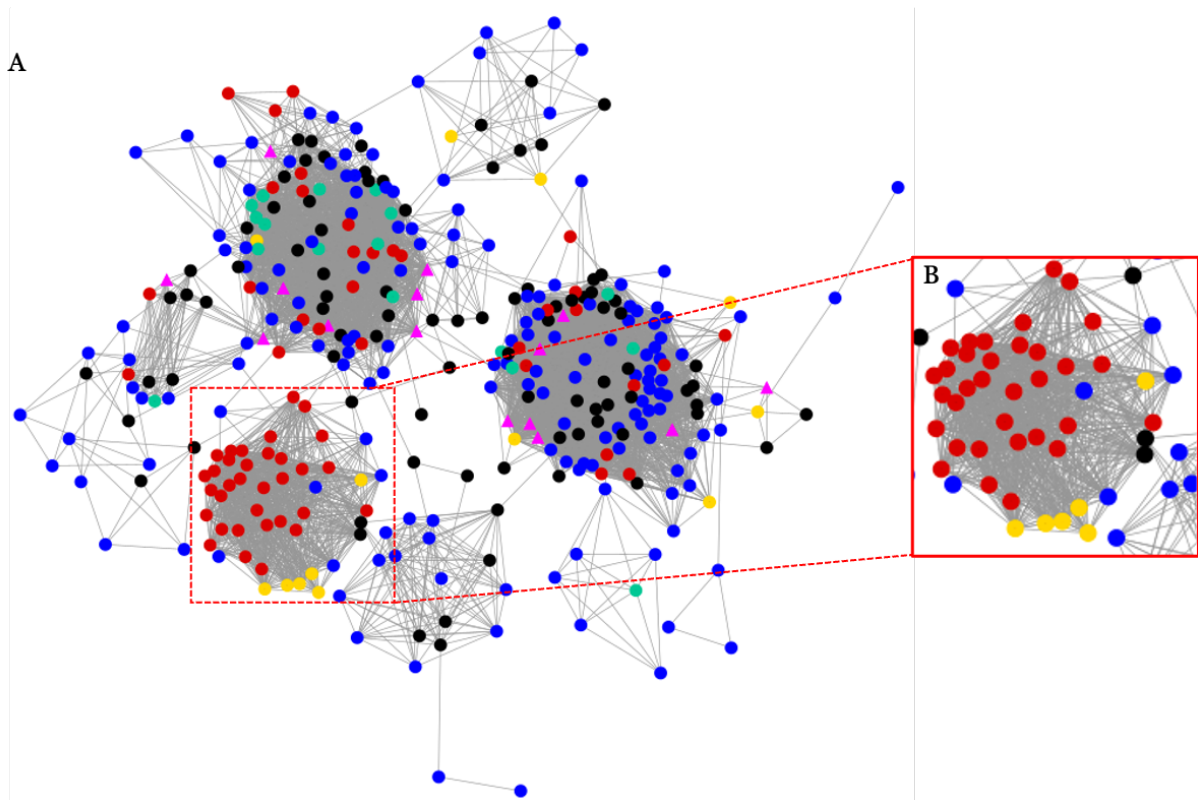


Abbildung 16: Netzwerkanalyse aller bekannten und charakterisierten Glycosyltransferasen aus der CAZY-Datenbank. Punkte stehen für bakterielle und Dreiecke für archeeälle Enzyme. Die weitere Einteilung erfolgte nach dem entsprechenden Phylum. Blau: Pseudomonadota, rot: Actinobacteria, cyan: Cyanobacteria, pink: Archaea, gelb: Bacteroidota. **A** Das gesamte Netzwerk mit seinen drei dominierenden Clustern. **B** Cluster, in dem alle in dieser Arbeit beschriebenen Glycosyltransferasen zu finden sind. Die Visualisierung des Netzwerkes erfolgte mittels der Cytoscape Software (Shannon et al. 2003).

Weitere Untersuchungen und Auswertungen ergaben, dass die einzelnen Enzyme in den drei dominierenden Clustern unterschiedlichen Funktionen zugeordnet werden konnten. Die Analysen und Auswertungen der Einträge des von Actinobacteria beherrschtem Clusters ergaben, dass diese Enzyme hauptsächlich Funktionen und Beteiligungen im Rahmen des Metabolismus der Zelle zuzuordnen sind. Dazu zählen oft Glykosylierungen von sekundären Pflanzenmetaboliten wie unter anderen Flavonoide. Die beiden weiteren Cluster beherbergen Enzyme, die beim Zellmembranaufbau und der Modifizierung dieser beteiligt sind (Abbildung 17).

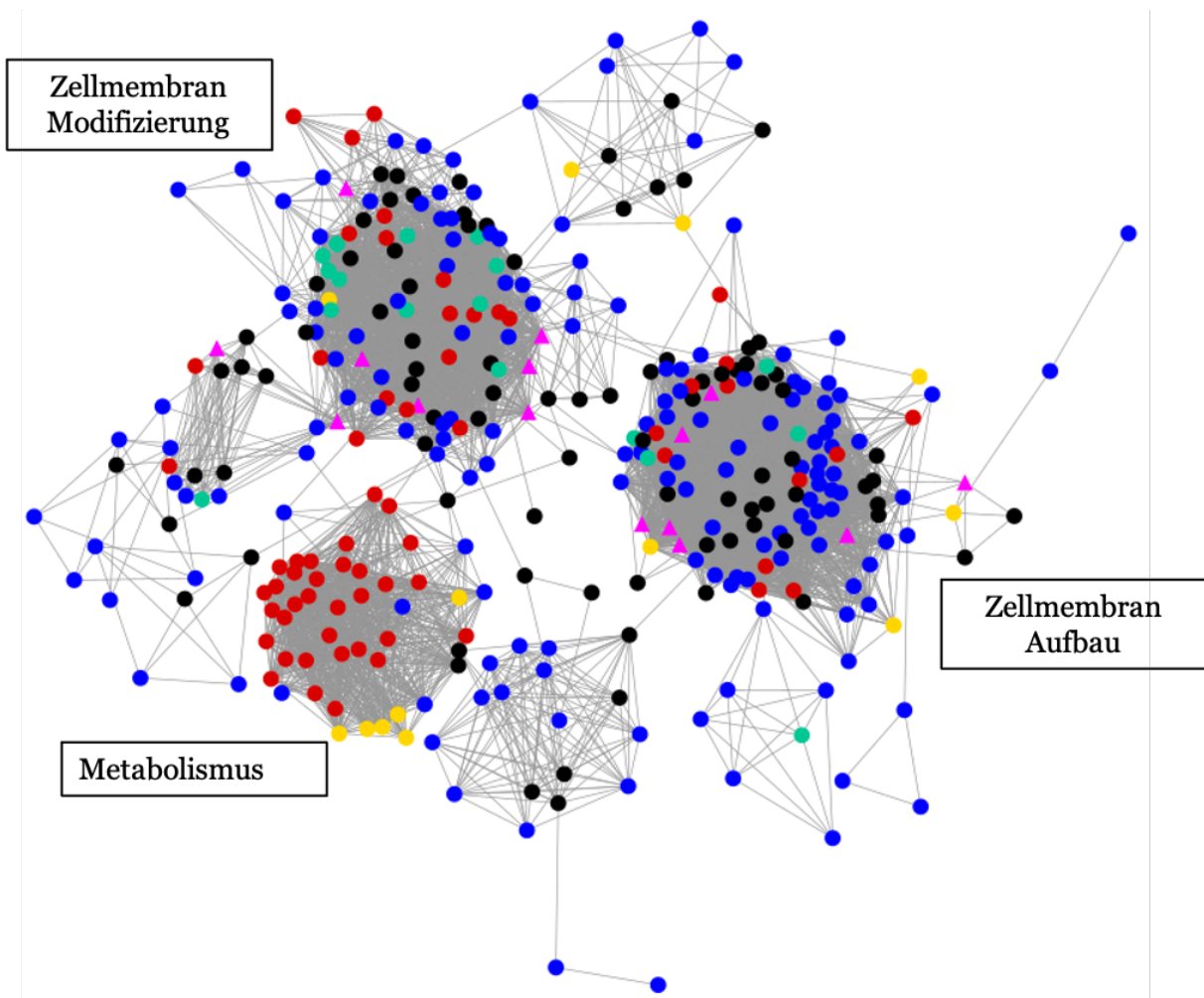


Abbildung 17: Netzwerkanalyse aller bekannten und charakterisierten Glycosyltransferasen aus der CAZY-Datenbank. Die einzelnen Einträge der Cluster lassen sich bestimmten Funktionen zuzuordnen. Diese beruhen auf der Auswertung der jeweiligen im Cluster versammelten Enzyme. Die Visualisierung des Netzwerkes erfolgte mittels der Cytoscape Software (Shannon et al. 2003).

4. Diskussion

Bisherige Veröffentlichungen und Studien fokussieren sich vorwiegend auf die Beschreibung von Glycosyltransferasen im Zusammenhang mit der Glykosylierung von sekundären Pflanzenmetaboliten oder Zellwandbestandteilen sowie im Kontext mit medizinischen Erkenntnissen (Yang 2023),(Lu et al.),(Wang 2024),(Grewal 2021),(Gharabli 2023),(Wang 2023). Den gegenüber werden Glycosyltransferasen und ihre Möglichkeiten im Zusammenhang mit der Nutzung von verschiedensten Molekülen wie Chitin oder Cellulose aus Phyla wie Bacteroidota bisher weniger beachtet. Daher lag im Rahmen dieser Arbeit der Fokus auf der Identifizierung und Charakterisierung neuartiger Glycosyltransferasen aus dem Phylum Bacteroidota.

Dafür wurden auf Proteinsequenzen basierende Datenbanksuchen durchgeführt. Die resultierenden Screeningergebnisse ausgewertet und vielversprechende Treffer experimentell untersucht und analysiert. Ausgangspunkt der Datenbanksuche stellte dabei die Erstellung eines Hidden Markov Modells auf Grundlage bereits veröffentlichter Sequenzen von Glycosyltransferasen dar. Das Screening lieferte 100 Treffer, von denen 11 für die Untersuchungen ausgewählt wurden. Die entsprechenden Gene der möglichen Glycosyltransferasen wurden amplifiziert oder synthetisiert, kloniert, und in *E. coli* exprimiert, um ihre Aktivität zu untersuchen. Als aktiv identifizierte Enzyme wurden anschließend weiter charakterisiert, wobei das Substratspektrum und die Exprimierbarkeit der Transferasen ermittelt wurde, sowie die phylogenetische Einordnung im Kontext bereits bekannter Glycosyltransferasen aus dem Phylum Bacteroidota erfolgte. Hierbei zeigte sich, dass dieses Phylum in Bezug auf charakterisierte Glycosyltransferasen noch sehr unbeachtet ist. So sind 3720 Sequenzen von Glycosyltransferasen kodierender Gene aus Bacteroidota bekannt, aber nur 0,22 % davon beschrieben und als charakterisiert eingestuft.

4.1 Sequenzbasiertes Screening nach neuen bacteroidalen Glycosyltransferasen

Als Ansatzpunkt bei der Suche nach neuen Glycosyltransferasen, die aus Vertretern des Phylum Bacteroidota stammen, wurde eine sequenzbasierende Methodik gewählt. Dafür wurde auf Grundlage von Sequenzdaten bereits beschriebener Glycosyltransferasen ein Hidden Markov Model erstellt. Mit diesem erfolgte im weiteren Verlauf der Arbeit die sequenzbasierende Datenbanksuche. Des Weiteren wurde das HMM mit den Sequenzdaten der neu identifizierten, aktiven Glycosyltransferasen verfeinert. Die Auswahl eines sequenzbasierten Screenings im Gegensatz zu einem funktionalen Screening geschah deshalb, weil mit diesem in kurzer Zeit eine große Menge an Daten analysiert werden kann und es für die Suche nach Enzymen mit spezifischen Eigenschaften gut geeignet ist (Güllert et al. 2016). Außerdem verspricht das sequenzbasierte Screening gegenüber eines Fosmid-Bibliothek Screenings, welches üblicherweise bei funktionalen Ansätzen genutzt wird (Streit and Daniel 2017), eine höhere Erfolgsaussicht und benötigt weniger vorbereitende Arbeit. So fallen das Anlegen der entsprechenden Fosmide, als auch ihre Überprüfung weg. Zudem liegen die Sequenzdaten bei einem sequenzbasierten Screening direkt vor und die Fosmide müssen nicht erst sequenziert werden. Weitere Vorteile eines auf Sequenzen basierendem Screening ist die Möglichkeit, verschiedene Datensätze (marine oder terrestrische) zu untersuchen. Dem gegenüber muss bedacht werden, dass diese Screenings immer auf Grundlage der jeweiligen verwendeten Sequenzen basieren und die Auswahl dieser einen enormen Einfluss auf die erhaltenen Ergebnisse hat. Oft findet man mit dem entsprechenden Screenings Sequenzhomologe in den Datensätzen, aber die Einträge strukturell anderer, ebenfalls potenziell aktiver, Enzyme werden nicht gefunden (Streit and Daniel 2017). Außerdem sind sequenzbasierte Methoden von den zur Verfügung stehenden Datensätzen und deren Akkuratheit abhängig. So ist es möglich, dass gleiche Enzyme in unterschiedlichen Datenbanken andere Namen tragen und es so zu Verwechslungen oder gar Irrtümern, bezüglich der Anzahl entsprechender Enzyme geben kann.

Für diese Arbeit wurden die entsprechenden Sequenzen aus Tabelle 25 verwendet und für das anfänglich erstellte Hidden Markov Model genutzt. Dieses wurden dann im Verlauf der Arbeit mit den Sequenzen der neu identifizierten, aktiven Glycosyltransferasen verfeinert (Tabelle 4). Als Datenbanken wurden die UniProt-Database sowie die nicht-redundante Datenbank des NCBI genutzt. Diese lieferten

insgesamt über 100 Einträge, die mittels des sequenzbasierten Screenings und dem HMM als potenzielle Einträge für Glycosyltransferasen kodierenden Genen betrachtet werden konnten. Davon konnten 66 Treffer dem Phylum Bacteroidota zugeordnet werden. Die Gegenüberstellung der Anzahl an gefundenen Treffer aus dem Phylum Bacteroidota verglichen mit den als beschrieben und charakterisiert eingestuft Einträgen aus der CAZY-Datenbank, zeigt, wie groß das Spektrum an noch zu erforschenden Glycosyltransferasen alleine aus diesem Phylum ist.

Auffällig bei der Auswertung der durch das HMM gefunden Treffer war, dass fast alle Sequenzen die gleiche Länge an Basenpaaren aufwiesen und die potenziell kodierte Proteine die gleiche Größe (kDa) besitzen. Hier lässt sich deutlich erkennen, dass die als Grundlage des Hidden Markov Model gewählten Sequenzen eindeutig die resultierenden Ergebnisse beeinflussen. Denn die Ausgangssequenzen (Tabelle 25) weisen ebenfalls eine Durchschnittslänge von 1335 bp sowie eine Größe von 50,2 kDa auf. Auch ist der Zusammenhang zwischen Input-Sequenzen und Output-Sequenzen eindeutig in den analysierten sekundären Proteinstrukturen der neuartigen Glycosyltransferasen zu erkennen.

4.2 *In silico* Analysen

Die *in silico* Analysen der sechs neu beschriebenen Glycosyltransferasen in dieser Arbeit sollten zum besseren Verständnis der einzelnen Enzyme beitragen und deren Bedeutung im Kontext mit dem Phylum Bacteroidota aufzeigen.

4.2.1 Strukturanalyse neuartiger Glycosyltransferasen

Die Analyse der Faltungstypen der neu beschriebenen Glycosyltransferasen ergab, dass alle sechs den Faltungstyp- B aufweisen. Dies ist unter anderem damit zu erklären, dass die für die Erstellung des Hidden Markov Model eingesetzten Sequenzen ebenfalls von GTs mit Faltungstyp- B stammen. Das Hidden Markov Model fungierte bei der Datenbanksuche als eine Art „Schablone“ und lieferte dadurch potenzielle Glycosyltransferasen die gleiche oder ähnliche Muster wie die

Ausgangssequenzen auswiesen. Wodurch mit der Auswahl der Ausgangsequenzen bereits teilweise vorgegeben war, was für Glycosyltransferasen als Treffer erhalten werden würden. Dieser Zusammenhang zwischen den Ausgangssequenzen und den erhaltenen Treffern in Datenbankabgleichen ließe sich für gezielte Suchen nach gewünschten Eigenschaften, wie zum Beispiel Glycosyltransferasen, die eine Glykosylierung über ein Kohlenstoffatom katalysieren (C-Glycosyltransferasen), oder Strukturmerkmalen aufweisen, wie eine GT-A oder GT-C Faltung, nutzen.

Einhergehend mit der Untersuchung des Faltungstypes der ganzen Glycosyltransferase erfolgte auch eine Analyse auf Aminosäureebene, welche mit den unter 3.1.1 identifizierten konservierten Motiven verglichen und in Zusammenhang gebracht wurden. Hierbei konnte gezeigt werden, dass das in Motiv B (aus Abbildung 9) markierte Histidin eine wichtige Rolle bei der Bindung der Aglykone übernimmt und als Brönsted-Base agiert. Dabei aktiviert dieses die Hydroxygruppe des Polyphenols, an der die Glykosylierung erfolgt, und bereitet dieses für den nukleophilen Angriff des Donors vor (Offen et al. 2006). An der Bindung des Donormoleküls und somit der Bildung der entsprechenden Bindungstasche sind die in Motiv A und D (Abbildung 9) gekennzeichneten Glycin-Reste mitbeteiligt (Modolo et al. 2009). Daneben wirkt auch Threonin, welches in Motiv C aus Abbildung 9 hervorgehoben ist, mit (Abbildung 18). Diesem fällt dabei eine Schlüsselrolle bei der Erkennung der Zuckerkomponente des Donorsubstrates zu (Offen et al. 2006). Neben dieser Funktion reagiert das Threonin wie bei Liu et al. (Liu et al. 2020) beschrieben mit dem Phosphat der UDP-Glucose. Grundlage dafür ist, dass Threonin das aufgrund seiner Hydroxygruppe in vielen Proteinen an Phosphorylierungen beteiligt ist (Seok 2021).

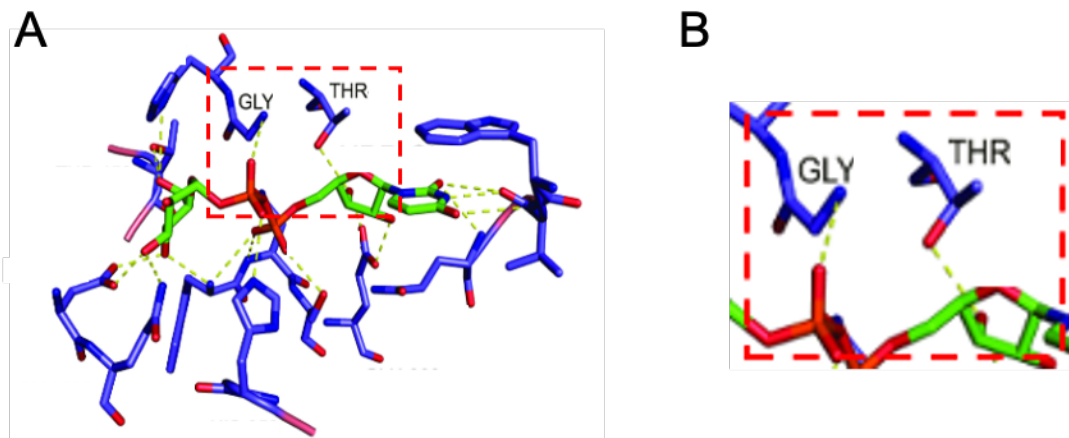


Abbildung 18: Darstellung der Interaktion von UGT7o8C1 mit UDP-Glucose.

A: Interaktion des aktiven Zentrums von UGT7o8C1 mit UDP-Glucose nach Liu et al. (Liu et al. 2020) modifiziert zur Hervorhebung (roter Rahmen) des Einflusses von Glycin und Threonin auf die Nukleotidzuckerbindung.

B: Darstellung der Interaktion von Threonin und Glycin mit dem entsprechenden Donormolekül. Deutlich wird, wie Threonin mit der Phosphatgruppe des Donors interagiert und Glycin zur Bindung des Moleküls an das aktive Zentrum beiträgt. Darstellung modifiziert nach Liu et al. (Liu et al. 2020).

Auffällig bei der weiteren Betrachtung der einzelnen Glycosyltransferasen war, dass bei GT-HHo4, wie in Abbildung 11 hervorgehoben, ein Glycin gegen ein Leucin ausgetauscht ist. Da Glycin bei der Bindung des Donormoleküls eine wichtige Rolle übernimmt, ist das eine mögliche Erklärung für das verminderte Substratspektrum von GT-HHo4. Bei GT-HHo6 dagegen ist ein Phenylalanin statt eines Tyrosins verbaut. Beide Aminosäuren unterscheiden sich in ihrer Polarität. Bei Phenylalanin handelt es sich um eine unpolare (hydrophobe) und bei Tyrosin um eine polare Aminosäure (Heinrich et al. 2014). Dieser Unterschied besteht, trotz ihrer fast identischen Struktur, aufgrund der beim Tyrosin vorhandenen Hydroxygruppe am Benzolring (Abbildung 19).

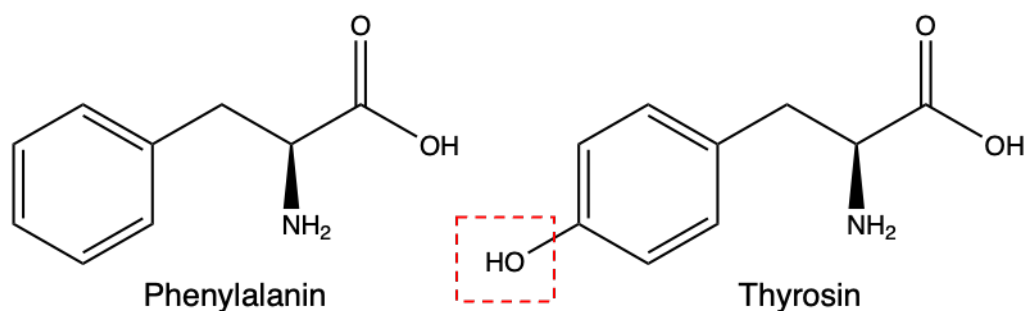


Abbildung 19: Darstellung von Phenylalanin und Thyrosin, sowie Hervorhebung (roter Rahmen) des Unterschiedes zwischen beiden Aminosäuren zur Verdeutlichung ihres Effektes auf das aktive Zentrum bei GT-HHo6. Erstellt mit ChemDraw 21.0.0 (ChemDraw 21.0.0)

Dies führt dazu, dass es bei GT-HHo6 zu einer Änderung der Polarität an der Außenseite des Proteins kommt. Hier würden, wenn kein Austausch der Aminosäure statt fände, eine hydrophile Aminosäure vorhanden sein, die dann mit dem meistens umgebenden wässrigen Milieu interagiert. Da es durch den Austausch zu hydrophoben Wechselwirkungen kommen kann, ist dies eine potenzielle Erklärung für die verminderte Aktivität. Infolgedessen können die zur Glykosylierung benötigten Aglykone bzw. Nukleotidzucker nicht mehr oder nur teilweise am aktiven Zentrum binden.

4.3 Charakterisierung von Glycosyltransferasen die durch sequenzbasierte Datenbanksuche identifiziert wurden

Die aus der Datenbanksuche resultierenden Treffer wurde anschließend weiter gefiltert. So konnten 66 Einträge dem Phylum Bacteroidota zugeordnet werden, wodurch sich der Pool an auszuwählenden Einträgen weiter reduziert. Anschließend wurden die 66 Einträge weiter gefiltert und 11 vielversprechende Sequenzen von potenziellen Glycosyltransferasen für die Charakterisierung ausgesucht. Daraus resultieren die neuartigen Glycosyltransferasen, namentlich GT-HHo1, GT-HHo2, GT-HHo3, GT-HHo4, GT-HHo5 und GT-HHo6, auf denen im Verlauf dieser Arbeit der Fokus lag. Aufgeführt sind diese in der Tabelle 4.

4.3.1 Aktivitätsbestimmung potenziell neuartiger Glycosyltransferasen

Der anfängliche Fokus bei der Charakterisierung der potenziellen Glycosyltransferasen lag auf der Bestimmung ihrer Aktivität. Diese war Voraussetzung für weiterführende Analysen.

Für die Aktivitätsbestimmung wurde die bereits bei Rabausch et al. (Rabausch et al. 2013) beschriebene Methode aus ganzzelliger Biokonversion und anschließender Dünnschichtchromatographie genutzt. Diese basiert auf der Grundlage, dass der verwendete Bakterienstamm, in dieser Arbeit *E. coli* BL21 DE3, selbst Nucleotidzucker produziert und diese von den untersuchten Enzyme für die Glykosylierung genutzt werden können. Die Produktion der notwendigen Nucleotidzucker erfolgt durch den Maltodextrin-Stoffwechselweg, den *E. coli* primär zur Gewinnung von Glukose-1-phosphat nutzt (Boos and Shuman 1998). Dieses spielt eine wichtige Rolle in weiteren Stoffwechselwegen wie dem Glykogenstoffwechsel, bei dem die benötigten Nukleotidzucker entstehen, und auch beim Aufbau von extrazellulären Polysacchariden. Weiterhin ist die Aufnahme des Aglykon in die Zelle und die Ausschleusung der Glykoside nach erfolgreicher Glykosylierung ein enormer Vorteil bei der Aktivitätsbestimmung. Dies vereinfacht die Probennahme für die anschließende Dünnschichtchromatographie, da kein aufwendiger Zellaufschluss nötig ist. Die zu untersuchenden Proben werden mit Ethylacetat versetzt und anschließend zentrifugiert. Dies hat zur Folge, dass es zur Phasentrennung zwischen dem wässrigen Medium und dem organischen Lösungsmittel kommt. Dabei gehen die Glykoside aufgrund ihrer schlechteren Wasserlöslichkeit aus der wässrigen in die organische Phase über. Hierdurch werden die entstandenen Glykoside aufkonzentriert und bereits kleinste Mengen sind durch die Dünnschichtchromatographie detektierbar (Lin et al. 2019). Eingesetzt wurde bei der Aktivitätsbestimmung das in Kapitel „3.3 Glycosyltransferasen Charakterisierung“ vorgestellte Modells substrat Quercetin und dessen mögliche Glykoside Quercitrin, Isoquercitrin sowie Rutin als Referenzen. Die Auswertung der Dünnschichtchromatographieplatten basierte auf dem Vergleich der Proben und den jeweils eingesetzten Referenzen. Dafür werden die Banden der Proben mit denen der Referenzen hinsichtlich Position auf der Platte (gleiche Höhe), UV-Spektrum (diese wurden mit dem TLC-Scanner 3 gemessen) und Färbung verglichen (siehe Abbildung 13). Hierbei ist darauf zu achten, dass die Auftrennung der einzelnen Proben gleichmäßig erfolgt und nicht die randständigen Proben der Platte durch den Kapillareffekt in der Chromatographiekammer schneller und weiter aufgetrennt

werden. Dieses erschwert im weiteren Verlauf die Auswertung, da die Vergleichbarkeit zwischen Proben und Referenzen nicht mehr gegeben ist. Auch ist eine sehr sorgsame Färbung der Platte Voraussetzung, um die Färbungen der einzelnen Banden vergleichen zu können, da es sonst zu Überlagerungen kommen kann, die zu falsch positiven oder falsch negativen Ergebnissen führen. Auf der in Abbildung 13 gezeigten Platte sind eindeutige Banden in den Spalten 5, 6 und 10 zu erkennen, die sowohl in Position und Färbung mit der Referenz Rutin übereinstimmen. Daneben sind ebenfalls in den Spalten 5 sowie 7 und 8 Banden mit gleicher Auftrennung und Färbung wie die Referenz Quercitrin zu erkennen. Dem gegenüber ist einzig in der Probe in Spalte 9 eine Bande für das Glykosid Isoquercetrin zu erkennen. Aus diesen Ergebnissen lässt sich schlussfolgern, dass es zur Glykosylierung des Aglykons (Quercetin) kam und die jeweilig untersuchten Glycosyltransferasen aktiv sind. Als weitere Bestätigung der Ergebnisse könnten die entstanden Glykoside aufgereinigt werden und anschließend mit Kernspinresonanzspektroskopie untersucht werden (Santos et al. 2021).

4.3.2 Substratspektrumsanalyse von Glycosyltransferasen

Zur Bestimmung der Substratspektren der neu identifizierten Glycosyltransferasen wurden neben dem Modells substrat für den Aktivitätsnachweis (Quercetin) andere Polyphenole genutzt, die sich in ihrer Struktur immer stärker vom Ausgangssubstrat (Quercetin) unterschieden. Angefangen wurde mit Naringenin, welches sich in seiner Struktur nur durch das Fehlen einer Hydroxygruppe am B-Ring von Quercetin unterscheidet. Dies konnte noch von allen sechs neuen Glycosyltransferasen glykosyliert werden. Das wurde auch erwartet, da beide Substrate sich nur unwesentlich unterscheiden und weiterhin eine Hydroxygruppe am B-Ring vorhanden ist, die zur Bindung des Substrates an das aktive Zentrum des Enzyms notwendig ist. In Abbildung 20 ist exemplarisch die Bindung von Quercetin an das aktive Zentrum vom GT-HH01 dargestellt. Hier wird deutlich, dass diese funktionelle Gruppe zur Anbindung des entsprechenden Substrates an das aktive Zentrum mitentscheidend ist. Weitere untersuchte Substrate im Rahmen der Substratspektrumsanalyse der neuen GTs waren Hesperetin, Kaempferol und Oxyresveratrol. Diese drei Substanzen konnten von den Glycosyltransferasen GT-HH01, GT-HH02 und GT-HH03 glykosyliert werden, wohingegen die GTs GT-HH04, GT-HH05 und GT-HH06 dazu nicht in der Lage waren. Ein Ansatzpunkt zur Erklärung dafür ist die für jedes Enzym

ganz eigne Akzeptorspezifität, welche auf die jeweilige Topologie der Akzeptorbindungsstelle zurückgeht (Wagner and Pesnot 2010). Diese spezifische Bindungsstelle wird bei den Glycosyltransferasen GT-HH04, 05 und 06 eventuell durch die beim Hesperetin vorkommende Methoxygruppe, welche größer ist als die beim Quercetin und Naringenin vorkommende Hydroxygruppe, blockiert, wodurch es zu keiner Glykosylierung kommen kann. Weiterhin kann auch der bei GT-HH04 gefundene Strukturunterschied an der Donorbindungsstelle, welcher in Abbildung 11 gezeigt wird, ein Grund dafür sein, dass diese Glycosyltransferase nicht in der Lage ist, alle getesteten Substrate zu glykosylieren. So bindet zwar das Akzeptormolekül (z.B. Hesperetin), aber ein entsprechendes Donormolekül kann aufgrund der andersartigen Donorbindungsstelle im Vergleich zu GT-HH01, welches alle getesteten Substrate verzuckern kann, nicht an das aktive Zentrum binden, wodurch es zu keiner Glykosylierung kommen kann.

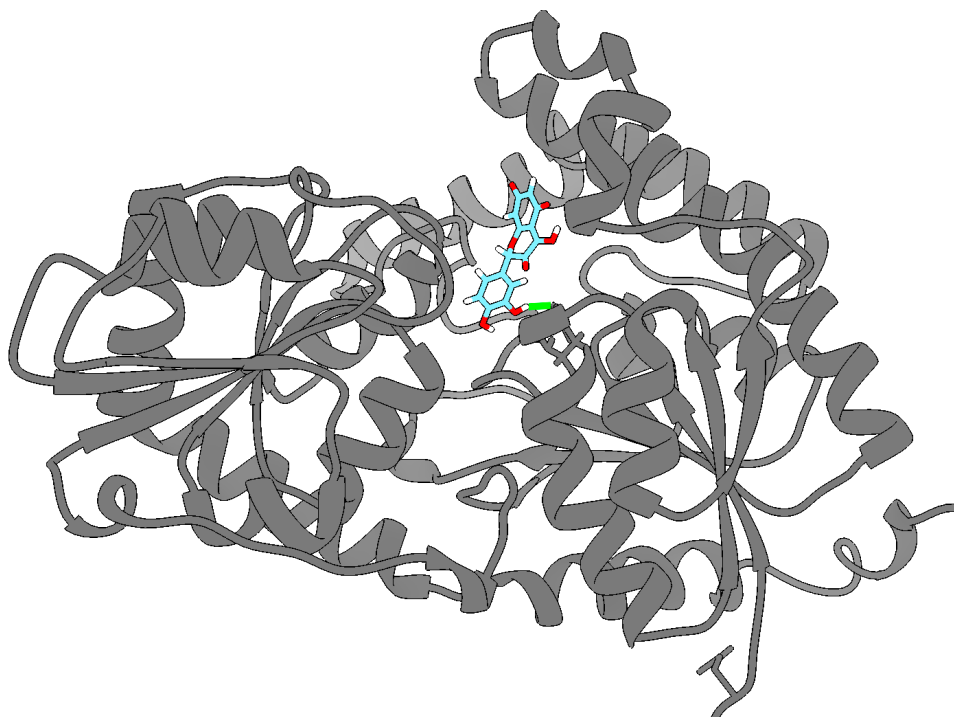


Abbildung 20: Dargestellt ist die Glycosyltransferase GT-HH01 mit Quercetin, welches an das aktive Zentrum bindet. Mit grün ist die Bindung der 3'-Hydroxygruppe des B-Ringes von Quercetin an das Enzym hervorgehoben. Ähnlich bindet auch Naringenin an die neuen GTs. Der Unterschied dabei ist, dass die Bindung von der 4'-Hydroxygruppe erfolgt. Die Struktur der Glycosyltransferase, sowie die Bindung von Quercetin an das Enzym wurden mittels Colabfold (Mirdita et al. 2022) berechnet und mit UCSF Chimera 1.16 (Pettersen et al. 2004) dargestellt.

4.4 Schlussfolgerungen und Ausblick

Durch die Identifizierung und Charakterisierung von sechs bisher unbekannter Glycosyltransferasen, GT-HH01 bis GT-HH06, gelang es, die Anzahl der beschriebenen und als charakterisiert eingestuften GTs aus dem Phylum Bacteroidota nahezu zu verdoppeln. Bisher waren lediglich sieben weitere als charakterisiert eingestufte Glycosyltransferasen aus diesem Phylum bekannt. Grundlage dafür war ein auf Proteinsequenzen bereits nachgewiesener aktiver Glycosyltransferasen erstelltes Hidden Markov Model. Dieses lieferte bei Datenbankabgleichen über 100 Treffer für potenziell aktive Enzyme. Bei rund 54 % der ausgewählten und anschließend untersuchten Enzyme konnte enzymatische Aktivität nachgewiesen werden. Im weiteren Verlauf der Arbeit wurde das Hidden Markov Model mittels der Sequenzen der nachweislich aktiven neuen GTs verfeinert. Auffällig bei weiteren Analysen der neuen Glycosyltransferasen war, dass alle mittels HMM gefunden Enzyme den Faltungstyp B aufweisen. Erklärbar ist dies dadurch, dass auch alle GTs, deren Sequenzen für die Erstellung des Hidden Markov Modelles genutzt wurden, ebenfalls diese für Glycosyltransferasen typische Faltung aufweisen. Dieser nachweisliche Zusammenhang zwischen den Eigenschaften der Ausgangsenzyme und den der neu gefundenen, zeigt eindeutig die Möglichkeit auf, gezielt das Hidden Markov Model zu nutzen, um nach vorher ausgewählten Eigenschaften von Enzymen in Datenbanken zu suchen. Dabei ist zu beachten, dass für die gesuchten Eigenschaften bereits bekannte Beispiele vorhanden sein müssen. Dadurch ergibt sich die Möglichkeit, dass Enzyme mit interessanten Eigenschaften bei der Datenbanksuche durch das Raster des Hidden Markov Models fallen, da sich deren Sequenzen von denen der Ausgangssequenzen zu stark unterscheiden. Erstellt man wiederum ein Hidden Markov Model auf der Basis eines breiten Spektrums an Sequenzen, kann dies dazu führen, dass man zwar viele Treffer bei der Datenbanksuche bekommt, aber die Erfolgsquote gewünschter Enzyme zu finden sinkt.

Der Nachweis über die glykosidische Aktivität der ausgewählten Enzyme aus der Datenbanksuche erfolgte mittels Dünnschichtchromatographie. Mittels dieser im Vergleich zur Hochleistungsflüssigkeitschromatographie deutlich sensitiveren Methode konnte Aktivität bei den Glycosyltransferasen GT-HH01, GT-HH02, GT-HH03, GT-HH04, GT-HH05 und GT-HH06 nachgewiesen werden. Bei der Wahl des

Modelsubstrates für den Aktivitätsnachweis fiel die Entscheidung auf das Polyphenol Quercetin. Ausgehend von diesem wurden für die Substratspektrumsanalyse der einzelnen GTs Substrate gewählt, die sich zunehmend von der Struktur des Quercetins unterschieden. Dabei zeigten die Enzyme GT-HH01, GT-HH02 und GT-HH03 das breiteste Spektrum an glykosidischer Aktivität und konnten alle im Rahmen dieser Arbeit getesteten Substrate glykosylieren. Zukünftige Untersuchungen könnten hier anschließen und noch ein weiteres Feld an Substraten testen, um einen genaueren Einblick über die Substratpromiskuität der einzelnen Enzyme zu erlangen. Von Interesse könnten dabei weitere polyphenolische Verbindungen wie beispielsweise Ellagsäure oder Fisetin sein, deren Glykoside eine nachweislich pharmakologische Wirkung haben (Fontaine et al. 2017; He et al. 2022). Ebenfalls beachtenswert wäre die Untersuchung von Substanzen, die nicht zu den Polyphenolen zählen. Dabei wären Substrate von Interesse, wie etwa Abbauprodukte von Recyclingprozessen, die anschließend mittels Glycosyltransferasen wieder aufbereiten werden könnten.

Neben der Identifizierung und Charakterisierung der einzelnen Glycosyltransferasen spielte auch deren gesamtheitliche Einordnung im Kontext des Phylum Bacteroidota eine wichtige Rolle in dieser Arbeit. Alle sechs neu identifizierten und nachweislich aktiven Glycosyltransferasen stammen aus Vertretern des Phylums Bacteroidota. Neben der Betrachtung der phylogenetischen Zugehörigkeit der Bakterien aus den die in dieser Arbeit neu identifizierten Glycosyltransferasen stammen wurde erstmalig in alle charakterisierten Glycosyltransferasen aus der CAZY-Datenbank in einem Interaktionsnetzwerk dargestellt. Dabei ließen sich drei große Cluster identifizieren, in denen fast alle charakterisierten Glycosyltransferasen einsortiert werden konnten. Neben 2 Clustern, die im Zusammenhang mit dem Zellmembranaufbau, sowie der Zellmembranmodifizierung stehen, konnte auch ein dritter Cluster festgestellt werden, der hauptsächlich GTs beinhaltet, die im Kontext mit Sekundärmetaboliten stehen. In diesem finden sich auch die in dieser Arbeit neu identifizierten Glycosyltransferasen GT-HH01 bis GT-HH06 wieder. Damit sind diese Enzyme die einzigen aus dem Phylum Bacteroidota, die sich in diesem Cluster einordnen lassen. Alle anderen zum Zeitpunkt der Arbeit als charakterisiert hinterlegten GTs in der CAZY-Datenbank ließen sich in den 2 anderen identifizierten Clustern einsortieren. Dies zeigt das enorme Potenzial, dass das Phylum Bacteroidota birgt, wenn es um die Erschließung neuartiger, für die Biotechnologie nutzbarer, Enzyme geht.

Literaturverzeichnis

- Ali MY, Liaqat F, Khazi MI, Sethupathy S, Zhu D (2023) Utilization of glycosyltransferases as a seamless tool for synthesis and modification of the oligosaccharides-A review. *Int J Biol Macromol* 249:125916. <https://doi.org/10.1016/j.ijbiomac.2023.125916>
- Bond SR, Naus CC (2012) RF-Cloning.org: an online tool for the design of restriction-free cloning projects. *Nucleic Acids Res* 40:W209–W213. <https://doi.org/10.1093/nar/gks396>
- Boos W, Shuman H (1998) Maltose/Maltodextrin System of Escherichia coli: Transport, Metabolism, and Regulation. *Microbiol Mol Biol Rev* 62:204–229. <https://doi.org/10.1128/MMBR.62.1.204-229.1998>
- Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Madden TL, Matten WT, McGinnis SD, Merezuk Y, Raytselis Y, Sayers EW, Tao T, Ye J, Zaretskaya I (2013) BLAST: a more efficient report with usability improvements. *Nucleic Acids Res* 41:W29–W33. <https://doi.org/10.1093/nar/gkt282>
- Braune A, Blaut M (2016) Bacterial species involved in the conversion of dietary flavonoids in the human gut. *Gut Microbes* 7:216–234. <https://doi.org/10.1080/19490976.2016.1158395>
- Brooks SA (2006) Protein glycosylation in diverse cell systems: implications for modification and analysis of recombinant proteins. *Expert Rev Proteomics* 3:345–359. <https://doi.org/10.1586/14789450.3.3.345>
- Chung MJ, Kang A-Y, Lee KM, Oh E, Jun H-J, Kim S-Y, Auh JH, Moon T-W, Lee S-J, Park K-H (2006) Water-Soluble Genistin Glycoside Isoflavones Up-Regulate Antioxidant Metallothionein Expression and Scavenge Free Radicals. *J Agric Food Chem* 54:3819–3826. <https://doi.org/10.1021/jf060510y>
- Danso D, Schmeisser C, Chow J, Zimmermann W, Wei R, Leggewie C, Li X, Hazen T, Streit WR (2018) New Insights into the Function and Global Distribution of Polyethylene Terephthalate (PET)-Degrading Bacteria and Enzymes in Marine and Terrestrial Metagenomes. *Appl Environ Microbiol* 84:e02773-17. <https://doi.org/10.1128/AEM.02773-17>
- Feng J, Qian Y, Zhou Z, Ertmer S, Vivas EI, Lan F, Hamilton JJ, Rey FE, Anantharaman K, Venturelli OS (2022) Polysaccharide utilization loci in Bacteroides determine population fitness and community-level interactions. *Cell Host Microbe* 30:200-215.e12. <https://doi.org/10.1016/j.chom.2021.12.006>
- Fernández-Gómez B, Richter M, Schüler M, Pinhassi J, Acinas SG, González JM, Pedrós-Alió C (2013) Ecology of marine Bacteroidetes: a comparative genomics approach. *ISME J* 7:1026–1037. <https://doi.org/10.1038/ismej.2012.169>
- Fontaine BM, Nelson K, Lyles JT, Jariwala PB, García-Rodríguez JM, Quave CL, Weinert EE (2017) Identification of Ellagic Acid Rhamnoside as a Bioactive

- Component of a Complex Botanical Extract with Anti-biofilm Activity. *Front Microbiol* 8. <https://doi.org/10.3389/fmicb.2017.00496>
- Gharabli H (2023) The function of UDP-glycosyltransferases in plants and their possible use in crop protection. *Biotechnol Adv*
- Grewal AK (2021) Mechanistic insights and perspectives involved in neuroprotective action of quercetin
- Güllert S, Fischer MA, Turaev D, Noebauer B, Ilmberger N, Wemheuer B, Alawi M, Rattei T, Daniel R, Schmitz RA, Grundhoff A, Streit WR (2016) Deep metagenome and metatranscriptome analyses of microbial communities affiliated with an industrial biogas fermenter, a cow rumen, and elephant feces reveal major differences in carbohydrate hydrolysis strategies. *Biotechnol Biofuels* 9:121. <https://doi.org/10.1186/s13068-016-0534-x>
- Hahnke RL, Meier-Kolthoff JP, García-López M, Mukherjee S, Huntemann M, Ivanova NN, Woyke T, Kyrpides NC, Klenk H-P, Göker M (2016) Genome-Based Taxonomic Classification of Bacteroidetes. *Front Microbiol* 7:2003. <https://doi.org/10.3389/fmicb.2016.02003>
- He B, Bai X, Tan Y, Xie W, Feng Y, Yang G-Y (2022) Glycosyltransferases: Mining, engineering and applications in biosynthesis of glycosylated plant natural products. *Synth Syst Biotechnol*
- Heinrich PC, Müller M, Graeve L, Löffler G, Petrides PE (eds) (2014) Löffler/Petrides Biochemie und Pathobiochemie, 9., vollständig überarbeitete Auflage. Springer, Berlin Heidelberg
- Koropatkin NM, Martens EC, Gordon JI, Smith TJ (2008) Starch Catabolism by a Prominent Human Gut Symbiont Is Directed by the Recognition of Amylose Helices. *Structure* 16:1105–1115. <https://doi.org/10.1016/j.str.2008.03.017>
- Lairson LL, Henrissat B, Davies GJ, Withers SG (2008) Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu Rev Biochem* 77:521–555. <https://doi.org/10.1146/annurev.biochem.76.061005.092322>
- Lee SS, Hong SY, Errey JC, Izumi A, Davies GJ, Davis BG (2011) Mechanistic evidence for a front-side, S_Ni-type reaction in a retaining glycosyltransferase. *Nat Chem Biol* 7:631–638. <https://doi.org/10.1038/nchembio.628>
- Liang D-M, Liu J-H, Wu H, Wang B-B, Zhu H-J, Qiao J-J (2015) Glycosyltransferases: mechanisms and applications in natural product development. *Chem Soc Rev* 44:8350–8374. <https://doi.org/10.1039/C5CS00600G>
- Lin L, Zhang Y-Y, Chen R, Yang L, Ni H, Li H-H (2019) Highly simplified preparation of tea flavonoids from surplus tea leaves by the novel three-phase extraction and purification. *Sep Sci Technol* 54:741–746. <https://doi.org/10.1080/01496395.2018.1515227>
- Liu M, Wang D, Li Y, Li X, Zong G, Fei S, Yang X, Lin J, Wang X, Shen Y (2020) Crystal Structures of the C -Glycosyltransferase UGT708C1 from Buckwheat Provide

- Insights into the Mechanism of C -Glycosylation. *Plant Cell* 32:2917–2931. <https://doi.org/10.1105/tpc.20.00002>
- Lu X, Huang L, Scheller HV, Keasling JD Medicinal terpenoid UDP-glycosyltransferases in plants: recent advances and research strategies
- Mackenzie AK, Pope PB, Pedersen HL, Gupta R, Morrison M, Willats WGT, Eijssink VGH (2012) Two SusD-Like Proteins Encoded within a Polysaccharide Utilization Locus of an Uncultured Ruminant Bacteroidetes Phylotype Bind Strongly to Cellulose. *Appl Environ Microbiol* 78:5935–5937. <https://doi.org/10.1128/AEM.01164-12>
- Mäkilä L, Laaksonen O, Alanne A-L, Kortenesniemi M, Kallio H, Yang B (2016) Stability of Hydroxycinnamic Acid Derivatives, Flavonol Glycosides, and Anthocyanins in Black Currant Juice. *J Agric Food Chem* 64:4584–4598. <https://doi.org/10.1021/acs.jafc.6b01005>
- Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Geer LY, Bryant SH (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 45:D200–D203. <https://doi.org/10.1093/nar/gkw1129>
- Mestrom, Przypis, Kowalczykiewicz, Pollender, Kumpf, Marsden, Bento, Jarzębski, Szymańska, Chruściel, Tischler, Schoevaart, Hanefeld, Hagedoorn (2019) Leloir Glycosyltransferases in Applied Biocatalysis: A Multidisciplinary Approach. *Int J Mol Sci* 20:5263. <https://doi.org/10.3390/ijms20215263>
- Mikkola S (2020) Nucleotide Sugars in Chemistry and Biology. *Molecules* 25:5755. <https://doi.org/10.3390/molecules25235755>
- Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M (2022) ColabFold: making protein folding accessible to all. *Nat Methods* 19:679–682. <https://doi.org/10.1038/s41592-022-01488-1>
- Mistry J, Finn RD, Eddy SR, Bateman A, Punta M (2013) Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res* 41
- Modolo LV, Li L, Pan H, Blount JW, Dixon RA, Wang X (2009) Crystal Structures of Glycosyltransferase UGT78G1 Reveal the Molecular Basis for Glycosylation and Deglycosylation of (Iso)flavonoids. *J Mol Biol* 392:1292–1302. <https://doi.org/10.1016/j.jmb.2009.08.017>
- Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H (1986) Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction. *Cold Spring Harb Symp Quant Biol* 51:263–273. <https://doi.org/10.1101/SQB.1986.051.01.032>
- Notredame C, Higgins DG, Heringa J (2000) T-coffee: a novel method for fast and accurate multiple sequence alignment 1 Edited by J. Thornton. *J Mol Biol* 302:205–217. <https://doi.org/10.1006/jmbi.2000.4042>

- Offen W, Martinez-Fleites C, Yang M, Kiat-Lim E, Davis BG, Tarling CA, Ford CM, Bowles DJ, Davies GJ (2006) Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *EMBO J* 25:1396–1405. <https://doi.org/10.1038/sj.emboj.7600970>
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera?A visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612. <https://doi.org/10.1002/jcc.20084>
- Rabausch U, Juergensen J, Ilmberger N, Böhnke S, Schubach B, Schulte M, Streit WR (2013) Functional Screening of Metagenome and Genome Libraries for Detection of Novel Flavonoid-Modifying Enzymes. *Appl Environ Microbiol* 79:4551–4563. <https://doi.org/10.1128/AEM.01077-13>
- Ruprecht C, Bönisch F, Ilmberger N, Heyer TV, Haupt ETK, Streit WR, Rabausch U (2019) High level production of flavonoid rhamnosides by metagenome-derived Glycosyltransferase C in *Escherichia coli* utilizing dextrins of starch as a single carbon source. *Metab Eng* 55:212–219. <https://doi.org/10.1016/j.ymben.2019.07.002>
- Sambrook J, Russell DW (2001) *Molecular cloning: a laboratory manual*, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y
- Santos AL, Soares MG, Medeiros LS, Ferreira MJP, Sartorelli P (2021) Identification of flavonoid-3-O-glycosides from leaves of *Casearia arborea* (Salicaceae) by UHPLC-DAD-ESI-HRMS/MS combined with molecular networking and NMR. *Phytochem Anal* 32:891–898. <https://doi.org/10.1002/pca.3032>
- Seok S-H (2021) Structural Insights into Protein Regulation by Phosphorylation and Substrate Recognition of Protein Kinases/Phosphatases
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* 13:2498–2504. <https://doi.org/10.1101/gr.1239303>
- Streit WR, Daniel R (eds) (2017) *Metagenomics: methods and protocols*, Second edition. Humana Press, New York, NY
- Sun C-C, Zhao W-J, Yue W-Z, Cheng H, Sun F-L, Wang Y-T, Wu M-L, Engel A, Wang Y-S (2023) Polymeric carbohydrates utilization separates microbiomes into niches: insights into the diversity of microbial carbohydrate-active enzymes in the inner shelf of the Pearl River Estuary, China. *Front Microbiol* 14:1180321. <https://doi.org/10.3389/fmicb.2023.1180321>
- Thomas F, Hehemann J-H, Rebuffet E, Czjzek M, Michel G (2011) Environmental and Gut Bacteroidetes: The Food Connection. *Front Microbiol* 2. <https://doi.org/10.3389/fmicb.2011.00093>
- van den Ent F, Löwe J (2006) RF cloning: A restriction-free method for inserting target genes into plasmids. *J Biochem Biophys Methods* 67:67–74. <https://doi.org/10.1016/j.jbbm.2005.12.008>

- Veith PD, Gorasia DG, Reynolds EC (2023) Characterization of the *O*-Glycoproteome of *Flavobacterium johnsoniae*. *J Bacteriol* 205:e00093-23. <https://doi.org/10.1128/jb.00093-23>
- Wagner GK, Pesnot T (2010) Glycosyltransferases and their Assays. *ChemBioChem* 11:1939–1949. <https://doi.org/10.1002/cbic.201000201>
- Wagner H, Bladt S, Zgainski EM (1983) Flavonoid-Drogen. In: Wagner H, Bladt S, Zgainski EM (eds) *Drogenanalyse: Dünnschichtchromatographische Analyse von Arzneidrogen*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 163–193
- Wallace IM (2006) M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res* 34:1692–1699. <https://doi.org/10.1093/nar/gkl091>
- Wang M (2023) Structure-function and engineering of plant UDP-glycosyltransferase. *Comput Struct Biotechnol J*
- Wang Y (2024) Insights into the role of glycosyltransferase in the targeted treatment of gastric cancer
- Wheeler TJ, Clements J, Finn RD (2014) Skyalign: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinformatics* 15:7. <https://doi.org/10.1186/1471-2105-15-7>
- Woo H-J, Kang H-K, Nguyen TTH, Kim G-E, Kim Y-M, Park J-S, Kim D, Cha J, Moon Y-H, Nam S-H, Xia Y, Kimura A, Kim D (2012) Synthesis and characterization of ampelopsin glucosides using dextransucrase from *Leuconostoc mesenteroides* B-1299CB4: Glucosylation enhancing physicochemical properties. *Enzyme Microb Technol* 51:311–318. <https://doi.org/10.1016/j.enzmictec.2012.07.014>
- Yang Q (2023) Genome-wide analysis of UDP-glycosyltransferases family and identification of UGT genes involved in abiotic stress and flavonol biosynthesis in
- Zheng R, Cai R, Liu R, Liu G, Sun C (2021) *Maribellus comscasis* sp. nov., a novel deep-sea *Bacteroidetes* bacterium, possessing a prominent capability of degrading cellulose. *Environ Microbiol* 23:4561–4575. <https://doi.org/10.1111/1462-2920.15650>
- Carbonhydrate Active Enzymes database <http://www.cazy.org/>
- Dimethyl-Sulfoxide-Mediated-Primer-Tm-Reduction--A-Method-for_1993_Analytica.pdf
- ChemDraw
- SnapGene software (www.snapgene.com)”

Anhang

Protein-Sequenzen

Glycosyltransferasen

GT-HHo1 (Acces: OJY95357)

N-
MQGRLKILVACFPADGHFKPLTGIAKYLDACGHDVRWYASKEYTNDIAGLGMKHYPFRIAKDLHISYMEVHYPER
AKIKNPVRKLNFDMIHVFVERGEEYMKDIDEIRQNFPPDIVICDIAFPGIPYIKDCMKLPVIAVSVFPFGSSCDLPP
YGLGLTPANTSAGKIFHKGLRWLSRKILFRKADSAMRKALEKYSIAHNNTGIFDIGYKKS DIVLQSGCPSFDYKRSD
LDKRVRFIGAVLPYRKSNQVTSWFDKRLEGYKIVVLVTQGTVEKDP SKLIPTLEAFKDDDEDVLVVCCTGNANTEAL
RAKYNAPNIIIEDFIPFENIMPYTNV FITNGGYGGVLLGIEHKMPMVVAGVHEGKNEICARVGYFKYGINLKTERPG
AARIKKAVKEIFETPVYKANVVRLAQELNSYATDALVEKYIDELINNRSPKCRQSLLPETVY-C

GT-HHo2 (Acces: OJW59638)

N-
MADAGASLAKKKMKILFANFPGEHFNPLTGLAAHLKRQGH DVRWYTSKIYADKLRR LQIARLKFDMINAFILRG
PEYYADILEIYKTRFDLLVADCAFTGIPFVKELIKVPVSVGLFPLTETSRDLPPAGLGMTPSQSFIGRVRQSLLRGI
ARRVIFGEPHRV MYQLLDQYKIPHQRESLFDMIRKSDLLQIGTPGFEYYRSDLGENIRFVGALLPHDGAKEGPR
WFDAGAYVTNGGYGGTMLGINLGVERPTSLQVRQAVEKVFSDDVYRRKV KELALEFSKYDAYERCSRHISSVVKK
ASINSSKVTDNSYVIPFSS-C

GT-HHo3 (Acces: GAO44619)

N-
MKQSTAKKILFANVPADGHFNPLTGIAVFLKDLGYDVRWYTSAKYAGKVKKLGLQYYPYKVALEVDTDNLSELF
DREKISSKIGKLNVDVQEFLKRGPEYYTDILGIYREFPFDLMISDVMFTGNAFVKNLMNIPVMAIGIMPLAEGSVD
LPPAGLGMHPSKSFFGRRKQAALRWLTDHVLFAKSIKVMYALLDKYGIRHGKENFFDMLYRQTDLV LQSGSPSFE
YERSDISDNVRFIGPLLPYQQPAGRDTWMDHRVHHYDRVILVTQGTVETDVNKL LVPTLEAYKDSNTLVVVTTGG
GGTAKLREAYPQANIIIEDFIPNDIMPYVDVYITNGGYGGVLLGIENQLPFV VAGVHEGKNEICARIGYFRLGIDLR
TETPKPSQIKGAVEKVLADGSYRKNVQQLSREFNTYKPLQLAAGYVQELLSNPQPRFIHPEAIVY-C

GT-HHo4 (Gen-Name: Sph21_4852)

N-
MKHPLQGKRILFASVPLDGHFNPLTGLAKYLQQVDCEVRWYTSKIYSEKLKALGIPHYPFVNAEELNGQN LALKLP
EIKLAPGPEKGKLYLKNLFINRAKEYYEDIRSIYSTFTFDLLIADSMFSGIPFIRYKLNVPVIAIGVVPLIEDSIDVAPA
GRALPPAKDAKHTHTAFAEMYRQKYTSINDLIEIYKADLYQHNTVAGSFIFDTLIKEANIY LQIGVPGFQYRSDLGK
NIRYIGALLPHSASGSRSIWHD PRLGKYDAVVVLVTQGTVESDATKLEPTLSAFAGTGTLIVVTTAGHGTEQLREKY
PHENVIIEDFIPFDDIMPYTSVYVTNGGYGGTIFSIMNGIPIVAAGVHELKNEVCARIDYFGLGIDLKTETPTPDAIFN
AAMSIMRND SYKKRVIALRDEM KQYDSLALCCRYISQSLDPS-C

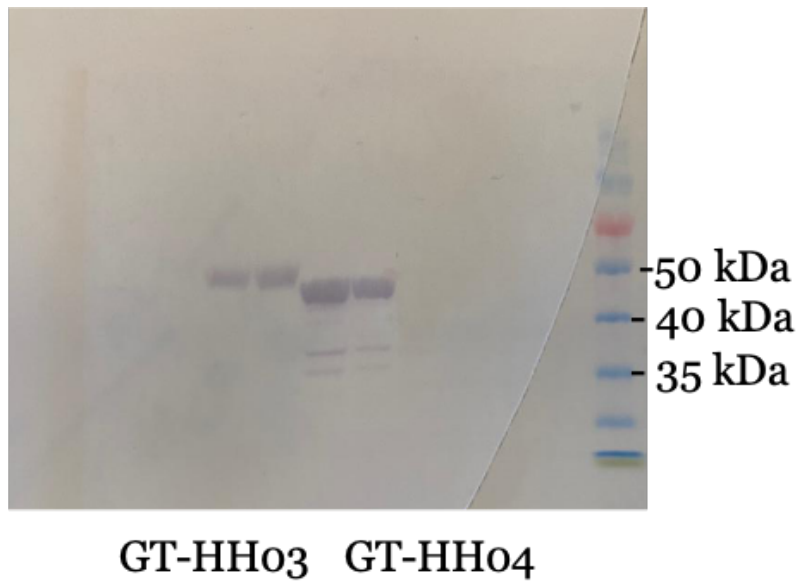
GT-HHo5 (Acces: WP_035084609)

N-
MHSIKILFACIPADGHFNPM TQLAVHLKKKGHDVRWYTGKVYEKKLEQMGIPIYIPFKKSKEIKVSELDQAYPERK
KLKGISHIKFDIINLFINRMKDYYLDI KEVSSSFSEILVFDNTFPGTIIKDKLNIPVVGIGVVPLALSSPDIPQYGLGH
QPATNFLSKRKQNFILKMAEKLLFRDTKEAYNELLENLNPKEKHNI FDISPLNSSIFLQNGIPEMDYPRYTIPKTIQ
YVGALDTSINHHNRKTKDWKKILDPSKKTILISQGTVEKNNKKLIIPSLEAFKNTNYN ILVATAHNETKLLQKKY
PHKNIHIEDFIPFDEIMPLTNIFITNGGYGSIMISIKHGVP MITAGINEGKNEICARIDY CGLGIDLKTERPRIISIQNA
ANKILEADIYYNNIKLIQKKLNSYDTLNICEKYITDLVFSKVK-C

GT-HHo6 (Acces: WP_149627120)

71

Zu 3.3.3 Nachweis der Expression



Espressionsnachweise für die Glycosyltransferasen GT-HH03 und GT-HH04

Danksagung

An dieser Stelle möchte ich meinen Dank aussprechen an all jene die zum Gelingen dieser Arbeit beigetragen haben und mir stets mit Rat und Tat zur Seite standen.

Zu allererst geht mein Dank an Prof. Dr. Wolfgang Streit für die Möglichkeit meine Dissertation in seiner Arbeitsgruppe durchzuführen. Gleichzeitig sage ich auch Danke für die unermüdliche Unterstützung sowie der entgegen gebrachten Geduld und das Vertrauen in allen Phasen dieses Projektes.

Daneben bedanke ich mich bei Frau Prof. Dr. Susanne Dobler sowie allen weiteren Mitgliedern meines Prüfungsausschusses dafür sich Zeit genommen zu haben meine Dissertation zu bewerten. Auch danke ich allen Mitarbeitern der Abteilung Mikrobiologie und Biotechnologie für das stets freundliche Arbeitsklima und den konstruktiven Austausch sowohl fachlicher als auf freundschaftlicher Art.

Einen großen Dank richte ich an meine Familie. Meinen Eltern, meiner Freundin Nina und allen Freunden für die Unterstützung und das Interesse an meiner Arbeit zu sowie jeglicher Hilfe ohne die diese Arbeit nicht möglich gewesen wäre.

Vielen Dank !!