

# Mechanism and Inhibition of Fucosyltransferases

---

Thesis

submitted in fulfillment of the requirements  
of the degree Doctor rer. nat. of the Department of Chemistry,  
Faculty of Sciences, University of Hamburg

by

Miriam P. Kötzler

from Hamburg

Hamburg, 2012



*Für meine Eltern*

*We are still confused, but on a higher level.*

Enrico Fermi

This thesis was conducted at the Institute for Organic Chemistry (managing director: Prof. Dr. Christian B.W. Stark) from June 2009 to September 2012.

I would like to thank Prof. Dr. Bernd Meyer for suggesting this highly interesting subject and for his continuous prolific, friendly and motivating support.

1<sup>st</sup> Reviewer: Prof. Dr. Bernd Meyer

2<sup>nd</sup> Reviewer: Priv.-Doz. Dr. Edzard Spillner

Date of defense: December 7<sup>th</sup> 2012

# Contents

<b>Contents.....</b>	<b>vi</b>
<b>1 Introduction.....</b>	<b>14</b>
1.1 Biological significance of fucosylated glycans .....	14
1.2 Human core $\alpha$ 1,6-fucosyltransferase (FUT8) .....	16
1.3 Honeybee core $\alpha$ 1,3-fucosyltransferase (FucTA) .....	19
<b>2 Objective .....</b>	<b>22</b>
<b>3 Introduction to the methods .....</b>	<b>23</b>
3.1 Preparation of fucosyltransferases.....	23
3.2 Preparation of complex type oligosaccharides .....	24
3.3 Analyzing kinetics and thermodynamics of binding processes .....	26
3.4 Analyzing enzyme kinetics of glycosyltransferases .....	30
3.5 Structural characterization of biomolecular complexes .....	33
3.5.1 Saturation transfer difference (STD) NMR .....	34
3.5.2 Transferred nuclear Overhauser enhancement NMR spectroscopy (trNOESY) .....	37
3.6 Molecular modeling of biomolecular complexes .....	38
3.6.1 Energy minimization techniques .....	39
3.6.2 Molecular dynamics simulation .....	40
3.6.3 Molecular docking.....	41
<b>4 Results and discussion.....</b>	<b>42</b>
4.1 Formation of the immunogenic $\alpha$ 1,3-fucose epitope: Elucidation of substrate specificity and of enzyme mechanism of core fucosyltransferase A .....	42
4.1.1 Preparation of complex type oligosaccharide for structural binding studies.....	42
4.1.2 FucTA recognizes GDP-Fuc via the nucleobase .....	44
4.1.3 FucTA recognizes a large epitope of the acceptor substrate .....	47
4.1.4 FucTA prefers unfucosylated oligosaccharides as acceptor substrate.....	50
4.1.5 Recognition of the immunogenic core $\alpha$ 1,3-fucosyl epitope by IgG .....	54
4.2 Donor substrate binding and enzymatic mechanism of human core $\alpha$ 1,6-fucosyltransferase (FUT8).....	56
4.2.1 A modified expression protocol gives enhanced yields of purified FUT8.....	56
4.2.2 SPR and STD NMR elucidate recognition of GDP-Fuc .....	58
4.2.3 Modeling of the FUT8–GDP-Fuc complex reveals enzyme–substrate contacts.....	61
4.2.4 Nucleoside and $\beta$ -phosphate direct binding of GDP-Fuc.....	69
4.2.5 Arg365 plays two key roles in catalysis .....	73
4.2.6 Function of the flexible loop .....	74

4.2.7	Comparison to $\alpha$ 1,3 fucosyltransferase of <i>H. pylori</i> .....	75
4.3	An unusually large acceptor binding site of human core $\alpha$ 1,6-fucosyltransferase is formed by donor assisted self-organization .....	77
4.3.1	The acceptor substrate binding shows moderate affinity and dissociation rate.....	77
4.3.2	STD NMR reveals the acceptor substrate epitope.....	79
4.3.3	A model of the ternary complex reveals FUT8–acceptor substrate contacts .....	80
4.3.4	FUT8 recognizes a large epitope of the acceptor substrate .....	89
4.3.5	The donor contributes significantly to acceptor binding .....	90
4.3.6	Recognition of the 3-branch, $\beta$ GlcNAc-2- $\alpha$ Man, at Man3 .....	92
4.3.7	The flexible loop adapts the 6-Man branch.....	92
4.3.8	FUT8 acts via an ordered bi–bi mechanism .....	93
4.3.9	FUT8 employs a substrate-assisted mechanism .....	94
4.4	Chitotriose derivatives as FUT8 inhibitors.....	96
4.4.1	Chitotriose is an acceptor substrate for FUT8 with low efficiency .....	97
4.4.2	Chitotriose shows a binding mode similar to the natural acceptor .....	99
4.4.3	A chitotriose derivative to specifically block Arg365.....	105
4.4.4	Towards a synthetic route with minimal protection strategy.....	106
<b>5</b>	<b>Summary .....</b>	<b>109</b>
<b>6</b>	<b>Zusammenfassung.....</b>	<b>111</b>
<b>7</b>	<b>Experimental procedures.....</b>	<b>114</b>
7.1	Recombinant expression of fucosyltransferase 8 .....	114
7.1.1	Cloning of human fucosyltransferase 8.....	114
7.1.2	Recombinant baculovirus production.....	114
7.1.3	Expression in baculovirus-infected insect cells.....	114
7.1.4	Protein purification.....	115
7.1.5	Immunoblot .....	115
7.1.6	Other methods .....	115
7.2	Preparation and characterization of complex type oligosaccharides as acceptor substrates	115
7.2.1	Preparation of asialo complex type N-glycans.....	115
7.2.2	Preparation of asialo agalacto complex type N-glycans.....	116
7.2.3	Preparation of 1- $\beta$ -N-acetyl asialo agalacto complex type N-glycans .....	116
7.2.4	Mass spectrometric characterization of oligosaccharides.....	117
7.2.5	NMR characterization of oligosaccharides.....	117
7.3	Synthesis of chitotriose derivatives .....	126
7.3.1	Deacetylation of peracetylated chitotriose (8a) .....	126

7.3.2	Synthesis of 1- $\beta$ - <i>N</i> -acetyl chitotriose 8b .....	127
7.3.3	1- $\beta$ - <i>N</i> -acetyl <i>N</i> -acetylglycosamine 19b .....	127
7.3.4	Synthesis of 1- $\beta$ - <i>N</i> -acetyl 4",6"- <i>O</i> -benzylidene chitotriose 11 .....	128
7.3.5	Synthesis of 3-vinyl benzoic acid <i>N</i> -imidazolyl amide 18 .....	129
7.3.6	Synthesis of 1- $\beta$ - <i>N</i> -acetyl (3-vinyl-)benzoyl <i>N</i> -acetylglycosamines .....	129
7.4	SPR Binding Studies .....	132
7.4.1	SPR affinity assays for ligands of FucTA .....	132
7.4.2	SPR affinity assays for ligands of FUT8 .....	132
7.5	STD NMR Experiments .....	133
7.5.1	STD NMR experiments of FucTA .....	133
7.5.2	STD NMR experiments of FUT8 .....	134
7.5.3	STD NMR experiments with anti-CCD IgG .....	136
7.6	TrNOESY experiments .....	137
7.7	Enzyme kinetic studies .....	137
7.7.1	Enzyme kinetic studies by progress curve analysis of FucTA .....	137
7.7.2	Enzyme kinetic studies by progress curve analysis of FUT8 .....	137
7.8	Molecular modeling .....	138
7.8.1	Determination of the GDP-Fuc binding site .....	139
7.8.2	Molecular dynamics simulations .....	139
<b>8</b>	<b>Hazards .....</b>	<b>153</b>
<b>9</b>	<b>References .....</b>	<b>155</b>
	<b>Acknowledgements .....</b>	<b>164</b>
	<b>Curriculum vitae .....</b>	<b>166</b>
	<b>Affidavits .....</b>	<b>168</b>

## Abbreviations

1D	1-dimensional
2D	2-dimensional
ADP	adenosine 5'-diphosphate
AP	alkaline phosphatase
BisTris	bis(2-hydroxyethyl)-amino-tris(hydroxymethyl)-methane
BSA	bovine serum albumin
CCD	cross-reactive carbohydrate determinant
CDI	1,1'-carbonyldiimidazole
<i>ce</i> POFUT1	<i>Caenorhabditis elegans</i> protein O-fucosyltransferase 1
CFG	Consortium for Functional Glycomics
COSY	correlation spectroscopy
DMF	<i>N,N</i> -dimethyl formamide
DNA	deoxyribnucleic acid
EGF	epidermal growth factor
EIC	extracted ion chromatogram
ELISA	enzyme-linked immunosorbent assay
ES	enzyme–substrate complex
es	excitation sculpting
ESI	electrospray ionization
FA	formic acid
Fc $\gamma$	fragment, crystallizable gamma
FID	free induction decay
FID	free induction decay
FRET	fluorescence resonance energy transfer
FucT	fucosyltransferase
FucTA	<i>Apis mellifera</i> fucosyltransferase A
FUT	fucosyltransferase gene
FUT8	human core $\alpha$ 1,6-fucosyltransferase
GDP	guanosine diphosphate
GDP-Fuc	6-deoxy- $\beta$ -L-galactopyranosylguanosine-5'-diphosphate

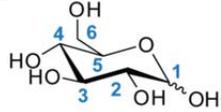
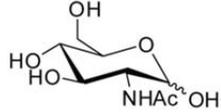
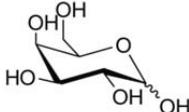
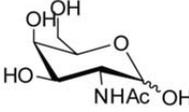
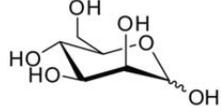
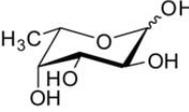
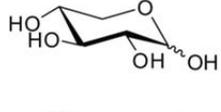
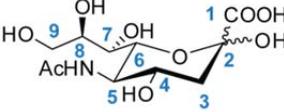
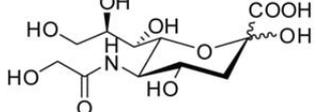
GTB	Human Blood Group Galactosyltransferase B
GT-B	glycosyltransferase B fold
HEK	human embryonic kidney
HILIC	hydrophilic interaction chromatography
HMBC	heteronuclear multiple-bond correlation
HPLC	high pressure liquid chromatography
HRP	horseradish peroxidase
HSQC	heteronuclear single-quantum correlation
IDP	inosine 5'-diphosphate
IgE	immunoglobulin E
IgG	immunoglobulin G
ITC	isothermal titration calorimetry
MALDI	matrix-assisted laser desorption/ionization
MD	molecular dynamics
Mes	2-( <i>N</i> -morpholino)ethanesulfonic acid
MS	mass spectrometry
MWCO	molecular weight cut-off
NMR	nuclear magnetic resonance
NodZ	nodulation factor Z (bacterial $\alpha$ 1,6-fucosyltransferase)
NOESY	nuclear Overhauser effect spectroscopy
OPLS	optimized potentials for liquid simulations
PBS	phosphate buffered saline
PDB	Protein Data Bank
PGC	porous graphitized carbon
PME	particle mesh Ewald
PNGase F	protein N-glycosidase F
RMSD	root-mean-square deviation
RMSF	root-mean-square fluctuation
S/N	signal-to-noise ratio
SPC	simple point charge
SPR	surface plasmon resonance

STD	saturation transfer difference
$T_1$	longitudinal relaxation time
TBDPS	<i>tert</i> -butyl diphenyl silyl
TFA	trifluoro acetic acid
TGF $\beta$	transforming growth factor beta
THF	Tetrahydrofuran
TIC	total ion current
TMSP	sodium-3-trimethylsilylpropionate
TOCSY	total correlation spectroscopy
TOF	time-of-flight
VEGF	vascular endothelial growth factor
XDP	xanthosine 5'-diphosphate

## Amino acids

<b>Amino acid</b>	<b>Abbreviation</b>	<b>Code</b>
alanine	Ala	A
arginine	Arg	R
asparagine	Asn	N
aspartate	Asp	D
cysteine	Cys	C
glutamine	Gln	Q
glutamate	Glu	E
glycine	Gly	G
histidine	His	H
isoleucine	Ile	I
leucine	Leu	L
lysine	Lys	K
methionine	Met	M
phenylalanine	Phe	F
proline	Pro	P
serine	Ser	S
threonine	Thr	T
tryptophane	Trp	W
tyrosine	Tyr	Y
valine	Val	V

## Monosaccharide Units

Name	Abbreviation	Code	CFG notation	Structure
D-glucopyranoside	Glc			
2-(acetylamino)-2-deoxy-D-glucopyranoside	GlcNAc	GN		
D-galactopyranoside	Gal			
2-(acetylamino)-2-deoxy-D-galactopyranoside	GalNAc			
D-mannopyranoside	Man	M		
L-fucopyranoside	Fuc	F		
D-xylopyranoside	Xyl	X		
D-N-acetyl neuraminic acid	NeuAc			
D-N-glycolyl neuraminic acid	NeuGc			

## 1 Introduction

### 1.1 Biological significance of fucosylated glycans

L-Fucose (6-deoxy-L-galactose) is a deoxyhexose that is present in a wide variety of organisms. It is a common component of glycolipids and of many N- and O-linked glycans carried by glycoproteins produced by eukaryotic cells. Two structural features distinguish fucose from other six-carbon sugars present in mammals: These are the lack of a hydroxyl group at C-6 and the L-configuration. Fucose frequently exists as a terminal modification of glycan structures that are synthesized by linking of a fucose residue to acceptor oligosaccharides. Fucosylation in asparagine-linked oligosaccharides (N-glycans) (cf. Figure 1-1) is known to confer unique functional properties on oligosaccharides. These specific fucosylated structures are involved in biological processes in the cell, including cell adhesion, cell differentiation and cell growth. [1-3] They were found to play a major role in clinically important events such as blood transfusion reactions, selectin-mediated leukocyte-endothelial adhesion, host-microbe interactions, and numerous ontogenic events, including signaling events by the Notch receptor family. [3] Also, prognosis and progression of tumor diseases, atherosclerosis, allergy and the activity of therapeutic antibodies have been found to depend on the presence of fucosylated structures. Presumably, many additional functions of fucosylated glycans still remain to be elucidated as biological functions of carbohydrates are challenging to study [4, 5]. This is due to their high structural diversity and the fact that their sequence and structure is not directly encoded in the genome.

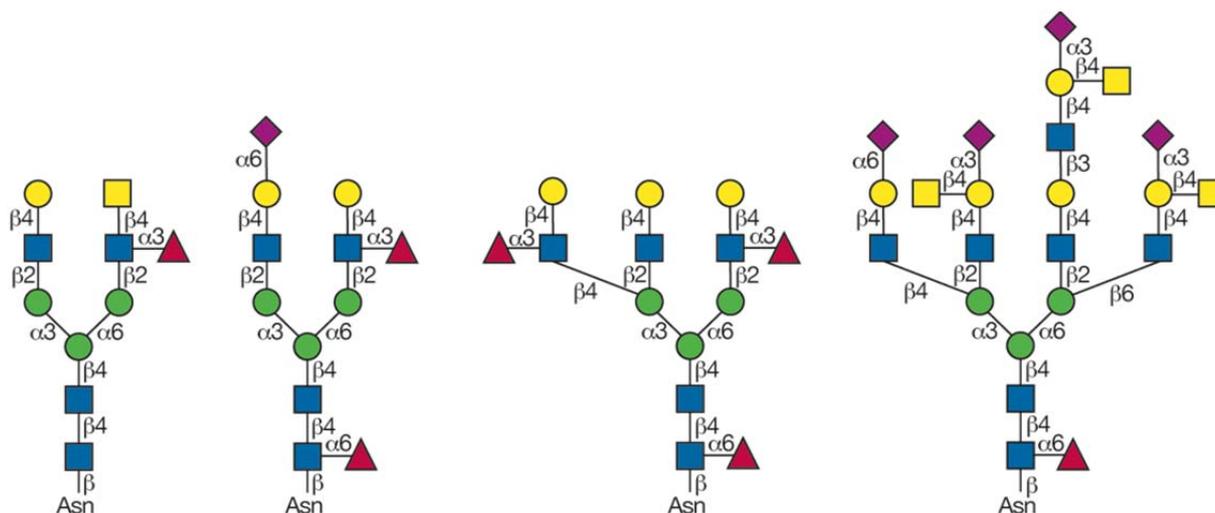


Figure 1-1: Some examples for complex-type oligosaccharides present in mammalian N-glycans illustrating their high structural diversity. The modifications include different capping structures (fucose and neuraminic acid), additional branches, poly-N-acetylglucosamine units and modifications of the core (fucose and bisecting GlcNAc).

## INTRODUCTION

In mammals, fucosylation is accomplished by the 14 members of the fucosyltransferase (FucTs or FUTs) family. [3, 6] All fucosyltransferases transfer a fucosyl residue from the activated sugar nucleotide diphosphate, guanine nucleotide diphosphate- $\beta$ -L-fucose (GDP-Fuc), to a wide array of acceptors with formation of an  $\alpha$ -linkage, i.e. under inversion of the anomeric configuration. Figure 1-2 depicts the different fucosylation sites present in N- and O-glycans. In complex type glycans, fucose is generally attached in 1,2-linkage to galactose or in 1,3-, 1,4- or 1,6-linkages to *N*-acetylglucosamine residues. [1-3]  $\alpha$ 1,2-fucosyltransferases (FUT1 and FUT2) are responsible for synthesis of the H blood group antigen and related structures. [7, 8]. The genes FUT3-FUT7 and FUT9-FUT11 encode fucosyltransferases that synthesize  $\alpha$ 1,3- and (in case of FUT3)  $\alpha$ 1,4- fucosylated glycans, such as the Lewis<sup>x</sup> antigens [9, 10]. FUT8 is an  $\alpha$ 1,6-fucosyltransferase that adds fucosyl residues to the innermost GlcNAc of N-glycans. In addition, there are two fucosyltransferases that catalyze the direct O-fucosylation of proteins, namely POFUT1 and POFUT2. The specific fucosylation pattern in tissues, developmental stages or under pathological conditions is predetermined by the expression level, the specific activity and the unique acceptor substrate specificity of each of these individual fucosyltransferases.

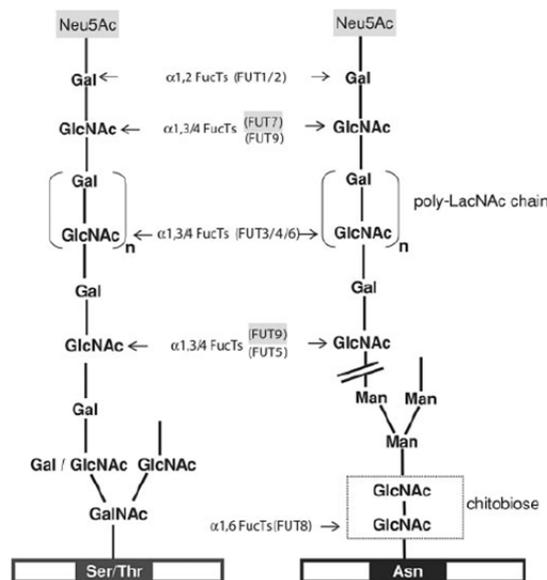


Figure 1-2: Fucosylation sites of human fucosyltransferases. Both O-glycans (left) and N-glycans (right) contain poly-LacNAc chains that are substrates to  $\alpha$ 1,2- and  $\alpha$ 1,3/4-FucTs. For these FucTs, only the preferred fucosylation sites are given. In the cases of FUT7 and FUT9 (shaded in grey), activity and site preference is dependent on the presence of NeuNAc. In contrast, the fucosylation site for  $\alpha$ 1,6-FucT (FUT8) is only found in N-glycans. (Figure reproduced from Ref. [11] with permission from Oxford University Press)

Elimination of distinct fucosylated structures by means of gene knock-out for single fucosyltransferases in mice has shed light on their biological function [12-14]. However, gene

knock-out in animals is a laborious procedure and only feasible in mice to date. More flexible tools for probing biological effects of fucosylation in life animals are therefore of high interest. However, despite the fundamental significance of the products of fucosyltransferases, specific, potent and membrane permeable inhibitors for these enzymes are lacking.

Developing such inhibitors is challenging as structural information and details about the mechanisms of mammalian fucosyltransferases are scarce. Recent progress in the development of inhibitors for the structurally well-studied Human Blood Group Galactosyltransferase B (GTB) [15] demonstrates the significance of structural information. As most glycosyltransferases exhibit high affinity to their family-wide shared donor substrate, but generally low affinity to their specific acceptor substrate, bi-substrate analog inhibitors are thought to be the most promising inhibitors [16]. Their successful development demands detailed knowledge of the binding process of donor and acceptor substrates.

Within this thesis, two core-modifying FucTs from two different organisms were studied. The specific properties and the biological significance of both enzymes, human  $\alpha$ 1,6-fucosyltransferase (FUT8) and core  $\alpha$ 1,3-fucosyltransferase (FucTA) from honeybee (*Apis mellifera*) are presented in the following sections.

## **1.2 Human core $\alpha$ 1,6-fucosyltransferase (FUT8)**

Human  $\alpha$ 1,6-fucosyltransferase (FUT8) catalyzes the transfer of a fucosyl residue from GDP-Fuc to the 6-hydroxy function of the innermost *N*-acetyl glucosamine (GlcNAc) residue of N-glycans with inversion of the anomeric configuration (cf. Figure 1-3). [17] The products of FUT8, glycoproteins carrying core  $\alpha$ 1,6-fucosylated N-glycans, are widely distributed in human and animal tissues and various biological functions are regulated by this common modification of glycoproteins. [1, 18]. The physiological significance of FUT8 has been demonstrated in FUT8-null mice [14, 19], of which 80% died within three days after birth. FUT8 knock-out mice exhibit severe growth retardation and emphysema-like changes in the lung. This phenotype is thought to be caused by inactivation of growth factor receptors due to the lack of core-fucosylation of their glycans. The results of these studies revealed that core-fucosylation is crucial for the activation of growth factor receptors like the transforming growth factor beta (TGF $\beta$ ) receptor [14], the epidermal growth factor (EGF) receptor [19] and the vascular endothelial growth factor (VEGF) receptor 2 [20]. Furthermore, core fucosylated N-glycans influence turnover and expression levels of E-cadherin [21] and modulate the activity of  $\alpha$ 3 $\beta$ 1 integrin [22]. FUT8 therefore influences cell adhesion and cell migration processes. Also, core-fucosylated N-Glycans were found to be of diagnostic value for tumor

diseases. For instance, the level of core fucosylation of  $\alpha$ -fetoprotein is a well-known tumor marker in hepatocellular carcinoma [23, 24], very likely due to high expression of the GDP-fucose transporter [25] and enhances cell–cell adhesion in human colon carcinoma cells [21]. Therapeutic antibodies show a 50–100fold increase of Fc $\gamma$ -mediated cytotoxicity if the core-fucosylation is deleted. [26, 27]

Some of the putative mechanisms by which core-fucosylation regulates the functions of glycoproteins have been elucidated. It has been demonstrated that core-fucosylation as well as bisecting GlcNAc shift the conformational equilibria of N-glycans. [28, 29] Also, the modification has been shown to affect the serum clearance of glycoproteins [28]. Prolonging the half-life of the glycoprotein has also been suggested as a mechanism for the high expression of the membrane protein E-cadherin upon enhanced core-fucosylation. [21]

Small molecules that could influence the activity of FUT8 *in vivo* with a high selectivity would facilitate further elucidation of the biological function of this biologically important carbohydrate epitope. However, the design of specific, potent and permeable inhibitors for fucosyltransferases and glycosyltransferases in general remains a challenging task. [15, 16]

FUT8 is a typical type-II transmembrane glycosyltransferase that resides in the medial Golgi [30] and was first described by Wilson *et al.* [17]. The enzyme has been purified and cloned from various human and animal tissues [30-32]. Detailed analyses of the FUT8 gene [33-35] have revealed that the overall homology of FUT8 to any other known glycosyltransferases is very low. However, distinct regions are highly conserved among  $\alpha$ 1,2-fucosyltransferases, bacterial  $\alpha$ 1,6FucT (NodZ) and protein O-fucosyltransferases [35-38]. The existence of these conserved motifs has led to site-directed mutagenesis studies of FUT8 in order to determine residues that are essential for the catalytic activity [38, 39]. These data presented a set of 15 mutants of highly conserved residues of which eight amino acids are essential for the enzymatic activity of FUT8. As all fucosyltransferases use the same donor substrate GDP-Fuc but differ in their acceptor substrates, the conserved regions are supposedly involved in donor substrate binding. In fact, two conserved arginine residues, Arg365 and Arg366, have been shown to be involved in binding the nucleotide, although the latter is not essential for activity of FUT8. [38]

The specificity of FUT8 towards its acceptor substrate, a branched N-glycan (cf. Figure 1-3), is highly unusual [40-43]. First, FUT8 requires an extremely large minimal structure of at least six monosaccharide units. Furthermore, the activity of FUT8 on the acceptor is strictly dependent on the presence of an unsubstituted GlcNAc residue attached to the  $\alpha$ 3-linked

mannose of the N-glycan. In contrast, the GlcNAc at the 6-mannose branch may be absent or substituted and even hybrid-type N-glycans are fucosylated by FUT8. Interestingly, the linkage of the reducing end to the asparagine is not essential. [39, 40] Other core modifications like core  $\alpha$ 1,3-fucosylation and bisecting GlcNAc were found to completely prevent fucosylation by FUT8. [32, 41-43] As all fucosyltransferases use the same donor substrate, the recognition process of the acceptor substrate is responsible for the specificity of the fucosyl transfer.

The high resolution crystal structure of FUT8 [44] gives further indications on the donor substrate binding mode. The 3D-structure revealed that the fold of FUT8 can be classified as glycosyltransferase-B (GT-B). The enzyme exhibits a Rossman fold, a structural motif that is widely distributed among nucleotide-binding proteins. A comparison to the crystal structure of NodZ [45, 46] revealed that the conserved motifs are also structurally similar to the bacterial  $\alpha$ 1,6 fucosyltransferase [47].

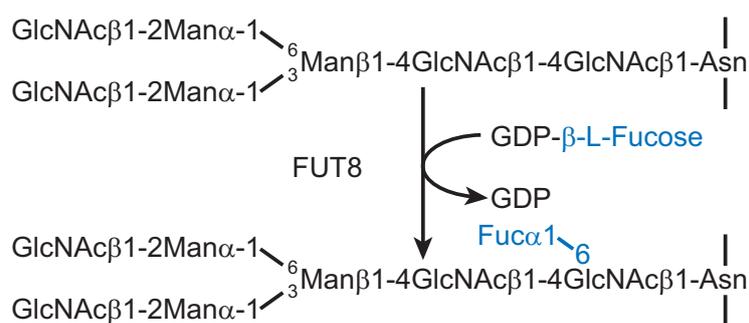


Figure 1-3: Enzymatic reaction catalyzed by FUT8. A fucosyl moiety is transferred from the activated sugar nucleotide GDP-Fuc to the 6-hydroxy group of the proximal GlcNAc of the acceptor with inversion of the anomeric configuration. The acceptor, a branched N-type oligosaccharide, has to present an unsubstituted GlcNAc residue at the 3-mannose to be recognized by the enzyme. In contrast, the linkage of the proximal GlcNAc to the asparagine side chain is not essential.

FUT8 acts via a bi-bi mechanism and both donor and acceptor substrates bind to the enzyme in order to form a ternary complex [39]. After transfer of the fucosyl residue, both GDP and the fucosylated acceptor are released [39, 41]. Potent inhibitors combine high efficiency, i.e. a low dissociation constant, with high specificity. Inhibitors that are either donor or acceptor substrate analogs exhibit only one of these properties. In the case of FUT8, donor substrate mimics will affect several fucosyltransferases in the organism. Blocking the acceptor substrate binding site is challenging because of the low binding affinity found in that region. Hence, bisubstrate analog inhibitors are most promising to feature both affinity and specificity.[15, 16] Yet, for the design of bisubstrate inhibitors, detailed knowledge of the recognition process of both substrates is essential. The general model for substrate binding in

glycosyltransferases involves a deep cavity which binds the sugar nucleotide with relatively high affinity. In contrast, the acceptor substrate binding site is shallow and the dissociation constant is generally in the mM range [48, 49].

Detailed information about the substrate binding mode and the catalytic mechanism as required for structure-based inhibitor design cannot be derived from the FUT8 structure, as it does not include any substrates or substrate analogs and all attempts to co-crystallize or soak FUT8 with its substrates have failed so far [44]. On the other hand, protein NMR-based methods are likewise demanding due to the size of FUT8 (62 kDa) and the fact that isotope labeling and particularly perdeuteration are difficult in eukaryotic expression systems required for the expression of functional FUT8. 3D structural data for fucosyltransferases in complex with GDP-Fuc is available for only three other fucosyltransferases: Pioneering, for the  $\alpha$ 1,3-fucosyltransferase of *Helicobacter pylori* [50], for the Protein O-fucosyltransferase 1 of *Caenorhabditis elegans* (*ce*POFUT1) [51] and, very recently, also for NodZ from *Bradyrhizobium sp.* [45]. In the case of the enzyme of *H. pylori*, the substrate binding site is fundamentally different to that of FUT8. *ce*POFUT1, however, exhibits conserved motifs that are also present in FUT8 and NodZ. These segments of all three proteins are also structurally related to some extent. [51]

For some non-mammalian fucosyltransferases, X-ray structures with bound donor substrate were successfully solved [45, 50, 51] Despite much effort, bound acceptor substrates have never been observed in X-ray structures of fucosyltransferases, probably due to their generic low binding affinity [48, 49, 52, 53] or other factors. It is therefore crucial to employ alternative methods in order to gain an insight into the molecular basics of acceptor substrate binding.

### **1.3 Honeybee core $\alpha$ 1,3-fucosyltransferase (FucTA)**

N-glycans attached to the glycoproteins of invertebrates exhibit a so-called pauci-mannosidic structure, i.e. they usually lack residues attached to the  $\alpha$ 1,3 and  $\alpha$ 1,6-linked mannose of the core pentasaccharide (cf. Figure 1-4). In addition, invertebrate N-glycans often exhibit both an  $\alpha$ 1,3 and an  $\alpha$ 1,6-linked fucose residue at the proximal GlcNAc, as for example found in drosophila. The former modification is the product of core  $\alpha$ 1,3-fucosyltransferase (FucTA), an enzyme that is unique to invertebrates and plants.

The xenobiotic nature of the core  $\alpha$ 1,3-fucosyl epitope renders it highly immunogenic, resulting in the induction of IgE as well as IgG antibodies, key components in different immune mechanisms. IgE reactivity directed against these cross-reactive carbohydrate

determinants (CCDs), present on venom glycoproteins of insects and other more common xenobiotics, like birch pollen, [54] severely hampers diagnosis of the causative agent and therefore its therapy [55-57]. Although immunomodulation and adjuvant effects have been reported, the impact of these glycotopes for biology, pathogenesis and clinical outcome of disease and therapy is not fully understood. Hence, avoidance of xenobiotic glycosylation and optimization by glyco-engineering are a central issue in biotechnology, in which a variety of cell types are increasingly used for the production of diagnostically and therapeutically relevant glycoproteins. [58]

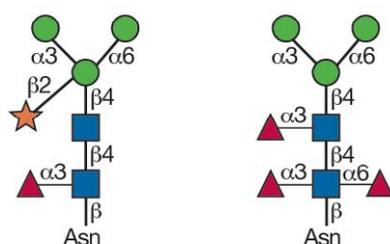


Figure 1-4: N-glycan structures typically found in plants (left) and invertebrates (right). In contrast to mammalian N-glycans, they are usually limited to the core pentasaccharide (pauci-mannosidic structure) that carries core modifications.

Honeybee (*Apis mellifera*) FucTA has recently been identified [59]. Similar enzymes are present in other invertebrates and have been cloned from *Caenorhabditis elegans* [60], *Vigna radiata* [61] and *Drosophila melanogaster* [62]. Figure 1-5 shows the reaction catalyzed by FucTA. The acceptor substrate, a heptasaccharide, is essentially the same used by FUT8. The  $\alpha$ 1,6-linked fucosyl residue is attached by core  $\alpha$ 1,6-fucosyltransferase (homolog to FUT8, cf. above). FucTA can act upon a fucosylated or non-fucosylated core oligosaccharide. Contrary, FUT8 can only attach the 6-linked fucosyl residue onto a core oligosaccharide that does not carry the  $\alpha$ 1,3 linked fucosyl residue. [42] In contrast to diptera, e.g. drosophila, hymenoptera, e.g. honeybee or wasp, show predominantly single  $\alpha$ 1,3-fucosylation on their glycoproteins, suggesting different acceptor substrate specificities for FucTA of both species [63, 64]. The pauci-mannosidic structure shown in Figure 1-4 is obtained *in vivo* by processing of the fucosylated glycan by specific hexosaminidases [42, 65, 66].

FucTA has not been characterized by X-ray or protein NMR methods yet. Because the sequence homology to structurally characterized fucosyltransferases is very low and none of them is related to FucTA from invertebrates, its 3D structure remains elusive. Consequently, the molecular basis underlying the specificity for both donor and acceptor substrates, is not known yet. Like other fucosyltransferases, FucTA is believed to operate in a bi–bi sequential

mechanism in which the donor sugar nucleotide binds to the enzyme together with the acceptor oligosaccharide and after transfer, the glycosylated product and the nucleotide are released from the ternary complex [48].

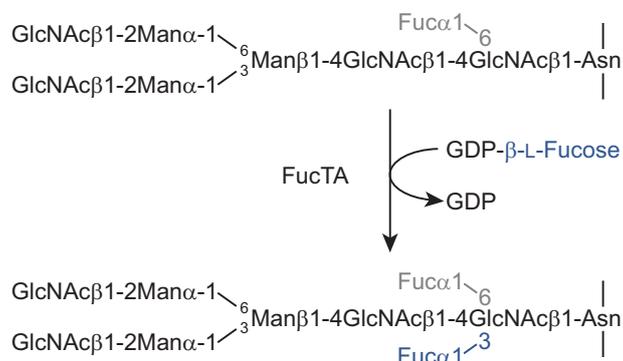


Figure 1-5: Reaction catalyzed by FucTA. Insect core α1,3-fucosyltransferase transfers a fucose residue from the donor substrate GDP-Fucose (GDP-Fuc) to the proximal *N*-acetylglucosamine of an *N*-glycan (acceptor substrate). Whereas the non-reducing GlcNAc at the 3-Man arm is required for activity, the linkage to the Asn side chain is not necessary and reducing oligosaccharides are accepted as substrates. *N*-glycans already carrying an α1,6-linked core fucose (shaded in grey) are preferred by drosophila FucTA over unfucosylated substrates. [62]

## 2 Objective

Fucosylation is a common modification of glycans and is known to be associated with various physiological functions like cell-cell recognition and numerous ontogenic events as well as pathophysiological problems like host-pathogen interactions, immune responses, selectin-mediated leukocyte-endothelial adhesion in inflammation and tumorigenesis. Fucosyl residues are transferred by fucosyltransferases and despite their importance in mammalian physiology little is known about the structural basis of their specificity or about their catalytic mechanism.

The objectives of this thesis are to elucidate the substrate binding mode and the specificity of the substrate recognition process of core-fucosyltransferases on a detailed structural basis. Based on these results, a structure-based strategy for the design of a specific inhibitor should follow and a synthetic route for this compound should be developed.

As a basis for these experiments, human FUT8 that transfers a  $\alpha$ 1,6 fucose residue should be cloned, expressed and purified in cooperation with the group of E. Spillner. In order to elucidate substrate specificity, the complex type oligosaccharides used by core FucTs as substrates should be synthesized. These substrates are necessary to analyze the binding by ligand-based NMR methods (STD NMR, trNOESY), SPR and enzyme kinetic assays. First, these binding analyses can be performed with another fucosyltransferase that was readily available in the group of E. Spillner, the FucTA from *Apis mellifera*.

The next goal was to analyze substrate binding and enzyme kinetics of FUT8 with ligand-based NMR methods, SPR and enzyme kinetic assays. The combined experimental information from these experiments and earlier data (i.e. X-ray crystal structure of the apo enzyme and site-directed mutagenesis studies) were supposed to serve as a reasonable basis to model substrate binding of FUT8 *in silico*. Molecular modeling methods, in particular molecular dynamics simulations, should subsequently be employed in order to elucidate detailed enzyme–substrate contacts. This information is hardly accessible through solely experimental methods as outlined in the introduction.

From the model, the mechanism of substrate binding, in particular the specific recognition of the unusually large acceptor should be derived. Based on these results, a specific inhibitor should be created by means of structure-based inhibitor design. As a last step, a synthetic route towards this inhibitor should be designed.

### 3 Introduction to the methods

#### 3.1 Preparation of fucosyltransferases

Experiments that give insight into structure or substrate binding processes require mg amounts of purified protein. Core fucosyltransferases have been isolated, purified and characterized from various human and animal tissues.[30-32, 41, 43, 59, 61, 67-69] This approach, however, poses the problem of different activities or splicing variants of the enzyme depending on the organism and the tissue the enzyme is isolated from. Also, purification of enzymes from natural sources requires affinity chromatography using the specific substrates of FucTs, a costly and laborious procedure. Finally, the yields of enzyme that can be obtained from natural tissues are limited. As FUT8 and FucTA have already been cloned from various species, [23, 30, 31, 59, 60, 62, 70], recombinant expression of these enzymes is the tool of choice in order to obtain mg amounts of purified fucosyltransferases. The preferred expression system for proteins that are to be subjected to structural analysis is *Escherichia coli*, a well characterized and established expression system for heterologous proteins. Expression in *E. coli* facilitates high yields, fast expression and low costs. Furthermore, the host allows expression of isotopic labeled proteins, a prerequisite for protein NMR-based methods. The detailed structural characterization of the glycosyltransferases GTB and bovine  $\alpha$ 1,4-GalT1 is mainly possible due to the fact that both enzymes can be functionally expressed in *E. coli*. However, none of the known eukaryotic FucTs has been successfully expressed in bacterial hosts. This is probably owed to the fact that eukaryotic FucTs require posttranslational modification, e.g. glycosylation [39], and/or proper folding in order to obtain enzymatic activity. Mammalian expression systems that guarantee nearly natural posttranslational modification of the desired protein yield in low protein amounts. Insect cells in combination with baculoviruses serve as a compromise between bacterial and mammalian hosts. The secretory pathway of these eukaryotic cells enables most posttranslational modifications. Furthermore, expression can be performed in serum-free media allowing for easy purification and yields up to tens of  $\mu$ g/mL suspension culture. Also, limited isotope labeling may be obtained as demonstrated for human kinases [71-73].

As basis for further studies of the enzyme mechanism, Blank *et al.* established a high-level expression and purification protocol for recombinant soluble honeybee (*A. mellifera*) FucTA [52, 74, 75]. The nearly autologous expression system in baculovirus infected Sf9 cells guarantees a natural posttranslational processing, including glycosylation, of FucTA lacking only the transmembrane domain. [59, 76-79] Analysis of purified sFucTA demonstrated that

the enzyme itself undergoes core  $\alpha$ 1,3-fucosylation and is therefore a substrate for its own enzymatic activity.[52] The immediate availability of a functional expression system of honey bee FucTA was due to a cooperation with the group of E. Spillner and provided a starting point for detailed mechanistic and specificity analysis addressed in this thesis. Furthermore, the presence of FucTA allowed for testing and development of NMR and SPR assays that were later successfully applied to human FUT8.

For human FUT8, functional expression employing the Sf9/baculovirus expression system was demonstrated by Ihara *et al.* [39]. They used a construct in which the transmembrane region (amino acids 1-67) was replaced by the sequence of the signal peptide of gp67 of the baculovirus for efficient secretion. Their large-scale expression and purification protocol utilizing ammonium sulfate precipitation and affinity chromatography via a C-terminal 6-fold histidine (His6) tag provided the basis for successful determination of the 3D structure of FUT8 by means of X-ray crystallography [44]. However, recombinant viruses were not available upon request from the group. Therefore, the full cloning, transfection and expression process of FUT8 was carried out within this study in cooperation with the research group of PD Dr. E. Spillner that also had established the expression protocol for FucTA.

### **3.2 Preparation of complex type oligosaccharides**

Core FucTs utilize complex type heptasaccharides as substrates (cf. Figure 1-3 and Figure 1-5). In order to study acceptor substrate recognition and specificity of FucTA and FUT8, mg amounts of purified oligosaccharides are required. Former activity studies on FucTA and FUT8 used either glycopeptides as isolated from glycoproteins [32, 40, 41, 80] or glycans/glycopeptides conjugated with fluorescent dyes [38, 39, 42, 62, 66, 81, 82]. Both approaches allow for separation of the products via reversed phase (RP) HPLC and detection of  $\mu$ g amounts with UV or MS detection. However, besides the fact that conjugation of glycans is tedious, the unnatural hydrophobic tag might bias binding constants and kinetic assays. Furthermore, the additional resonances of peptide or aromatic protons complicate the interpretation of NMR spectra in ligand-based NMR methods.

Synthesis of oligosaccharides by means of organic chemistry is very time consuming, as synthetic routes require many steps, extensive protecting group strategies and glycosylation reactions with high anomeric selectivity. For complex type N-glycans containing natural monosaccharide units and linkage, isolation from appropriate glycoprotein sources is an established procedure. Such glycoproteins that are available in high amounts and exhibit limited microheterogeneity at their glycosylation sites include fetuin [83], transferrin, various

glycoproteins from hen's egg yolk [84] or fibrinogen. The general approach involves denaturation followed by tryptic digest of the protein in order to render the glycosylation sites accessible to glycan-cleaving enzymes. The mixture of peptides and glycopeptides is then digested by protein N-glycosidase F (PNGase F) to release reducing oligosaccharides from asparagine side chains. Further digestion with sialidase guarantees structurally uniform asialo biantennary N-glycans, as attached *N*-acetylneuraminic acid (NeuAc) and *N*-glycolylneuraminic (NeuGc) acid are a source of microheterogeneity. Furthermore, the uncharged asialo glycans can easily be separated from the peptide mixture by ion exchange chromatography. For utilization as substrates for core-FucTs, galactose units attached to the non-reducing GlcNAc residues have to be cleaved by galactosidase. The resulting asialo agalacto biantennary glycan still contains significant amounts of other glycan compounds and requires purification.

Underivatized glycans require special phases for chromatographic purification because their hydrophilic nature prevents retention on reversed phase columns. Phases suitable for separation of oligosaccharides by HPLC are hydrophilic interaction phases (HILIC), i.e. silica phases conjugated with ionizable functional groups or porous graphitized carbon (PGC). The latter is used in this study for semi-preparative purification of N-glycans, as it does not bleed and offers high separation performance. However, reducing oligosaccharides are separated by their anomeric configuration if the rate of mutarotation is not enhanced by additives in the mobile phase. If combined with ESI MS detection, the desired glycan-containing fractions can readily be identified.

Core FucTs were found to process reducing oligosaccharides with nearly the same rate and affinity as the corresponding glycopeptides. Hence, the use of reducing oligosaccharides as substrates is convenient. For unambiguous characterization of complex type N-glycans, NMR analysis is required in order to assign composition, linkage and anomeric configuration. For reducing oligosaccharides, NMR analysis of the structural reporter groups, i.e. the well dispersed resonances outside the bulk region of the 1D <sup>1</sup>H NMR spectrum, [85] is usually adequate to obtain this information. In addition, only a few ng of oligosaccharide are required with modern instruments in order to obtain a proton spectrum. [5] However, for identifying ligand epitopes or ligand conformations with ligand-based NMR methods, full assignment of all resonances is very helpful. Therefore, β-*N*-acetyl derivatives of the glycans are of value because their spectra contain only one set of signals due to uniform anomeric configuration. Furthermore, the anomeric linkage of these derivatives resembles the natural linkage in N-linked glycans. As the 1D proton spectra of oligosaccharides still suffer from severe signal

overlap, 2D correlation spectra are required for complete assignment. These include  $^1\text{H}$ - $^1\text{H}$  COSY spectra, in which resonances of proton pairs are correlated via their  $^2J$  or  $^3J$  scalar coupling and therefore allow for identification of neighboring protons.  $^1\text{H}$ - $^1\text{H}$  TOCSY on the other hand correlates all protons belonging to one spin system by applying a strong spin-lock field subjecting all scalar coupling to Hartman-Hahn conditions. Thus, TOCSY is well-suited for biomolecules that are built up of discrete building blocks like peptides, nucleotides and oligosaccharides. A problem with the latter, however, arises from disruption of magnetization transfer due to small scalar coupling constants as found in oligosaccharides with equatorial protons like mannose (H-2), galactose (H-4) and fucose (H-4) and in  $\alpha$ -linked sugars. Recording several TOCSY spectra with increasing mixing time (i.e. spin-lock time) has proven useful for assignment of proton chemical shifts in oligosaccharides. In NOESY experiments, pairs of proton chemical shifts are correlated by their nuclear Overhauser effect (NOE). The NOE is a relaxation phenomenon that arises due to dipolar coupling between two protons with a distance of less than 4 Å. NOESY spectra therefore give information about linkage between carbohydrates. Furthermore, they facilitate assignment of proton chemical shifts in pyranoses due to specific NOEs arising from 1,3-diaxial protons.  $^1\text{H}$ - $^{13}\text{C}$  HSQC correlates carbon chemical shifts with proton chemical shifts via their  $^1J$  coupling. The large dispersion of chemical shifts on the  $^{13}\text{C}$  channel causes well-separated resonances even with larger molecules. Furthermore,  $^{13}\text{C}$  chemical shifts add information on the chemical environment of the attached proton and facilitate the assignment. Due to the low natural abundance of  $^{13}\text{C}$ , HSQC spectra require a few hundred  $\mu\text{g}$  of oligosaccharide in order to be recorded within reasonable instrument time.

### 3.3 Analyzing kinetics and thermodynamics of binding processes

Fucosyltransferases convert two substrates into two products. Prior to the actual process of substrate conversion, one or both substrates have to bind to the enzyme in order to form a binary or ternary complex. After substrate conversion, the products have to be released from the product-enzyme complex and allow the enzyme to enter a further catalytic cycle. Therefore, the kinetic and thermodynamic analysis of the process of substrate binding is crucial to elucidate substrate recognition and the enzymatic mechanism of FucTs.

In its simplest form, the formation of a complex between two molecules, e.g. enzyme and substrate, follows a bimolecular association reaction with second-order kinetics. The equilibrium equation given in Equation (3-1) contains the rate constants for the association ( $k_{on}$ ) and dissociation ( $k_{off}$ ) reaction. The equilibrium constant for the dissociation reaction

given in Equation (3-1) calculates from the law of mass action given in Equation (3-2) and is called the dissociation constant  $K_D$ . The dissociation constant is related to the Gibb's free energy of the dissociation reaction as shown in Equation (3-3). Therefore,  $K_D$  is a measure for the affinity of the enzyme to the substrate.



$$K_D = \frac{[E] \cdot [S]}{[ES]} = \frac{k_{off}}{k_{on}} \quad (3-2)$$

$$\Delta G^0 = RT \cdot \ln K_D \quad (3-3)$$

Where  $[E]$  = enzyme concentration;  $[S]$  = substrate concentration;  $[ES]$  = concentration of enzyme-substrate complex;  $k_{on}$  = rate constant of association;  $k_{off}$  = rate constant of dissociation,  $K_D$  = dissociation constant;  $\Delta G^0$  = free enthalpy;  $R$  = gas constant;  $T$  = absolute temperature.

Assuming that the association is diffusion-controlled ( $k_{on} = 10^7 \text{ s}^{-1} \text{ M}^{-1}$ ), values for  $k_{off}$  can directly be estimated from  $K_D$  employing Equation (3-2) as shown in Table 3-1. However,  $k_{on}$  may vary several orders of magnitude from  $10^4$  to  $10^{11} \text{ s}^{-1} \text{ M}^{-1}$ . Especially for large and flexible ligands or flexible binding sites in the enzyme, slow association rates are often observed due to conformational rearrangements upon the binding process.

Table 3-1: Estimations of  $k_{off}$  assuming diffusion-controlled association.

$K_D$	$k_{off} [\text{s}^{-1}]$
<b>1 mM</b>	10,000
<b>1 <math>\mu\text{M}</math></b>	10
<b>1 nM</b>	0.01

Several experimental methods allow for determination of dissociation constants. Isothermal titration calorimetry (ITC) allows for direct determination of the thermodynamic parameters of a binding process in solution by measuring the release of thermal energy of the exothermic reaction. ITC however requires large amounts of enzyme (tens of mg), especially when the affinity to ligands is low. Other techniques measure the dissociation constant by measuring a signal which intensity is proportional to the fraction of receptor-ligand complexes present in solution with varying ligand concentration. The most wide-spread techniques include FRET-based assays, ELISA, surface plasmon resonance (SPR) and various NMR experiments (cf. below).

Within this study, SPR was chosen to analyze the thermodynamic parameters of the binding process of FucTs with their substrates for the following reasons: First, SPR is a sensitive method that requires minimal amounts of enzyme. Second, labeling with fluorescent tags is not required as opposed to ELISA or FRET. Finally, the technique additionally allows for kinetic characterization of binding processes. In the following, the basics of affinity and kinetic assays by SPR are discussed in detail.

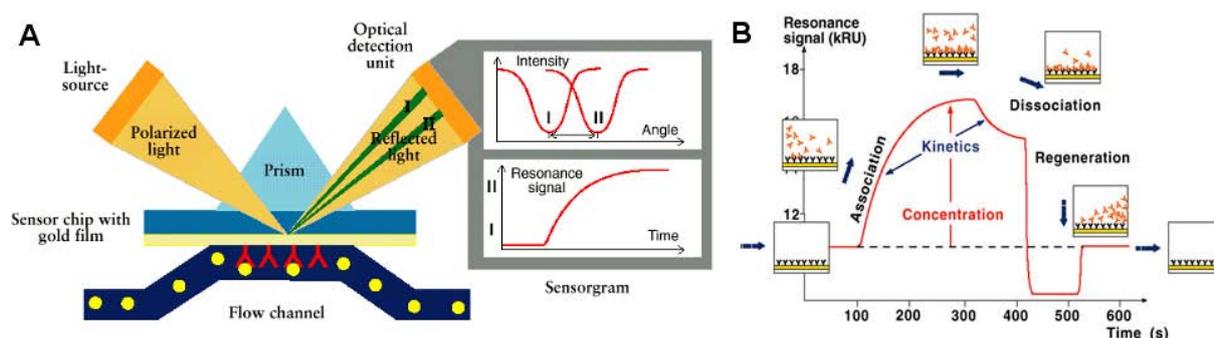


Figure 3-1: Principle of analyzing biomolecular interactions with SPR. A: Schematic representation of an SPR experiment. B: Idealized representation of an SPR sensorgram.

Surface plasmon resonance (SPR) is an optical method for detection of binding events. Basically, SPR is the resonant, collective oscillation of valence electrons in a solid stimulated by incident light. [86] The resonance condition is met when the frequency of the light matches the natural frequency of surface electrons oscillating against the restoring force of positive nuclei. In order to excite surface plasmons in a resonant manner, p-polarized light (polarization occurs parallel to the plane of incidence) is employed and resonance occurs at a given angle. Practically, SPR is archived by passing linear polarized light through a glass block that is coated with a thin metal film, typically gold or silver. The light illuminates the glass block and is totally reflected at the glass–metal interface. The evanescent wave present at the interface penetrates the metal film and hence excites the plasmons, leading to a reduction of the intensity of reflected light. Technically, the angle of the absorption maximum is measured, as shown in Figure 3-1A. The resonance conditions of the surface plasmon waves are highly sensitive to changes of the boundary conditions. Molecules adsorbing at the metal film cause changes of these boundary conditions leading to changes in the order of  $0.1^\circ$  of the angle of minimum reflection. SPR instruments usually output these changes in resonance units (RU), where 1 RU is defined as the change of angle caused by an adsorbent mass of 1 pg (ca.  $0.0001^\circ$  change in angle).

For measuring binding events between biomolecules, one interaction partner is immobilized at the gold surface of a chip while the other partner is in solution and passed over the surface

in a flow cell. Among different immobilization techniques, standard amide coupling is most wide-spread and was also used in this study. Standard amide coupling uses a gold chip linked with carboxymethyl dextran (CM5 chip). The carboxy functions are converted into active ester functions and consequently covalently linked to amino functions of the analyte. For proteins, this method is convenient, as most proteins carry lysine residues at their surface. Since SPR is very sensitive, only a few hundred fmol of protein are required to analyze interactions with small molecules.

Figure 3-1B shows an idealized SPR sensorgram, i.e. the signal obtained from the experiment when the mobile interaction partner is injected into a flow cell with an immobilized interaction partner. The actual sensorgram is normalized by subtraction of a sensorgram simultaneously measured at a reference flow cell that has no binding partner immobilized on its surface. The so-called double referencing procedure guarantees that bulk contribution is cancelled out. The SPR sensorgram subdivides into four stages: The association phase starts when the mobile analyte is injected and ligands successively occupy the binding sites of the immobilized molecule. In the simplest case of a 1:1 interaction, SPR response follows the integrated rate equation of the association phase processes.

$$RU_t = RU_{max}[1 - e^{-k_{obs}t}] + RU_0 \quad (3-4)$$

Where  $RU_t$  = response at a given time point;  $RU_{max}$  = maximal response at a certain concentration;  $R_0$  = response at  $t = 0$ ;  $k_{obs} = k_{on} \cdot c_L + k_{off}$

$k_{obs}$  (and consequently  $k_{on}$ ) can therefore be revealed by non-linear curve fitting of Equation (3-4) to measured data points. Notably, very fast association rates, i.e. when the steady state is reached within a few seconds, cannot be resolved by SPR sensorgrams at the employed sample rate of 10 Hz.

If equilibrium is achieved, the rates of association and dissociation reactions are equal. In the so-called steady state, no net change of the SPR response is observed (plateau of the sensorgram) and the SPR response has its maximum value ( $RU_{max}$ ) at the measured ligand concentration. From the  $RU_{max}$  value at different ligand concentrations, the dissociation constant  $K_D$  of the molecular complex can be calculated. Therefore,  $RU_{max}$  is plotted against the ligand concentration and a binding isotherm described by the one-site-binding model (cf. Equation (3-5)) is fitted to the data by non-linear regression.

$$RU = \frac{RU_{max} \cdot c_L}{K_D + c_L} \quad (3-5)$$

After the injection of ligand has ended, buffer is injected to the flow cell and the complexes dissociate. The dissociation phase of the sensorgram obeys the integrated rate equation of the dissociation process as shown in Equation (3-6). Hence, non-linear regression of equation (3-6) to the dissociation phase directly yields  $k_{off}$ .

$$RU_t = RU_0 \cdot e^{k_{off}(t-t_0)} \quad (3-6)$$

Regeneration as depicted in Figure 3-1B is necessary when the dissociation process is very slow, e.g. at very low dissociation constants. With the systems analyzed within this study, regeneration was found to be not needed.

SPR sensorgrams can be biased by additional effects that have to be taken into account when analyzing binding events. Association and dissociation may have contributions by mass transport effects, especially at low analyte concentrations, low flow rates or high density of immobilized binding partner at the chip surface. Thus, these contributions may be prevented by adequate experimental set-up. In addition, unspecific binding or other binding events, for instance resulting from denatured protein may give rise to more complex sensorgrams. In these cases, it can be helpful to fit two-site-binding models to the data.

### 3.4 Analyzing enzyme kinetics of glycosyltransferases

Analysis of enzyme kinetics yields key parameters for understanding the functionality and the mechanism of biocatalyzed reactions. Provided that the reaction involves one substrate, is irreversible and is not inhibited by the forming product, the enzymatic reaction is described by Equation (3-7):



In an initial bimolecular reaction, the specific ES complex is formed. Although the enzymatic mechanism for the unimolecular reaction can be quite complex, there is typically one rate-determining enzymatic step that allows this reaction to be described as a single catalytic step with an apparent unimolecular rate constant ( $k_2$ , also referred to as turnover number  $k_{cat}$ ). However, even for this simplified enzymatic reaction, the rate equations can only be solved assuming a quasi-steady-state, namely that the concentration of the substrate-bound enzyme (and hence also the unbound enzyme) changes much more slowly than those of the product

and substrate and thus the change over time of the complex can be set to zero. The resulting rate equation is called Michaelis–Menten equation and given in Equation (3-8). The Michaelis–Menten equation involves two key kinetic constants  $K_M$  (cf. Equation (3-9)) and  $V_{max}$  (cf. Equation (3-10)), whereas  $V_{max}$  is the maximum velocity of the enzymatic reaction that is achieved when the enzyme is saturated with substrate and the reaction rate is only dependent on  $k_2$  (reaction proceeds with 0. order).  $K_M$  is the substrate concentration where the velocity is  $V_{max}/2$  and therefore a measure for the affinity of the enzyme towards its substrate under the condition of substrate conversion.  $K_M$  approximates  $K_D$  of the enzyme–substrate complex if the rate-determining enzymatic step is slow compared to substrate dissociation.

$$v = \frac{d[P]}{dt} = k_2[ES] = \frac{k_1[S][E]_0}{[S] + (k_{-1} + k_2)/k_1} \quad (3-8)$$

$$K_M = (k_{-1} + k_2)/k_1 \quad (3-9)$$

$$V_{max} = k_2[E]_0 \quad (3-10)$$

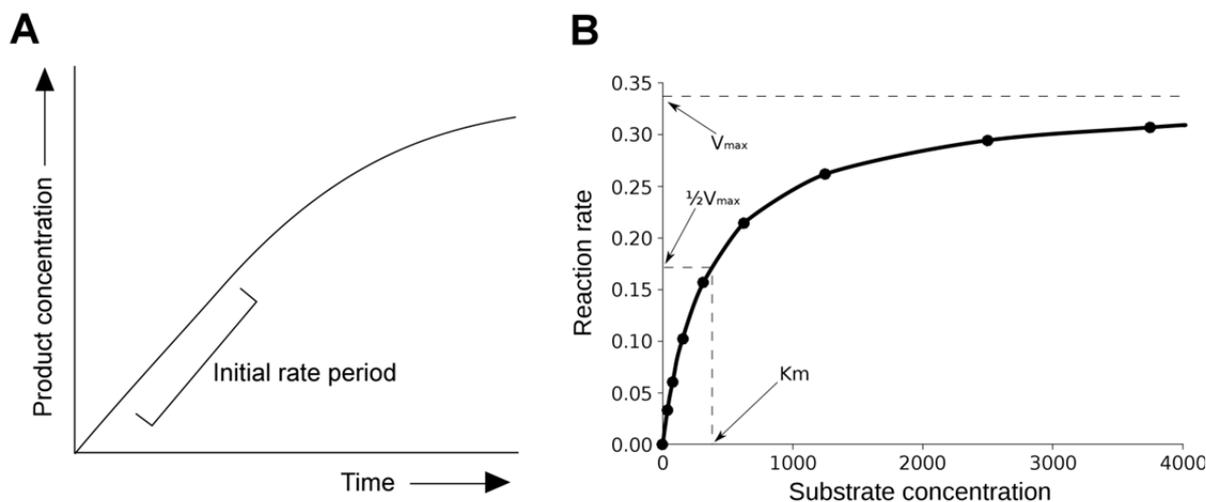


Figure 3-2: Graphical representation of enzyme kinetics. A: Time course of the formation of product (progress curve). B: Rate of the initial reaction versus initial substrate concentration as given by the Michaelis–Menten equation illustrating the origin of  $K_M$  and  $V_{max}$ .

Figure 3-2A shows the progress curve of an enzymatic reaction, i.e. the course of product formation with time. The enzyme produces product at an initial rate that is approximately linear for a short period after the start of the reaction as long as substrate concentration is at saturating levels. As the reaction proceeds and substrate is consumed, the rate continuously slows. Most enzyme kinetic studies use the initial, approximately linear part of the progress curve measured at different initial substrate concentrations. Plotting the reaction rates against

substrate concentration yields an isotherm as shown in Figure 3-2B from which  $K_M$  and  $V_{max}$  can be directly obtained with non-linear fitting algorithms using Equation (3-8).

Measurements of the initial rate require a fast detection in the concentration of either substrate or product, as the linear part of the curve may take less than a second. If fast online detection is not feasible due to the characteristics of the compounds, quenching of the reaction, work-up and quantification of the substrate or product is necessary. This approach is most often used with glycosyltransferases employing labeled substrates (fluorescence dyes or stable isotopes).

An alternative is provided by the direct analysis of the time course (progress curve) of the enzymatic reaction. [87-89] Fast detection is not necessary because the first part of the curve is linear as long as the initial substrate concentration much higher than  $K_M$ . The progress curve is mathematically described by the integrated Michaelis–Menten equation (Equation (3-11)) and contains information on all reaction rates at every substrate concentration that is passed through the enzymatic reaction. Therefore, only one reaction has to be performed in order to determine kinetic parameters. As Equation (3-11) is implicit concerning substrate concentration, the solution had to be found numerically in the past, a fact that rendered the analysis of progress curves inconvenient. Only in 2004 the Lambert- $W$  function, (cf. Equation (3-13) for definition), was described as analytical solution for the integrated Michaelis–Menten equation. [90] Employing this explicit solution shown in Equation (3-12),  $K_M$  and  $V_{max}$  can directly be obtained from the progress curves with standard non-linear fitting algorithms [91]. Practical aspects that have to be taken into account with the analysis of progress curves are product inhibition and reversibility of the enzymatic reaction.

$$V_{max}t = K_M \ln \frac{[S]_0}{[S]} + [S]_0 - [S] \quad (3-11)$$

$$[S] = K_M W \left\{ \frac{[S]_0}{K_M} \exp \left( \frac{[S]_0 - V_{max}t}{K_M} \right) \right\} \quad (3-12)$$

$$W(z)e^{W(z)} = z \quad (3-13)$$

Glycosyltransferases catalyze reactions involving two substrates and two products (bi–bi mechanism). Multi-substrate reactions follow complex rate equations that depend on the sequence of substrate binding. Their kinetics therefore generally does not obey the Michaelis–Menten equation. However, if one of both substrates is applied at such high concentrations that its concentration may be regarded as constant during the reaction, a bisubstrate reaction behaves just like a single-substrate reaction.

In ternary-complex mechanisms as employed by glycosyltransferases, the order of binding of the two substrates can either be random (in a random mechanism) or substrates have to bind in a particular sequence (in an ordered mechanism). Random and ordered bi–bi mechanism can be experimentally distinguished through product inhibition studies. Rapid equilibrium bi–bi mechanisms show competitive inhibition for each of both products towards each of both substrates. In contrast, for ordered mechanisms, competitive inhibition is only observed for the second product to be released towards the leading substrate. In all other combinations, the inhibition is mixed.

### **3.5 Structural characterization of biomolecular complexes**

The first step of an enzymatic reaction involves the formation of an enzyme–substrate complex that typically has dissociation constants in the range of  $10^{-2}$ – $10^{-8}$  M (corresponds to free enthalpies of binding from –13 to –50 kJ/mol). The catalytic activity of enzymes is largely based on their ability to bring the substrates together in a favorable special arrangement and to facilitate the transition state of the reaction. This is accomplished by catalytic residues in the active site of the enzyme. These residues facilitate the formation of the transition state by means of specific interactions with the substrates. The specificity for a substrate depends on the defined arrangement of atoms in the active site. Substrates are bound via multiple weak binding forces, e.g. electrostatic interactions, hydrogen bonds, van-der-Waals forces and hydrophobic interactions. Because these forces have a short range, the active site preferentially has a shape complementary to that of the substrate. Often, this complementary shape is only completely formed upon substrate binding. For some glycosyltransferases, formation of the binding site for the acceptor substrate has been found to be dependent on the occupation of the binding site of the donor substrate.

In order to exploit substrate specificity to design specific inhibitors, detailed knowledge on the active site and the structural basics of the enzyme–substrate complex at atomic resolution is crucial. In principle, methods that are established for the determination of the 3D structure of proteins at high resolution are also suited for structural elucidation of substrate–enzyme complexes. Hence, much of the knowledge about the structure of these complexes is based on X-ray structures where enzymes have been co-crystallized or soaked with substrate. However, solving X-Ray structures of non-mammalian fucosyltransferases in complex with GDP-Fuc or GDP has been successful in a few cases [45, 50, 51]. Despite much effort, bound acceptor substrates have never been observed in these experiments, probably due to their generic low binding affinity [48, 49, 52, 53]. The other powerful technique to obtain high resolution structures is protein NMR, which is performed in solution. Provided that enzyme is

completely assigned, substrate binding residues can be identified by the change in the chemical shift of the corresponding resonances, e.g. in a  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum in presence of substrate. Yet, protein NMR-based methods require isotope labeling which is highly challenging in the eukaryotic expression systems necessary for expression of functional mammalian fucosyltransferases. Furthermore, complete assignment of protein resonances becomes a difficult task with increasing size of the protein. Even for glycosyltransferases with average size (around 40 kDa) that can be expressed in bacterial expression systems, no characterization by NMR has been published to date.

In contrast to protein NMR-based methods, ligand-based methods do not require isotopic labeling of the enzyme. Furthermore, much lower amounts of protein are required than in protein-based methods, as the ligand is detected in these experiments. Contrary to protein NMR-based methods, ligand-based NMR-experiments yield information about the ligand's binding mode, e.g. the ligand epitope or the bound conformation of the ligand. Both techniques employed in this study, STD NMR and trNOESY, are based on the nuclear Overhauser effects (NOE). NOEs are extremely useful for the characterization of binding events as NOEs undergo drastic changes for ligands binding to large receptors leading to transferred NOEs (trNOEs). The observation of trNOEs relies on different tumbling rates  $\tau_c$  of free and bound molecules. Depending on the applied field strength, shape and size of the molecule, small to medium-sized compounds exhibit small positive NOEs, no NOEs or small negative NOEs. Large molecules like proteins, in contrast, show large negative NOEs due to their slow tumbling rate. Ligands binding to proteins behave as a part of a large molecule and thus acquire the corresponding NOE properties. The large negative NOEs observed for ligands under these conditions are called trNOEs. The observation of intermolecular trNOEs that occur between protein and ligand is exploited various experimental schemes including relaxation filtered NMR experiments and STD NMR experiments. Intramolecular trNOEs, on the other hand, are the key to elucidate bound-ligand conformations.

### **3.5.1 Saturation transfer difference (STD) NMR**

Saturation transfer difference (STD) NMR is a powerful and robust method for characterization of binding events between large receptors (e.g. proteins) and small molecules. [92] STD NMR is used for screening of entire libraries or mixtures for binding compounds [93] as well as for group epitope mapping of ligands. [94] Furthermore, STD NMR allows for determination of dissociation constants and inhibition constants of protein–ligand complexes. A typical sample contains protein and ligand in a ratio around 1:100.

Hence, depending on the particular system, a successful STD NMR experiment often requires only around 10  $\mu\text{g}$  of protein.

STD NMR is based on selective saturation of protein resonances. Proteins give rise to very broad proton resonances due to very short transversal relaxation times ( $T_2$ ). In addition, chemical shift anisotropy in ordered regions of the protein yields extreme chemical shifts of protein resonances. Hence, signal intensity of aliphatic resonances is found beyond 0 ppm in proton spectra of proteins. These resonances are saturated by a cascade of soft pulses that is typically applied around -1 ppm. Since spin diffusion is very efficient in large molecules ( $\text{MW} > 10 \text{ kDa}$ ), magnetization spreads out the entire protein within 2-100 ms. Ligand resonances, on the other hand, are not affected as they give rise to sharp resonances within the normal chemical shift range. Thus, only resonances of ligands binding to the saturated protein acquire saturation via dipolar cross relaxation. Consequently, the signal intensities of such ligands are reduced compared to a spectrum acquired in presence of unsaturated protein.

Provided that the longitudinal relaxation times ( $T_1$ ) of the ligand are longer than the dissociation rate constant  $k_{\text{off}}$  of the protein-ligand complex, saturated ligand accumulates in solution if an excess of ligand is given. This scenario normally applies to small ligands with dissociation constants in the range of nM to mM and yields a concentration of saturated ligand in solution that corresponds to a multitude of the protein concentration.

Practically, STD experiments are performed by alternating acquisition of two types of 1D proton NMR spectra. Afterwards, the STD spectrum is generated from these two spectra by means of subtraction as shown in Figure 3-3. During the acquisition of the on resonance spectrum, the protein is saturated. For the off resonance spectrum (also referred to as reference spectrum) the protein is not saturated as the saturation pulse is applied outside the spectral window of protein signal intensity (for instance at 40 ppm). The off resonance spectrum therefore corresponds to a standard 1D proton NMR spectrum of the sample and the saturation pulse is only applied to guarantee identical conditions for both partial experiments. Further pulse sequences for suppression of residual protein resonances ( $T_{1\rho}$  filter) or suppression of solvent (i.e. excitation sculpting sequence) can easily be integrated in STD experiments. The difference (STD) spectrum, that is generated by subtracting the on resonance spectrum from the off resonance spectrum after processing, contains signal intensity only for protons with a distance of less than 4 Å in the protein–ligand complex. The size of the STD effect is inversely proportional to the sixth power of the distance of the

protons participating in the cross relaxation process. Hence, quantification of the observed STD effects allows for mapping of the ligand epitope (cf. Figure 3-3).

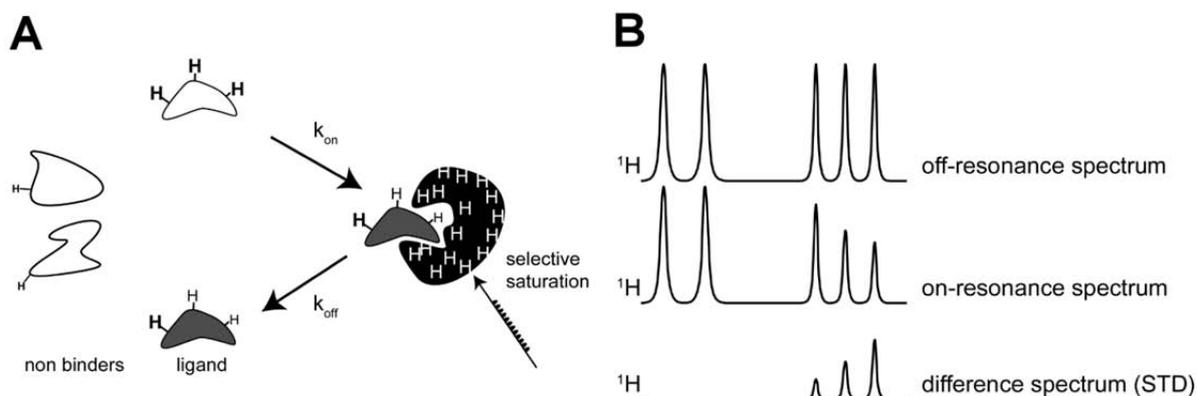


Figure 3-3: Principle of the STD NMR experiments. A: Only ligands that bind to the acceptor acquire saturation. B: As a result, only resonances of these ligands give rise to signal intensity in the STD spectrum.

Some other effects have to be considered for group epitope mapping (GEM). It is important for GEM that the residence time of the ligand in the bound state is significantly shorter than the  $T_2$  time of the ligand in the bound state. Otherwise, spin diffusion erases the differences between protons within one ligand. Such long residence times (slow  $k_{off}$ ) are often observed with ligands with dissociation constants below  $0.1 \mu\text{M}$ . Very slow dissociation rates also lead to very small absolute STD effects because not enough ligands dissociate from the binding sites for the length of saturation time. Extremely short residence times associated with very fast off rates, on the other hand, do not leave enough time for efficient saturation transfer to build up.

STD NMR can also be used to determine dissociation constants. Therefore, STD spectra are recorded a protein sample with different ligand concentrations. From the excess of ligand and the size of the STD effect, the STD amplification factor ( $STD_{amp}$ ) is calculated according to Equation (3-14).  $STD_{amp}$  quantifies the throughput of ligand molecules and is therefore a measure for the concentration of protein–ligand complexes. Plotting of  $STD_{amp}$  against total ligand concentration yields a binding isotherm (cf. Equation (3-15)) from which  $K_D$  is obtained. Similarly, inhibition constants ( $K_I$ ) for ligands competing for the same binding site can be determined when the competing binder is titrated to a sample where the ligand is present at fixed concentration and  $STD_{amp}$  of the ligand is plotted against the inhibitor concentration. In order to span the whole range from  $STD_{amp(max)}$  (no inhibitor present) to  $STD_{amp} = 0$  (all binding sites are occupied by the inhibitor), the inhibitor concentration may have to cover several orders of magnitude. Hence, for better visualization, a logarithmic representation of inhibitor concentration is often preferred. The inflection point of the

sigmoidal binding isotherm corresponds to the concentration of inhibitor yielding the half maximal inhibition ( $IC_{50}$  value). The  $K_I$ , which is independent on the ligand concentration, is obtained from  $IC_{50}$  and  $K_D$  of the ligand using the Cheng-Prusoff equation [95] given in Equation (3-16).

$$STD_{amp} = \frac{I_0 - I_{sat}}{I_0} \cdot \frac{[L]_0}{[P]_0} \quad (3-14)$$

Where  $I_0$  = Intensity of resonance in the off resonance spectrum;  $I_{sat}$  = Intensity of resonance in the on resonance spectrum;  $I_0 - I_{sat}$  = Intensity of resonance in the STD spectrum;  $[L]_0$  = ligand concentration;  $[P]_0$  = protein concentration.

$$STD_{amp} = \frac{STD_{amp,max} \cdot [L]_0}{K_D + [L]_0} \quad (3-15)$$

Where  $STD_{amp}$  = STD amplification factor;  $STD_{amp,max}$  = maximal STD amplification factor.

$$K_I = \frac{IC_{50}}{1 + \frac{c_L}{K_D}} \quad (3-16)$$

Where  $K_I$  = inhibition constant;  $IC_{50}$  = inhibitor concentration at half maximal inhibition;  $c_L$  = total ligand concentration.

### 3.5.2 Transferred nuclear Overhauser enhancement NMR spectroscopy (trNOESY)

trNOESY allows for determination of conformations of bound ligands. The technique has successfully been applied to elucidate many biological questions. In the field of glycosyltransferases, the active conformation of UDP-Gal bound to GTB was characterized by means of trNOESY experiments. [49] The observation of intramolecular trNOEs is most often experimentally accomplished by acquiring a 2D transient NOE spectrum (NOESY) of a protein sample containing an excess of ligand (typically 10–40 fold). From the intensity of the strong negative cross peaks (same phase as diagonal peaks), NOEs between proton pairs of the ligand can be quantified and used to derive the conformation of the bound ligand. As mentioned above, trNOEs can be differentiated from the NOEs of free ligands by their sign and the size. In addition, the build-up rate is four to ten times shorter for trNOEs (50 to 100 ms) than for NOEs from free ligands. Hence, trNOEs are observed at significantly shorter mixing times.

$$\langle \sigma \rangle = N_F \sigma_F + N_B \sigma_B \quad (3-17)$$

Where  $\langle \sigma \rangle$  = average cross relaxation rate;  $N_F, N_B$  = fractions of free and bound ligand;  $\sigma_F, \sigma_B$  = cross relaxation rates of free and bound ligand.

Compared to STD NMR experiments, the successful observation of trNOEs is much more sensitive to the kinetic and thermodynamic parameters of the binding process. The trNOESY

spectrum reflects the average NOEs of free and bound form of the ligand as shown in Equation (3-17). Hence, ligand excess has to be chosen carefully and with respect the exact values of  $\sigma_F$  and  $\sigma_B$  (and hence the tumbling rates of protein and ligand) and to the dissociation constant in order to meet an appropriate population density of binding sites with ligands. A crucial issue for the observation of trNOEs is the exchange rate. Off rates in the same order than the difference in chemical shift difference yield broadened signals, rendering the signal of the bound form unobservable in many cases. Only for exchange that is fast on the chemical shift scale it is easy to produce a situation where the observed NOEs are dominated by cross-relaxation in the bound state that are efficiently transferred into the free state (hence the name trNOE). In this scenario, the only signals observed are sharp and virtually identical to those of free ligand and thus can readily be assigned. Finally, relaxation rates affect trNOESY. As for STD experiments, the off rate may not be much slower than the longitudinal relaxation rate of the ligand (particularly in the bound form) as no information of the bound state can ever be transferred into solution when lost in relaxation.  $T_1$  is typically around 0.1 s, off rates slower than  $1 \text{ s}^{-1}$  usually mean that trNOEs are not observable. The cross-relaxation rate is responsible for the build-up rate of the NOE and is typically around  $1\text{-}10 \text{ s}^{-1}$ . Only if the cross-relaxation rate is slower than the dissociation rate, NOEs can build up in one state and be transferred across to the other state without significant loss of magnetization from longitudinal relaxation.

In order to prevent quantification of trNOEs to be biased by spin diffusion it is useful to employ perdeuterated proteins for trNOE measurements. If labeling of the protein with deuterium is not possible, monitoring NOE build-up helps to distinguish protein-mediated NOEs from true intramolecular NOEs.

### **3.6 Molecular modeling of biomolecular complexes**

The term molecular modeling encompasses all theoretical methods and computational techniques that are used to model or mimic the behavior of molecules and is also referred to as computer-aided molecular design (CAMD). Besides representation of complex molecules, molecular modeling techniques allow for calculation and prediction of physico-chemical properties of a molecule on an atomic level. Molecular modeling methods are routinely used to investigate structure, dynamics, surface properties and thermodynamics of inorganic, polymeric and biological systems, including protein folding, enzyme catalysis, protein stability, conformational changes associated with biomolecular function and molecular recognition processes. [96]

Molecular modeling methods can be assigned to two different general approaches: Semi-empirical and statistical methods treat atoms as the smallest individual unit (molecular mechanics approach). Quantum chemical methods explicitly model electrons of each atom. Although the latter are very precise, the high computational effort renders quantum chemical methods inappropriate for large systems like biomolecules. Thus, molecular modeling techniques based on molecular mechanics are state of-the-art in the fields of drug design, computational biology and materials science. The term molecular mechanics refers to the use of classical mechanics to describe the physical basics behind the models. Atoms in these models are characterized as point charges with an associated mass. Chemical bonds between the atoms are modeled as spring-like interactions and van der Waals forces (commonly described by the Lennard-Jones potential). Electrostatic forces are calculated using Coulomb's law. Atoms are assigned to coordinates and therefore velocities can be calculated in dynamic simulations. The kinetic energy is related to the temperature of the system. The collective mathematical expression containing the system internal energy, i.e. the sum of kinetic and potential energy, is known as a potential function. This function computes the molecular potential energy as a sum of energy terms that describe the deviation of bond lengths, bond angles and torsion angles from equilibrium values. Furthermore, it contains terms for pairs of atoms that interact non-covalently (van der Waals and electrostatic interactions). The potential function uses parameters consisting of equilibrium bond lengths, bond angles, partial charge values, force constants and van der Waals parameters. This collective set of parameters is designated as force field. Common force fields utilized today have been developed using high-level quantum mechanics calculations and fitting to experimental data. As biomolecules contain a limited set of atom types and bond types present in their building blocks (e.g. amino acids, nucleobases, carbohydrates), users of molecular mechanics implementations can choose from a variety of sophisticated force fields suitable for modeling of biological systems. Molecules can either be modeled in vacuum or in presence of a solvent. Explicit solvent simulations include solvent molecules whereas implicit solvent simulations estimate solvent effects employing an empirical mathematical expression. [96]

### **3.6.1 Energy minimization techniques**

Methods which minimize the potential energy are known as energy minimization techniques (e.g., steepest descent and conjugate gradient algorithms). The technique is used to find a local energy minimum by approaching positions of zero gradient for all atoms. Energy minimization calculations are convenient for obtaining a static picture for comparing states of similar systems. Hence, it is a versatile tool for estimating, for instance, the effect of

derivatizations of an inhibitor on the relative binding energy and as such frequently used in computer-aided drug design. [96]

### 3.6.2 Molecular dynamics simulation

A molecular dynamics (MD) simulation, on the other hand, computes the behavior of a system with propagation of time. In the most common version, Newton's laws of motion are numerically solved for a system of interacting particles. In MD implementations, force on an atom is defined as the negative gradient of the potential energy function. Integration of Newton's laws of motion leads to atomic trajectories in space and time. Hence, MD simulations provide information about the dynamic processes with the intrinsic inclusion of temperature effects. [97]

The available computational power is a main constraint that has to be taken into account when designing MD simulations. The required CPU time is dependent on the number of particles, time step and total duration time. The duration should match the time scale of the natural process to be studied. Biomolecular MD simulations typically span durations from ns to  $\mu$ s requiring several CPU-days to CPU-years. Using explicit solvent enhances simulation size by roughly ten times and must be calculated expensively in the potential function. However, the granularity and viscosity of explicit solvent is essential to reproduce certain properties of the solute molecules. Furthermore, the simulation box size must be large enough to avoid boundary condition artifacts. Employing periodic boundary conditions in which one side of the simulation box loops back to the opposite side mimics a bulk phase without increasing the particle number with additional solvent molecules. Within the MD simulation, the most CPU intensive task is the evaluation of energy as a function of the particles' coordinates. As the non-covalent part is the most CPU expensive (scales roughly by  $n^2$  if all pair-wise interactions are accounted for explicitly) within the potential function, electrostatics methods such as Particle Mesh Ewald (PME) or spherical cut-off techniques are employed. Finally, the size of the integration time step, i.e. the time between evaluations of the potential function, impacts total CPU time. In order to avoid discretization errors, value has to be chosen smaller than the fastest vibrational frequency in the system (typically around 1 fs). [97]

As biomolecules operate in open systems at well-defined temperatures, they can be simulated as canonical ensemble (NVT ensemble, i.e. moles ( $N$ ), volume ( $V$ ) and temperature ( $T$ ) are conserved). In NVT, the energy of endothermic and exothermic processes is exchanged with a thermostat. In order to guarantee a stable temperature, algorithms that simulate a thermostat are employed.

### 3.6.3 Molecular docking

Docking is a method to predict the preferred orientation of a ligand when bound to a protein in a stable complex and hence plays an important role in the rational design of drugs. The aim of a molecular docking procedure is to simulate the molecular recognition process by achieving an optimized conformation for both the protein and ligand in order to find a relative orientation between protein and ligand associated with minimal free energy of the overall system. Two approaches are used for molecular docking: Geometric matching methods describe protein and ligand as complementary surfaces. In simulation techniques, on the other hand, the ligand finds its position into the protein's active site after a certain number of conformational changes. The latter approach is closer to physical reality and is more amenable to incorporate ligand flexibility. However, exploring a rather large energy landscape takes far more simulation time depending on the number of flexible bonds of the ligand. The success of a docking algorithm depends on two components: the search algorithm and the scoring function. As it is impossible to calculate the entire search space, conformational search strategies create an ensemble of snapshots of ligand and protein referred to as poses. Search algorithms in use are based on statistical searches, MD simulations or genetic algorithms. The scoring function takes a pose as input and returns a number indicating if the pose represents a favorable binding interaction. Most scoring functions are based on molecular mechanics force fields that estimate the energy of the pose. [98]

The reliability of results from all molecular modeling techniques is highly dependent on the quality of the input data. Hence, files containing coordinates from experimental data (e.g. X-ray or NMR structures) have to be carefully checked for incorrect atom types or bond types and angles. In order to prepare the structure, hydrogen atoms not present in X-ray structures have to be added and partial charges have to be assigned with the employed algorithm.

## 4 Results and discussion

### 4.1 Formation of the immunogenic $\alpha$ 1,3-fucose epitope: Elucidation of substrate specificity and of enzyme mechanism of core fucosyltransferase A

In the following section, a structure based characterization of the substrate specificity of honeybee FucTA is presented. The high-level expression system of soluble FucTA (sFucTA) established by Blank *et al.* [52] allowed for STD NMR and SPR studies on molecular recognition of both donor and acceptor substrates. In addition the system was successfully used to develop an NMR-based enzyme kinetic assay suitable for the characterization of glycosyltransferases by progress curve analysis. Text and figures presented in this section have been published in *Insect Biochemistry and Molecular Biology* (DOI: 10.1016/j.ibmb.2011.11.004; cf. Ref. [52]) and are reproduced with permission from Elsevier.

#### 4.1.1 Preparation of complex type oligosaccharide for structural binding studies

The detailed structural analysis of the substrate binding mode of sFucTA by NMR and SPR required about 1  $\mu$ mol of acceptor substrate. Therefore, reducing asialo N-glycans were prepared by tryptic digestion of denatured bovine and porcine fibrinogen and subsequent deglycosylation and desialylation of the tryptic peptides as described before [83, 99]. Table 1 gives an overview of oligosaccharide compounds referred to in this thesis. Bovine fibrinogen yielded a complex N-glycan mixture that prior to degalactosylation needed to be separated by porous graphitized carbon (PGC) HPLC. [100, 101] Digesting the pure asialo nonasaccharide with  $\beta$ -galactosidase resulted in heptasaccharide **1a**. Acceptor substrate **2a** was prepared from porcine fibrinogen that dominantly carries core  $\alpha$ 1,6-fucosylated N-glycans. 5 g fibrinogen yielded 2-5 mg pure asialo agalacto N-glycans. Identity and purity of PGC HPLC purified oligosaccharides were unambiguously confirmed by MS,  $^1\text{H}$  and  $^{13}\text{C}$  NMR. The reducing oligosaccharides **1a** and **2a** served as substrates for honeybee FucTA.

Generally, reducing oligosaccharides give rise to complex NMR spectra due to two sets of signals originating from the  $\alpha$ - and the  $\beta$ -anomers, respectively. Since 1D  $^1\text{H}$  STD NMR epitope mapping experiments require unambiguous assignment of ligand resonances 1-amino-acetylated derivatives were prepared [102]. The *N*-acetate in the 1-position of the oligosaccharides is a very close related mimic of the natural linkage to the side chain of asparagine. The 1- $\beta$ -*N*-acetyl oligosaccharides **1b** and **2b** were characterized by MS, NMR, and 2D NMR techniques.

## RESULTS AND DISCUSSION

Table 4-1: Overview of prepared oligosaccharides. The oligosaccharides **1a** and **2a** were prepared by enzymatic digestion of bovine and porcine fibrinogen, respectively. [83] The core  $\alpha$ 1,3-fucosylated structures **3a**, **4a** and **4b** are products of FucTA using acceptors **1a**, **2a** and **2b** respectively. Other  $\beta$ 1-aminoacetylated oligosaccharides (marked by an asterisk) were synthesized from the reducing oligosaccharides. All compounds were purified by PGC HPLC, retention times ( $R_t$ ) are listed in the table. Compounds **5-7** were analyzed with a PGC HPLC MS method using different conditions, hence, their  $R_t$  values are not comparable and not listed.

Number	Structure	Mass	$R_t$ [min]	Characterization
<b>1a</b>		1316.487	$\beta$ 39.4 $\alpha$ 45.6	NMR, LC-ESI-MS
<b>1b*</b>		1357.547	$\beta$ 40.1	NMR, LC-ESI-MS
<b>2a</b>		1462.544	$\beta$ 49.4 $\alpha$ 55.7	NMR, LC-ESI-MS
<b>2b*</b>		1503.604	$\beta$ 51.3	NMR, LC-ESI-MS
<b>3a</b>		1462.544	$\beta$ 26.4 $\alpha$ 27.2	NMR, LC-ESI-MS
<b>3b*</b>		1503.604	$\beta$ 27.8	NMR, LC-ESI-MS
<b>4a</b>		1608.602	$\beta$ 30.6 $\alpha$ 36.2	NMR, LC-ESI-MS
<b>4b</b>		1649.662	$\beta$ 33.4	NMR, LC-ESI-MS
<b>5</b>		1113.407		NMR, LC-ESI-MS
<b>6</b>		1113.407		NMR, LC-ESI-MS
<b>7</b>		1259.465		NMR

### 4.1.2 FucTA recognizes GDP-Fuc via the nucleobase

Like all fucosyltransferases, honeybee FucTA uses GDP- $\beta$ -L-fucose as donor substrate (Figure 1-5). Commonly, donor substrates bind to glycosyltransferases with dissociation constants in the 10  $\mu$ M range. Here, the interaction of the GDP-fucose donor with FucTA was characterized using surface plasmon resonance (SPR) by analyzing binding of the building blocks of GDP-Fuc, GDP and guanosine to sFucTA (Figure 4-4C). All binding assays of donor substrates were performed in absence of acceptor substrate.

Figure 4-1 to Figure 4-5 show SPR data that was used for analysis of the binding constants of complexes of FucTA and its substrates. Data that did not show specific binding (e.g. methylfucoside, chitobiosylasparagine and octasaccharide **2a**) is not shown.

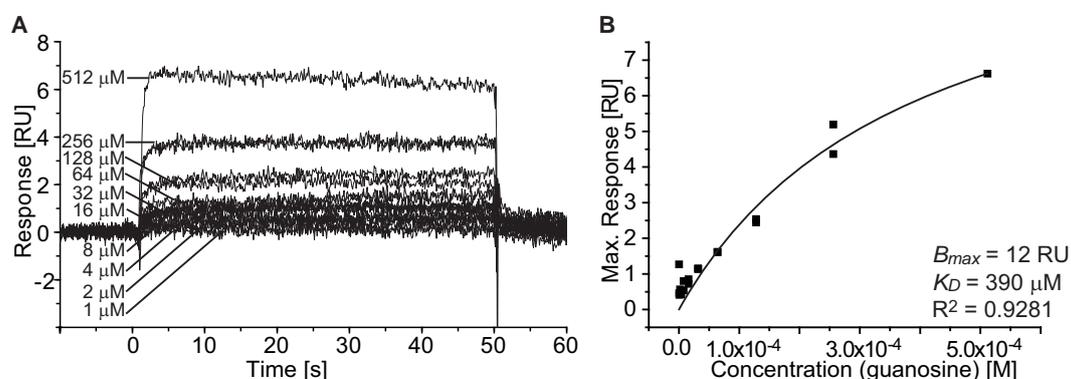


Figure 4-1: Analysis of the thermodynamic binding constant of the guanosine-FucTA complex by SPR. A: SPR sensorgrams at different guanosine concentrations. B: Affinity plot and analysis by means of the one-site-binding model using the response 5 seconds prior to the end of the injection. Concentrations higher than 512  $\mu$ M could not be measured due to the low solubility of guanosine in water.

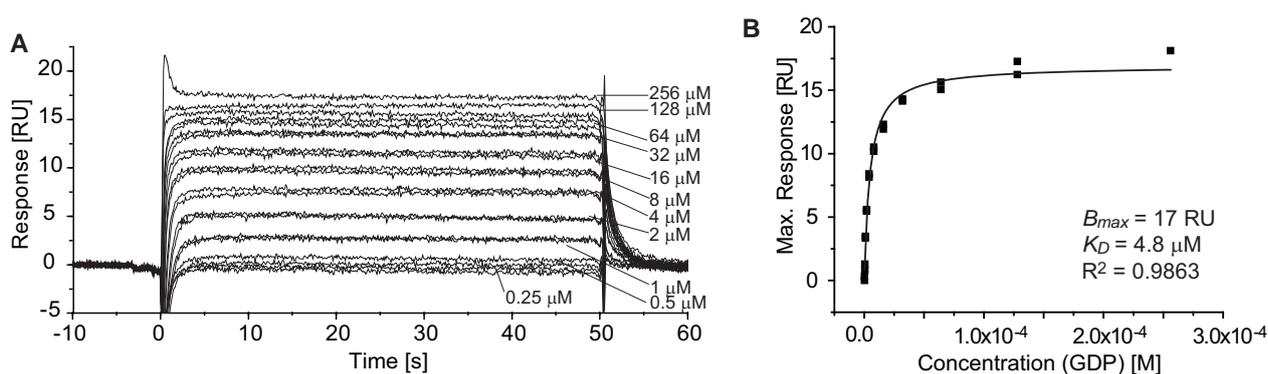


Figure 4-2: Analysis of the thermodynamic binding constant of the GDP-FucTA complex by SPR. A: SPR sensorgrams at different GDP concentrations. B: Affinity plot and analysis by means of the one-site-binding model using the response 5 seconds prior to the end of the injection.

## RESULTS AND DISCUSSION

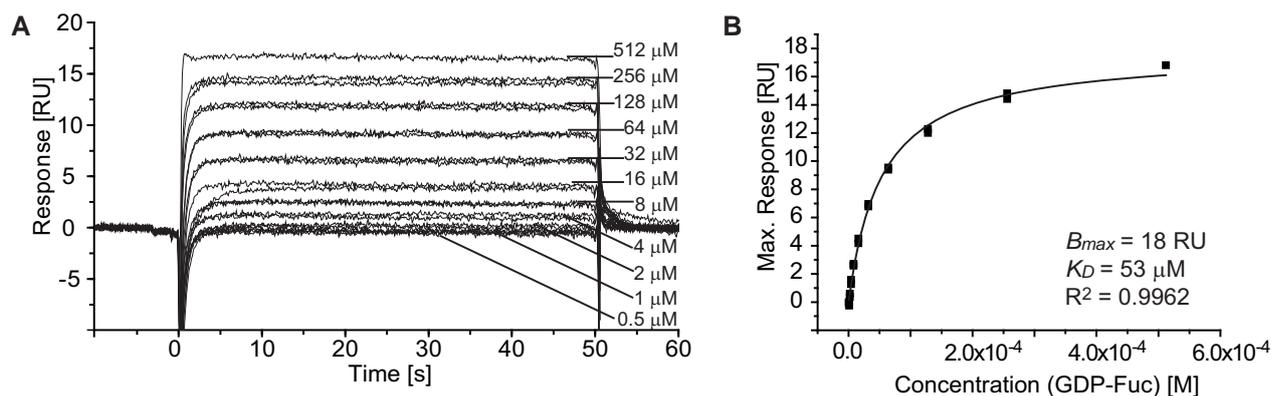


Figure 4-3: Analysis of the thermodynamic binding constant of the GDP-Fuc-FucTA complex by SPR. A: SPR sensorgrams at different GDP-Fuc concentrations. B: Affinity plot and analysis by means of the one-site-binding model using the response 5 seconds prior to the end of the injection.

For *s*FucTA the nucleoside part contributes dominantly to binding affinity with a  $K_{D,\text{guanosine}} = 410 \text{ μM}$  (corresponding to a binding energy  $E = -19.5 \text{ kJ/mol}$ ). GDP gave  $K_{D,\text{GDP}} = 7.0 \text{ μM}$ , equivalent to a binding energy  $E = -29.8 \text{ kJ/mol}$ , indicating that the pyrophosphate bridge accounts for  $\sim -10 \text{ kJ/mol}$  binding energy. GDP-fucose binds to *s*FucTA with a slightly weaker dissociation constant compared to GDP ( $K_{D,\text{GDP-Fuc}} = 37.5 \text{ μM}$ , equivalent to  $E = -25.7 \text{ kJ/mol}$ ).

The detailed epitope of the substrate was determined using STD NMR [92-94]. The results supported the data obtained by SPR (Figure 4-4A and B). Protons of the ribose show the largest relative saturation transfer suggesting that the nucleoside moiety binds closest to the enzyme surface in the binding pocket. STD effects of the guanine cannot be evaluated, as the protons attached to nitrogen atoms exchange too quickly with deuteriumoxide. However, the proton at C-8 showed only low STD effects. Also, the contribution of the pyrophosphate cannot be assessed by STD NMR due to the lack of covalently linked protons. However, the fucosyl residue clearly gains saturation transfer indicating that it is in close contact to FucTA in the bound state.

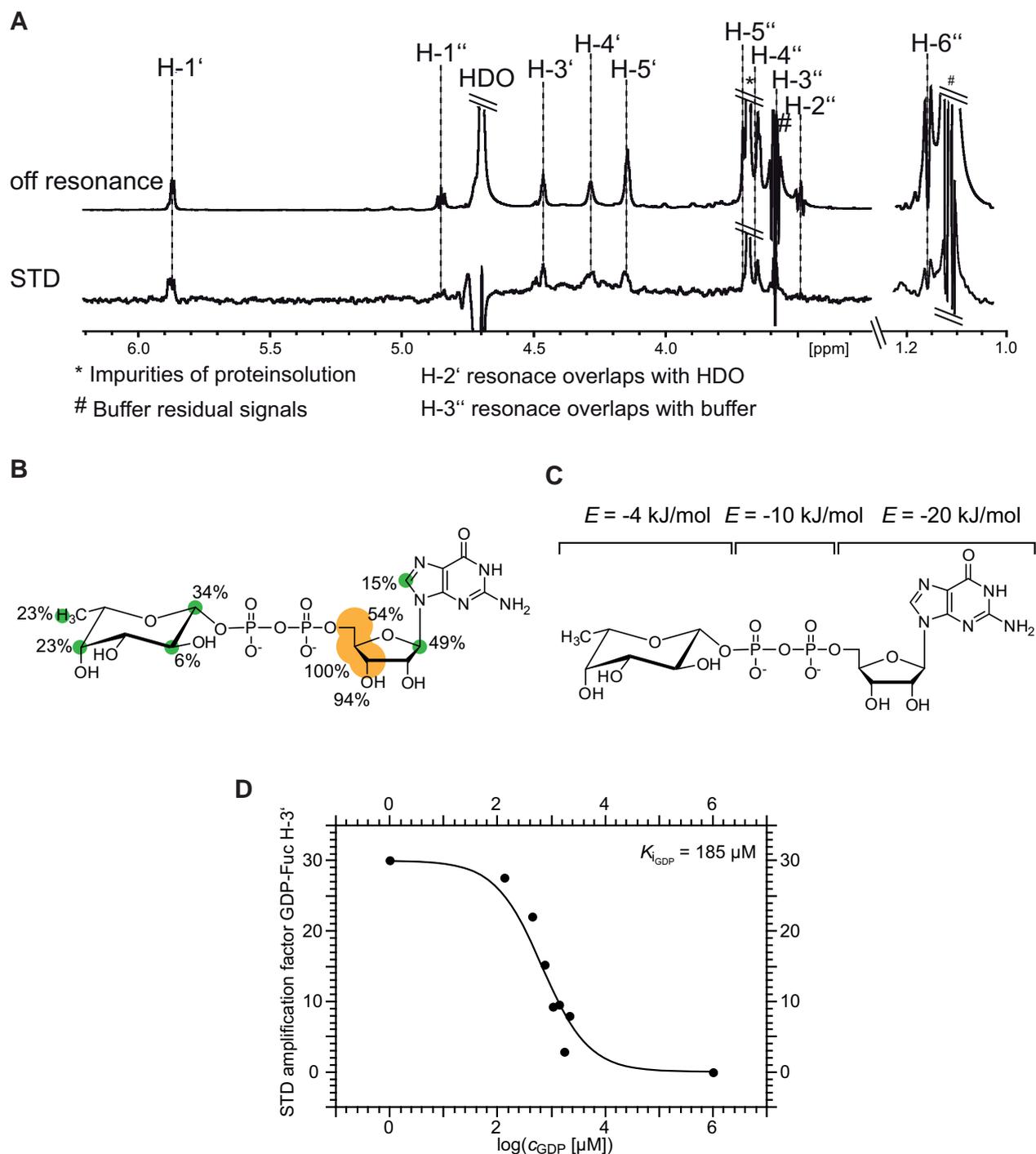


Figure 4-4: Interaction of sFucTA with its donor substrate. A: 1D  $^1\text{H}$  NMR spectrum (top) and STD NMR spectrum (bottom) of GDP-Fucose and FucTA. B: Binding epitope of GDP-Fucose as determined by STD NMR. Circles indicate the relative STD effects which are proportional to the distance of the respective protons to the enzyme's surface. The ribose is clearly closest to the enzyme surface. C: Fragment based screening of the donor substrate binding pocket of FucTA. The brackets indicate the fragments used to elucidate donor substrate specificity by SPR. D Product inhibition of FucTA by GDP. The STD amplification factor for each inhibitor concentration is plotted against the inhibitor concentration. A fit using the one-site competition model yields the  $\text{IC}_{50}$  value from which the  $K_i$  of 185  $\mu\text{M}$  was calculated.

To further elucidate the role of the fucose moiety for the binding substrate, a competitive binding assay monitored by STD NMR was performed (Figure 4-4D). For this analysis GDP-

Fuc was added to a solution of the enzyme and then titrated with GDP. For quantitative analysis of the inhibition constant  $K_I$  [95], eight titration points were determined. STD amplification factors were recorded as a function of the total concentration of the inhibitor GDP ( $c_{GDP}$ ). Here, the STD signal of H-3' of GDP-fucose was used to track the occupancy of the donor binding site with GDP-fucose in presence of different GDP concentrations. A fit of the one-site competition model to the semi-logarithmic affinity plot yielded a value of  $K_I = 185 \mu\text{M}$  for GDP using a  $K_D = 37.5 \mu\text{M}$  for GDP-fucose as determined by SPR experiments. This fact indicates that the fucose enhances the affinity by 4 kJ/mol which is in contrast to the SPR data, where the fucosyl residue weakened binding by about 2 kJ/mol. There are several possible explanations for this small discrepancy: first, the enzyme is in different states in the NMR and SPR based assays – dissolved vs. immobilized. Here especially the immobilization may alter the enzyme dynamics and in consequence the binding affinity. Second, in the NMR experiment, a deuterated was used buffer which might affect the binding strength of hydrogen bonds. Third, SPR experiments are performed under flow conditions which may affect the binding contribution. However, the overall contribution of the fucosyl residue to binding only can be designated as small and almost negligible.

These findings are in agreement with data from other glycosyltransferases and, thus, like in other glycosyltransferases, the nucleobase directs the binding process of the donor substrate to FucTA [39, 48, 103]. However, a quantitative analysis of the  $K_I$  value of GDP revealed that GDP is the weaker binder compared with the natural donor GDP-fucose and the fucosyl residue increases the relative binding energy by a small contribution.

#### 4.1.3 FucTA recognizes a large epitope of the acceptor substrate

Generally, glycosyltransferases show very low affinity towards their acceptor substrates [48, 49]. Determination of binding constants of the acceptor substrates **1a** and **2a** by SPR yielded a dissociation constant of 10 mM for heptasaccharide **1a**. Up to the concentrations used for **1a** binding could not be detected for **2a**. Thus, octasaccharide **2a** has an even higher dissociation constant. These extremely high ligand concentrations (e.g., 10 mM and higher) were accompanied by detrimental effects on the chip surface resulting in negative sensorgrams. Therefore, the binding affinity of **2a** to FucTA could not be determined. Obviously, oligosaccharides with core  $\alpha$ 1,6-linked fucose residues do not favorably bind to the acceptor region of FucTA. Asn-linked chitobiose and  $\alpha$ -methylfucoside gave no SPR response even at concentrations as high as 20 mM. This is in agreement with the fact that the core pentasaccharide of complex type N-glycans is not sufficient to bind to FucTA and to undergo

fucosylation [59]. These data suggest that in contrast to other FucTs and other glycosyltransferases [49, 104] core-modifying fucosyltransferases such as FucTA require a much larger oligosaccharide as acceptor.

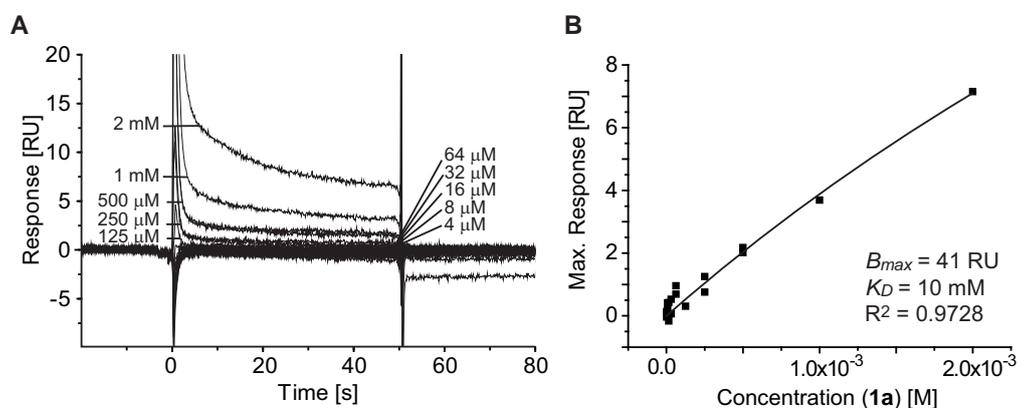
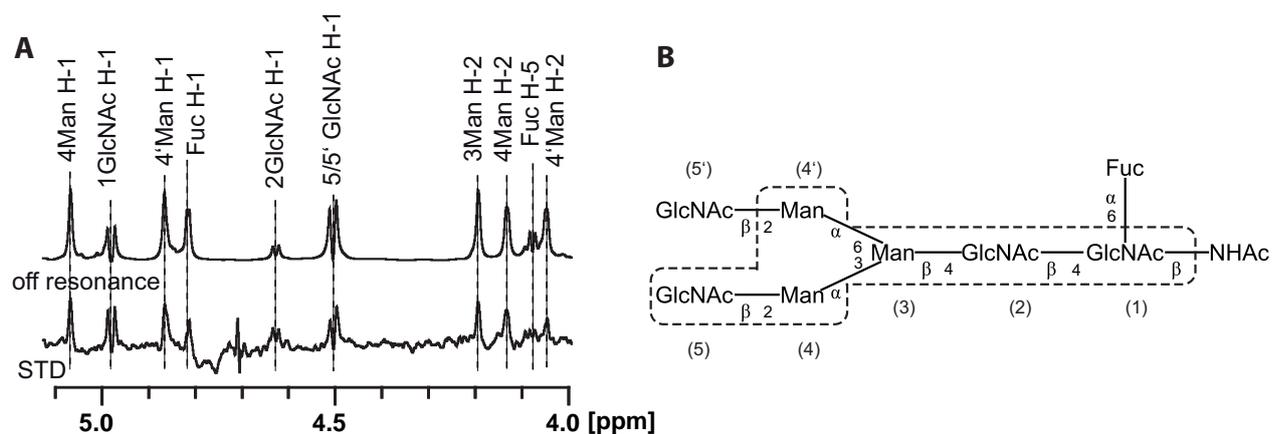


Figure 4-5: Analysis of the thermodynamic binding constant of the heptasaccharide **1a**-FucTA complex by SPR. A: SPR sensorgrams at different concentrations of **1a**. B: Affinity plot and analysis by means of the one-site-binding model using the response 5 seconds prior to the end of the injection. Sensorgrams acquired at concentrations of 5 and 10 mM of the heptasaccharide **1a** were dominated by adverse effects on the chip surface and were not used for analysis.



**C**

Resonance	4Man-H1	1GlcNAc-H1	4'Man-H1	Fuc-H1	2GlcNAc-H1	5/5'GlcNAc-H1	3Man-H2	4Man-H2	Fuc-H5	4'Man-H2
Abs.STD effect	0.29	0.39	0.27	0.12	0.60	0.32	0.33	0.37	0.35	0.24
Rel.STD effect	48	64	46	21	100	53	54	62	59	40

Figure 4-6: Interaction of sFucTA with its acceptor substrate. A: 1D  $^1\text{H}$  NMR spectrum (top) and STD NMR spectrum (bottom) of FucTA and the acceptor substrate **2b**. B: Epitope of the acceptor substrate **2b** as determined by 1D  $^1\text{H}$  STD NMR of the structural reporter groups. The structural reporter groups of the chitobiose core, the  $\beta$ -linked mannose and the  $\alpha$ 1,3-linked mannose residue (dashed boxes) receive relative STD percentages of >50%. C: Absolute and relative STD effects of the octasaccharide binding to FucTA.

STD NMR experiments with octasaccharide **2b** showed only very small absolute STD percentages of 0.3–1.0%. These low absolute STD effects are most likely due to the transient binding and short residence time of the acceptor substrate in the pocket as indicated by nearly rectangular SPR sensorgrams (cf. Figure 4-5A). Virtually all structural reporter groups [85] in the 1D  $^1\text{H}$  STD NMR spectrum of acceptor substrate **1b** (Figure 4-6A) were found to exhibit STD effects indicating that large portions of the molecule are in contact with the enzyme. However, core *N*-acetylglucosamines and the 3-mannose branch seem to be closer to the enzyme surface compared to the fucose and the 6-mannose branch, which show relative STD percentages below 50%. Since the resonances of the two terminal GlcNAc residues cannot be distinguished in the spectra, it is unclear which GlcNAc residue gives rise to the corresponding STD signal. It is likely that the GlcNAc of the 3-branch has a shorter distance to the protein surface. Table 4-2 shows STD effects from the methyl groups of the GlcNAc acetates of heptasaccharide **1b**. At a relaxation delay of 2 s, these resonances are biased by slower longitudinal relaxation and therefore their absolute STD effects cannot directly be compared to the other structural reporter groups. This data was used to elucidate the relative distance of the two GlcNAcs at the antennae to the protein surface. Analysis of the STD effects of the *N*-acetate groups reveals that both GlcNAc residues of the branches have similar distances to the protein surface. The *N*-acetate groups of the ‘reducing’ GlcNAc1 have the highest STD effect overall.

Table 4-2: STD effects of methyl groups of compound **2b** in presence of 10  $\mu\text{M}$  FucTA.

$\delta$ [ppm]	Assignment	Abs. STD effect	Rel. STD effect
2.031	2Gn Ac	1.74%	79%
1.997	5Gn Ac	1.45%	65%
1.994	5'GnAc	1.40%	63%
1.947	1Gn Ac	1.69%	77%
1.943	1Gn Ac	2.20%	100%
1.140	Fuc H-6	0.79%	36%

Further information on the enzyme's acceptor substrate specificity was obtained by fucosylation of a mixture of heptasaccharide **1a** and two hexasaccharides lacking the respective GlcNAc residues at the non-reducing end that originate from cleavage of either terminal GlcNAc residue. NMR analysis revealed that the hexasaccharide lacking the GlcNAc in the 3-branch is not fucosylated (data not shown).

#### 4.1.4 FucTA prefers unfucosylated oligosaccharides as acceptor substrate

Several methods to study enzyme kinetics of core fucosyltransferases have been published [80-82, 84]. In many cases, complex type asialo agalacto N-glycans have been derivatized with fluorescence dyes in order to apply reversed phase HPLC separation employing UV detection. [39, 62] However, fluorescence labeling of oligosaccharides is a laborious procedure and the hydrophobic scaffold might give rise to unwanted protein-ligand interactions. Radiolabeled substrates are the basis for the second established method which offers sensitive detection without changing chemical properties of the substrate. However, all steady state initial rate assays based on detecting substrate conversion after a given time interval are relatively material and time consuming, as several samples with starting concentrations have to be prepared. In addition, conditions for the steady state have to be found in separate experiments.

NMR spectroscopy presents an attractive analytical method because substrate conversion can be directly detected from specific resonances of substrates and products without special probing, quenching or work-up procedures during the enzyme reaction. [5] Direct analysis of the time course (progress curve) of the enzymatic reaction [105] yields  $K_M$  and  $k_{cat}$ . It can be realized by solving the integrated Michaelis–Menten equation (Equation (3-12)) [87-89]. Recently, the Lambert- $W$  function, defined as the function that satisfies Equation (3-13), has been described as solution for the integrated Michaelis–Menten equation (Equation (3-11)) [90]. A non-linear fitting procedure employing the Lambert- $W$  function (Eq.(3-12)) allows for determining  $K_M$  and  $V_{max}$  directly from the progress curve [91].  $^1\text{H}$  NMR provides an ideal tool to transiently record the concentrations of all substrates and products, respectively, as a function of time through integration of their resonances (Figure 4-7A). Product inhibition might influence the velocity of the reaction and thus has to be considered when analyzing progress curves. In line with data obtained from the binding assays (cf. above), samples containing 10  $\mu\text{M}$  FucTA, 1 mM acceptor **1a** and 2 mM GDP-fucose reached equilibrium after 1.5 hours. However, only 60% of the substrate was converted documenting strong product inhibition by GDP (Figure 4-7B). The kinetic model of such a reaction cannot be described by a Michaelis–Menten equation. To determine kinetic constants of the FucTA reaction, product inhibition was circumvented by adding alkaline phosphatase to a sample of 1.5  $\mu\text{M}$  FucTA, 2 mM GDP-fucose and 1.1 mM acceptor substrate **1a** (Figure 4-7D) or 1.6 mM acceptor substrate **2a**, respectively (Figure 4-7C). The excess of donor substrate ensured the reaction to proceed under pseudo first order conditions required to apply Michaelis–

Menten kinetics. No GDP resonances appeared at any time of reaction indicating the high efficiency of the phosphatase.

Like other glycosyltransferases, honeybee FucTA hydrolyses its donor substrate in absence of acceptor substrate [106]. This hydrolysis reaction corresponds to a transfer of a fucosyl residue to a water molecule. For FucTA, a  $k_{cat} = 2 \text{ min}^{-1}$  was found for GDP-fucose hydrolysis, which is somewhat slower than the transfer of the fucosyl residue to its natural acceptor ( $6 \text{ min}^{-1}$ ). In presence of the acceptor, however, no hydrolysis of the substrate is observed.

In this study the NMR-based characterization of glycosyltransferase reactions based on progress curves was established to gain insights into molecular and kinetic characteristics of fucosyltransferases and their interaction with the corresponding substrates. The data illustrate that direct observation of enzymatic processes by NMR spectroscopy allows a label-free observation of all substrates and products and it has the potential for obtaining information on molecular interplay of substrates and products with the enzyme. When run as part of an STD NMR experiment one has additional molecular information on binding processes.

FucTA binds its acceptor with a high dissociation constant. The thermodynamic binding constant determined for the heptasaccharide **1a** by SPR is 10 mM, which is a typical binding constant of acceptor substrates to glycosyltransferases. For the core  $\alpha$ 1,6-fucosylated acceptor substrate **2a**, the dissociation constant was even higher suggesting that the enzyme, unlike its homologue from *Drosophila melanogaster* [62], prefers the unfucosylated N-glycan over the fucosylated structure. This finding is supported by results of enzyme kinetics: The  $K_M$  value for the heptasaccharide **1a** was determined to 400  $\mu\text{M}$ , whereas the octasaccharide **2a** showed  $K_M = 1 \text{ mM}$ .

The discrepancy between the  $K_D$  and  $K_M$  values for each acceptor substrate is remarkable. This has been shown for other glycosyltransferases already [53]. The acceptor substrate binding seems to be dependent on the presence of donor substrate. Thus, thermodynamic binding constant  $K_D$  is often found to be much weaker than the kinetic constant  $K_M$  [107]. SPR in presence of GDP-Fuc was also considered, but the idea was discarded because under these conditions, substrate conversion will occur and thus SPR results will be meaningless and could not be interpreted on the basis of a one-site-binding model. Importantly, SPR assays with the acceptors in presence of GDP yielded the same results as the assay with acceptor substrate only (data not shown). The weak  $K_D$  values are also not uncommon for

protein carbohydrate interaction as it is known from lectin interactions with oligosaccharides. Still lectin chromatography is widely used in purification protocols of carbohydrates.

The substrate specificity is one possible explanation for the rather high proportion of mono  $\alpha$ 1,3-fucosylated N-glycans in comparison to other species found in the glycoproteins of honeybees [63]. However, for detailed elucidation of the order of fucosylation in honeybees, knowledge on substrate specificity of the involved glycosidases is required. Furthermore, no binding affinity of the enzyme to  $\alpha$ -methylfucoside was observed and GlcNAc $\beta$ 1,4-GlcNAc $\beta$ 1,4-Asn by SPR. As the proximal GlcNAc is the focus of action of the enzyme, it is unexpected that no detectable affinity is found for this structure. To dissect which molecular features of the acceptor are actually recognized by the enzyme, the binding epitope of ligand **2b** was determined by STD NMR. Due to the high dissociation constant and the resulting short residence time of **2b** in the enzyme pocket, the absolute STD percentages detected were low. Core GlcNAc-1 and -2,  $\beta$ -Man and  $\alpha$ 1,3-linked Man residues receive significantly more saturation than other monosaccharide units. Also, the GlcNAc residue of the 3-branch shows considerable saturation, which, however, cannot be quantified because of overlap with the corresponding residue of the 6-branch. Taken together with the data from SPR experiments with acceptor substrates and its fragments, recognition of the pentasaccharide from GlcNAc-1 to the GlcNAc of the 3-branch is proposed. All STD contacts of the ring protons are very weak which is to be expected because of the low affinity of the heptasaccharide. It is therefore concluded that the oligosaccharide included in the dashed line in Figure 4 is specifically recognized by the enzyme. It is known from previous studies of the specificity of the acceptor substrate of honeybee FucTA that at least one of the GlcNAc residues in the antennae is essential for activity [59]. For *Drosophila* FucTA and vertebrate core  $\alpha$ 1,6-FucT, the GlcNAc at the 3-Man arm is required [62] and according to the above mentioned STD NMR data, the same is true for honeybee FucTA. In this context, it should be mentioned that an unusual enzymatic activity was observed in the FucTA sample. In presence of its acceptor substrate **2b** and absence of the donor substrate, the terminal GlcNAc of the 1,6-Man branch was selectively hydrolyzed as indicated by a time-dependent decay of the corresponding  $\alpha$ 1,3-Man H-1 and H-2 resonances. Also, products that showed the fucosylation and that were lacking the GlcNAc residue in the 6-branch were found even though no such starting material had been present. This indicates presence of a GlcNAc-ase activity specific for the 6-branch. This observation is also in agreement with the binding epitope shown in Figure 4.

## RESULTS AND DISCUSSION

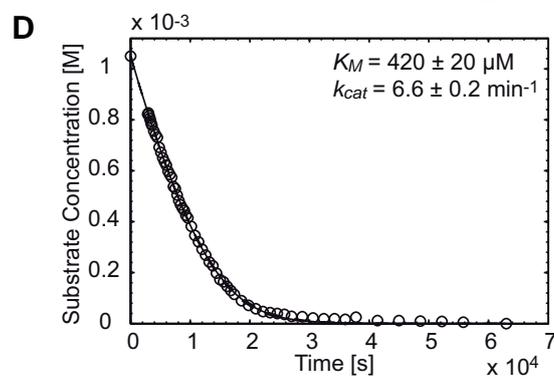
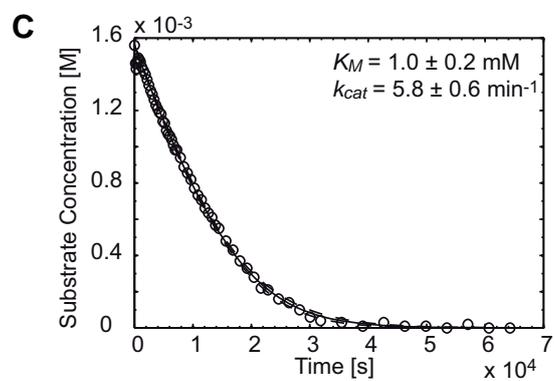
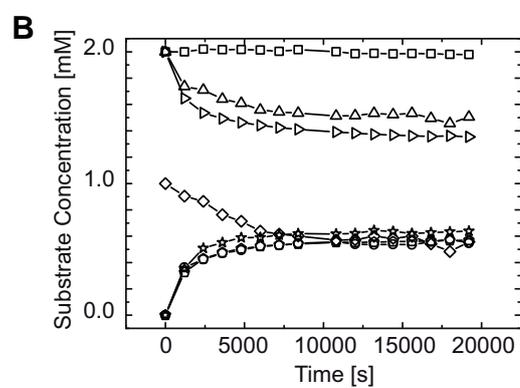
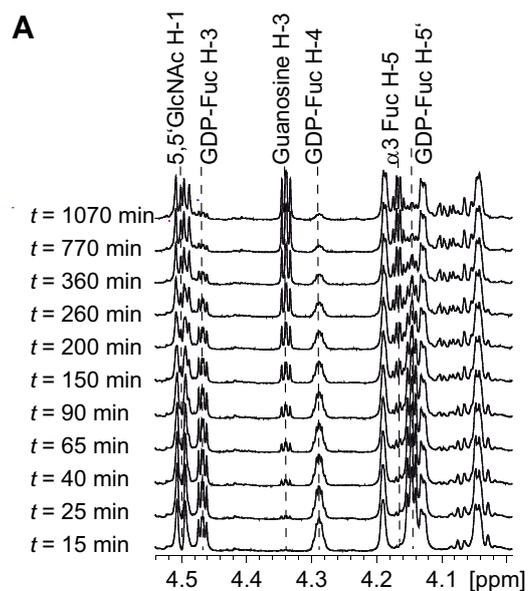


Figure 4-7: Kinetic characterization of sFucTA. A: Observation of the progress of the transfer of fucose residues. 1D  $^1\text{H}$  NMR spectra were recorded at different points in time to monitor the change of concentration of all four compounds (dashed lines) involved in the progress of the reaction (bi-bi enzyme). B: Product inhibition of sFucTA by Guanosine diphosphate. Shown is the time course of product generation (circles: H-1' of GDP, pentagons: H-6 of  $\alpha$ 1,3-Fuc, stars: H-1 of  $\alpha$ 1,3-Fuc) and consumption of substrate (triangles up: H-1' of GDP-Fuc, triangles right: H-6'' of GDP-Fuc, diamonds: GlcNAc H-1 $\alpha$  of heptasaccharide **1a**) as measured by 1D  $^1\text{H}$  NMR. Squares indicate the sum integral of H-1' of GDP and GDP-Fuc that was used as reference integral to determine concentrations from integrals. C, D Time course of the FucTA reaction in presence of alkaline phosphatase using octasaccharide **2a** (C) and heptasaccharide **1a** (D) as acceptor substrate. The integrated resonance of the H-1' of guanosine was used and subtracted from the initial substrate concentration (circles). The solid line indicates the non-linear Lambert-W fit, dashed lines indicate the standard deviation.

#### 4.1.5 Recognition of the immunogenic core $\alpha$ 1,3-fucosyl epitope by IgG

As outlined in the introduction, the serum of patients that show pronounced immune reactions against a certain antigen exhibiting core  $\alpha$ 1,3 fucose (and possibly additional  $\alpha$ 1,2 xylose) also reacts with other antigens carrying these so-called carbohydrate cross-reactive determinants (CCDs). However, the structural basis for the existence of carbohydrate cross-reactive determinants (CCDs) remains elusive. This is certainly to a large part owed to the difficult accessibility of appropriate amounts of purified core  $\alpha$ 1,3-fucosylated complex type oligosaccharides. Yet, with the tools of recombinant FucTA from honeybee and of purified, 1- $\beta$ -*N*-acetylated substrates, enzymatic synthesis of the desired structures was convenient. From the two 1- $\beta$ -*N*-acetylated and core  $\alpha$ 1,3-fucosylated compounds presented here, the bifucosylated structure **4b** proved to be easier available in mg amounts for two reasons: First, the yields of precursor **3b** are higher, as the source (porcine fibrinogen) seems to exhibit a more uniform glycosylation than the source of **1b** (bovine fibrinogen, cf. Ref. [4]). Second, purification of **4b** by means of PGC HPLC from the precursor is much easier due to a large difference in retention times. Hence, despite the complication of an additional carbohydrate unit that renders complete assignment of resonances challenging, the nonasaccharide **4b** was chosen for epitope mapping by means of STD NMR. The respective IgGs were prepared by the group of E. Spillner: As a first step, rabbits were immunized with horse radish peroxidase that is a plant glycoprotein with uniform glycosylation carrying core  $\alpha$ 1,3 fucose. The serum from these rabbits was subjected to affinity purification with a bromelain-conjugated column yielding a fraction that is enriched with polyclonal anti-bromelain IgG. The epitope of **4b** to these polyclonal antibodies was analyzed by STD NMR and is shown in Figure 4-8. Although STD effects could be quantified only for a part of all resonances due to severe overlap of signals, at least one resonance serving as probe could be evaluated for each carbohydrate unit. The STD effects showed pronounced differences. In an analog experiment involving serum of

rabbits immunized with a different antigen, no STD effects were observed. The overall epitope shows a clear emphasis at the core region, in particular at GlcNAc1. Additionally, the  $\alpha$ 1,3-linked fucose receives higher saturation than the  $\alpha$ 1,6-fucose. Interestingly, also the branches exhibit differences in their proximity to the IgG binding site: The  $\alpha$ 1,6-linked branch receives higher saturation, although located opposite of  $\alpha$ 1,3-linked fucose.

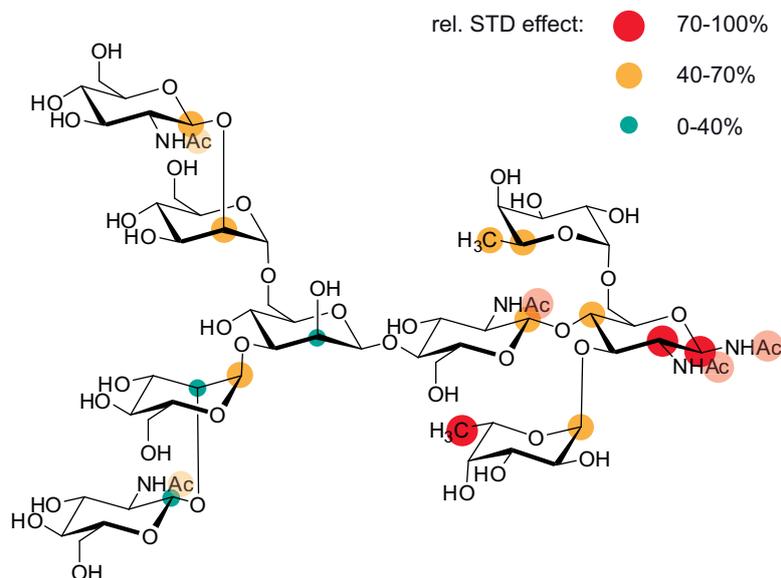


Figure 4-8: Ligand epitope of the bifucosylated complex type saccharide **4b** binding to affinity purified, polyclonal anti-bromelain IgG mapped by means of STD NMR. Due to severe signal overlap in the 1D  $^1\text{H}$  spectrum of the nonasaccharide, only the STD effects of a part of all assigned resonances could be analyzed. The relative STD effects of the acetate functions are indicated in pale colors. The STD effects of these resonances are generally higher than those of the ring protons due to their longer  $T_1$ . Hence, their relative STD effects were calculated with respect to the largest STD effect of all acetates whereas relative STD effects of the ring protons refer to the largest STD effect of all ring protons. Therefore, relative STD effects of the acetates cannot directly be compared to those of the ring protons. The distribution of STD effects from the acetate protons over the molecule support an emphasis at the core region that supports the epitope derived from the ring protons.

These results show that IgGs that are directed against glycoproteins are able to bind the pure carbohydrate compound and do not necessarily require the protein parts for binding. In addition, the core region is specifically recognized by these IgGs and it shows the highest STD effects. Because the serum used is polyclonal, the epitope represents an average of collective binding events from all IgG clones binding **4b** in the serum. Further studies are aimed to characterize the binding epitope of monoclonal anti-CCD IgGs. Comparison to each other and to the epitope with polyclonal serum as presented here will help to further resolve the mechanism behind cross reactivity of the immunogenic core  $\alpha$ 1,3-fucose epitope.

## **4.2 Donor substrate binding and enzymatic mechanism of human core $\alpha$ 1,6-fucosyltransferase (FUT8)**

In this section, a model of donor substrate binding for FUT8 is presented. The model is based on the crystal structure of FUT8 and the crystal structure of *ce*POFUT in complex with GDP-Fuc together with results from STD NMR and SPR studies of the ligands' binding. The experimental information was combined with *in silico* studies to create the first model of FUT8 in complex with GDP-Fuc. The results provide detailed information about enzyme–substrate contacts and about the catalytic mechanism of FUT8. Text and figures presented in this section have been published in *Biochimica et Biophysica Acta* (DOI: 10.1016/j.bbagen.2012.08.018.; cf. Ref. [108] ) and are reproduced with permission from Elsevier.

### **4.2.1 A modified expression protocol gives enhanced yields of purified FUT8**

In cooperation with the group of E. Spillner, FUT8 was cloned and expressed as soluble enzyme lacking the transmembrane region with minor modifications as described [6]. In contrast to the recombinant FUT8 published by Ihara *et al.*, the construct carries the 10fold histidine affinity tag as well as the V5-epitope at the N-terminus as it is known that the catalytic domain is located near the C-terminus in most glycosyltransferases. Furthermore, the expression applying the baculovirus-mediated infection of Sf9 insect cells was conducted in serum-free medium. The modified protocol yielded soluble and secreted recombinant protein and allowed for direct His-tag affinity purification of FUT8 from the supernatant after exchange of the buffer. (Figure 4-9) Thus, no precipitation step was necessary and enhanced yields of purified FUT8 at 5  $\mu$ g/mL culture supernatant were obtained. The identity of recombinant FUT8 was verified by Mascot MS analysis, SDS-PAGE and Western blot analysis.

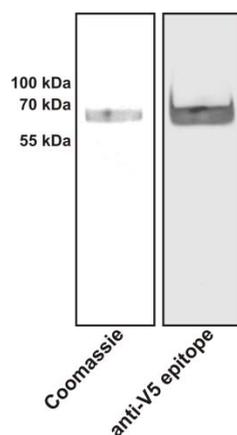


Figure 4-9: SDS-PAGE analysis of purified recombinant FUT8. The expression and purification protocol yielded one band in the coomassie stained gel (left) at the expected mass (62 kDa). The identity of the expression product was verified by Western-Blot analysis using a monoclonal antibody against the V5 epitope (right).

Analysis by MALDI TOF MS gave no evidence for the presence of dimers or oligomers of FUT8. This is in agreement with earlier data, including gel filtration analysis, on FUT8 isolated from human blood platelets [40]. Therefore, it can be assumed that FUT8 is present as a monomer in solution. Many other glycosyltransferases are known to be present as dimer or tetramer. The MS analysis here was only intended to confirm that the construct presented here has the same behavior as the native protein.

The enzymatic characterization with progress curve analysis using  $^1\text{H}$  NMR [52, 105] allows the determination of  $K_M$  and  $k_{cat}$  without special probing, quenching or work-up procedures. The change of substrate concentration with time is described by the integrated form of the Michaelis–Menten kinetics [87-89] that has the Lambert- $W$  function as solution.[90, 91]. The Lambert- $W$  function can be fitted to the data using a non-linear fitting routine. The kinetic constants were obtained as  $K_{M,acceptor} = 12 \mu\text{M}$  and  $k_{cat} = 0.45 \text{ s}^{-1}$ . (Figure 4-10) These values are in perfect agreement with those reported by Ihara *et al.* [39] using fluorescently labeled substrates. It can therefore be concluded that the His-tag at the N-terminus present in the construct has no impact on the activity of FUT8. The data presented here were obtained in less than an hour using 0.3 nmol (= 19  $\mu\text{g}$ ) of enzyme.

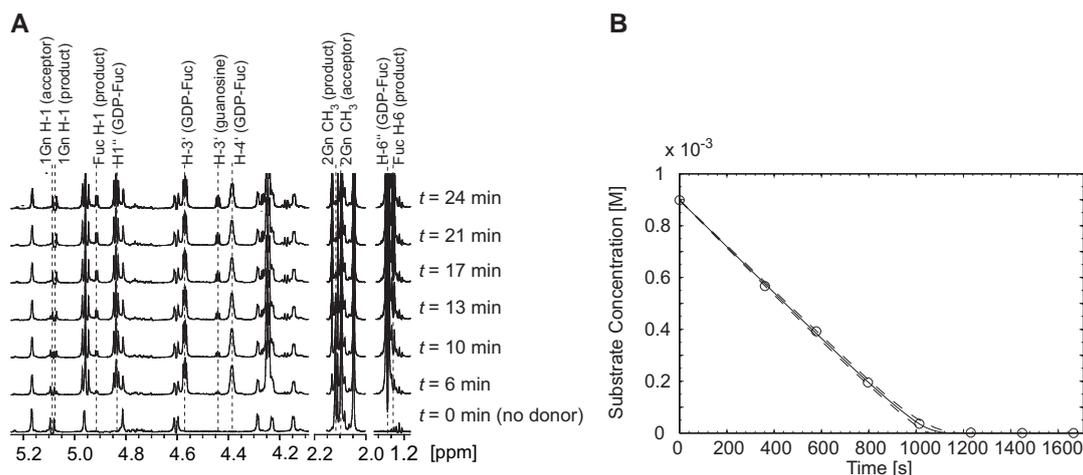


Figure 4-10: Enzyme kinetic characterization of FUT8 by means of progress curve analysis using NMR detection of products and substrates. A: Selected sections of <sup>1</sup>H NMR spectra recorded at different time points of the reaction (for clarity, the spectrum at 27 min is not shown). Resonances that were used to determine substrate and product concentrations in the region between 1.0 and 5.2 ppm are indicated by dotted lines. The well-dispersed resonances of GDP-Fuc (H-8, H-1', H-3', H-4', H-1'', H-6''), guanosine (H-8, H-1', H-3'), heptasaccharide **1b** (GlcNAc1 H-1, GlcNAc2 CH<sub>3</sub>) and the fucosylated product **2** (GlcNAc1 H-1, Fuc H-1, GlcNAc2 CH<sub>3</sub>, Fuc H-6) were integrated and used for determination of substrate concentrations. The sample contained 2 μM FUT8, a starting concentration of 0.9 mM acceptor substrate and 3 mM donor substrate, respectively, 10 U alkaline phosphatase and 1 mg/mL BSA in 160 μL deuterated Mes-NaOH buffer, pH 7.0. B): Plot of the concentration of acceptor substrate versus time, i.e. progress curve, of the FUT8 reaction using the integral of the 2-*N*-acetyl resonance of GlcNAc1. The Lambert-*W* fit yields the enzyme kinetic constants  $K_{M,acceptor} = 12 \mu\text{M}$  and  $k_{cat} = 0.45 \text{ s}^{-1}$ . The dashed lines give the confidence interval obtained from the fit.

#### 4.2.2 SPR and STD NMR elucidate recognition of GDP-Fuc

In order to gain insight into substrate recognition of FUT8, ligand-based NMR experiments were performed. Saturation transfer difference (STD) NMR [92-94] allows the determination of ligand epitopes without requirement of large protein amounts or labeling of the protein. Figure 4-11 shows the epitope of GDP-Fuc in complex with FUT8. The different longitudinal relaxation constants of the different protons of GDP-Fuc may bias the absolute STD percentages measured and thus affect the epitope. Therefore, the STD percentages that were directly obtained from the experiment were corrected by the respective  $T_1$  constants as described. [109] Clearly the nucleoside protons acquire relatively high saturation, indicating close proximity to the protein surface in the complex. The fucose moiety, in contrast, shows only minor saturation. Due to the lack of covalently linked protons, the contribution of the pyrophosphate group to the epitope remains elusive. These findings complement the results of inhibition studies by Ihara *et al.* They found that GDP is a strong inhibitor for FUT8 ( $K_I = 3.6 \mu\text{M}$ ), whereas the absence of the β-phosphate group weakens the inhibitor potential almost by a factor of 1000 ( $K_{I,GMP} = 2.8 \text{ mM}$ ). Taken together, the results suggest a binding mode where the nucleobase and the ribose direct the binding process to the enzyme. The

fucosyl part, per contra, is positioned farther from the enzyme surface and is supposedly located towards the shallow acceptor binding site.

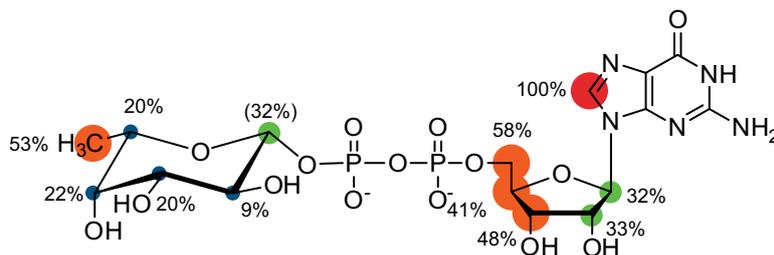


Figure 4-11: Epitope of GDP-Fuc in complex with FUT8 as determined by STD NMR. The size of the relative STD effect of a resonance is proportional to the distance of the corresponding proton to the enzyme surface. Protons of the guanosine moiety receive highest saturation. In contrast, the protons of the fucose moiety show only minor STD effects, indicating that the fucose is located farther away from the FUT8 surface and thus is surrounded by solvent.

The dissociation constants for the complexes of FUT8 with its substrates and parts of the donor substrate (GDP, GMP and guanosine) were determined with SPR (c.f. Figure 4-12 and Figure 4-13). The obtained  $K_D$  values are listed in Table 4-3. The data shows that the affinity of GDP is 10-fold higher compared to that of GDP-Fuc. These findings are in good agreement with the inhibition constants determined by Ihara *et al.* [39] A fit of the one-site-binding model to the SPR data of guanosine and GMP yields  $K_D$  values of approximately 2 mM for both ligands, indicating very weak binding (data not shown). From these data it becomes clear that the  $\beta$ -phosphate moiety has a major contribution to the binding affinity of the donor substrate. This is in stark contrast to other glycosyltransferases where the base is the dominant part of the donor substrate [8,10].

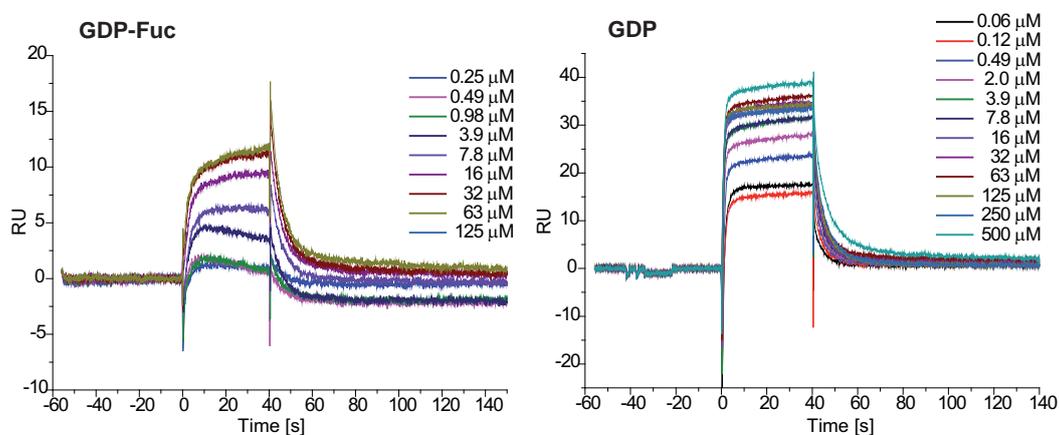


Figure 4-12: Sensorgrams for GDP and GDP-Fuc. The data shown are the native response curves minus the curve of the buffer signal and was multiplied by  $-1$ .

## RESULTS AND DISCUSSION

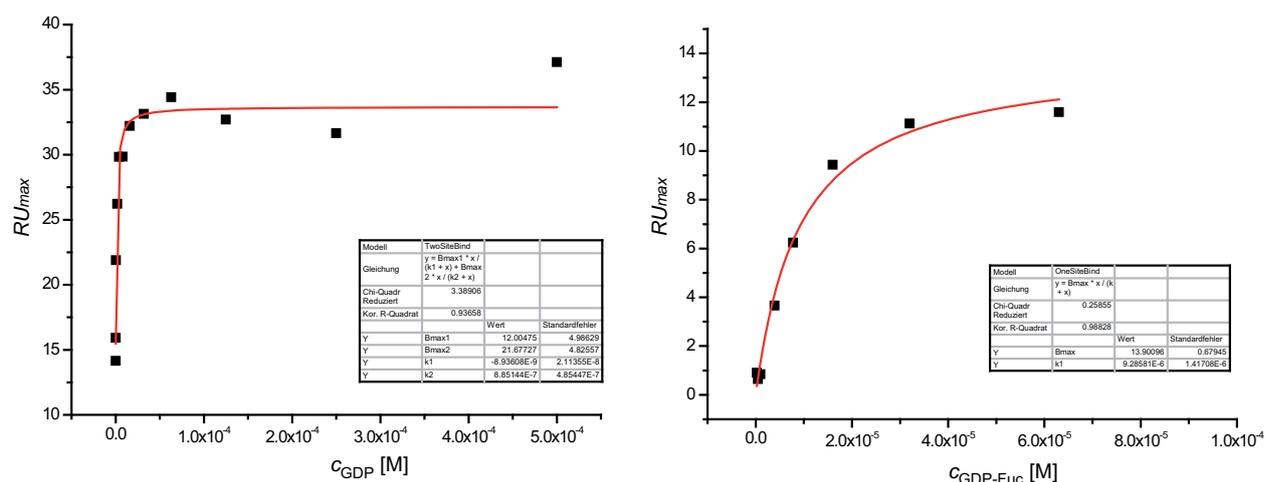


Figure 4-13: Affinity plots of the SPR binding analysis GDP (left) and GDP-Fuc (right) (note the different scales of the x-axes). All plots were fitted with a two-site binding model in order to accommodate contributions of unspecific binding in the assay. The data plotted for guanosine and GMP represent dominantly unspecific binding.

Table 4-3:  $K_D$  values for FUT8 and ligands occupying the binding site of the donor substrate as determined by SPR.

Ligand	$K_D$ [mM]
<b>GDP-Fuc</b>	0.0093 ( $\pm 0.001$ )
<b>GDP</b>	0.00089 ( $\pm 0.0005$ )
<b>GMP</b>	Unspecific binding
<b>Guanosine</b>	Unspecific binding

In order to obtain information about the active conformation of GDP-Fuc bound to FUT8, transferred NOE experiments were performed at various protein to ligand ratios, mixing times, temperatures and pH values. However, the conditions that allow observation of transferred NOEs were not met. The problems that prevent observation of trNOEs are the very fast hydrolysis of the donor substrate by FUT8 even at moderate temperatures combined with a slow off-rate of  $0.2 \text{ s}^{-1}$  of GDP-Fuc (cf. Figure 4-14 and Table 3-1).

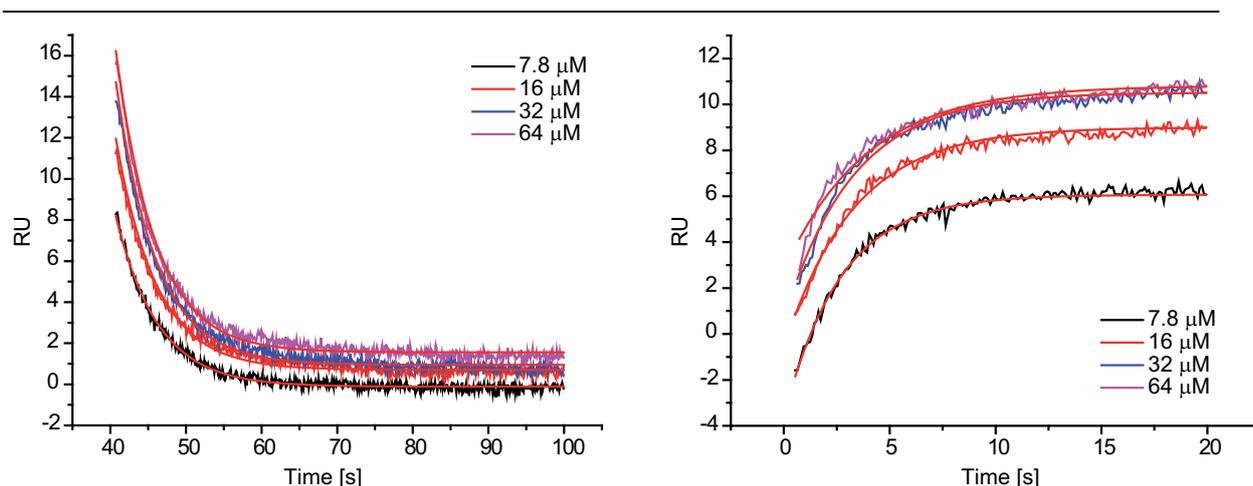


Figure 4-14: Dissociation (left) and association (right) part of SPR sensorgrams of GDP-Fuc.  $k_{\text{off}}$  was obtained by fitting an exponential decay to the curves. Fitting the association curves yields  $k_{\text{obs}}$ , from which  $k_{\text{on}}$  can be calculated.

Table 4-4: Kinetic parameters of GDP-Fuc binding to FUT8 obtained from the fits in Figure 4-14.  $k_{\text{on}}$  was calculated by fitting the association phase of the sensorgrams to yield  $k_{\text{obs}}$ , from which  $k_{\text{on}}$  was calculated by subtracting  $k_{\text{off}}$  and subsequent division of the result by the ligand concentration. The  $K_{\text{D}}$  from kinetic data was calculated by division of  $k_{\text{off}}$  by  $k_{\text{on}}$ .

$c$ [ $\mu\text{M}$ ]	$t_{\text{off}}$ [s]	$k_{\text{off}}$ [1/s]	$t_{\text{on}}$ [s]	$k_{\text{obs}}$ [1/s]	$k_{\text{on}}$ [L/(s·mol)]	$K_{\text{D,from kinetics}}$ [ $\mu\text{M}$ ]
64	5.26	0.190	3.90	0.256	1050	180
32	5.37	0.186	3.27	0.306	3740	49.8
16	5.39	0.186	3.29	0.304	7400	25.1
7.8	5.57	0.180	3.90	0.256	9610	18.7
Average	5.40	0.185	3.59		5450	68.6

### 4.2.3 Modeling of the FUT8–GDP-Fuc complex reveals enzyme–substrate contacts

The crystal structure of the apo enzyme FUT8 [44] was used as a starting point for an *in silico* model of FUT8 in complex with its two substrates. In addition to the usual preparation steps of the structure for *in silico* calculations, five amino acids (Asn368–Thr372) were modeled at the tip of a loop that were not resolved in the crystal structure. In order to obtain a reliable position for the donor molecule GDP-Fuc, model construction was started from the X-ray crystal structure of *ce*POFUT in complex with its donor GDP-Fuc. The binding sites of the donor molecule in *ce*POFUT and FUT8 are structurally homologous [51]. The amino acids of the donor binding site in *ce*POFUT were defined by the distance of the amino acids to the donor molecule of less than 4 Å. Four peptide segments with a total of 48 residues were obtained and their structural motifs were structurally aligned with the backbone of the crystal structure of FUT8 (cf. Figure 4-15). The peptides Arg40-Leu58, Pro233-Arg240, Leu348-Asn352 and Val354-Gly369 of *ce*POFUT were simultaneously aligned with peptides Gly219-

Thr237, Pro358-Arg365, Phe462-Thr466 and Ser468-His483 of FUT8). The two long sequences are forming  $\alpha$ -helices and the two short  $\beta$ -strands. The backbones of the two enzymes matched within these five peptide segments with an RMSD of 1.49 Å (cf. Figure 4-15).

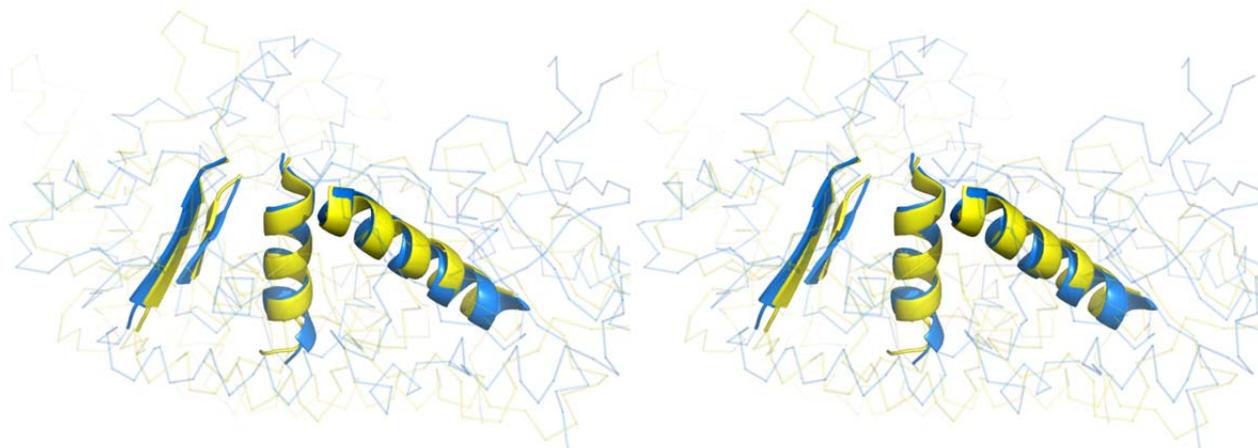


Figure 4-15: Ribbon representation of structural alignment of the 3D structures of hFUT8 (yellow ribbon) and *ce*POFUT (blue ribbon). The sections that were selected for the alignment are emphasized by opaque cartoon ribbons.

Now the donor molecule was placed into the FUT8 by using the exact same positioning relative to the five peptide segments as was found in *ce*POFUT. As a result, GDP-Fuc was prepositioned in the putative binding pocket of the donor substrate of FUT8. The side chain of Arg365 is folded onto the enzyme surface in the FUT8 structure, probably because there is no substrate present. The conformation of the side chain of Arg365 in the apo structure of FUT8 results in clashes of the atoms of Arg365 with the donor molecule. This explains why all previous attempts to place GDP-Fuc into the FUT8 structure via docking algorithms were unsuccessful. However, if the donor substrate is present and the side chain of Arg365 of FUT8 is rotated into the conformation of the Arg240 of *ce*POFUT, a perfect fit is obtained. Subsequent energy minimization of the FUT8/GDP-Fuc complex yielded a ‘homology’ model of GDP-Fuc bound to FUT8 derived from the structure of the *ce*POFUT complex as shown in Figure 4-16. The guanine is bound via hydrogen bonds to two key residues, Asp453 and His363, which form a conserved binding motif as these key interactions are also found in the crystal structures of *ce*POFUT [51] and NodZ [45]. In the cases of *ce*POFUT and NodZ, the ribose is mainly bound via hydrophobic interaction with a phenylalanine residue. In FUT8, this residue is substituted by valine residues forming the hydrophobic pocket for the ribose part. FUT8 also exhibits a tyrosine residue Tyr250 corresponding to the Tyr45 in NodZ. Tyr45 in NodZ forms H-bonds with the 2' and 3' OH groups of GDP-Fuc. [45] In FUT8 a

similar interaction between the OH of Tyr250 and ribose OH-2' and OH-3' is found during the MD simulation. The pyrophosphate moiety in the FUT8 model is bound via multiple hydrogen bonds to the backbone NH atoms of the peptide chains forming two  $\alpha$ -helices of the Rossman fold and the side chain hydroxy group of the essential Ser469. This binding mode of the pyrophosphate is very similar to that observed in the X-ray crystal structure of *ce*POFUT. Arg365 forms hydrogen bonds to the  $\beta$ -phosphate oxygen atoms and the 5'' and 1'' oxygen atoms of the fucose of GDP-Fuc and is also very close to the O-4''. (Figure 4-16) The fucose moiety lies outside the donor substrate binding cavity and points towards a shallow area formed by the peptide segment from Asp494 to Gly501. The methyl group of Thr367 is close enough to interact with the methyl group of the fucose and probably gives rise to the high STD effects observed for these protons.

### Molecular dynamics simulation

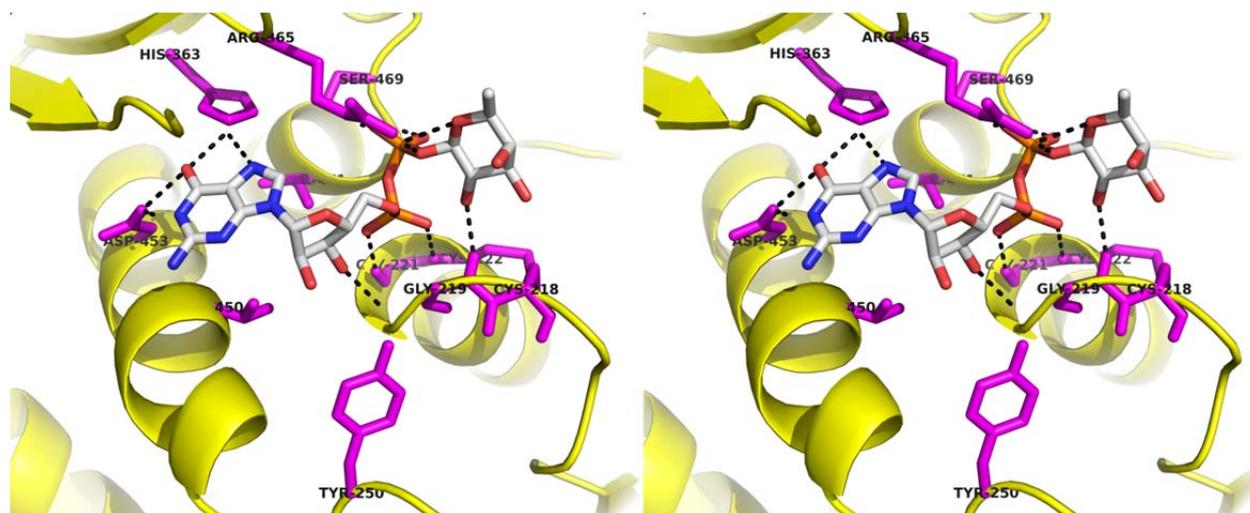


Figure 4-16: Crossed-eye stereo plot of the initial placement of GDP-Fuc (sticks by atom color) into FUT8 (yellow cartoon ribbon) as described in the text. Residues that directly interact with GDP-Fuc according to the model are shown in magenta. The structure was used as a starting point for the MD simulation.

The positioning of the donor substrate was validated by a molecular dynamics simulation. The complex of donor substrate and protein was embedded into water box containing also 200 mM sodium chloride. After minimization of the water box a 1.5 ns MD simulation time was run. During that time the position of the donor did not change significantly, indicating that the positioning was reasonable (cf. Figure 4-23). For the donor substrate, the distances between the key amino acids (cf. Figure 4-22 and Table 4-5) and their hydrogen bonding partners in the nucleotide part of GDP-Fuc remained perfectly constant. Especially the four hydrogen bonds binding the  $\beta$ -phosphate constantly kept their distance to the corresponding donor atoms between 2.5 and 3.0 Å, underlining the significance for binding of the donor

## RESULTS AND DISCUSSION

substrate that was also found experimentally (cf. above). Remarkably, new unpredicted interaction established after 200 ps simulation time. Namely, hydrogen bonds between the nitrogen atoms of the guanidinium group of Arg365 and the O-1" and O-5" of the fucosyl moiety of GDP-Fuc built up and remained for the rest of the simulation time, locking the fucosyl residue into a position perfectly suited for the nucleophilic attack.

Table 4-5: Interactions of atoms of FUT8 with atoms of GDP-Fuc that are observed during the MD simulation.

Residue	Atom	GDP-Fuc Atom	Type of Interaction	Average distance [Å]	Fraction of Frames with distance < 3.6 Å
Asp453	O $\gamma$	N-1	H-bond	2.9 $\pm$ 0.3	96%
Asp453	O $\gamma$	N-2	H-bond	2.9 $\pm$ 0.3	99%
Gly449	CO	N-2	H-bond	4.0 $\pm$ 0.4	19%
His363	N $\epsilon$	O-6	H-bond	3.0 $\pm$ 0.4	93%
His363	N $\epsilon$	N-9	H-bond	3.5 $\pm$ 0.4	68%
Val471	C $\gamma$	C-1'	Hydrophobic	4.5 $\pm$ 0.2	0%
Tyr250	O $\eta$	O-2'	H-bond	3.8 $\pm$ 0.4	32%
Gly221	NH	O-3'	H-bond	3.3 $\pm$ 0.3	81%
Tyr250	O $\eta$	O-3'	H-bond	4.3 $\pm$ 0.4	3%
Cys222	NH	P $\alpha$ O-1	H-bond	3.3 $\pm$ 0.2	86%
Gly221	NH	P $\alpha$ O-1	H-bond	3.3 $\pm$ 0.2	86%
Tyr220	NH	P $\alpha$ O-1	H-bond	3.3 $\pm$ 0.2	88%
Gly219	NH	P $\alpha$ O-2	H-bond	2.9 $\pm$ 0.2	100%
Gly221	NH	P $\alpha$ O-2	H-bond	4.4 $\pm$ 0.3	0%
Arg365	N $\epsilon$	P $\beta$ O-1	H-bond	2.7 $\pm$ 0.1	100%
Ser469	O $\gamma$	P $\beta$ O-1	H-bond	2.7 $\pm$ 0.1	100%
Gln470	NH	P $\beta$ O-2	H-bond	2.8 $\pm$ 0.1	100%
Gln470	N $\epsilon$	P $\beta$ O-2	H-bond	2.8 $\pm$ 0.2	100%
Arg365	N $\eta$	O-1"	H-bond	2.6 $\pm$ 0.1	100%
Gln470	N $\epsilon$	O-2"	H-bond	3.1 $\pm$ 0.2	96%
Arg365	N $\eta$	O-4"	H-bond	4.8 $\pm$ 0.3	0%
Arg365	N $\eta$	O-5"	H-bond	3.3 $\pm$ 0.3	87%
Thr367	C $\gamma$	C-6"	Hydrophobic	4.5 $\pm$ 0.6	1%

## RESULTS AND DISCUSSION

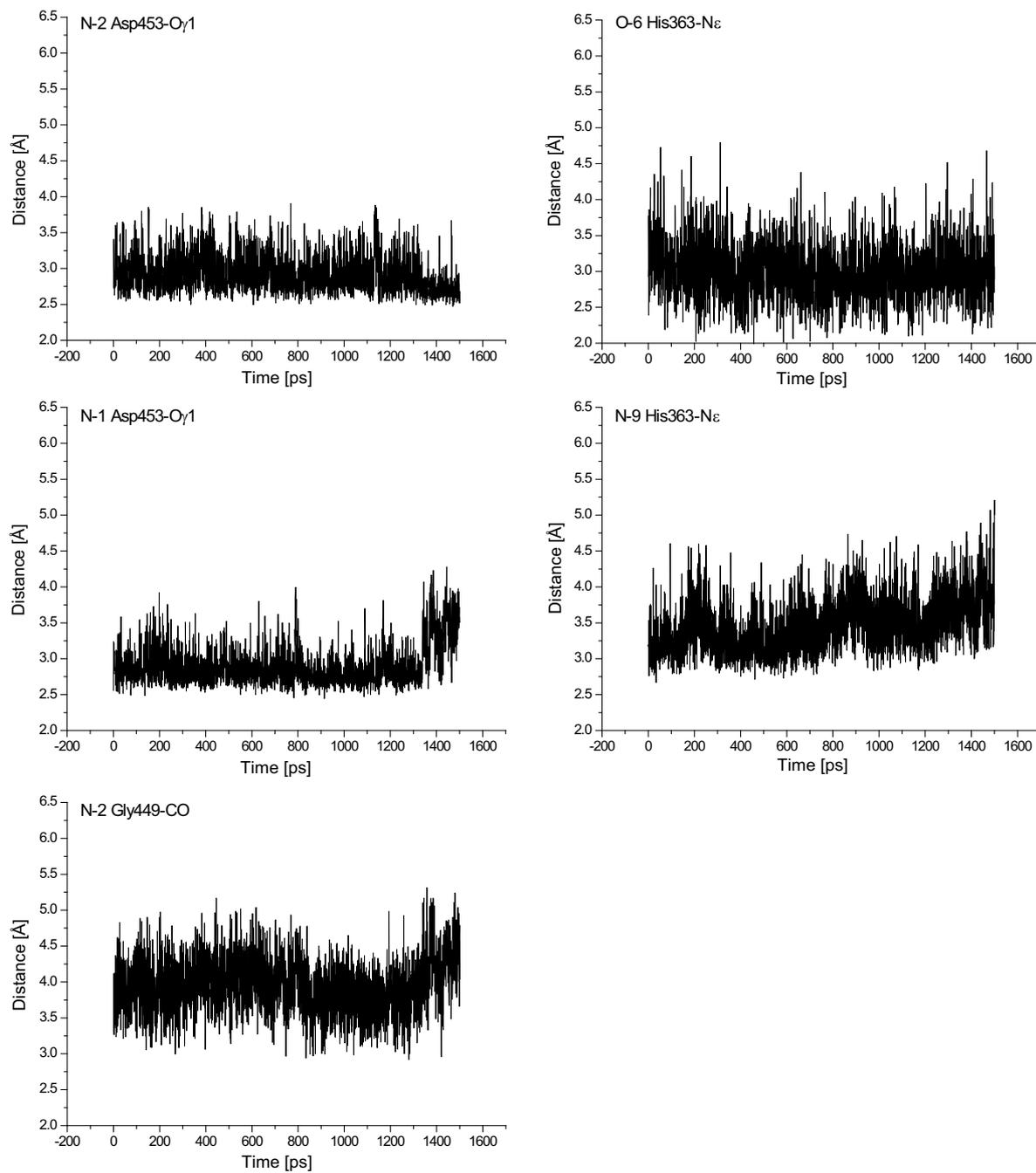


Figure 4-17: Trajectories of distances between atoms of the guanine part of GDP-Fuc and FUT8 during the MD simulation.

## RESULTS AND DISCUSSION

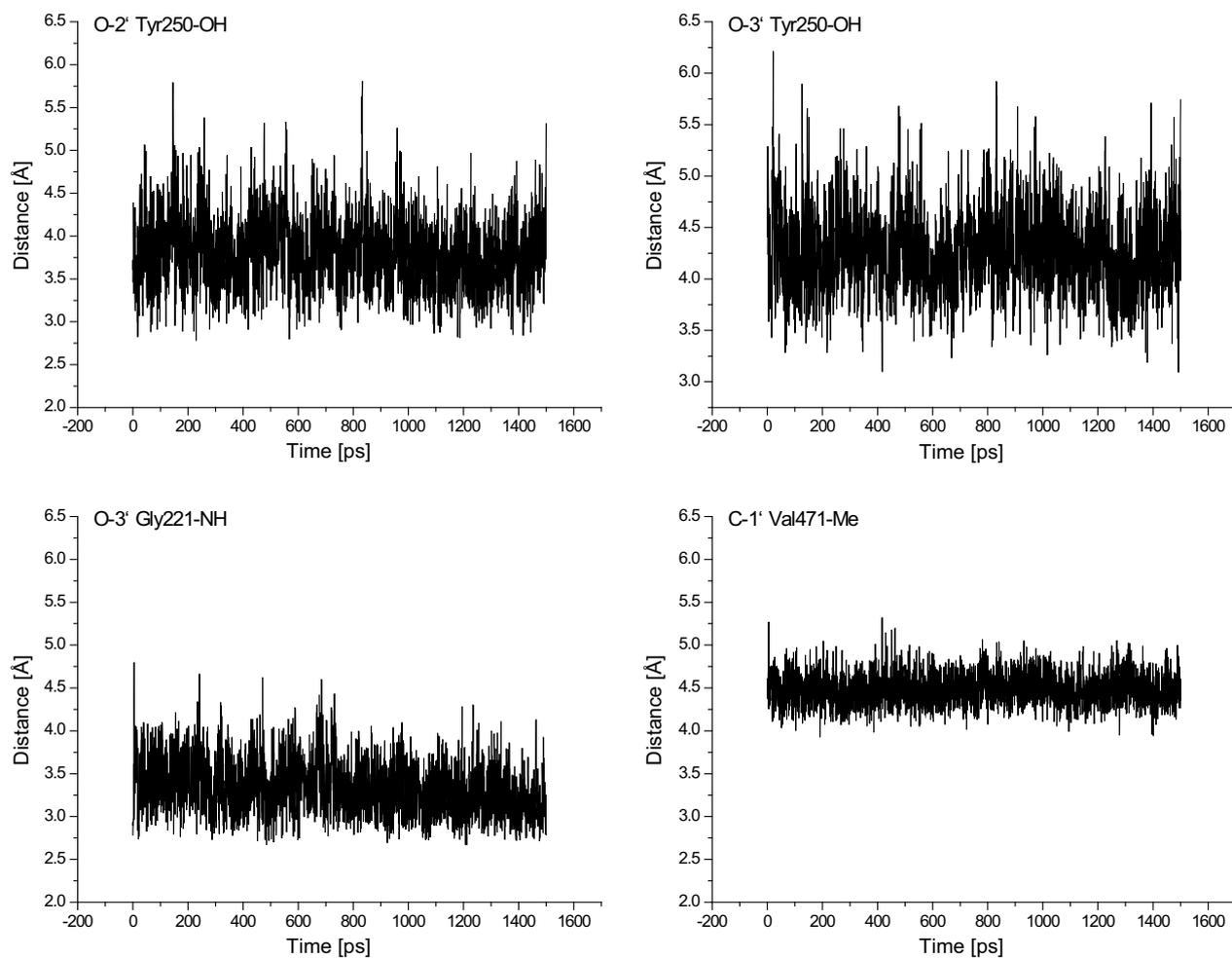


Figure 4-18: Trajectories of distances between atoms of the ribose part of GDP-Fuc and FUT8 during the MD simulation.

## RESULTS AND DISCUSSION

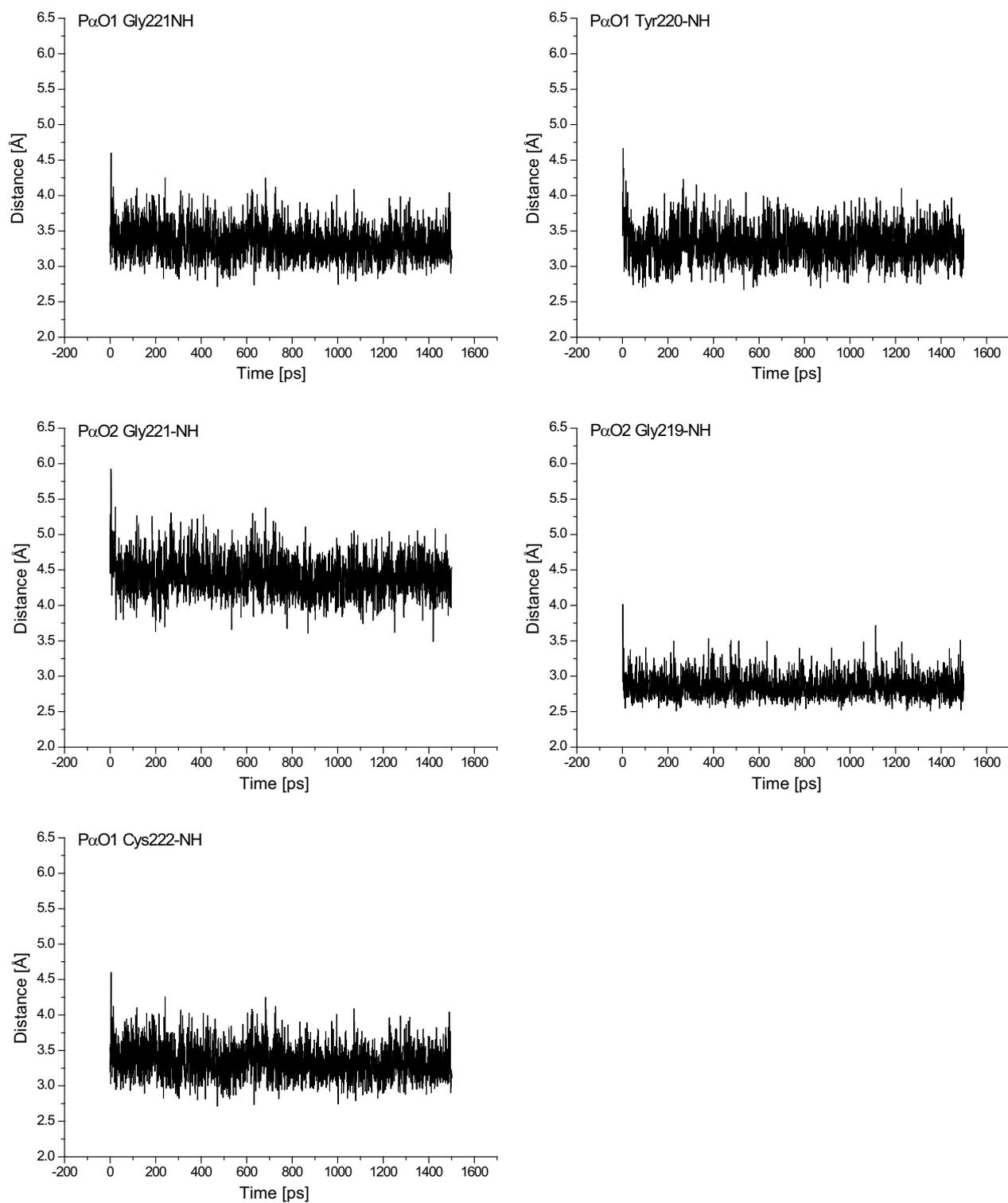


Figure 4-19: Trajectories of distances between atoms of the  $\alpha$ -phosphate part of GDP-Fuc and FUT8 during the MD simulation.

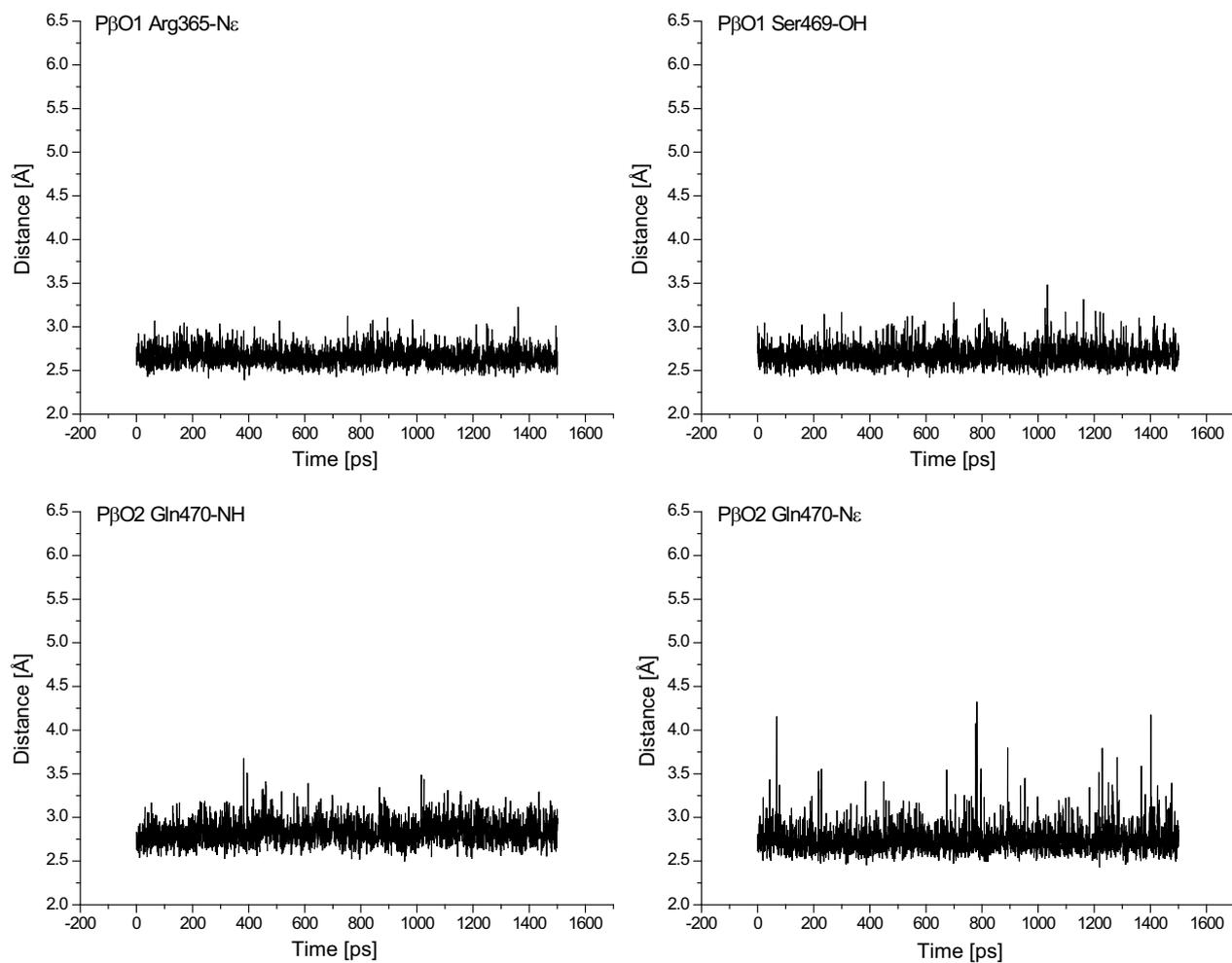


Figure 4-20: Trajectories of distances between atoms of the  $\beta$ -phosphate part of GDP-Fuc and FUT8 during the MD simulation.

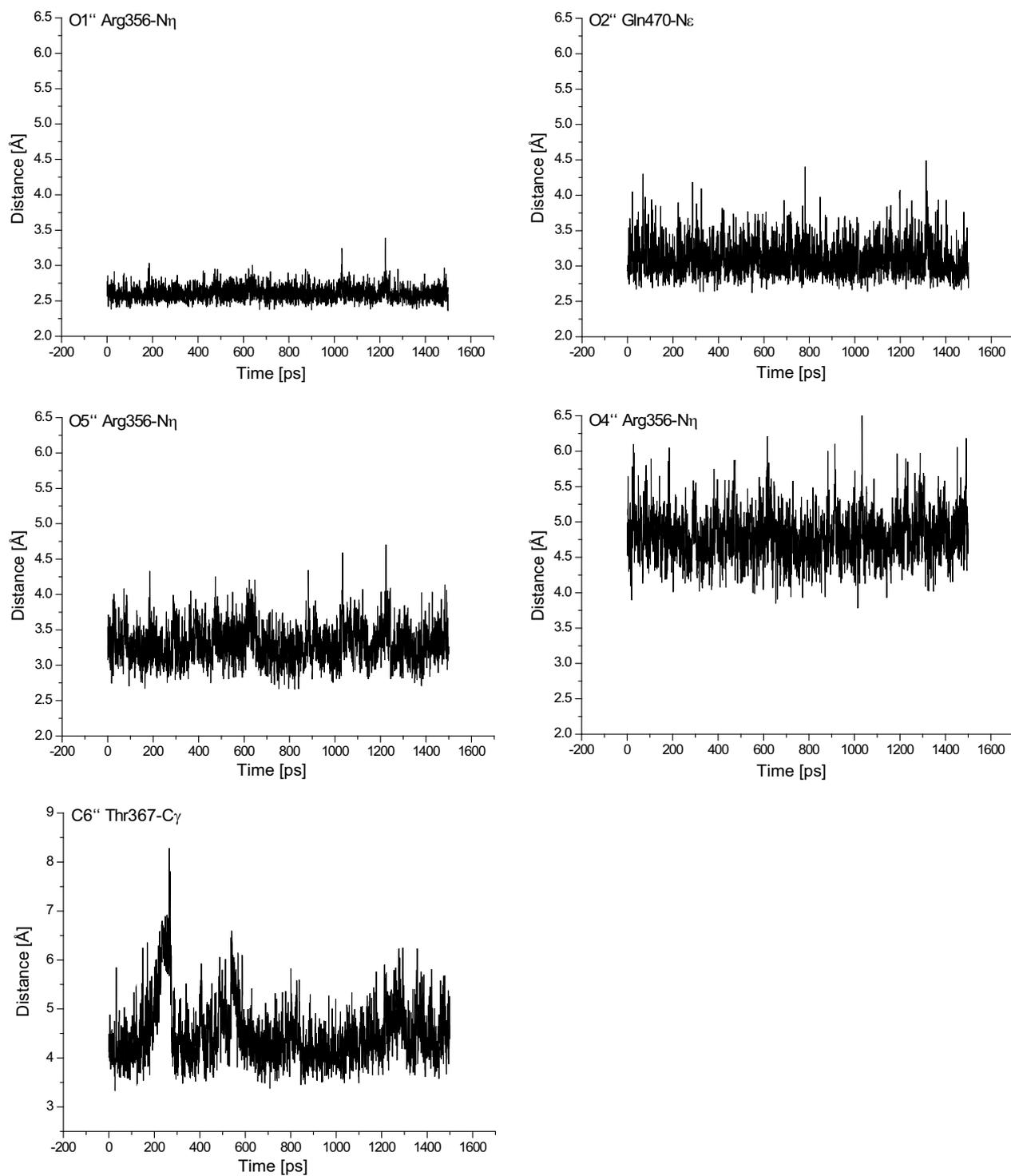


Figure 4-21: Trajectories of distances between atoms of the fucosyl part of GDP-Fuc and FUT8 during the MD simulation.

#### 4.2.4 Nucleoside and $\beta$ -phosphate direct binding of GDP-Fuc

The recognition of the guanine is maintained by the side chain functions of two key residues: Asp453 and His363 both form stable hydrogen bonds with amino groups of the guanine (cf. Figure 4-22). In addition, hydrophobic contacts to the base of the binding pocket formed by two valine residues are also involved in binding. This recognition process is highly conserved

among structural relatives of FUT8, i.e. *ce*POFUT and NodZ. Surprisingly, inhibition studies [39] and the SPR results clearly indicate that the guanine part of GDP-Fuc has only a minor contribution to the total affinity of the donor substrate. Furthermore, it was previously shown that Asp453 is essential for activity of FUT8, but not His363, although a drop in activity by 50% is observed when His363 is mutated to alanine [44]. It can therefore be assumed that these residues are important for the specific recognition of the nucleotide, e.g. for the discrimination from other nucleobases. This finding is also reflected by inhibition studies with other nucleic acid diphosphates [39]. The purine diphosphates adenosine 5'-diphosphate (ADP) and xanthosine 5'-diphosphate (XDP) do not inhibit FUT8 in a competitive manner. Adenine lacks a hydrogen bond donor in position 1 as well as the amine in position 2. Both groups of guanosine interact with Asp453. Xanthine, on the other hand, has a carbonyl group at position 2 that may interfere with Asp453. However, inosine 5'-diphosphate (IDP) is a weak millimolar competitive inhibitor. The inhibition constant in the mM range underlines the significance of the amine residue at position 2 present in guanosine and its interaction with Asp453. The data indicate that the base has approximately a millimolar interaction strength.

The interactions between amide functions of the backbone of the Rossmann fold, Ser469 and Arg365 with the  $\beta$ -phosphate unit provide the other center for affinity generation. Contacts with the  $\beta$ -phosphate do not allow discrimination from other pyrophosphates and thus do not add to the specificity towards the donor molecule. Sulfate or phosphate ions that are found at the location of the  $\beta$ -phosphate in the crystal structures of the apo enzymes of *ce*POFUT and NodZ underline this conclusion as well as the fact that GDP has a higher affinity to FUT8 than GDP-Fuc. Pyrophosphate itself was found to inhibit FUT8 in the mM range [39]. The facts that pyrophosphate and nucleobase have similar affinities in the millimolar range expectedly add up to the observed affinity of GDP of about 1  $\mu$ M.

The third part of the recognition is formed with the fucose moiety that interacts with the guanidinium group of Arg365 and with the side chain of Gln470. It has been suggested that both Arg365 and Arg366 play a major role in binding of GDP-Fuc and GDP via hydrogen bonding to the  $\beta$ -phosphate [38]. From the model, the structural basis for the importance of Arg365 is explained: The guanidinium group is found to exert binding in a twofold way throughout the MD simulation. The positively charged side chain binds the  $\beta$ -phosphate and the guanidinium group is in part responsible for binding of the fucose residue via hydrogen bonds. The bivalent binding mode that was observed is possible because of the size of the guanidinium group that allows for a large distribution of the positive charge. It was shown

previously that the residue Arg365 is essential for activity and even conservative substitutions like Arg365Lys lead to a complete loss of enzymatic action [44]. It remains elusive to which extent Arg365Lys mutants of FUT8 are able to bind GDP-Fuc. Corresponding data from *ce*POFUT indicate that the binding affinity towards GDP-Fuc is reduced accompanied by a complete abolishment of activity like in FUT8 [51]. The loss of catalytic activity induced by the mutation of arginine to lysine is discussed below.

The role of Arg366 seems to be completely different: Arg365, Asp309, Arg366 and Asp410 form an Arg-Asp ladder, which stabilizes the binding site and orient the guanidinium group of Arg365. No direct interaction of Arg366 with the substrate was observed in the model and the distance and orientation of Arg366 relative to Arg365 in the FUT8 crystal structure renders a direct interaction of Arg366 with the substrate very unlikely. In addition, Arg366Lys and Arg366Ala mutants exhibit both residual enzymatic activity as well as binding affinity [38]. This interpretation is supported by the fact that Arg366 is not conserved in *ce*POFUT (Asn) and NodZ (His). Since these enzymes differ significantly in 3D structure of the direct environment of the binding site but not in the residues interacting directly with GDP-Fuc, Arg366 as well as the referred aspartate residues probably have structural functions. In *ce*POFUT, the structural aspects of the amino acids Asp409, Arg366 and Asp410 are taken over by other amino acid residues involving a tryptophan residue.

The interactions of Arg365 and Gln470 with the fucose of GDP-Fuc are indispensable for specific recognition of GDP-Fuc compared to GDP-mannose (GDP-Man). The discrimination of GDP-Fuc and GDP-Man is accomplished via specific hydrogen bonds of Gln470 to the hydroxyl functions OH-2" and of Arg365 to O1", OH-4" and the O-5" of the fucose, respectively. GDP-Man and GDP-Glc were found to inhibit FUT8 in a competitive manner with a  $K_I$  tenfold of the  $K_M$  of GDP-Fuc but both substrates are not converted. [39] Because the mannose in GDP-Man differs in the configuration of the positions 2 and 4, possesses  $\alpha$ -configuration and is a D-sugar, the interactions with Arg365 and Gln470 are not possible.

The general binding model of FUT8 therefore suggests that recognition of GDP-Fuc is driven in a twofold manner: The high affinity to the nucleotide and the discrimination of the pyranose. In this respect, a comparison to other glycosyltransferases is highly interesting.

RESULTS AND DISCUSSION

Table 4-6: Glycosyltransferases for which the binding mode of the donor substrate has been elucidated via epitope mapping by STD NMR. In most cases additional binding studies have been employed in order to estimate fragments contributing highly to affinity or to determine the active conformation.

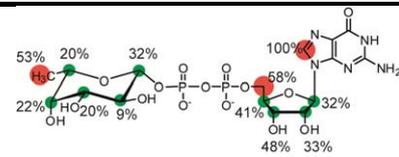
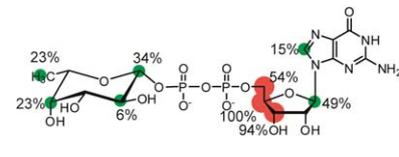
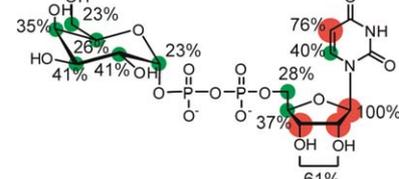
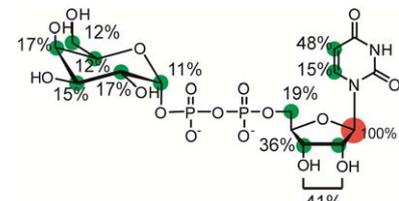
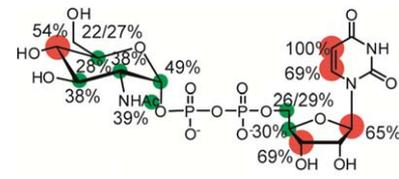
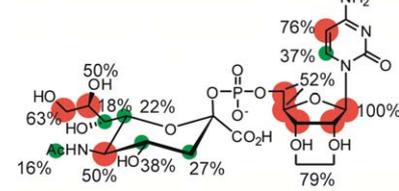
Glycosyltransferase	Donor Substrate	Epitope	Fragments Generating Affinity*	Reference
<b>FUT8 (human)</b>	GDP-Fuc		β-phosphate	This study
<b>FucTA (honeybee)</b>	GDP-Fuc		Nucleobase, ribose, pyrophosphate	This study
<b>GTB (human)</b>	UDP-Gal		Nucleobase, ribose, β-phosphate	[103]
<b>β4Gal-T1 (bovine)</b>	UDP-Gal		–	[110]
<b>GnT-V (human)</b>	UDP-GlcNAc		–	[111]
<b>ST6Gal-I (human)</b>	CMP-Sia		–	[112]

Table 4–6 lists studies that employed STD NMR and (in some cases) thermodynamic binding assays to elucidate the binding mode of donor substrates. Generally, the nucleobase is recognized strongly and comparison to the respective X-ray structures, if present, supports these findings. Interestingly, the results from STD NMR and measuring binding affinities of fragments lead to similar models of a ‘locked’ conformation of the pyranose residue of the

\* Determined by means thermodynamic binding studies (STD NMR, SPR or ITC) of fragments of the donor substrate

donor substrate. For the two galactosyltransferases Human Blood Group Galactosyltransferase B (GTB) [103] and bovine  $\beta$ 1,4-Galactosyltransferase I ( $\beta$ 4Gal-T1) [110] as well as for core  $\alpha$ 1,3-Fucosyltransferase (FucTA) [52] from honeybee, sugar nucleotides bind slightly weaker than the respective nucleotide diphosphate. As an explanation, it has been suggested that the flexible sugar moiety is trapped in the enzyme in its active conformation, leading to a substantial loss of conformational entropy.[103] If this loss is not compensated by binding enthalpy, a drop in affinity results i.e. for non-matching donor sugars like GDP-Man. From the epitopes, this general mechanism is visible from the very low STD effects at the protons of the pyranose units if a suitable nucleotide with a 'wrong' sugar are measured, indicating that the sugar has no contact to the enzyme in these complexes (data not shown in table). Interestingly, the two galactose transferring enzymes exhibit almost identical binding epitopes on the donor side even though one enzyme is inverting and the other is non-inverting, the anomeric configuration during the transfer. In contrast, the two fucosyltransferases show very different binding epitopes for the respective donor molecules with respect to binding of the base and the fucose residue (cf. Table 4–6). In FucTA, also a core-fucosyltransferase from insects, the nucleotide binding pocket seems to be substantially different from that of FUT8, as guanosine is bound as a minimal fragment with medium affinity. In FucTA, the affinity is more pronounced towards the guanine and less to the  $\beta$ -phosphate. In addition, the differences in binding kinetics between the two FucTs are striking. For FUT8, SPR sensorgrams indicate a considerably slower association rate of GDP-Fuc compared to GDP (cf. Figure 4-12). This observation is in agreement with a putative conformational change of GDP-Fuc upon binding to FUT8. For FucTA, in contrast, both GDP and GDP-Fuc show a fast interchange indicating a different binding mode (cf. Figure 4-2 and Figure 4-3).

#### 4.2.5 Arg365 plays two key roles in catalysis

The model suggests that Arg365, besides binding of the donor via the  $\beta$ -phosphate, locks the fucosyl residue in a position that allows the nucleophilic attack of the acceptor substrate. The correct orientation of the fucose is essential, as sugar nucleotides may adopt many different conformations in solution [113] and stretched conformations of the  $\beta$ -phosphate fucose linkage would impede the accessibility for an incoming nucleophile. Besides Arg365, the Gln470 is involved in stabilization of the active GDP-Fuc conformer. The second task of Arg365 is to neutralize the developing negative charge upon cleavage of the O-1"-C-1" bond of GDP-Fuc. Glycosyltransferases use different mechanisms to enhance the quality of the

nucleotide diphosphate as a leaving group, i.e. a bivalent metal cation or positively charged amino acid residues. FUT8 does not require metal cations for activity. As Arg365 was shown to be involved in the binding of GDP [38], it is reasonable to assume that Arg365 facilitates the release of GDP. The lysine residue in the mutant Arg365Lys cannot both bind the  $\beta$ -phosphate and stabilize the fucosyl residue at the same time. GDP was also found to exhibit a much shorter residence time in FUT8 than GDP-Fuc as indicated by the dissociation rates in the SPR assay (cf. Figure 4-12). This guarantees a fast exchange of the product with a new donor substrate molecule while the GDP concentration is low compared to the concentration of GDP-Fuc.

Inverting glycosyltransferases require a base catalyst in order to enhance nucleophilicity of the attacking hydroxyl group for a direct-displacement  $S_N2$ -like mechanism. [114] For FUT8, the essential residues Asp409 and Asp453 have been suggested for the role of the basic catalyst. [44] The results presented here indicate that both residues are not suitable as base catalysts due to their location. Asp453 is essential for guanine recognition (c.f. above) and Asp409 rather seems to be critical for the correct orientation of Arg365. Aspartate or glutamate residues acting as base catalysts in inverting glycosyltransferases, in contrast, are usually found in hydrogen bonding distance to the acceptor and the anomeric carbon of the sugar nucleotide in crystal structures.[50, 114] Similarly, the location of essential Asp368 as part of the flexible loop is not suitable for a base catalyst although the high flexibility of the loop makes a reliable prediction with *in silico* methods difficult.

#### 4.2.6 Function of the flexible loop

FUT8 exhibits a flexible loop adjacent to the donor substrate binding pocket and five residues (368–372) within this loop are not resolved in the X-ray structure. The MD simulation reflects this mobility well. Yet, there is no evidence from the simulation that the loop is directly involved in binding of GDP-Fuc. In many glycosyltransferases, a flexible loop serves a lid that covers the bound nucleotide. [48] In FUT8, however, this mechanism seems rather unlikely as the flexible part of the loop is comparatively short and the operating distance is insufficient to cover the nucleotide. Furthermore, both NodZ and *ce*POFUT lack such a loop despite the profound analogy in the recognition of GDP-Fuc. Nevertheless, residues in this loop were shown to be important for catalysis. [44] One could therefore speculate that the flexible loop plays a role in the recognition of the acceptor substrate that is fundamentally different than the acceptors of NodZ and *ce*POFUT and hence is not found in the structures of these related enzymes.

#### 4.2.7 Comparison to $\alpha$ 1,3 fucosyltransferase of *H. pylori*

Substrate recognition and the catalytic mechanism of FUT8 is significantly different from that proposed for the  $\alpha$ 1,3 fucosyltransferase from *H. pylori*. Both enzymes have GT-B fold and a Rossman fold for binding of the nucleotide. They differ, however, profoundly in their overall 3D structure and in their sequence. Guanine is bound and recognized by backbone carbonyl functions instead of an essential aspartate. Furthermore, two positively charged residues (Arg195 and Lys250) bind the pyrophosphate in addition to the backbone amides of helices of the Rossman fold. Finally, an essential glutamate residue (Glu95) is properly positioned to take over the function of the base catalyst. Such a residue is not present in FUT8.

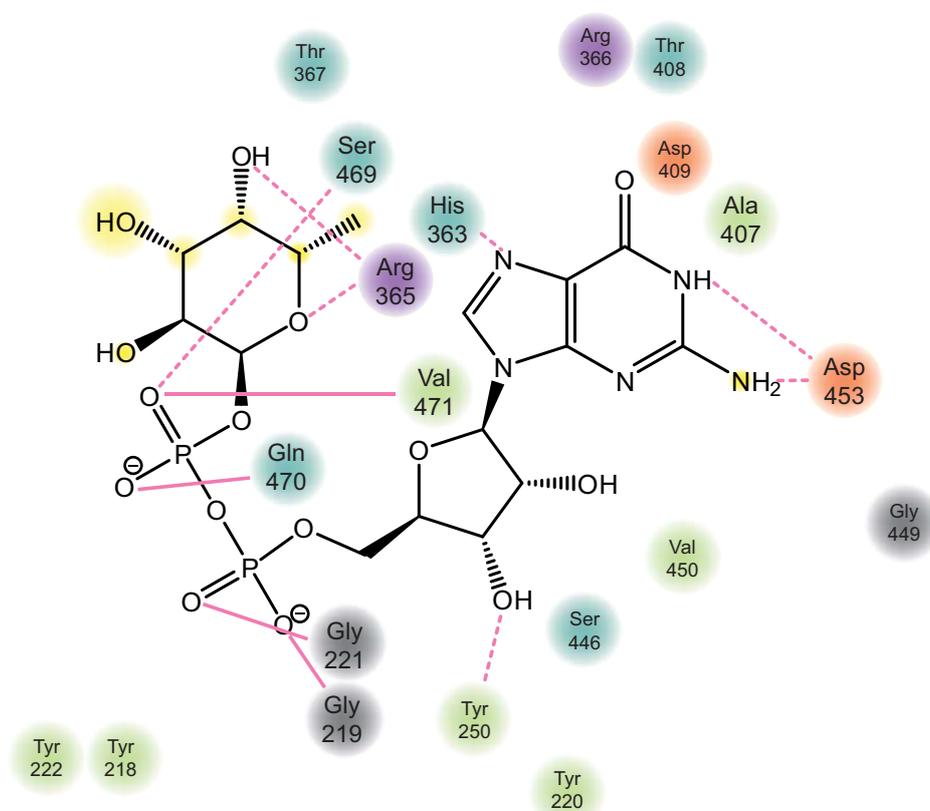


Figure 4-22: 2-dimensional plot taken from a representative frame of the MD simulation (frame 2278) illustrating the interactions of the donor substrate GDP-Fuc with residues in the active site of FUT8. Dotted lines show hydrogen bonds with functions from side chains and solid lines such with functions of the backbone. Other residues shown are close to the substrate and participate in hydrophobic interactions.

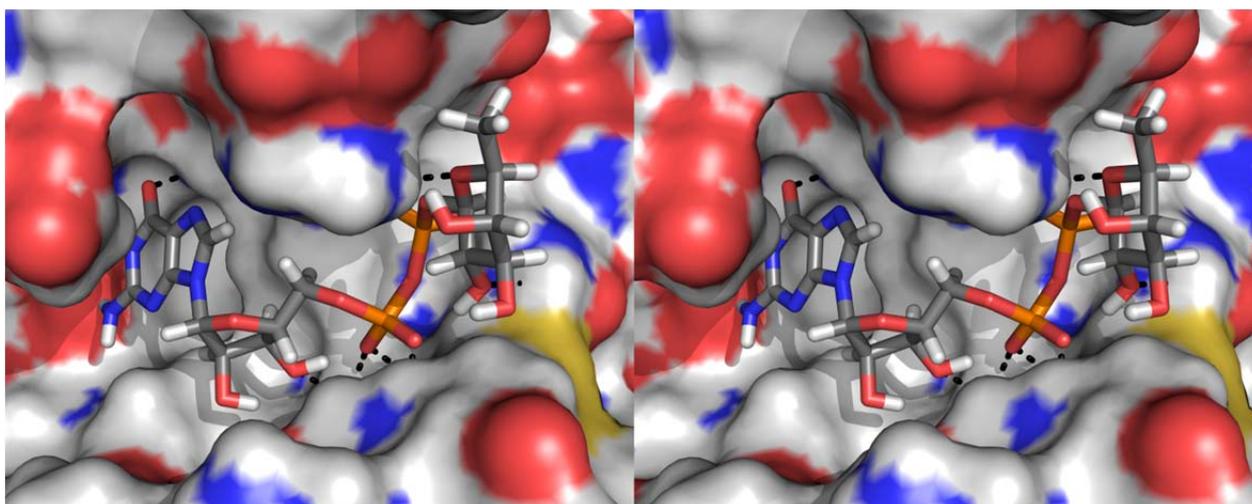


Figure 4-23: Last frame of the MD simulation of GDP-Fuc in complex with FUT8. Polar contacts are indicated as black dotted lines. It is clearly visible that the pattern of hydrogen bonds as well as the positioning of the ligand has not changed throughout the MD.

### 4.3 An unusually large acceptor binding site of human core $\alpha$ 1,6-fucosyltransferase is formed by donor assisted self-organization

The structural model of GDP-Fuc binding to GDP-Fuc developed in section 4.2 provides detailed insight into the mechanism of FUT8 on the side of the donor substrate. The structural basis for the unusual specificity of FUT8 for its acceptor substrate, however, remained elusive and is addressed in this section.

#### 4.3.1 The acceptor substrate binding shows moderate affinity and dissociation rate

GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-*N*-Ac (heptasaccharide **1b**, cf. Section 4.1.1) was used as acceptor substrate throughout the experiments [52, 108] This structure is a close mimic to the natural acceptor substrate having also an amide function in  $\beta$ -configuration at the reducing end.

The thermodynamic dissociation constant of heptasaccharide **1b** with FUT8 was determined by means of an SPR affinity assay and found an affinity of  $K_D = 390 \mu\text{M}$  (cf. Figure 4-25A). The recorded SPR sensorgrams also allow for analysis of the kinetic parameters of the binding process, e.g. the association rate constant  $k_{on}$  and the dissociation rate constant  $k_{off}$ . A moderate association rate of  $k_{on} = 4000 \text{ L s}^{-1}\text{mol}^{-1}$  and a moderate dissociation rate of  $k_{off} = 2 \text{ s}^{-1}$  was observed (cf. Figure 4-24, Figure 3-1 and Table 4-8). The dissociation constant from the kinetic on and off rates is  $K_D = k_{off}/k_{on} = 500 \mu\text{M}$  which is in excellent agreement with the thermodynamic determination of the  $K_D$  value described above.

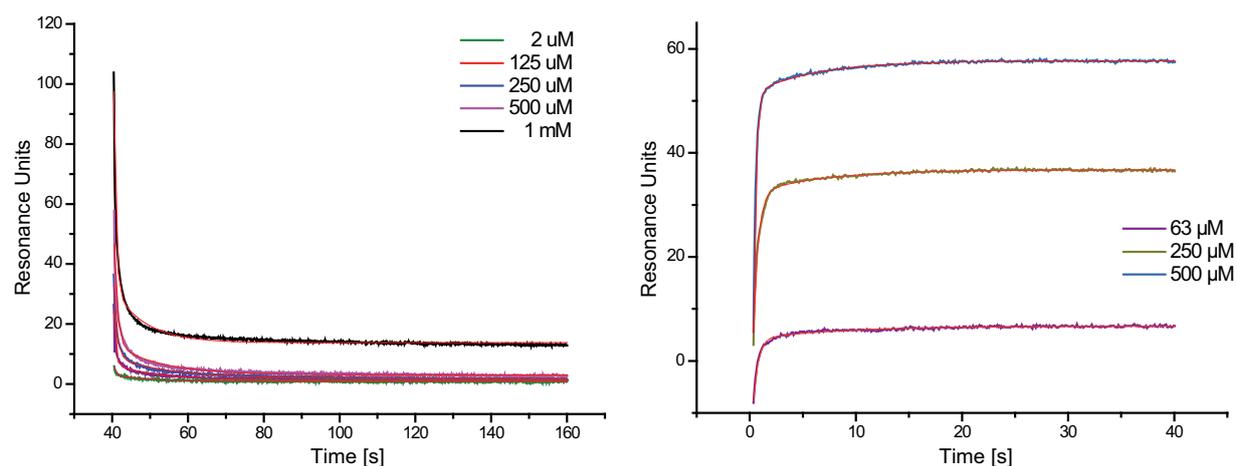


Figure 4-24: Dissociation (left) and association (right) part of SPR sensorgrams of the acceptor substrate.  $k_{off}$  was obtained by fitting an exponential decay to the curves. Fitting the association curves (red lines) yields  $k_{obs}$ , from which  $k_{on}$  can be calculated.

## RESULTS AND DISCUSSION

Table 4-7:  $k_{off}$  values of the acceptor substrate as determined by fitting a double exponential decay function to the dissociation part of the sensorgrams.

$c$ [ $\mu\text{M}$ ]	$t_f$ [s]	$k_{off}$ [1/s]
1000	0.558	1.79
500	0.881	1.13
250	0.805	1.24
125	0.464	2.15
63	0.477	2.10
2.0	0.332	3.01
Average	0.586	1.71

Table 4-8: Kinetic parameters of GDP-Fuc binding to FUT8 obtained from the fits in Figure 4-14.  $k_{on}$  was calculated by fitting the association phase of the sensorgrams to yield  $k_{obs}$ , from which  $k_{on}$  was calculated by subtracting  $k_{off}$  and subsequent division of the result by the ligand concentration. The  $K_D$  from kinetic data was calculated by division of  $k_{off}$  by  $k_{on}$ .

$c$ [ $\mu\text{M}$ ]	$t_{obs}$ [s]	$k_{obs}$ [1/s]	$k_{on}$ [L/(s*mol)]	$K_{D,from\ kinetics}$ [ $\mu\text{M}$ ]
500	0.237	4.22	6170	184
250	0.453	2.21	3860	322
63	0.459	2.18	1270	1650
Average	0.383	2.87	3770	719

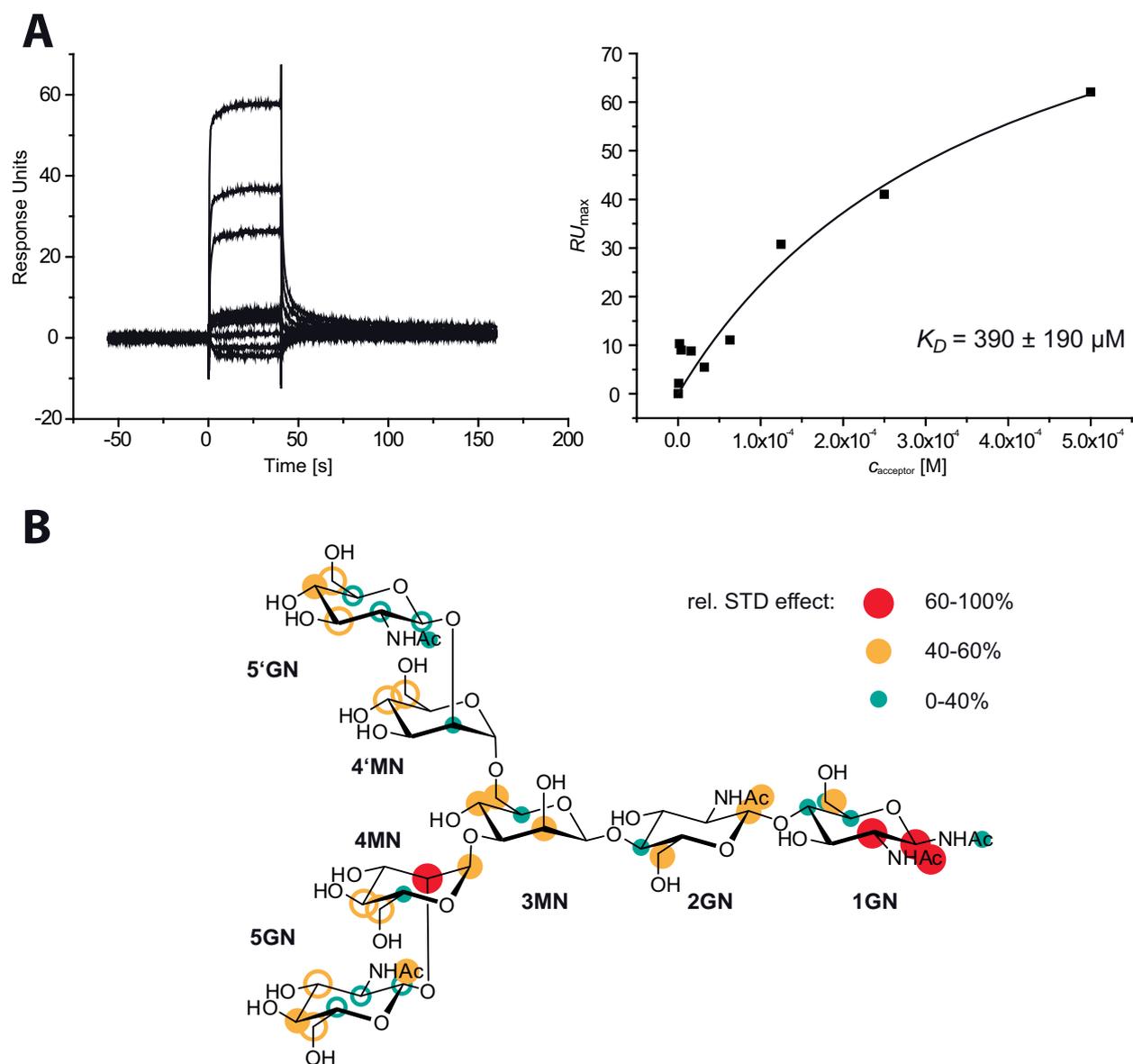


Figure 4-25: A: Sensorgrams (left) for the heptasaccharide **1b** with immobilized FUT8. The data shown are the response curves minus the curve of the buffer signal. The isothermal binding curve shown in the right hand panel was obtained by fitting a one-site binding model to the maximum response units ( $RU_{max}$ ) from the SPR association curves (left). B: Acceptor substrate epitope of FUT8. Filled circles indicate the size of the relative STD effect (red: high, orange: medium, green: low). Open circles are used for resonances at the branches that overlap with resonances of the corresponding monosaccharide unit at the other branch and therefore cannot be distinguished from each other.

#### 4.3.2 STD NMR reveals the acceptor substrate epitope

STD NMR allows for mapping the binding epitope of ligands binding to large molecules without the need for labeling of high amounts of protein.[93, 94] The absolute and relative STD percentages measured for the heptasaccharide **1b** when binding to FUT8 are listed in Table 7-13 and the resulting epitope is shown in Figure 4-25B. Higher STD effects indicate a shorter distance to the enzyme's surface and are designated by larger circles in Figure 4-25B.

Clearly, STD effects were observed for every part of the heptasaccharide. The size of the relative STD effects, however, differs for the respective regions of the acceptor substrate. The protons at positions 1 and 2 of the proximal GlcNAc receive highest saturation transfer, as well as the H-2 of the mannose-4. Furthermore, the other two core residues GlcNAc-2 and Man-3 experience moderate saturation transfer at numerous positions. The overall epitope is dominantly formed by the three core saccharide units and the  $\alpha$ 1,3-linked mannose unit. In addition, saturation transfer effects on resonances originating from protons of the two non-reducing GlcNAc residues were observed. In some cases the resonances could not be unambiguously assigned to one of these residues due to severe signal overlap in the 1D  $^1\text{H}$  NMR Spectrum. These protons are indicated by open circles in Figure 4-25B and are recorded as a sum of STD effects of the respective proton pairs.

### 4.3.3 A model of the ternary complex reveals FUT8–acceptor substrate contacts

The model of the GDP-Fuc in complex with FUT8 from the previous section [108] represented the starting point for *in silico* modeling of the ternary complex. Since there is no indication from protein crystallographic data for the position of the acceptor substrate, for initial positioning the information of the STD NMR epitope was used in combination with four additional boundary conditions that are given by: 1) the 6-hydroxy group of GlcNAc-1 of the acceptor substrate has to be in close proximity to the anomeric position of the fucosyl residue of the bound donor in order to allow a transfer of the fucosyl residue to the 6-position of the first GlcNAc residue, 2) GlcNAc-5 at the 3-mannose branch of the acceptor is essential for the catalytic process and thus, presumably, has to have contact to the protein, 3) the shape and size of the acceptor has to match a complementary surface on the protein and 4) the peptide part of the nascent glycopeptide chain has to be accommodated by the protein–acceptor complex. Taking all these facts together, only one position of the heptasaccharide was possible (Figure 4-26A). GlcNAc-1 is near the anomeric center of the donor, the 3-mannose branch is located between the beta sheets  $\beta$ 9 and  $\beta$ 11 and the 6-mannose branch is positioned near the flexible loop between Asp368 and Phe376.

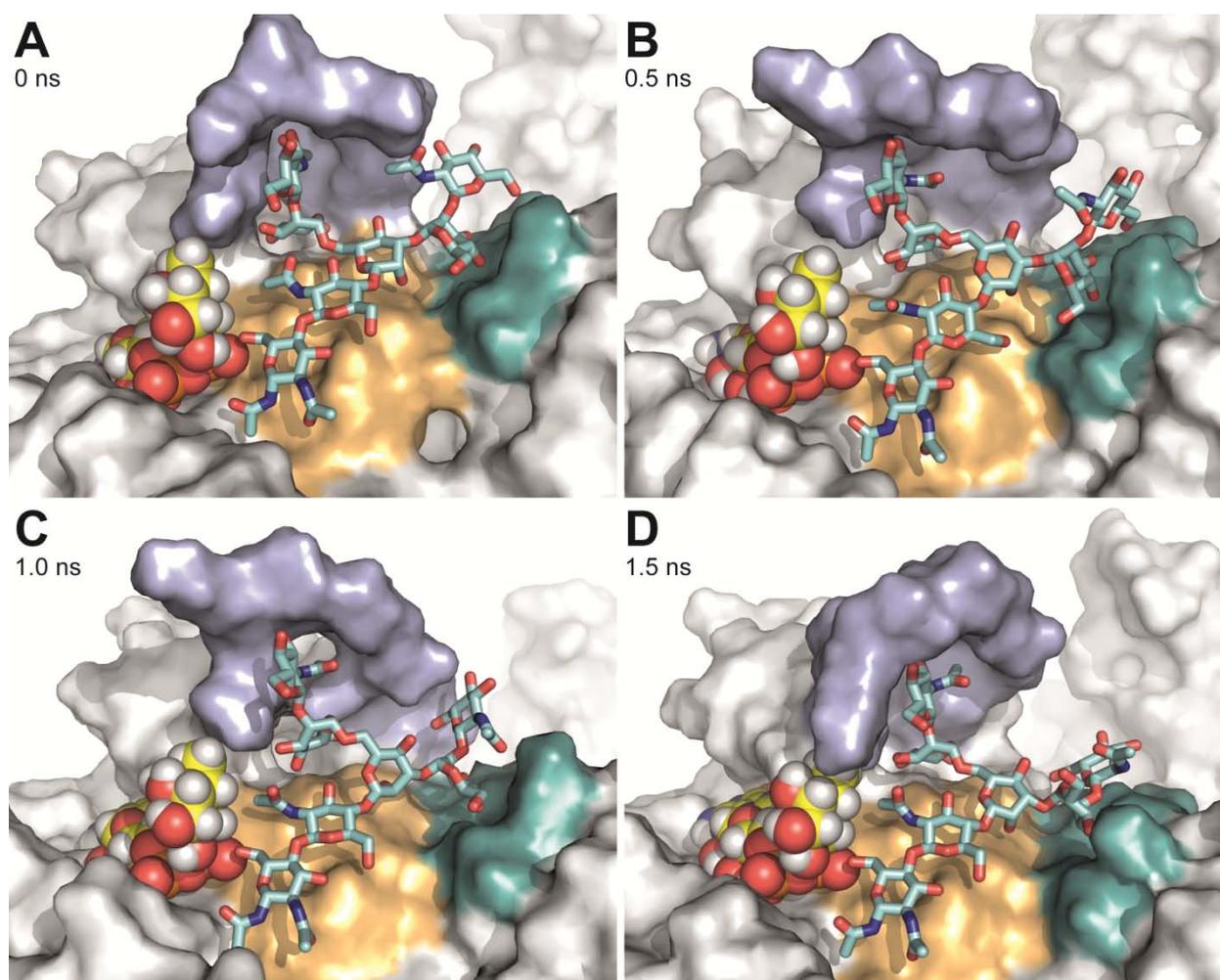


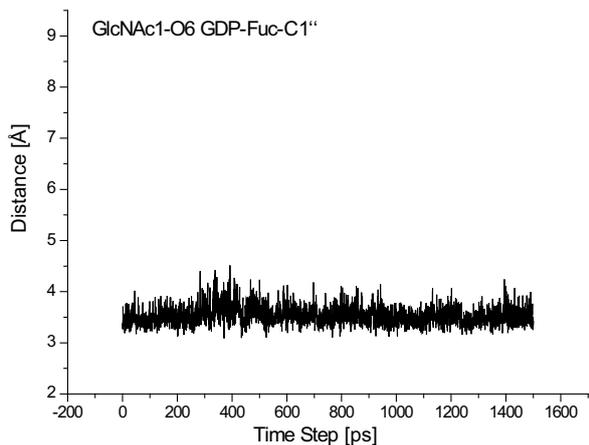
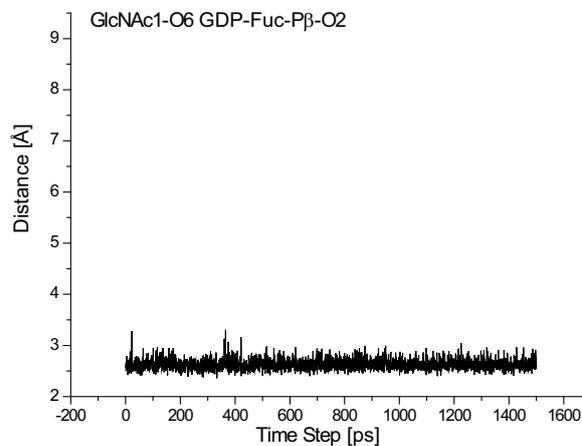
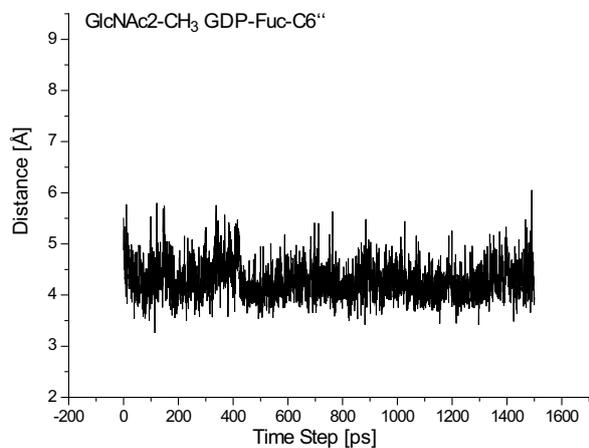
Figure 4-26: Acceptor binding site of FUT8 at four time steps of the MD simulation. The acceptor oligosaccharide is shown in sticks with cyan carbon atoms, FUT8 with a grey surface and GDP-Fuc as space filling model with yellow carbon atoms and grey hydrogen atoms. The peptides of FUT8 that form the acceptor binding site are indicated by different colors of the surface: The three regions that interact with the heptasaccharide are the flexible loop (Thr367-His377, blue, top), the peptides forming the moderately hydrophobic cavity (Ser468-Arg473 and Asp494-Gly501, orange, under the chitobiose) and the C-terminal  $\beta$ -sheets  $\beta$ 10 and  $\beta$ 11 (Ala532-Lys541, teal, right). The four frames clearly show the flexibility of the branches of the acceptor substrate and of the interacting parts of the acceptor substrate binding site.

The positioning of both donor and acceptor substrates was validated by a molecular dynamics simulation (cf. Figure 4-26A-D). The ternary complex of donor, acceptor and protein was embedded into a water box containing 200 mM sodium chloride. After minimization of the water box a MD simulation was run for 1.5 ns. During that time neither the position of the donor nor that of the acceptor changed significantly showing that the initial positioning was correct.. The tight binding of the donor substrate by several hydrogen bonds did not change compared to a simulation of the donor and FUT8 alone. [108] The branches of the heptasaccharide show however a much more dynamic behavior throughout the molecular dynamics simulation than the donor substrate. The higher dynamics observed in the binding

## RESULTS AND DISCUSSION

of the acceptor becomes obvious from the trajectories of the distances of selected atom pairs (cf. Figure 4-28) that participate in hydrogen bonds or hydrophobic contacts between enzyme and acceptor.

### A



### B

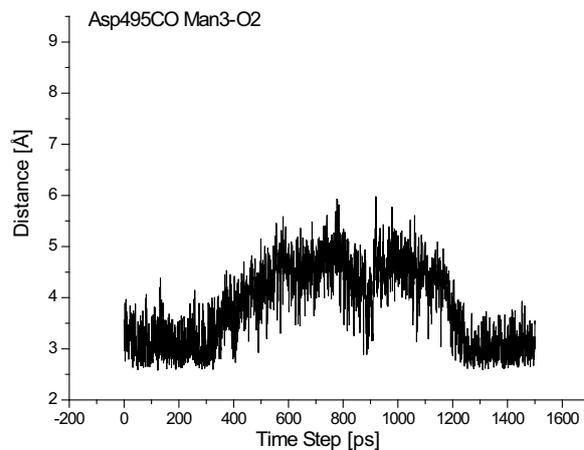
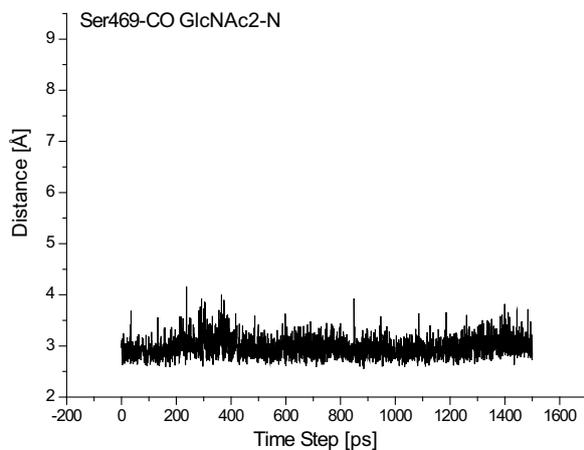
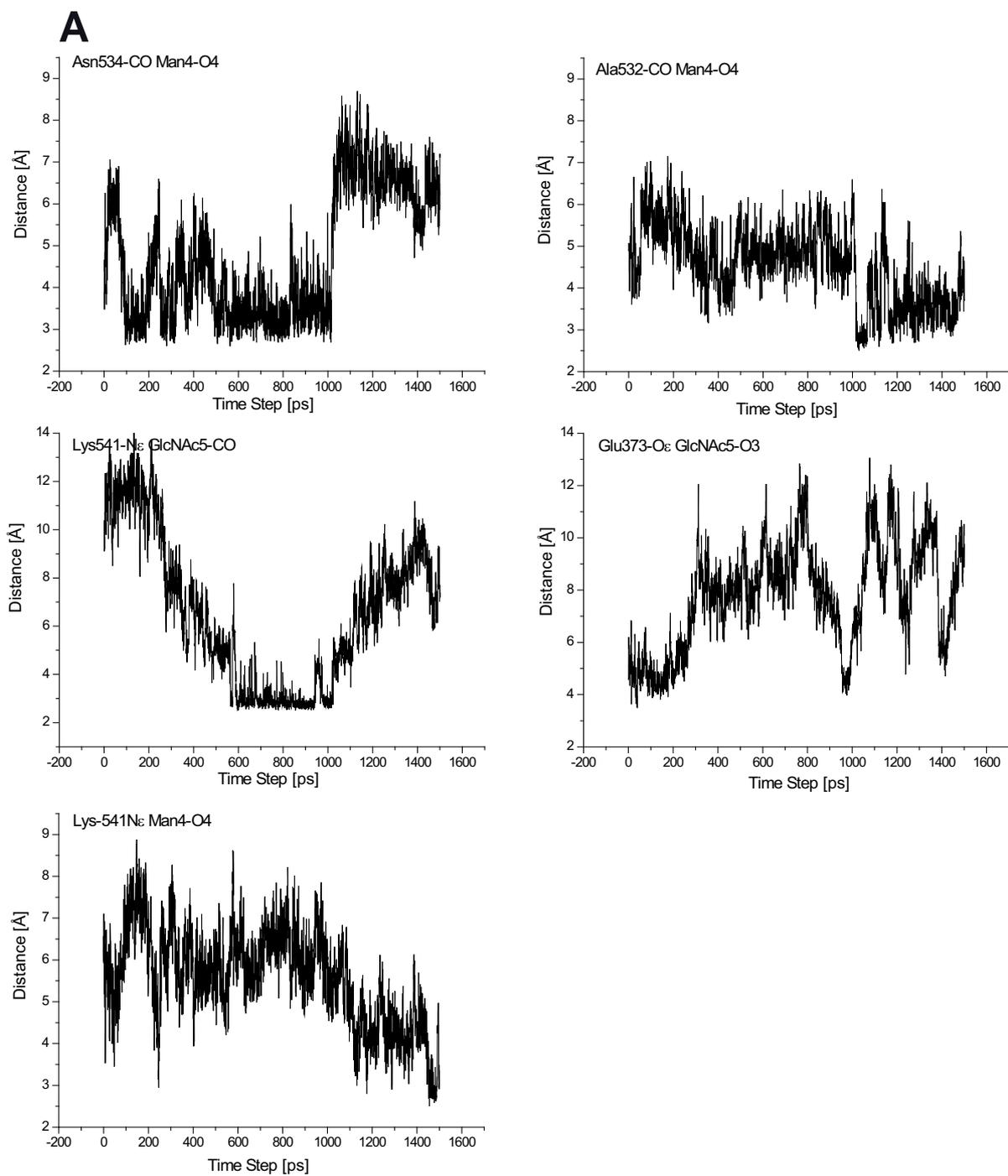


Figure 4-27. Selected interactions of the core trisaccharide. A: Trajectories of the distance between atoms of the core trisaccharide and atoms of GDP-Fuc clearly indicate the extremely strong interactions of the acceptor with the donor molecule stabilizing the binding of the acceptor dramatically by approximately a factor of 10 to 100 due to stable hydrogen bonds and hydrophobic interactions. B: The panels shown at the bottom indicate the stability of the FUT8 interactions with the core trisaccharide; mannose-3 already shows some flexibility that is even more pronounced in the branches (cf. Figure 4-28). (for all trajectories cf. experimental section)

## RESULTS AND DISCUSSION



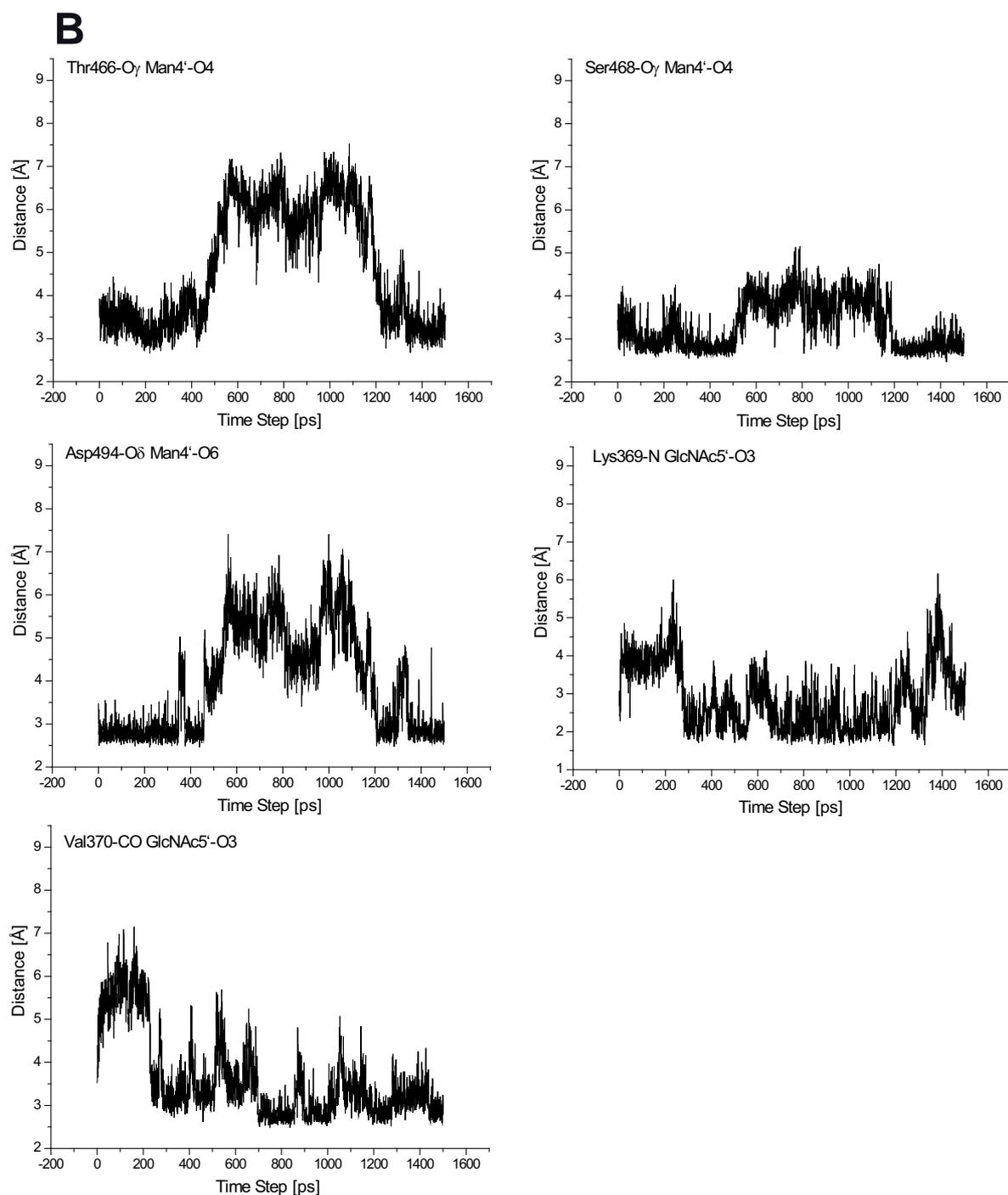


Figure 4-28. Dynamic interactions of the 3-Man branch (A) and the 6-Man branch (B). The trajectories show the flexible behavior of the non-reducing ends and some flexibility of the binding site: Hydrogen bonds build up and break rapidly and exist only for a fraction of the simulation time. However, a switching is observed such that if one hydrogen bond is broken another hydrogen bond takes over the function of the former H-bond. As an example, compare the trajectories in the top two panels of A, where at 1.05 ns a switch of the respective H-bonds of the O-4 of Man4 occurs from Asn534 to Ala532.

The interactions of GlcNAc-1 and GlcNAc-2 shown in Figure 4-27 are very stable and do not change over the simulation time. All other distances of acceptor atoms from Man-3 and the two branches to enzyme atoms show a much higher variability and most of them exist only

for a fraction of the simulation time (cf. Figure 4-28). The detailed analysis of the acceptor substrate binding during the MD simulation yields the following results: First, the core pentasaccharide of the acceptor is mainly bound via a shallow, large hydrophobic epitope established by the loops between Ser468-Arg473, Asp494-Gly501 (interactions with core trisaccharide) and Ala532-Lys541 (interactions with Man4). Hence, less solvent exposure is found for these carbohydrate residues resulting in relatively high STD values. Polar residues and backbone amide functions in these regions stabilize acceptor substrate binding via hydrogen bonding. The three saccharide units of the core GlcNAc1, GlcNAc2 and Man3 participate in few, but relatively stable hydrogen bonds compared to the residues at the antennae. In detail, for the proximal GlcNAc, two hydrogen bonds are observed: the carbonyl oxygen of the side chain of Gln470 recognizes the amide nitrogen at the anomeric position and secondly, the hydroxyl function at the 3-position is bound via the side chain hydroxyl function of Tyr498. GlcNAc2 binds also via the nitrogen of the amide function to the backbone carbonyl function of Ser469 and via its 6-hydroxy function (to the carboxy function of Asp495 or to the hydroxyl group of Tyr 498. For Man-3, only the 2-hydroxy function is bound by the backbone carbonyl function of Asp495.

The branch at the 3-position of mannose-3 is bound via multiple hydrogen bonds. All three unsubstituted hydroxy functions of Man4 are bound by backbone carbonyl functions of Ala532 and Asn534 throughout the simulation. The hydrogen bonding of GlcNAc5, in contrast, is less persistent. The 6-hydroxy function binds to the carbonyl function of His535. The other interactions of GlcNAc5 are established by side chain functions of residues of the flexible loop and exist only for a minor fraction of the simulation time. The carbonyl group of the amide function of GlcNAc5 is bound by the hydroxyl group of Thr372 and the 3- and 4-hydroxy group is bound by the carboxy function of the essential Glu373.

The more flexible 6-mannose branch is bound between the loops Ser468-Arg473, Asp494-Gly501 and the flexible loop (Thr367-Glu373). The first two peptides are also involved in recognition of the core trisaccharide. The main interactions of Man4' are all established between the side chain functions of Asp494, Ser468 and Thr466 and the 6- and 4-hydroxy group of the mannose residue. The backbone amide function of Thr367 as part of the flexible loop only plays a minor role. In contrast, all hydroxyl functions of GlcNAc5' are exclusively bound via hydrogen bonds to the backbone amide functions of residues of the flexible loop. Remarkably, many of these residues (Thr367, Lys369, Val370, Thr372 and Glu373) can either function as donor or as acceptor in the hydrogen bond established to GlcNAc5.

Additionally, strong interactions between the donor and the acceptor substrate were observed during the MD simulation. A hydrogen bond between one of the oxygen atoms of the  $\beta$ -phosphate and the 6-OH of the GlcNAc1 was in effect throughout the simulation. This inter-substrate interaction is the only hydrogen bond with an average distance of 2.63Å that remained in effect for 100% of the time. Secondly, the methyl group of the acetyl function of GlcNAc2 is pointed towards the hydrophobic face of the fucose residue and exhibits a stable distance to the C-5 of fucose of  $3.9 \pm 0.3\text{Å}$ .

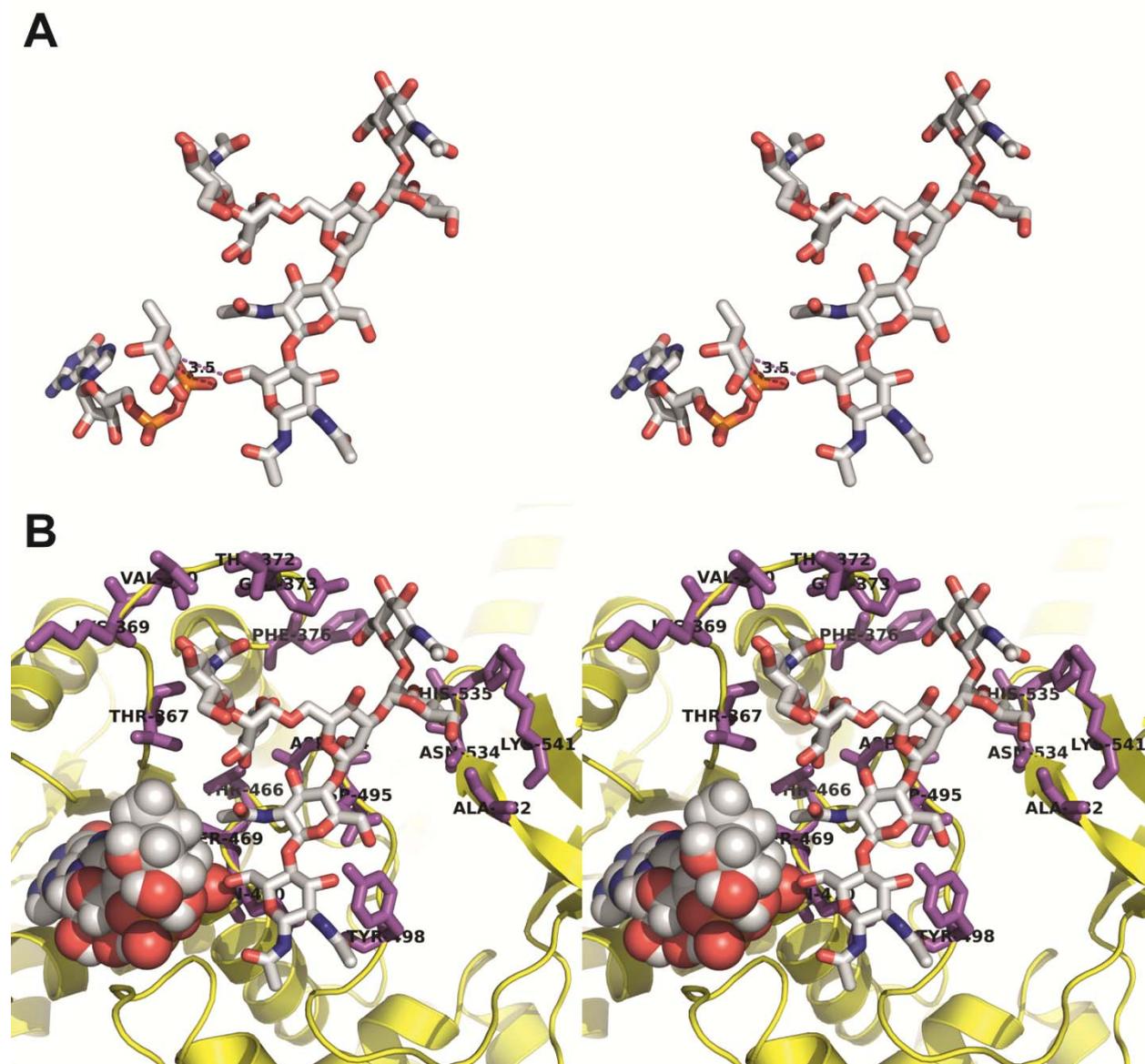


Figure 4-29. A: Crossed-eye stereo plot of the relative position of donor (left, space fill model colored by atom type) and acceptor substrate (right, stick model colored by atom type) to each other when bound to FUT8 in the model of the ternary complex (1.0 ns, frame 1975). The atoms involved in the process of the nucleophilic attack, the 6-oxygen of the proximal GlcNAc and the anomeric carbon of the fucose, have a distance of 3.5Å (magenta dotted line). B: The same frame including the substrate binding site of FUT8. The enzyme's backbone is represented by a yellow cartoon. Side chains interacting with the acceptor substrate are shown as magenta sticks. GDP-Fuc is shown as space fill model in blue and the acceptor heptasaccharide as stick model colored by atom type.

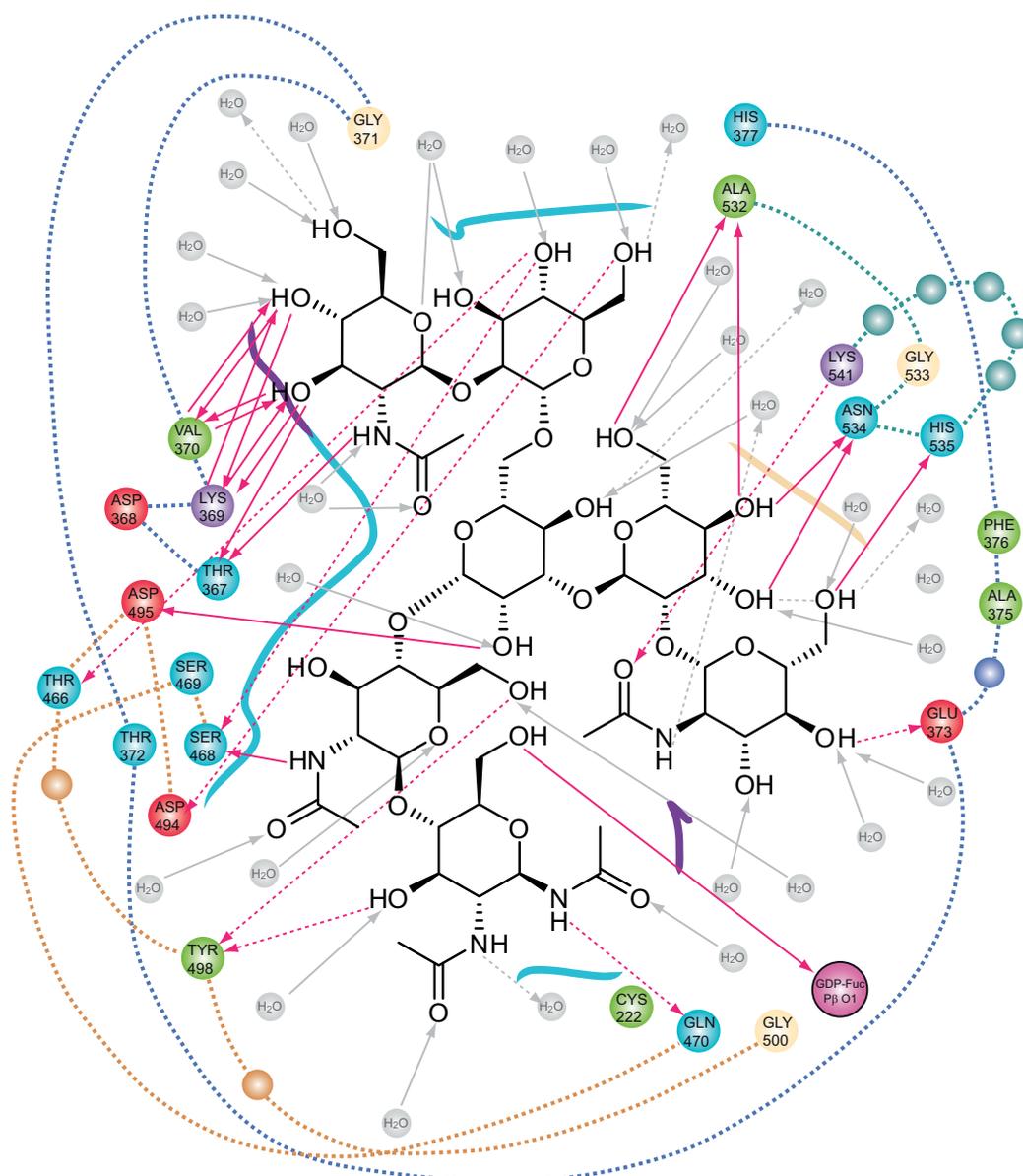


Figure 4-30. 2-dimensional plot taken from a representative frame of the MD simulation (1.0 ns) illustrating the interactions of the acceptor substrate with residues in the active site of FUT8. Dotted arrows show hydrogen bonds with functions from side chains and solid arrows such with functions of the backbone. Other residues shown are close to the substrate and likely to participate in hydrophobic interactions.

#### 4.3.4 FUT8 recognizes a large epitope of the acceptor substrate

The first central question addressed in this section is the molecular basis of the recognition of the unusually large acceptor substrate. A large, shallow cavity of approximately  $18 \times 23 \text{ \AA}$  that is adjacent to the donor substrate binding site was identified. The site is formed by the segments Ser468-Arg473 and Asp494-Gly501 at the bottom and flanked by the flexible loop (Thr367-Glu373) at one site and the  $\beta$ -strands  $\beta 11$  and  $\beta 12$  (Ala532-Lys541) at the opposite site (cf. Figure 4-26 and Figure 4-29B). Identification of the acceptor substrate binding site was straightforward, as 1) the donor substrate binding cavity and the donor substrate binding mode are known [44, 108] and as 2) the acceptor substrate has an extensive size that has to be

accommodated by its binding cavity. All terminal sugar residues, i.e. GlcNAc1, GlcNAc5 and GlcNAc5', have to be involved in the recognition of the acceptor by the enzyme. [40-43] These boundary conditions leave only one possibility to position the acceptor molecule relative to the enzyme. The reducing end is oriented with the 6-OH function close to the fucose of the bound donor in order to allow the transfer. The hydroxymethyl group of GlcNAc1 has a favorable *gt* conformation. The core trisaccharide has contact to the bottom of the cavity, as epitope mapping by means of STD NMR indicates that protons of these carbohydrate units are closer to the enzyme surface than those attached to the antennae.

The experimental and theoretical results shown here demonstrate that FUT8 recognizes a large epitope comprised of all monosaccharide units of the heptasaccharide. Yet, binding modes are different for the core and the two branches of the acceptor. The interactions of the core trisaccharide are more of hydrophobic nature combined with few, but constant hydrogen bonds to amino acids at the bottom of the shallow acceptor binding pocket. Notably, one of these stable hydrogen bonds is established between the O-3 of GlcNAc1 and O $\eta$  of Tyr498. This interaction could be the reason why core  $\alpha$ 1,3-fucosylated structures are not recognized although the binding site is large enough for such a substrate. The general binding of the acceptor is therefore fundamentally different from that of the donor substrate, which is tightly bound in a network of hydrogen bonds and lipophilic interactions in a narrow fold of FUT8. [108] This observation is also supported by the weak thermodynamic dissociation constant as determined by SPR. In addition, the experimental results from SPR and STD NMR show that the apo form of FUT8 is capable to bind the acceptor substrate with  $K_D = 390 \mu\text{M}$ , although the affinity is low compared to that of the donor substrate, whose  $K_D = 9.3 \mu\text{M}$  [108].

#### 4.3.5 The donor contributes significantly to acceptor binding

The MD simulation shows the most stable interactions of the acceptor with GDP-Fuc indicating that the hydrogen bonds and hydrophobic interactions of these components contribute a significant proportion to the binding of the acceptor (cf. Figure 4-27, Figure 4-29 and Table 4-9). The  $K_D$  of the acceptor with the apo enzyme is much larger than the  $K_M$ . However, the interactions listed above explain the increased affinity of the acceptor towards the GDP-Fuc–FUT8 complex compared to the apo enzyme very well. Based on the strength of hydrogen bonds in aqueous environment, [115] an increase of the binding affinity of the acceptor to the FUT8/GDP-Fuc complex by a factor of 10 to 100 compared to the apo enzyme is estimated.

RESULTS AND DISCUSSION

Table 4-9: Contacts between FUT8 or donor substrate with the acceptor substrate during the MD simulation.

FUT8 or donor Residue	Atom	Acceptor Substrate Residue	Atom	Average distance [Å]	Std Dev of distance [Å]	Fraction of Frames with distance < 3.6 Å
GDP-Fuc	PβO-2	GlcNAc1	O-6	2.63	0.1	100.0%
GDP-Fuc	C-1''	GlcNAc1	O-6	3.53	0.2	70%
GDP-Fuc	C-6''	GlcNAc2	CH <sub>3</sub>	4.28	0.4	0.8%
GDP-Fuc	C-5''	GlcNAc2	CH <sub>3</sub>	3.93	0.3	7.7%
Gln470	Oε	GlcNAc1	N-1	3.93	0.6	32.9%
Tyr498	Oη	GlcNAc1	O-3	3.07	0.3	91.8%
Asp495	Oδ	GlcNAc2	O-6	4.70	0.5	2.7%
Ser468	CO	GlcNAc2	N	2.97	0.2	99.1%
Tyr498	Oη	GlcNAc2	O-6	3.72	0.5	46.4%
Glu373	Oε	GlcNAc5	O-3	7.81	2.0	0.1%
Glu373	Oε	GlcNAc5	O-4	7.20	1.7	3.4%
His535	CO	GlcNAc5	O-6	4.61	1.2	17.5%
Lys541	Nε	GlcNAc5	CO	6.51	3.1	26%
Thr372	Oγ	GlcNAc5	CO	9.91	3.6	3.1%
Glu373	CO	GlcNAc5'	O-3	6.43	2.0	15.0%
Lys369	CO	GlcNAc5'	O-3	5.02	1.0	14.7%
Lys369	CO	GlcNAc5'	O-4	5.50	1.1	13.0%
Lys369	N	GlcNAc5'	O-4	4.26	0.9	23.2%
Lys369	N	GlcNAc5'	O-3	3.57	0.8	64.4%
Thr367	CO	GlcNAc5'	N	3.76	1.0	60.1%
Thr372	Oγ	GlcNAc5'	O-5	6.87	1.6	2.8%
Val370	CO	GlcNAc5'	O-3	3.60	1.0	67.8%
Val370	CO	GlcNAc5'	O-4	3.97	1.0	44.9%
Val370	N	GlcNAc5'	O-3	3.58	1.0	53.0%
Val370	N	GlcNAc5'	O-4	3.38	1.4	67.1%
Asp495	CO	Man3	O-2	3.83	0.8	43.5%
Ala532	CO	Man4	O-4	4.52	0.9	19.2%
Ala532	CO	Man4	O-6	4.97	1.6	24.5%
Asn534	CO	Man4	O-3	3.76	0.6	48.6%
Asn534	CO	Man4	O-4	4.75	1.5	34.5%
Lys541	Nε	Man4	O-4	5.56	1.2	4.6%
Asp494	Oδ	Man4'	O-6	3.97	1.2	47.5%
Phe376	Cβ	Man4'	C-6	6.75	1.6	0.1%
Ser468	Oγ	Man4'	O-4	3.33	0.6	64.2%

FUT8 or donor Residue	Atom	Acceptor Substrate Residue	Atom	Average distance [Å]	Std Dev of distance [Å]	Fraction of Frames with distance < 3.6 Å
Thr367	CO	Man4'	O-3	5.73	0.9	4.5%
Thr367	N	Man4'	O-4	5.74	0.8	0.9%
Thr466	O $\gamma$	Man4'	O-4	4.68	1.4	35.3%

#### 4.3.6 Recognition of the 3-branch, $\beta$ GlcNAc-2- $\alpha$ Man, at Man3

The binding behavior of GlcNAc-5 at the 3-mannose branch is of major interest, as this carbohydrate unit is essential for substrate activity although positioned far from the location of fucosyl transfer. The less flexible 3-branch of Man3 lies between the flexible loop and the C-terminal  $\beta$ -sheets. Surprisingly, the essential GlcNAc5 shows comparatively few interactions both experimentally and theoretically. GlcNAc5 exhibits considerable flexibility in the MD simulation and shows short-lived interactions with the C-terminal  $\beta$ -strands and with the carboxy function of Glu373 that was identified to be essential for catalytic activity [44]. The latter interaction is observed as hydrogen bond with the O-3 and sometimes with O-4 of GlcNAc5. Although this interaction is only observed for a small fraction of the simulation time, it clearly shows that Glu373 is able to bind O-3 and O-4 of GlcNAc5. The interaction seems to be mediated by water molecules most of the time. One might suggest that the Glu373-GlcNAc 5 OH-3 hydrogen bond is responsible for recognition of the 3-Man branch. This mechanism also explains the fact that structures that are  $\beta$ -3/4-galactosylated at the 3-mannose branch are hardly recognized although the binding pocket could spatially accommodate further carbohydrate units. GlcNAc5 is also recognized by Lys541 and by His535 that form hydrogen bonds to the carbonyl group of the N-acetate and OH-6, respectively.

The epitope from STD NMR studies and the MD simulation also indicate strong binding of Man4 to the enzyme. Both Ala532 and Asn534 exert long lasting hydrogen bonds with the OH-groups of Man4 (cf. Figure 4-28A). Lys541 switches its hydrogen bond between the carbonyl of GlcNAc5 and the OH-4 of Man4.

#### 4.3.7 The flexible loop adapts the 6-Man branch

FUT8 has been shown to be tolerant towards substitutions at the 6-mannose branch and the lack of GlcNAc5'. [32, 41, 43]. The 6-mannose branch and particularly GlcNAc5' is bound via numerous hydrogen bonds with the flexible loop, whereas Man4' also interacts with the more rigid peptides that also bind the core trisaccharide. Remarkably, the observed hydrogen bonds

of GlcNAc5' are almost exclusively established by the backbone amide functions at the bottom of the flexible loop. Three peptide bonds were observed to interact time alternating with their carbonyl oxygen as hydrogen bond acceptor or with their nitrogen as hydrogen bond donor. This finding illustrates a highly accommodative nature of the flexible loop and reflects the variety of substitutions that can be recognized at the 6-Man branch. From the observed adaptive binding of the 6-Man branch the question arises why FUT8 developed a mechanism to bind the 6-Man branch at all, as no specific substitution is preferred or excluded. A possible explanation can be the back-folding of the 6-Man branch that is observed for N-Glycans in solution. In these conformers, GlcNAc5' is very close to the reducing end and the O-6 of GlcNAc1. A high population of the back-folded conformer could therefore slow the FUT8 reaction significantly down and it seems likely that the enzyme binds only conformers without back-folding. Promiscuous binding observed for the 6-branch allows the enzyme to accommodate both complex type structures and hybrid structures.

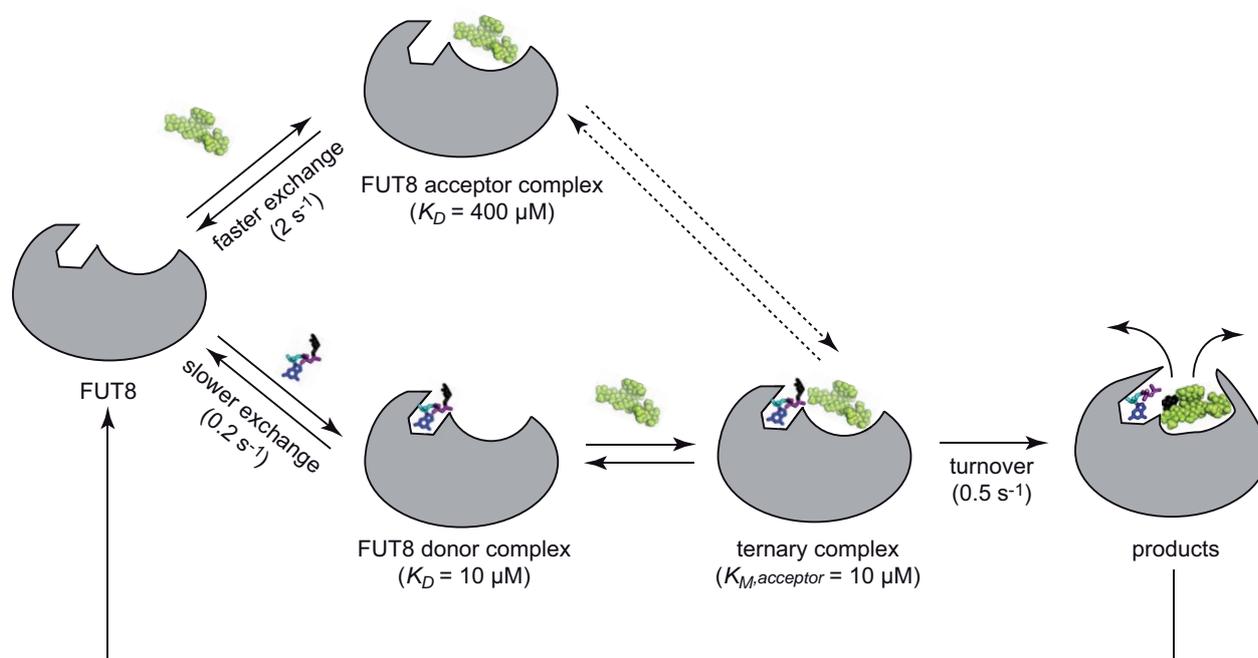
The more flexible 6-branch of Man3 occupies a *gg* conformation of the C5-C6 bond of Man3. This orientation is normally occupied to about 60% in solution [116] and allows the acceptor molecule to fit into the cavity without clashing with the atoms of the flexible loop.

#### 4.3.8 FUT8 acts via an ordered bi–bi mechanism

FUT8 was proposed to act via a rapid equilibrium random bi–bi mechanism, where donor and acceptor substrate bind to the enzyme in random order to form the ternary complex.[39] This conclusion was drawn from the  $K_M$  values of donor and acceptor, which are very similar. The results from thermodynamic and kinetic binding studies give clear evidence for an ordered bi–bi mechanism (cf. Scheme 4-1), like it was already shown for other glycosyltransferases. [49] From the exchange rates of donor (0.2 Hz) and acceptor (2 Hz), it is concluded that the donor substrate binds first to FUT8. The FUT8–GDP-Fuc complex has a tenfold higher life time (5.0 s) than the FUT8–acceptor complex (0.5 s). Therefore, the scenario that GDP-Fuc is already present in the binding pocket of FUT8 when the acceptor substrate binds is much more likely than the reverse case. The same applies for the thermodynamic point of view: As the affinity of the apo form of FUT8 is 50 fold higher towards GDP-Fuc than the affinity towards the acceptor, the occupancy of the GDP-Fuc binding site is much higher, especially at lower substrate concentrations. Another indication for the ordered bi–bi mechanism is the discrepancy between the  $K_M$  and  $K_D$  values of the acceptor substrate: Whereas the  $K_M$  and  $K_D$  value are the same for the donor substrate, the  $K_M$  is 50 fold lower for the acceptor substrate than the  $K_D$ . This difference shows that the affinity of FUT8 towards the acceptor substrate is

much higher under the conditions of substrate conversion, i.e. when the donor is present. This mechanism also prevents the hydrolysis of the donor even at low concentrations of the acceptor substrate. Interestingly, GDP-Fuc is hydrolyzed by FUT8 with a  $k_{cat}$  that is approximately 1/100 of the  $k_{cat}$  observed for the transfer to the natural acceptor (cf. section 4.4.1). However, with a fivefold excess of donor substrate in the NMR assay no hydrolysis of the donor substrate is observed.

The acceptor substrate is readily converted once it has bound to the FUT8–GDP-Fuc complex. This conclusion is indicated by the fact that the conversion rate  $k_{cat}$  is similar to the  $k_{off}$  of the acceptor. However, the actual mechanism of FUT8 *in vivo* depends on the concentration of both substrates in the Golgi apparatus, which is difficult to estimate.

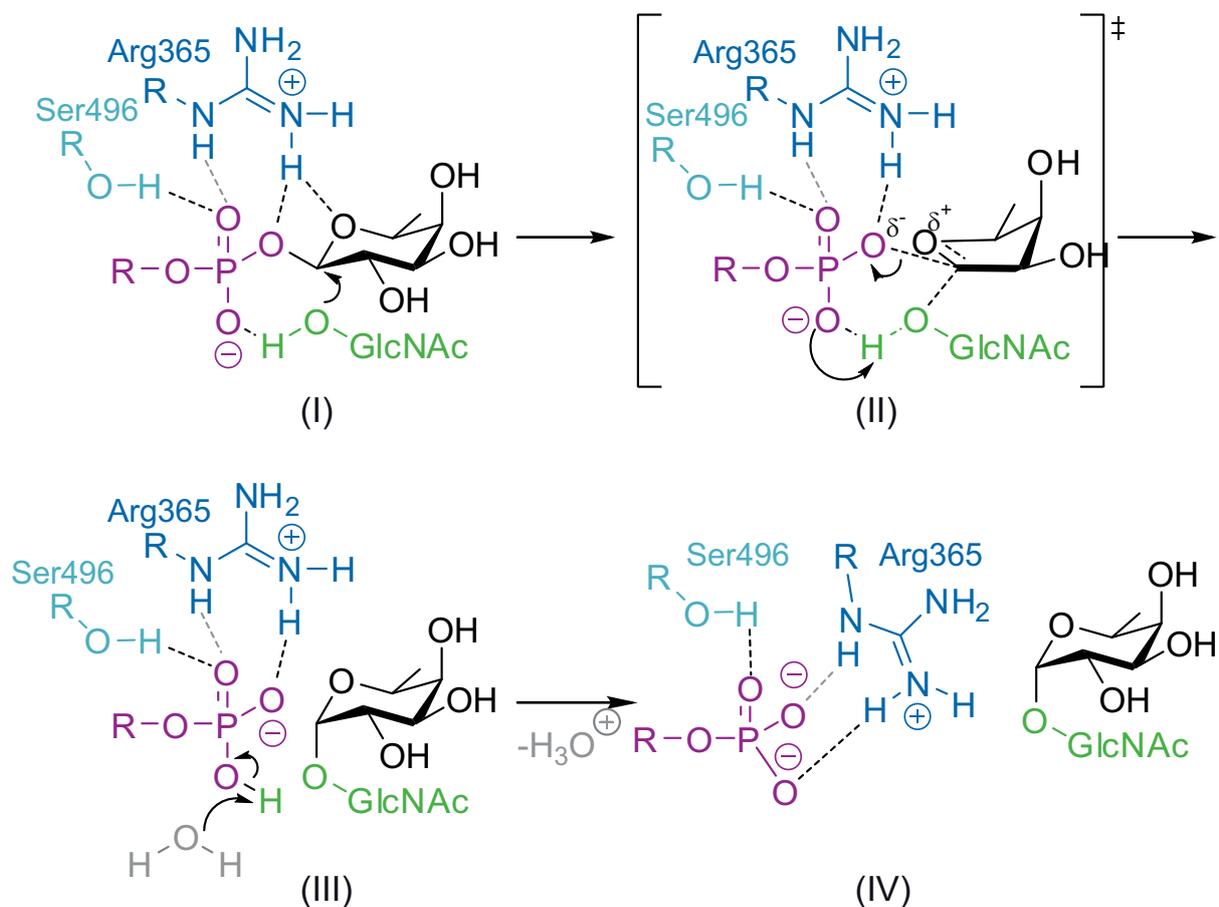


Scheme 4-1. Sequence of the catalytic cycle of FUT8. According to the exchange rates, the donor binds first to FUT8 (cf. text). After binding of the acceptor, which is assisted by the bound donor, the acceptor substrate is rapidly converted and both products are released in unknown order. Hypothetically, FUT8 undergoes one or more conformational changes upon substrate binding. It is unclear if the ternary complex can also form from the FUT8/acceptor complex (dotted lines).

#### 4.3.9 FUT8 employs a substrate-assisted mechanism

The mechanistic strategy applied by inverting glycosyltransferases is that of a direct displacement  $S_N2$ -like reaction. [114, 117] In many inverting glycosyltransferases an active-site basic side chain, often glutamate or aspartate, serves as a base catalyst. [114] The base catalyst increases the nucleophilicity of the hydroxyl group of the acceptor by deprotonation and facilitates the direct displacement of the phosphate leaving group. The identity of the base

catalyst and the method used to facilitate departure of the leaving group are therefore key questions in examining the catalytic mechanism of inverting glycosyltransferases. [114]



Scheme 4-2. Proposed catalytic mechanism for the reaction catalyzed by FUT8. An oxygen of the  $\beta$ -phosphate group of GDP-Fuc assists in deprotonation of the 6-hydroxyl group of GlcNAc1 of the acceptor substrate in order to promote a nucleophilic attack to the anomeric carbon of the fucose (I). The direct nucleophilic displacement mechanism involves a single transition state with oxocarbenium ion-like character (II). Protonation of the phosphate improves the leaving group qualities of GDP (III). The resulting negative charge is neutralized by the side chain of Arg365 (IV).

In the structure of FUT8 there is no basic residue present that could act as an assisting basic residue in hydrogen bond distance to the acceptor. The results presented here strongly suggest the  $\beta$ -phosphate of the bound GDP-Fuc acts as a base in the catalytic process of FUT8. The length of the hydrogen bond established between the oxygen of the  $\beta$ -phosphate of GDP-Fuc and the O-6 of GlcNAc1 is  $2.6 \text{ \AA}$  for the entire simulation time, indicating a very tight interaction due to a strong hydrogen bond. For a similar reason, Lira-Navarrete *et al.* already proposed a related mechanism for *ce*POFUT. [51] Furthermore, the importance of the  $\beta$ -phosphate group for catalysis is an explanation for the very strong binding of the  $\beta$ -phosphate group by FUT8 [39, 108], an observation that is rather unusual for glycosyltransferases. The tight binding of the  $\beta$ -phosphate is necessary to guarantee a precise orientation of the

incoming nucleophile with respect to the anomeric carbon of the fucose and to the residues that are active in catalysis, like Arg365.

The complete catalytic mechanism as deduced from the data for FUT8 is shown in Scheme 4-1. The key catalytic residue Arg365 binds the donor molecule and orients the fucose moiety by forming a bifurcated hydrogen bond with its N $\eta$  to the O1" and O5" as described. [108] The hydrogen bond between the oxygen of the  $\beta$ -phosphate of GDP-Fuc and the O-6 of GlcNAc1 ensures the nucleophile to be in close proximity to the anomeric carbon (3.6 Å). After the nucleophilic attack of O-6 to the anomeric carbon of fucose, a single transition state with delocalization of a partial positive charge to the ring oxygen is traversed. The phosphate acts as leaving group and the developing negative charge at the former O-1" is further stabilized by Arg365. The simultaneous protonation of the  $\beta$ -phosphate further enhances its qualities as a leaving group. After the transfer of the fucosyl residue is completed, Arg365 binds exclusively to the  $\beta$ -phosphate of the product GDP, that is quickly released as fast off-rates indicate. [108] The mechanism proposed here does not involve an oxocarbenium ion intermediate but rather a highly activated pyrophosphate as leaving group in a S<sub>N</sub>2 type mechanism.

Some other FucTs have been proposed to act via an S<sub>N</sub>1-like mechanism involving a more or less separated oxocarbenium–phosphate ion pair. [51, 118] However, for FUT8, there is evidence that favors an S<sub>N</sub>2-like mechanism. First, catalytic efficiency is easier to achieve for the enzyme selecting a mechanism with a lower energy intermediate species. Second, the enzyme cannot stabilize an oxocarbenium ion with a basic residue itself. In accordance with the Hammond postulate this corresponds to a (stabilized) transition state of the lowest possible free energy.

#### 4.4 Chitotriose derivatives as FUT8 inhibitors

Specific inhibition of FUT8 will require exploiting its unique specificity for the natural acceptor molecule. The data presented on FUT8 in previous sections of this thesis has revealed much of the enzyme's mechanism for substrate recognition and catalysis. The results, however, clearly indicate that acceptor substrate recognition is based on many weak contacts. It is undoubtedly very difficult to mimic the large epitope of the acceptor that is recognized by FUT8 with small molecules. Modifying the natural acceptor oligosaccharide in order to increase binding affinity is likewise extremely challenging due to numerous functional groups exhibited by such compounds. Besides, complex type oligosaccharides are expensive and only available in very limited amounts. As illustrated in the introduction, FUT8 is also not active

on smaller fragments of the natural acceptor, e.g. the core pentasaccharide or chitin. However, a secondary activity on chitooligosaccharides was recently observed for FUT8. [47] This activity is probably based on its structural similarity to NodZ, a bacterial enzyme that fucosylates chitooligosaccharides via an  $\alpha$ 1,6-linkage. Chitotriose was identified as a minimal structure with a reported activity of 6% of that on the natural acceptor oligosaccharide. [47] The trisaccharide exhibits a manageable number of hydroxyl groups for modification by means of organic synthesis. In addition, chitotriose is available from chitin in g amounts and hence offers a cheap alternative as a starting material. Therefore, an approach to inhibit FUT8 based on its activity on chitotriose was chosen. In the following, the acceptor substrate qualities of chitotriose are further evaluated by means of kinetic analysis and *in silico* studies. Based on the knowledge about the mechanism of FUT8, a chitotriose derivative is designed as putative selective inhibitor for FUT8. Finally, the development of a synthetic route for the inhibitor is presented.

#### 4.4.1 Chitotriose is an acceptor substrate for FUT8 with low efficiency

As the enzyme kinetics of FUT8 reacting with chitotriose are not known, the reaction was analyzed by progress curves that were acquired by means of proton NMR as described above. In order to simplify the spectra, 1- $\beta$ -*N*-Acetyl chitotriose **8b** was employed as acceptor substrate instead of the reducing trisaccharide. The starting concentration of 1- $\beta$ -*N*-Acetyl chitotriose was 2 mM, as the  $K_M$  was expected to be much higher than that of the natural acceptor. Analysis of the spectra clearly showed that hydrolysis of GDP-Fuc, a reaction that can be considered as an unspecific transfer reaction to a water molecule as acceptor substrate, is in strong competition to the transfer of the fucosyl residue to chitotriose. Figure 4-31 shows the complete progress curves that had been measured over 60 h. A second immediate observation is that the transferase reaction becomes very slow when the concentration of the acceptor has reached approximately 1 mM. As no product inhibition is expected, the slow reaction rate indicates a  $K_M$  value for 1- $\beta$ -*N*-Acetyl chitotriose significantly above 1 mM. The competing reactions in addition to the high  $K_M$  prevent a quantitative analysis of the progress curves. Considering these findings, it is questionable if it is possible to characterize the pure transfer reaction to chitotriose quantitatively. The high  $K_M$  value requires starting concentrations of approximately 10-20 mM for both acceptor and donor substrate. The initial rates obtained from this assay therefore are not expected to correspond to  $k_{cat}$  in case of the transferase reaction. Hence, their initial reaction rate constant, indicated as  $k_{ini}$ , was calculated from the data.

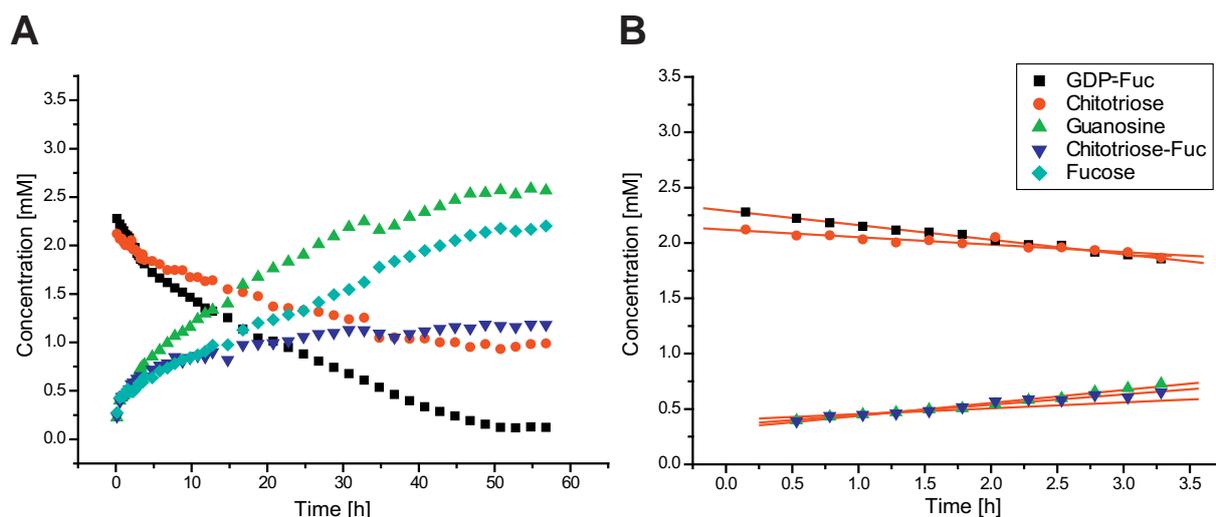


Figure 4-31: Progress curves of the transfer of a fucosyl moiety to the O-6 of chitotriose catalyzed by FUT8. The sample contained 3 mM GDP-Fuc, 2 mM  $\beta$ -1-*N*-Acetyl chitotriose **8b**, 2.4  $\mu$ M FUT8, 1 mg/mL BSA and 10 U AP in *Mes-d*<sub>13</sub>-D<sub>2</sub>O, pH 7.0. A: Plot of complete progress curves of all compounds reveals that the fucosyl transfer to chitotriose is in competition to hydrolysis of GDP-Fuc (i.e. fucosyl transfer to a water molecule). B: Analysis of initial rates allows estimating  $k_{ini}$  for both reactions. Up to 15 h the formation of Chitotriose-Fuc is somewhat faster than the hydrolysis. After that time point hydrolysis rate exceeds fucosyl transfer.

The apparent catalytic constants for the two reactions are shown in Table 4-10. Analysis of the initial rates of the formation of products and depletion of substrates allows estimating the rates for both reactions separately. Formation of guanosine and depletion of GDP-Fuc reflects with the sum of both hydrolysis and transfer. Formation of Fucose, on the other hand, proceeds with the rate of pure hydrolysis whereas formation of fucosylated chitotriose and depletion of chitotriose show the rate of the pure transferase reaction. Furthermore, the rate of hydrolysis was estimated from the second linear part of GDP-Fuc consumption that predominantly corresponds to pure hydrolysis when the transferase reaction had slowed down. The observed rate constant for the transferase reaction is  $k_{ini} = 0.009 \text{ s}^{-1}$  and hence 50 times slower than the  $k_{cat}$  with the natural substrate. The value, however, cannot directly be compared to  $k_{cat}$  as saturating substrate concentrations are not achieved in the assay due to the high  $K_M$  value. The observed 2% of activity is in agreement with the value estimated by Ihara *et al.* Pure hydrolysis proceeds with  $k_{cat} = 0.005 \text{ s}^{-1}$  and is 100 times slower than the natural transferase reaction. The apparent  $k_{cat}$  values for both reactions is  $k_{cat} = 0.014 \text{ s}^{-1}$  and consistent with the values determined for the single reactions.

Table 4-10: Apparent catalytic constants extracted from the progress curves shown in Figure 4-31. Values marked with an asterisk \* were not determined at saturating concentration of acceptor substrate. Therefore, the real  $k_{cat}$  is expected to be higher.

Reaction	Data	$k_{ini}$ [ $s^{-1}$ ]
Hydrolysis and transfer to chitotriose	Initial rate of depletion of GDP-Fuc	0.015
	Initial rate of guanosine formation	0.013
Transfer to chitotriose	Initial rate of depletion of chitotriose	0.0078*
	Initial rate of product formation	0.011*
Hydrolysis	Initial rate of fucose formation	0.0061
	Rate of GDP-Fuc depletion after 3.5 h	0.0042

The analysis of the fucosyl transfer to chitotriose by means of progress curve analysis with  $^1H$  NMR illustrates well how valuable this method is for characterizing complex enzyme reactions.

#### 4.4.2 Chitotriose shows a binding mode similar to the natural acceptor

In order to evaluate how chitotriose interacts with the acceptor binding site of FUT8, a MD simulation based on the model of the ternary complex presented in Section 4.3 was performed. Therefore, a model of 1- $\beta$ -*N*-acetyl chitotriose was overlaid with the heptasaccharide from the last frame of the MD simulation of the ternary complex by fitting the carbon atoms of chitotriose to the core trisaccharide of the natural acceptor. Afterwards, the heptasaccharide was deleted from the model and the chitotriose ternary complex was minimized prior to the MD simulation. As expected, the average potential energy of the complex lies between that of the FUT8–GDP-Fuc complex (highest energy) and that of the FUT8 ternary complex (lowest energy). Like the natural receptor, chitotriose maintains its general position over the simulation time. Evaluation of the trajectories (cf. Figure 4-33, Figure 4-34 and Figure 4-35) of distances between interacting atoms of the enzyme and chitotriose reveals a more detailed picture, illustrating the mechanism for the rather weak recognition of chitotriose: Most importantly, the strong hydrogen bond between the oxygen atom of the  $\beta$ -phosphate of GDP-Fuc and the O-6 of GlcNAc1 is also present with chitotriose as acceptor molecule. However, this hydrogen bond is much weaker (average distance:  $3.3 \pm 0.8$  Å, cf. Table 4-11) than the corresponding interaction in the complex with the heptasaccharide (average distance:  $2.6 \pm 0.1$  Å). Thus, one of the interactions identified as crucial for affinity towards the natural acceptor is attenuated with chitotriose, giving an

account of the more than 1000fold increased  $K_M$  value. Additionally, this softened hydrogen bond has grave consequences on the turnover of the substrate, as deprotonation of the acceptor is impaired and hence nucleophilicity drops. Two reasons may underlie this weakened key hydrogen bond: First, for obvious reasons, chitotriose is not able to undergo most of the interactions of the natural acceptor leading to a less accurate positioning of GlcNAc1. Second, the partly altered pattern of hydrophilic interactions of amino acid residues with chitotriose (compared to chitobiose as part of the natural acceptor) may result in a slightly different preferred position of GlcNAc1 in the ternary complex. Whereas all of the most persistent hydrogen bonds with the chitobiose unit of the natural acceptor are maintained by chitotriose with a generally larger equilibrium distance (e.g. Tyr498-O $\eta$  GlcNAc1-O3, Gln470-O $\epsilon$  GlcNAc1 N1, Ser468CO GlcNAc2 N2, Asp495-O $\delta$  GlcNAc2-O6), other essential amino acids undergo alternative interactions. Altered interactions are obviously observed with GlcNAc3 to a large part. The hydrogen bonds of GlcNAc3 exhibit a similar intermittent and dynamic behavior like the monosaccharide units of the branches of the natural acceptor (cf. Figure 4-35). Interestingly, GlcNAc3 was found to interact with Asp494, Lys541 and Lys369. With the natural acceptor, these residues bind exclusively monosaccharide units in both branches. These ‘surrogate’ interactions with GlcNAc3 might lead to the altered positioning of the molecule. They likewise explain why chitotriose is recognized as a substrate at all, implying key roles of the involved residues Asp494, Lys541 and Lys369. The latter has already been proven to be essential to activity of FUT8 [44].

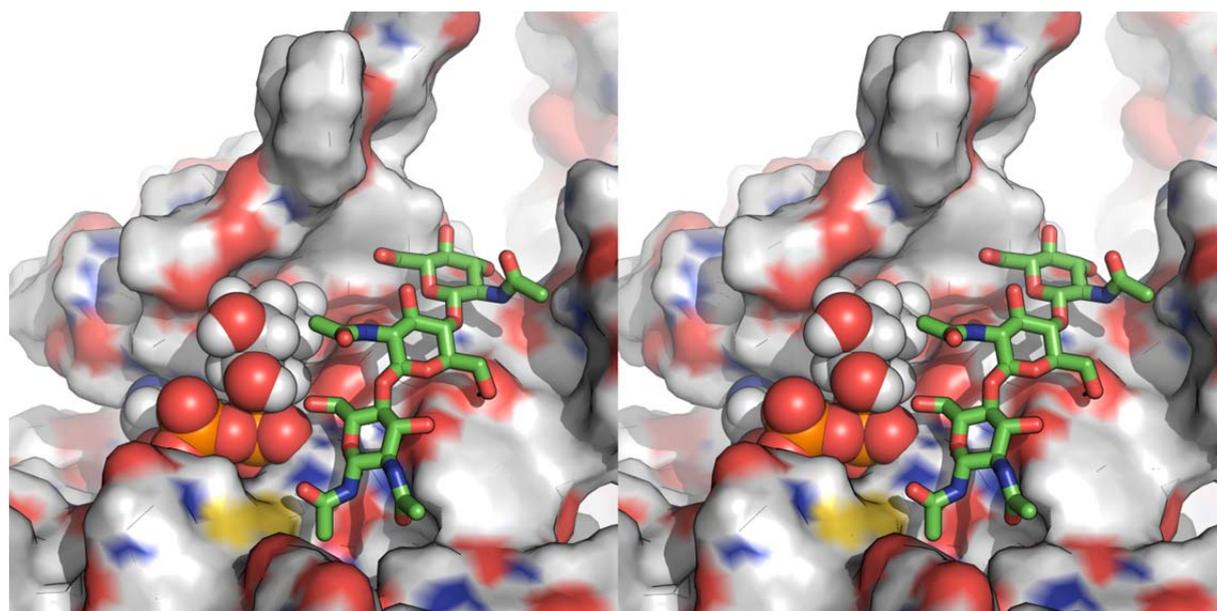


Figure 4-32: Representative frame (frame 466) taken from the MD simulation illustrating binding of chitotriose to the FUT8–GDP-Fuc *in silico*.

## RESULTS AND DISCUSSION

Table 4-11: Contacts between FUT8 or donor substrate with chitotriose as acceptor substrate during the MD simulation.

<b>FUT8 or donor Residue</b>	<b>Atom</b>	<b>Chitotriose residue</b>	<b>Atom</b>	<b>Type of Interaction</b>	<b>Average distance [Å]</b>	<b>Fraction of Frames with distance &lt; 3.6</b>
GDP-Fuc	Pβ O-2	GlcNAc1	O-6	H-bond	3.34 ± 0.82	69.7%
GDP-Fuc	C5"	GlcNAc2	CH <sub>3</sub>	hydrophobic	5.05 ± 0.53	0.0%
GDP-Fuc	O-2"	GlcNAc1	O-6	H-bond	3.10 ± 0.46	88.3%
Tyr498	Oη	GlcNAc1	O-3	H-bond	3.60 ± 0.6	47.6%
Gln470	Nε	GlcNAc1	O-5	H-bond	2.96 ± 0.54	88.1%
Gln470	Oε	GlcNAc1	NH-1	H-bond	4.09 ± 0.70	28.8%
Arg473	Nη	GlcNAc1	NH-2	H-bond	6.09 ± 0.61	0.4%
Cy222	Sγ	GlcNAc1	CH <sub>3</sub>	hydrophobic	4.46 ± 0.69	7.8%
Asp495	Oδ	GlcNAc2	O-6	H-bond	3.77 ± 0.95	60.3%
Ser468	CO	GlcNAc2	NH	H-bond	5.51 ± 0.95	8.0%
Asp495	Oδ	GlcNAc2	O-6	H-bond	3.77 ± 0.95	60.3%
Asp494	CO	GlcNAc3	O-3	H-bond	4.03 ± 0.81	31.1%
Lys541	Nε	GlcNAc3	O-4	H-bond	7.82 ± 2.5	3.7%
Asp459	Oδ	GlcNAc3	NH	H-bond	4.65 ± 0.45	0.9%
Lys369	Nε	GlcNAc3	O-6	H-bond	8.53 ± 2.62	5.9%

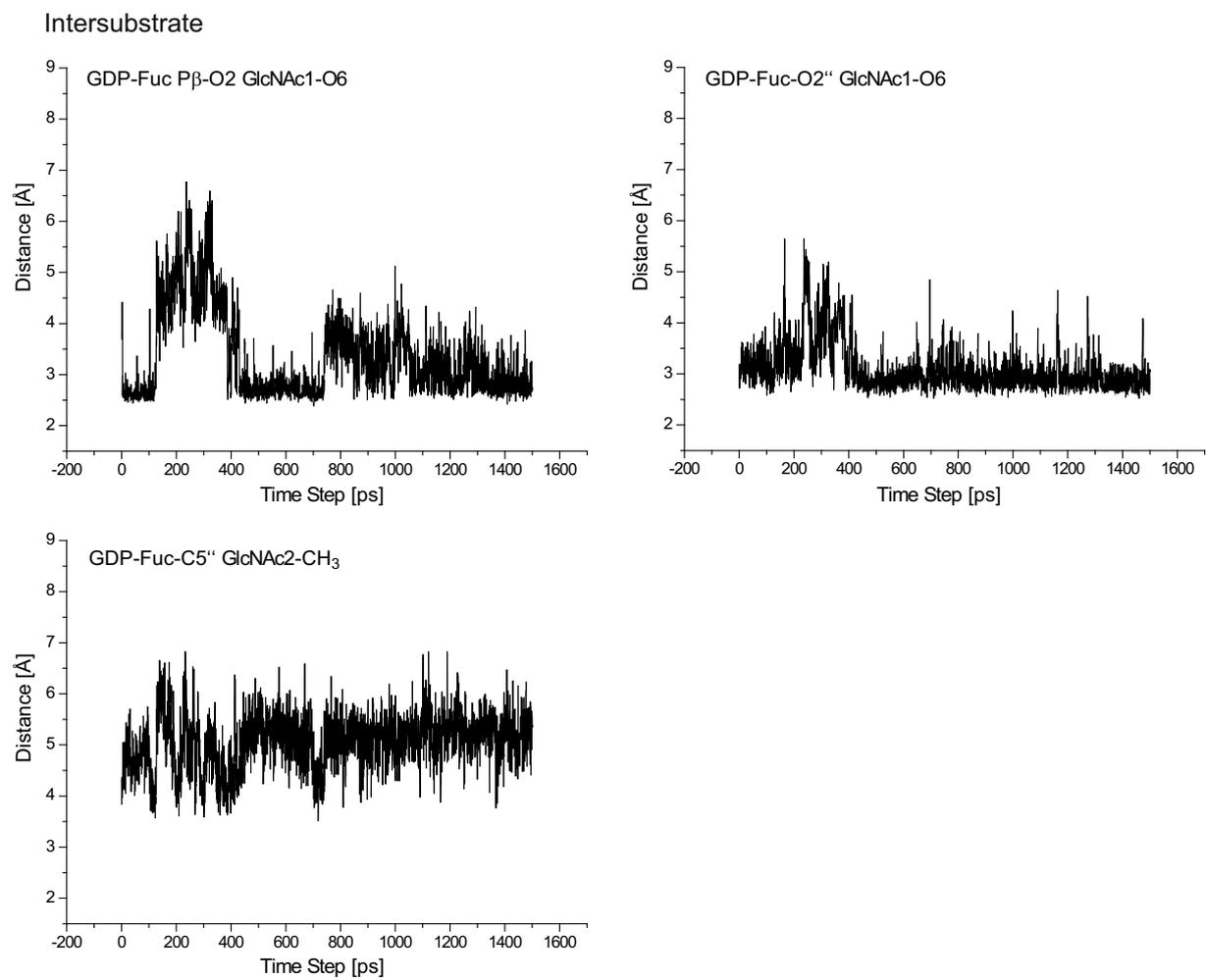


Figure 4-33: Trajectories of distances of atoms the donor substrate and of chitotriose as acceptor substrate.

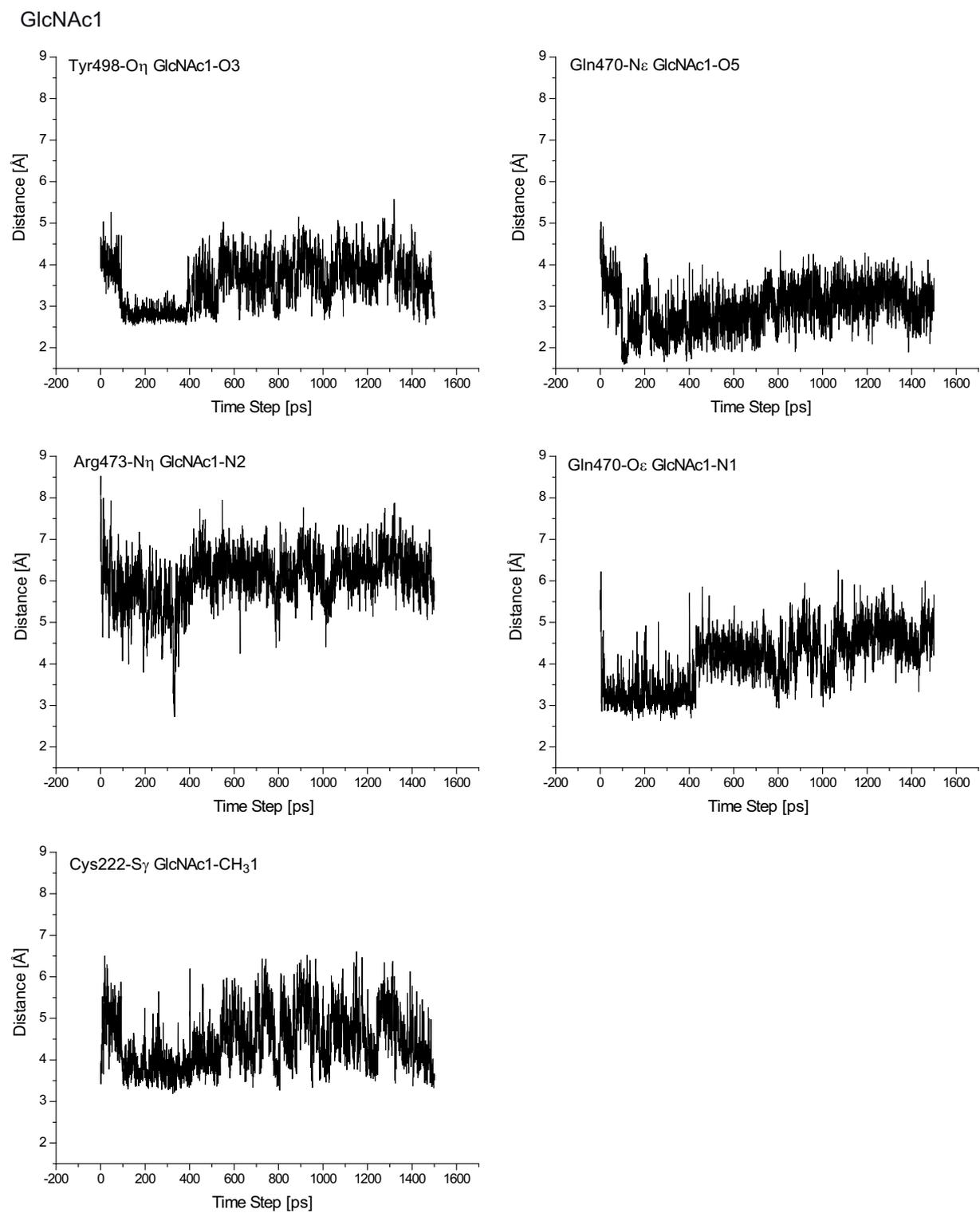


Figure 4-34: Trajectories of distances of atoms of FUT8 to atoms of GlcNAc1 of chitotriose as acceptor substrate.

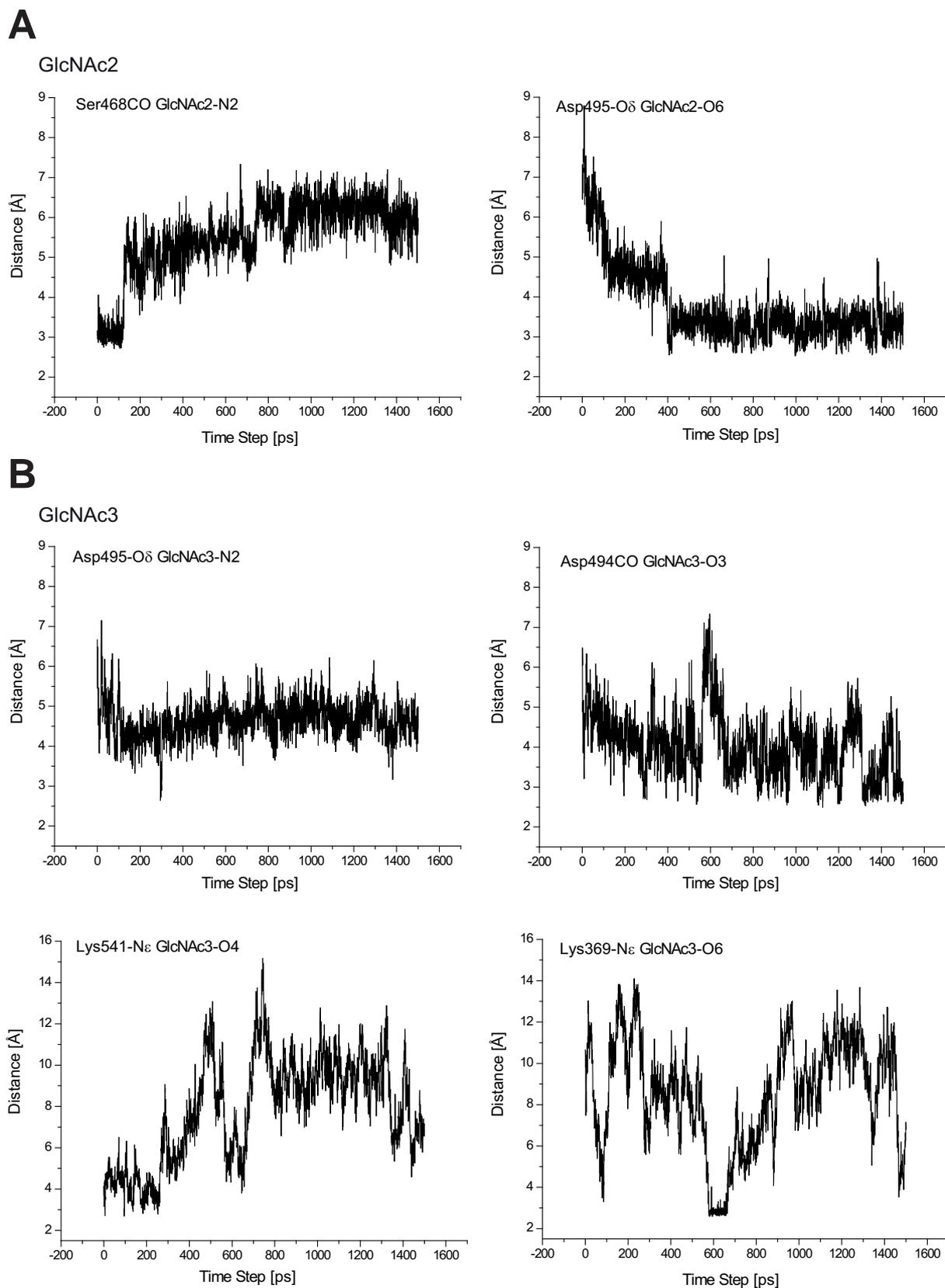


Figure 4-35: Trajectories of distances of atoms of FUT8 to atoms of GlcNAc2 (A) and GlcNAc3 (B) of chitotriose as acceptor substrate.

#### 4.4.3 A chitotriose derivative to specifically block Arg365

Chitotriose has been identified as an alternative acceptor substrate specific to FUT8 among mammalian FucTs (cf. Ref. [47] and Section 4.4.2). Its low affinity, however, will not suffice for generating a potent inhibitor against the natural acceptor. As illustrated above, gaining affinity in the acceptor binding region is wasteful as the most stable interactions are found in the donor substrate binding site. The key role of Arg365 in catalysis and the position of Arg365 adjacent to the acceptor binding site suggest a modification of the acceptor molecule that enables it to bind the guanidinium group of Arg365. As electrostatic interactions are known to exhibit highest gain in free enthalpy upon binding compared to other interaction types, it seems appropriate to offer a negatively charged group at this position. The modification would logically be realized at the O-6 of GlcNAc because of its proximity to Arg365 and the fact that it interacts with the donor substrate in the ternary complex. The linker between this group and the GlcNAc1 O-6 of chitotriose ought to be designed in such a way that allows for additional hydrophobic interactions in the fucose binding region of FUT8. Furthermore, precursors for the derivatization should be readily available.

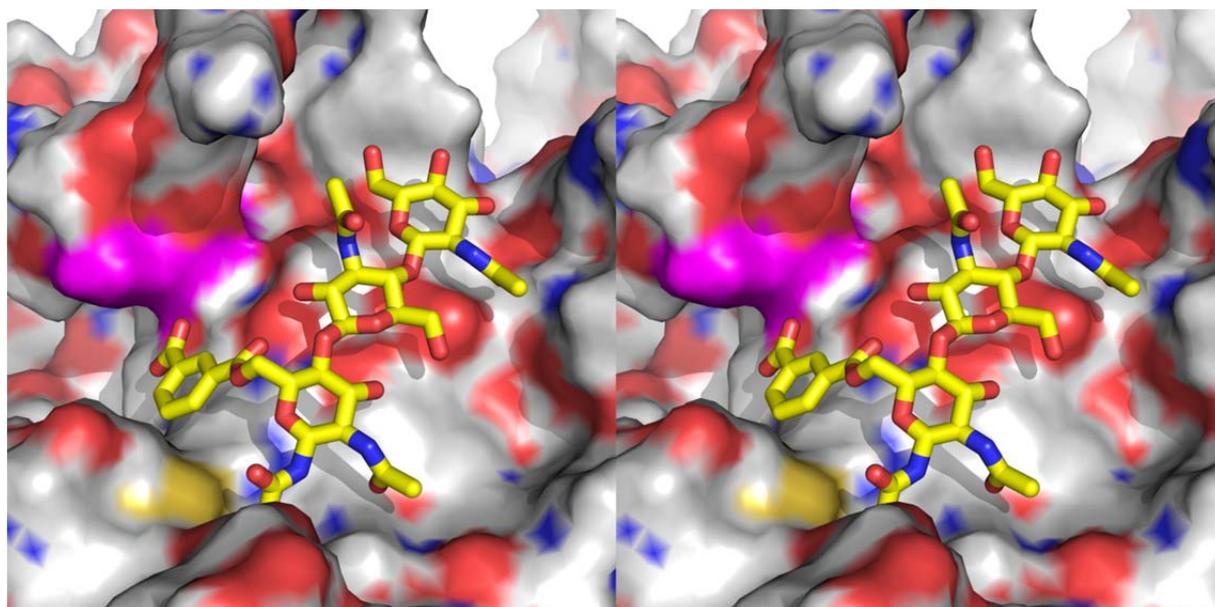


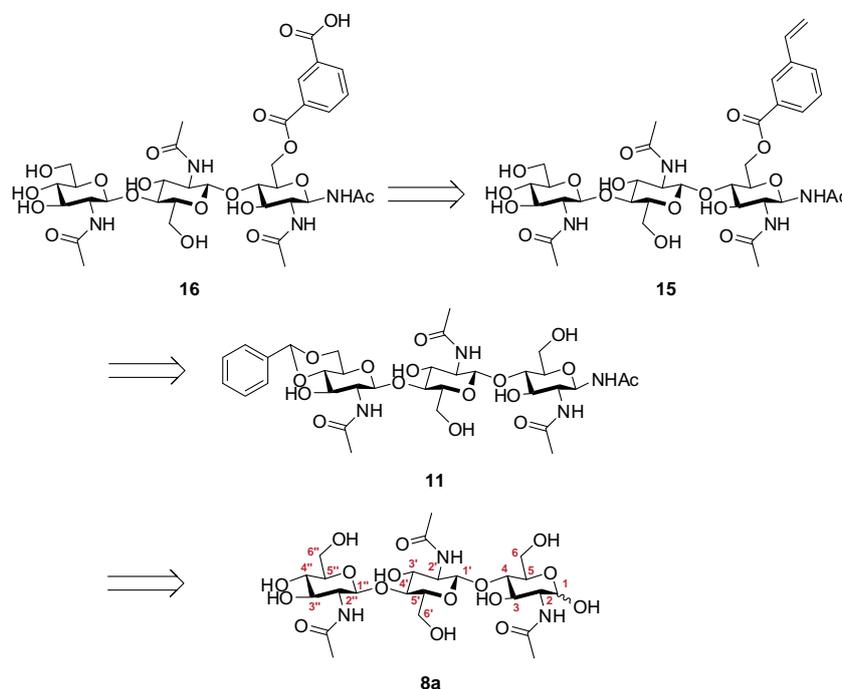
Figure 4-36: Crossed-eye stereo view of the proposed inhibitor 6-(2-carboxy)benzoyl-1- $\beta$ -N-acetyl chitotriose **16** (stick model with yellow carbon atoms) in complex with FUT8 (surface representation in atom color). The key catalytic residue Arg365 is highlighted in magenta.

Based on these requirements, several linkers have been designed and tested *in silico* via energy minimization. Among these linkers, aromatic species were judged to be most promising, as their rigidity is accompanied by low entropic cost in the binding process. A carboxy function in *meta* position to the carbohydrate-linked carbon atom was found to exhibit appropriate orientation and positioning to bind the guanidinium group. The resulting

molecule is shown in Figure 4-36. The linkage to the O-6 of chitotriose was chosen to be an ester (although an ether linkage yielded slightly better results) due to the reduced synthetic effort associated with esterification.

#### 4.4.4 Towards a synthetic route with minimal protection strategy

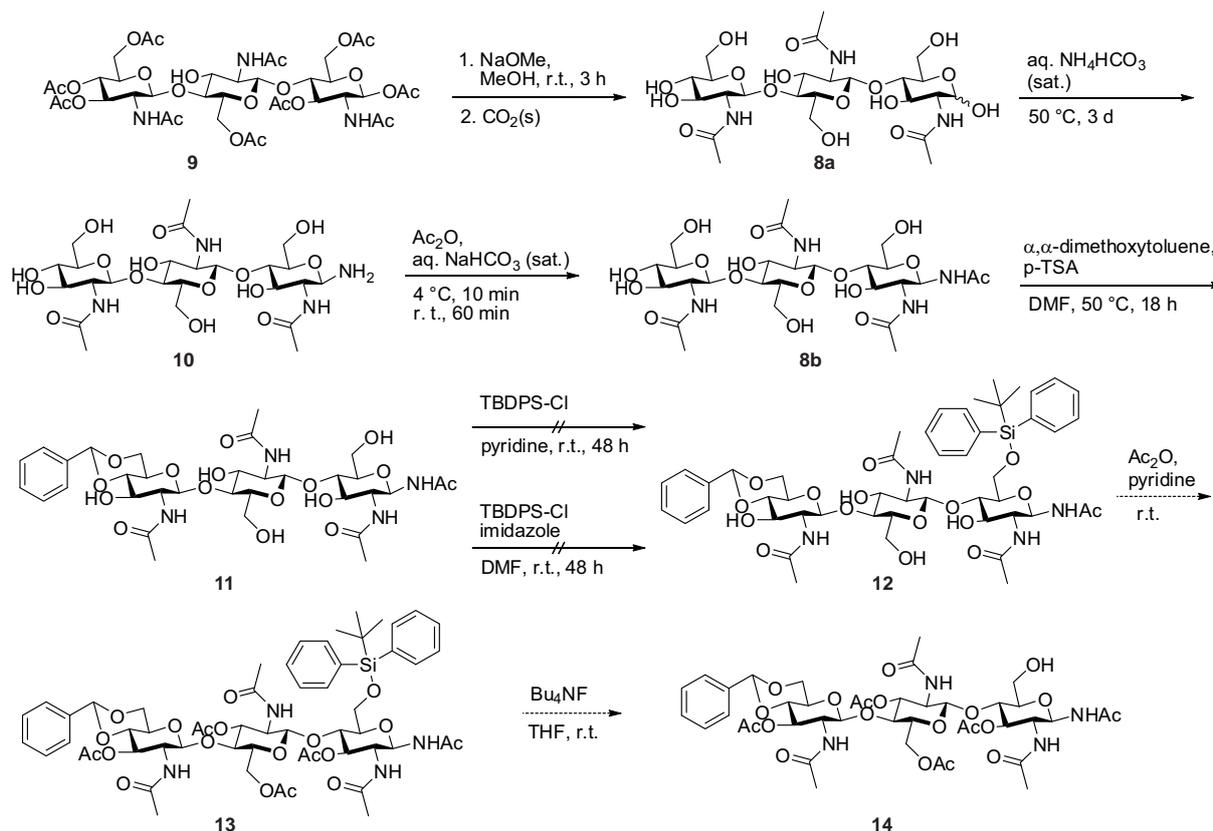
The retro synthetic route for **16** is depicted in Scheme 4-3. In order to esterify the 6-hydroxyl function of GlcNAc1 selectively, the naturally higher accessibility of 6-hydroxy groups as compared to secondary hydroxyl groups at other positions can be exploited. However, as chitobiose exhibits three primary hydroxyl groups, a protection group strategy has to be developed. Protection of the 6''-OH group is straight forward: As GlcNAc3 is the only monosaccharide unit having a 4-hydroxyl group, both 4''- and 6''-hydroxyl group can be selectively protected by means of a 4'',6''-*O*-benzylidene group. Differentiation between the 6'- and 6-hydroxy function may be accomplished employing a bulky protection group such as *tert*-butyl diphenyl silyl (TBDPS) because it is likely that the 6-position of GlcNAc1 is more accessible than that of GlcNAc2, subsequent peracetylation and selective deprotection of the 6-hydroxy group. Finally, esterification with 3-vinyl benzoic acid and subsequent ozonolysis yield 6-(2-carboxy)benzoyl-1- $\beta$ -*N*-acetyl chitotriose **16**



Scheme 4-3: Retrosynthetic approach towards 6-(2-carboxy)benzoyl-1- $\beta$ -*N*-acetyl chitotriose **16**.

The detailed protection strategy is shown in Scheme 4-4. The reaction sequence proceeds from chitotriose octa-*O*-acetate (**9**), which is obtained from acidic hydrolysis of chitin. Deacetylation with methanolate yields reducing chitotriose (**8a**). In order to obtain uniform

anomeric configuration, **8a** is aminated (**10**) and 1- $\beta$ -acetylated to yield 1- $\beta$ -*N*-acetyl chitotriose **8b**. 4'',6''-*O*-benzylidation of **8b** yielding 1- $\beta$ -*N*-acetyl 4'',6''-*O*-benzylidene chitotriose **11** could be accomplished using standard conditions. However, due to the low solubility of unprotected chitotriose in DMF,  $\alpha,\alpha$ -dimethoxytoluene had to be applied in large excess in order to achieve a reasonable reaction rate. As a consequence, the reaction time had to be adjusted in such a way that generation of open acetals at positions 6' and 6 was kept at a minimum. Purification of **11** from these byproducts and unreacted **8b** turned out to pose a major challenge due to the amphiphilic character of **11**. Its low solubility in water and in all tested organic solvents that are suitable for chromatography had the consequence that purification by means of RP HPLC was inefficient. Peracetylation of **11** in order to subject it to normal-phase chromatography proved to be as inefficient, probably due to the slow reaction rates of the hydroxyl groups at positions 4' and 4. Hence, RP HPLC purified **11** was used for further reaction steps although the purification protocol clearly needs to be optimized for larger synthetic approaches.

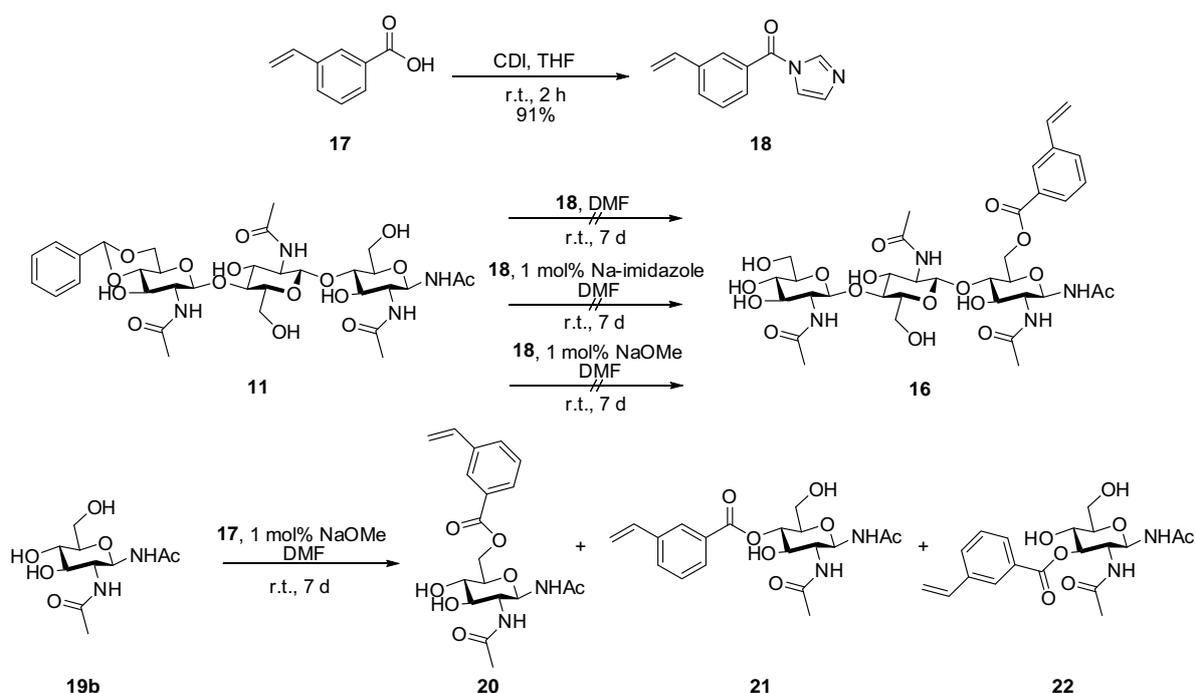


Scheme 4-4: Synthetic route for preparing selective esterification of the 6-hydroxy group involving an explicit protection strategy.

Attempts to protect OH-6 of **11** with TBDMS to yield **12** failed in all cases. The reasons lie probably in the extreme steric requirement of TBDMS indicating that smaller silane

protecting reagents may be more appropriate. However, instead of a smaller protection group, direct esterification seems to be the more elegant approach. As the accessibility of OH-6 turned to be less pronounced than initially expected, the size of 3-vinyl benzoate might suffice in order to achieve discrimination of OH-6 from OH-6'. In addition, even if the reaction yields a product mixture with a ratio of 1:1 in the worst case, both products are expected to give significantly different retention times on a (normal phase) chromatographic column.

Esterification of hydroxyl groups from carbohydrates is straight forward using acid chlorides from the respective carboxylic acids. The high reactivity, however, renders selectivity for OH-6 more improbable. Therefore, the imidazolyl-derivative of 3-vinyl benzoate was synthesized from 3-vinyl benzoic acid (**17**) according Scheme 4-5 in order to give a moderately activated benzoic acid derivative **18** [119]. Esterification with the less reactive aromatic carboxylic imidazolides can be accelerated by adding catalytic amounts of methanolate or sodium imidazolidate [120]. However, all approaches to esterify **11** applying these protocols were unsuccessful (c.f. Scheme 4-5). Esterification of 1- $\beta$ -*N*-acetyl GlcNAc as a testing compound indeed showed the desired esterification of the 6-position (and only minor esterification at the other positions that are much less accessible in **11**). However, despite extremely long reaction times of about 7 d, unreacted precursor **19b** was still left over. These observations strongly imply that reactivity of the imidazolide **18** is not sufficient for the chosen substrate **11**. Therefore, further attempts aim to esterification employing 3-vinyl benzoic acid chloride.



Scheme 4-5: Synthetic steps towards esterification of OH-6 of chitotriose derivative **11** in order to directly give the desired product **16**.

## 5 Summary

The products of fucosyltransferases are fucosylated oligosaccharides that are involved in many physiological and pathophysiological processes. Here the  $\alpha$ 1,3-fucosyltransferase FUTA and the  $\alpha$ 1,6-fucosyltransferase FUT8 are analyzed in detail. Information on the structure and detailed mechanism of fucosyltransferases is, however, scarce. Because of the biological importance specific inhibitors for these enzymes would be of high interest.

Both fucosyltransferases studied here act on the proximal GlcNAc of N-glycans. Honeybee (*A. mellifera*) FucTA, the key enzyme in biosynthesis of the immunogenic core- $\alpha$ 1,3-fucosyl epitope, was studied to get – without X-ray structure analysis – information on the enzyme's mechanism, the binding of the substrate and the kinetics of the enzyme. Both structure and reaction mechanisms that lead to the formation of the epitope were analyzed by biophysical methods. Donor (GDP-Fuc) and acceptor (a complex type heptasaccharide) binding to FucTA were elucidated with STD NMR, SPR and enzyme kinetic methods. These results demonstrate that FucTA binds GDP-Fuc mainly via the guanosine part and also via the pyrophosphate. As acceptor substrates FucTA prefers N-glycans that are not core  $\alpha$ 1,6-fucosylated. Furthermore, a large binding epitope with emphasis on the core pentasaccharide was found for the acceptor substrate, indicating that FucTA recognizes a large part of the heptasaccharide. Progress curve analysis of FucTA by NMR was established to determine the enzyme kinetics of FucTA. FucTA was used to prepare and characterize mg amounts of core  $\alpha$ 1,3-fucosylated N-glycans. With these compounds, that are difficult to access otherwise, recognition of the immunogenic structure by polyclonal antibodies directed against CCD was analyzed by means of STD NMR and it was found that these antibodies bind mainly the core structure Fuc<sub>2</sub>GlcNAc of the N-glycans.

In order to design specific inhibitors, structural information on substrate recognition with atomic resolution and on the catalytic mechanism is required. Human FUT8 was chosen since a crystal structure of the apo enzyme was available. Ligand based NMR analysis in combination with *in silico* methods yielded a model of donor substrate binding of human FUT8. The model of the complex provides detailed insight into substrate–enzyme contacts and is consistent with experimental data from NMR and SPR studies and also with prior mutation studies. The results suggest that GDP-Fuc is recognized in a highly specific manner by hydrogen bonds of guanine with the side chain of the essential Asp453 and the side chain of His363. Binding affinity of the donor molecule to FUT8 is dependent on the  $\beta$ -phosphate group as found by SPR experiments in this study. Compared to FucTA, binding of the

guanine is less pronounced, indicating a different recognition process. The model provides the structural basis for this finding with multiple very stable hydrogen bonds that the pyrophosphate moiety of the donor forms with the amide functions of the backbones in the Rossman fold. The fucose moiety points towards a shallow region that is able to accommodate the acceptor molecule of FUT8. Finally, the model presented on the FUT8–GDP-Fuc complex gives implications for the role of the essential Arg365: Besides binding of the  $\beta$ -phosphate, Arg365 assists in GDP release and proper orientation of the fucose residue for nucleophilic attack of the acceptor. The specific interactions of Arg365 and of Gln470 with the fucose also ensure discrimination between GDP-Fuc and GDP-Man.

In the next step, the model was extended to the binding of the acceptor substrate. Like for FucTA, an unusually large acceptor is required as minimal structure. Acceptor substrate binding of FUT8 was experimentally studied by STD NMR and SPR and theoretically by molecular dynamics simulations. The acceptor binding site was identified as shallow cavity between the Rossman fold, a very flexible loop and the C-terminal  $\beta$ -strands. Compared to the donor, the acceptor has a faster kinetics of dissociation. In addition, binding of the acceptor was found to be much faster and stronger if the donor is present. This is due to strong hydrogen bonding between the O-6 of the proximal GlcNAc and an oxygen atom of the  $\beta$ -phosphate of GDP-fucose. From these data, an ordered bi-bi-mechanism for FUT8 is proposed where the donor molecule is the leading substrate. No specific amino acid is present that could act as base during catalysis. The experimental and theoretical results indicate a substrate-assisted mechanism, where a  $\beta$ -phosphate oxygen deprotonates the acceptor. The specific recognition of the complex acceptor substrate is also provided by the flexible loop which binds the 6-mannose branch and by specific amino acids of  $\beta$ 11 and the flexible loop that are in contact with the 3-mannose branch of the acceptor substrate.

Finally, the knowledge on substrate recognition and enzymatic mechanism of FUT8 was used to design a specific inhibitor based on chitotriose, which represents a weakly active acceptor substrate for FUT8 that can more easily be chemically modified than the natural acceptor. Enzyme kinetics of the transfer of fucose to chitotriose revealed that the affinity to this substrate is very low. These results were further supported with *in silico* methods, indicating less persistent key interactions for the core chitobiose unit. With the help of this model, a derivative of chitobiose was developed that is able to block the key catalytic residue Arg365. Ultimately, a synthetic strategy for this compound was developed.

## 6 Zusammenfassung

Die Produkte der Fucosyltransferasen sind an einer Vielfalt von wichtigen biologischen Prozessen beteiligt. Es ist daher von großem Interesse, spezifische Inhibitoren für diese Enzyme zu entwickeln. Die dafür benötigten detaillierten Informationen über Struktur und Mechanismus sind jedoch größtenteils noch nicht verfügbar. Daher ist das strukturbasierte Design spezifischer Inhibitoren für einzelne Fucosyltransferasen sehr schwierig.

In dieser Arbeit wurden zwei Fucosyltransferasen untersucht, die das proximale GlcNAc von N-Glycanen modifizieren. Das bestehende Expressionssystem für die FucTA der Honigbiene, das Schlüsselenzym in der Biosynthese des immunogenen *core- $\alpha$ 1,3-Fucosylepitops*, bildete die Basis um die komplexen Mechanismen der Biosynthese und der Erkennung dieser Strukturen durch das Immunsystem zu untersuchen. Sowohl die Strukturen als auch die Reaktionsmechanismen, die zu der Bildung des Epitops führen, wurden mit biophysikalischen Methoden untersucht. Dabei wurden die Bindung von Donorsubstrat (GDP-Fuc) und Akzeptorsubstrat (ein Heptasaccharid vom komplexen Typ) an FucTA mit STD-NMR-Spektroskopie, SPR und enzymkinetischen Methoden untersucht. Die Ergebnisse zeigen dass FucTA GDP-Fuc hauptsächlich über das Guanosin und zusätzlich über das Pyrophosphat erkennt. Im Falle des Akzeptorsubstrats wird eine Struktur bevorzugt, die nicht bereits am Kern  $\alpha$ 1,6-fucosyliert ist. Außerdem wurde ein ausgedehntes Bindungsepitop bestimmt, das zeigt, dass FucTA einen großen Teil seines Akzeptorsubstrats bindet. Diese Daten beleuchten die allgemeinen Prinzipien, die der Biosynthese und Modifikation von Glycanstrukturen zugrunde liegen. Außerdem konnten mit der Untersuchung von FucTA wertvolle Werkzeuge, wie die Progresskurvenanalyse für Glycosyltransferasereaktionen mit NMR, etabliert werden. Weiterhin konnten die Substrate, also reduzierende und 1- $\beta$ -N-acetylierte N-Glycane, die für die Untersuchung von FucTA in dieser Arbeit dargestellt wurden, für die enzymatische Synthese von mg-Mengen der sonst schwierig zugänglichen *core- $\alpha$ 1,3-fucosylierten N-Glycane* in genutzt werden. Mit diesen Verbindungen konnte die Erkennung der immunogenen Struktur durch polyklonale anti-CCD-Antikörper mit Hilfe von STD NMR untersucht werden und gezeigt werden, dass diese hauptsächlich die Kernregion des N-Glycans erkennen.

Um spezifische Inhibitoren zu entwickeln wird strukturelle Information der Substraterkennung und des Katalysemechanismus mit atomarer Auflösung benötigt. Es wurde humane FUT8 gewählt, da zu dieser bereits eine Kristallstruktur des Apoenzyms veröffentlicht war. Die Kombination aus ligandbasierten NMR-Methoden und *in-silico-*

Studien lieferte ein Modell der Donorsubstratbindung von FUT8. Dieses Modell zeigt detaillierte Kontakte zwischen Enzym und Substrat und stimmt sowohl mit den experimentellen Ergebnissen der NMR- und SPR-Studien als auch mit denen aus früheren Mutationsstudien überein. Die Ergebnisse zeigen dass GDP-Fuc hochspezifisch über Wasserstoffbrückenbindungen des Guanins mit den Seitenketten von Asp453 und His363 erkannt wird. Weiterhin konnte mit SPR herausgefunden werden, dass die Bindungsaffinität des Donormoleküls von der  $\beta$ -Phosphatgruppe abhängig ist. Im Modell wird diese Erkenntnis strukturell durch die Anwesenheit mehrerer fester Wasserstoffbrückenbindungen zwischen dem Pyrophosphat und den Amidfunktionen des Peptidrückgrats in der Rossmannfalte erklärt. Die Fucoseinheit von GDP-Fuc zeigt in eine flache Region, die groß genug ist, um den Akzeptor unterzubringen. Schließlich liefert das Modell Hinweise auf die essentielle Funktion von Arg365: Außer der Bindung des  $\beta$ -Phosphats hilft dieser Rest bei der Aktivierung von GDP als Austrittsgruppe und sorgt für die richtige Orientierung des Fucoserestes für einen nukleophilen Angriff durch den Akzeptor. Die spezifischen Wechselwirkungen von Arg365 und Gln470 mit der Fucose garantieren außerdem die Unterscheidung zwischen GDP-Fuc und GDP-Man.

Im nächsten Schritt wurde das Modell auf den ternären Komplex erweitert. Wie FucTA benötigt FUT8 einen außergewöhnlich großen Akzeptor als Minimalstruktur. Die Akzeptorsubstratbindung wurde experimentell mit STD NMR und SPR und theoretisch mit MD-Simulationen untersucht. Die Akzeptorsubstratbindungstasche wurde als eine flache Vertiefung zwischen Rossmannfalte, einer flexiblen Schleife und den C-terminalen  $\beta$ -Strängen identifiziert. Verglichen mit dem Donor weist der Akzeptor eine schnellere Assoziations- und Dissoziationskinetik auf. Außerdem wurde herausgefunden, dass die Akzeptorbindung in Anwesenheit des Donors deutlich stärker ausfällt. Der Grund dafür ist wurde eine starke Wasserstoffbrückenbindung zwischen dem O-6 des proximalen GlcNAcs und einem Sauerstoffatom des  $\beta$ -Phosphats im ternären identifiziert. Aus diesen Daten ergibt sich ein geordneter Bi-Bi-Mechanismus in dem der Donor das führende Substrat ist. FUT8 weist keinen Aminosäurerest auf, der als Base während der Katalyse infrage kommt. Die experimentellen und theoretischen Ergebnisse dieser Arbeit legen einen substratunterstützten Mechanismus nahe, bei dem ein Sauerstoff der Phosphatgruppe den Akzeptor deprotoniert. Die spezifische Erkennung des Akzeptorsubstrats wird weiterhin über Aminosäuren der flexiblen Schleife, die den 6-Man-Arm binden, und der  $\beta$ -Stränge, die im Kontakt mit dem 3-Man-Arm stehen, erzielt.

Schließlich wurden die detaillierten Erkenntnisse über die Substraterkennung und den enzymatischen Mechanismus genutzt um einen spezifischen Inhibitor, basierend auf Chitotriose, zu entwerfen. Chitotriose ist ein schwach aktives Akzeptorsubstrat für FUT8, der einfacher chemisch modifiziert werden kann als der natürliche Akzeptor. Es wurde herausgefunden, dass die Affinität für Chitotriose sehr schwach ist und dieses im *in-silico*-Modell weniger dauerhafte Schlüsselinteraktionen aufweist als die entsprechende Chitobioseeinheit im natürlichen Akzeptor. Mithilfe dieses Modells wurde ein Derivat der Chitotriose entworfen, die den für die Katalyse essenziellen Rest Arg365 blockiert. Schlussendlich wurde eine Synthesestrategie für diese Verbindung ausgearbeitet.

## 7 Experimental procedures

### 7.1 Recombinant expression of fucosyltransferase 8

#### 7.1.1 Cloning of human fucosyltransferase 8

Total RNA was isolated from HEK293 (human embryonic kidney) cells using peqGold TriFast (Peqlab Biotechnologie, Erlangen, Germany). SuperScript III RT (Invitrogen, Karlsruhe, Germany) and the gene-specific primer 5'-GATCGCGCCGCTTATTTCTCAGCCTCAGGATATGTGGG-3' were used to synthesize cDNA of FUT8 from the isolated total RNA. The FUT8 coding region (Arg68 to Lys575) was amplified from cDNA with *Pfu* DNA polymerase (Fermentas, St. Leon-Rot, Germany) using the primers 5'-GATCTCTAGA GGGATCGAGGGAAGGCGGATACCAGAAGGCCCTATTG-3' adding an XbaI restriction site and 5'-GATCGCGCCGCTTATTTCTCAGCCTCAGGATATGTGGG-3' adding a NotI restriction site. The PCR product was subcloned via XbaI and NotI into the digested baculovirus transfer vector pAcGP67B (BD Pharmingen, Heidelberg, Germany) which was modified by addition of an N-terminal 10-fold His-tag, a V5 epitope, a factor Xa cleavage site as well as an XbaI restriction site. The sequence of selected subclones was verified by sequencing.

#### 7.1.2 Recombinant baculovirus production

*Spodoptera frugiperda* cells (Sf9) (Invitrogen, Karlsruhe, Germany) were grown at 27 °C in serum-free medium (Express Five SFM; Lonza, Verviers, Belgium containing 10 µg/mL gentamycin; Invitrogen, Karlsruhe, Germany). Cell density was determined by hemocytometer counts, cell viability was evaluated by staining with Trypan Blue. Recombinant baculovirus was generated by cotransfection of Sf9 cells with BaculoGold bright DNA (BD Pharmingen, Heidelberg, Germany) and the baculovirus transfer vector pAC-GP67-B containing *sFucTA*. High titer stocks were produced by three rounds of virus amplification and optimal MOI for protein expression was determined empirically by infection of Sf9 cells in 100 mL suspension flasks ( $1.5\text{--}2.0 \cdot 10^6$  cells/mL in 20 mL suspension culture) with serial dilutions of high titer virus stock.

#### 7.1.3 Expression in baculovirus-infected insect cells

The high titer stock of recombinant baculovirus was used to infect 400 mL suspension cultures of Sf9 cells ( $1.5 \cdot 10^6$  cells per mL) in 2000 mL flasks. For protein production the cells were incubated at 27 °C and 110 rpm for 72 h.

#### **7.1.4 Protein purification**

The supernatant of baculovirus-infected cells was collected, adjusted to pH 8 and centrifuged at  $4000 \times g$  for 5 minutes. The buffer was exchanged to PBS, pH 8 applying the QuixStand Benchtop System (GE Healthcare, Munich, Germany), and applied to a nickel-chelating affinity matrix (NTA-agarose, Qiagen, Hilden, Germany). The column was washed with binding buffer (50 mM sodium phosphate, pH 8, 500 mM NaCl) and pre-eluted with NTA binding buffer containing 20 mM imidazole. The recombinant protein was eluted from the matrix using NTA-binding buffer containing 300 mM imidazole. Purification was confirmed by SDS-PAGE. Protein concentration was determined by absorption at 280 nm.

#### **7.1.5 Immunoblot**

For immunoblot procedures purified recombinant enzyme was separated by SDS-PAGE and immobilized onto nitrocellulose membranes. Anti-V5 epitope mAb (Invitrogen) was applied according to the recommendations of the manufacturer and bound antibodies visualized via anti-mouse IgG conjugated to alkaline phosphatase (Sigma, Taufkirchen, Germany) and nitrotetrazolium blue chloride/5-bromo-4-chloro-3-indoyl phosphate according to the recommendations of the manufacturer.

#### **7.1.6 Other methods**

SDS-PAGE, Western blotting and molecular biology standard procedures such as PCR, DNA-restriction, ligation, transformation, and plasmid-isolation were performed according to established protocols [121].

### **7.2 Preparation and characterization of complex type oligosaccharides as acceptor substrates**

#### **7.2.1 Preparation of asialo complex type N-glycans**

Asialo N-glycans were isolated from bovine and porcine fibrinogen (Sigma) as described previously [83] with minor modifications. Briefly, 5 g of fibrinogen was denatured in 40 mL 8 M guanidinium hydrochloride in 200 mM Tris-HCl, pH 8.3, and 1.1 g of dithiothreitol (Fluka) was added. After 1h incubation at RT, 5.1 g of iodoacetamide (Sigma) was added followed by 0.5 h incubation at RT. The alkylated glycoprotein was dialyzed in an MWCO 10,000 Da tubing (Sigma) against  $3 \times 10$  L ultrapure water (SGWater, 18 $\mu$ S), centrifuged to remove precipitated protein and freeze dried. The retentate was dissolved in 12 mL 8 M guanidinium hydrochloride and diluted by adding 80 mL 200 mM Tris-HCl, pH 8.3. The pH

of the fibrinogen solution was adjusted to 7.6 by adding 1 M HCl and the protein was digested by adding 50 mg of trypsin (Sigma) and incubating at 37°C for 24 h. Afterwards, the digest was dialyzed in an MWCO 1000 Da tubing (SpectrumLabs) against 3 × 10L ultrapure water, centrifuged at 6000g to remove precipitated protein and freeze dried. The glycopeptide mixture was digested with PNGase F by dissolving into 25 mL of 50 mM ammonium acetate buffer, pH 7.5, containing 0.05% sodium azide, and the pH was adjusted by adding 1 M ammonium hydroxide. 10 U (10 µL) PNGase F (Roche) was added and the reaction was allowed to proceed for 24 h at 37 °C. A second portion of 10 U PNGase F was added and the incubation was continued for further 48 h. Desialylation was performed by adjusting the pH to 5.0 by adding 1 M acetic acid followed by addition of 2 U sialidase from *Clostridium perfringens* (Roche) and incubation for 48 h at 37 °C. The digest was freeze dried and the retentate was dissolved into 6 mL of ultrapure water. Precipitated protein was removed by centrifugation at 6000 g and the supernatant was chromatographed in two equal parts of a mixed bed ion exchange column (3 × 30 cm each; top: AG50W-X2 (Biorad), H<sup>+</sup> form; bottom AG1-X2 (Biorad), acetate form). Asialo N-glycan-containing fractions were identified by ESI-TOF MS analysis, pooled and freeze dried.

### 7.2.2 Preparation of asialo agalacto complex type N-glycans

For degalactosylation of the asialo N-glycan, 10 mg asialo N-glycan were dissolved in 1 mL of 50 mM BisTris–HCl buffer, pH 7.2, containing 100 mM sodium chloride, 5 mM magnesium chloride, 0.05% sodium azide. 1000 U (1 mg) β-galactosidase from *E. coli* overproducer (Sigma), dissolved in 100 µL of the same buffer were added. The reaction was allowed to proceed for 48 h at 37 °C. At completion of the reaction as verified by <sup>1</sup>H NMR, galactosidase was separated from the solution by ultrafiltration (MWCO 50,000) and the filtrate was dialyzed against ultrapure water. The freeze dried oligosaccharide was purified by porous graphitized carbon (PGC) HPLC (Thermo Fisher, Hypercarb 150 × 4.6 mm; particle size 3 µm), flow rate 0.8 mL/min, eluent A: ultrapure water; eluent B: acetonitrile; gradient (share of eluent B): 0 min: 0%; 10 min: 2%; 22 min: 9%; 75 min: 17%; 95 min: 40%; 105–115 min: 100%.

### 7.2.3 Preparation of 1-β-N-acetyl asialo agalacto complex type N-glycans

1-β-N-acetyl oligosaccharides were synthesized as described previously [102]. Briefly, 1-5 mg of reducing oligosaccharide were dissolved into 500 µL of saturated ammonium bicarbonate and incubated at 50 °C for 72 h. Solid ammonium bicarbonate was added several times during incubation to maintain saturation of the solution. After the reaction had finished,

ammonium bicarbonate was removed by repeated freeze-drying cycles. The glycosyl amine was dissolved into 200  $\mu\text{L}$  of saturated sodium bicarbonate solution on ice and 10  $\mu\text{L}$  of acetic anhydride were added. The reaction was allowed to proceed for 10 minutes on ice. Then, another 10  $\mu\text{L}$  of acetic anhydride were added and the solution was incubated at room temperature for 1 h. After dialysis against ultrapure water and freeze drying, the reaction product was subjected to PGC HPLC as described above. The reaction yielded 60-80% 1- $\beta$ -*N*-acetyl oligosaccharide as determined by PGC HPLC (UV detection, 215 nm).

#### 7.2.4 Mass spectrometric characterization of oligosaccharides

Mass spectrometric data were acquired on an ESI-TOF (Agilent 6224). 1  $\mu\text{L}$  of the fractions from the PGC-LC was injected directly into the ion source. MS data acquisition was achieved in the positive ion mode with a capillary voltage of 4000 V. The ion source was set to a gas temperature of 300  $^{\circ}\text{C}$ , a gas flow of 10 L/min and a nebulizer pressure of 30 psi. The acquisition time was 0.95 s with a mass range of 500–3200 m/z. The MS data of all fractions were analyzed using the software Masshunter (Agilent). An automatic compound search including the most frequent ions for glycans (i.e.  $[\text{M}+\text{H}]^+$ ,  $[\text{M}+\text{Na}]^+$ ,  $[\text{M}+\text{K}]^+$ ,  $[\text{M}+2\text{H}]^{2+}$ ,  $[\text{M}+2\text{Na}]^{2+}$ ,  $[\text{M}+\text{H}+\text{Na}]^{2+}$ ,  $[\text{M}+\text{H}+\text{K}]^{2+}$ ) was performed resulting in the EICs and a spreadsheet carrying the integrals of each glycan in each fraction. The data represent the reconstructed-EICs (rEIC) of the PGC-LC separation and the spreadsheet was exported to Microsoft Excel (.xls) format.

#### 7.2.5 NMR characterization of oligosaccharides

Table 7-1 to Table 7-11 contain NMR chemical shift assignments of the oligosaccharide compounds in this study. For reducing oligosaccharides, only structural reporter groups were assigned. For the  $\beta$ -1-*N*-acetylated oligosaccharides **1b**, **2b** and **4b** proton and carbon chemical shifts were assigned completely.

##### 7.2.5.1 GlcNAc $\beta$ -2Man $\alpha$ -1-3(GlcNAc $\beta$ -1-2Man $\alpha$ -1-6)Man $\beta$ -1-4GlcNAc $\beta$ -1-4GlcNAc (**1a**)

Table 7-1:  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts compound **1a** relative to acetone (2.22 ppm) at 700 MHz,  $\text{D}_2\text{O}$ , 300 K. [122].  $\alpha:\beta = 70:30$ .

Assignment	$\delta[\text{ppm}] ^1\text{H}$	$\delta[\text{ppm}] ^{13}\text{C}$	Multiplicity	$J_{\text{H,H}} [\text{Hz}]$
<b>1Gn H-1a</b>	5.186	100.0	d	2.9
<b>4Mn H-1</b>	5.122	100.2	d	1.2
<b>4'Mn H-1</b>	4.924	97.6	d	1.3
<b>3Mn H-1</b>	4.772	101		
<b>1Gn H-1b</b>	4.637	95.4	d	7.7
<b>2Gn H-1(<math>\alpha</math>)</b>	4.614	102	d	8.5

## EXPERIMENTAL PROCEDURES

Assignment	$\delta[\text{ppm}]^1\text{H}$	$\delta[\text{ppm}]^{13}\text{C}$	Multiplicity	$J_{\text{H,H}}[\text{Hz}]$
2Gn H-1( $\beta$ )	4.606	102	d	8.3
5,5'Gn H-1	4.561	100.2	d	9
3Mn H-2	4.247	70.9	dd	1.8, 0.8
4Mn H-2	4.187	77.1	dd	3.4, 1.6
4'Mn H-2	4.106	77	dd	3.7, 1.4
2Gn Ac	2.077	23	s	
5Gn Ac	2.053	23.1	s	
5'Gn Ac	2.051	23.1	s	
2Gn Ac	2.036	22.6	s	

### 7.2.5.2 GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-N-Ac (1b)

Table 7-2: Complete  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts of compound **1b** relative to acetone (2.22 ppm) at 700 MHz,  $\text{D}_2\text{O}$ , 300 K.

Assignment	$\delta[\text{ppm}]^1\text{H}$	$\delta[\text{ppm}]^{13}\text{C}$	Multiplicity	$J_{\text{H,H}}[\text{Hz}]$
4Mn H-1	5.109	100.3	d	1
1Gn H-1	5.039	79.0	d	9.7
4'Mn H-1	4.914	97.8	d	1
3Mn H-1	4.767	101.2	Overlap with HDO resonance	
2Gn H-1	4.611	102.1	d	8.1
5,5'Gn H-1	4.549	100.4	d	8.5
3Mn H-2	4.244	70.9	d	2.8
4Mn H-2	4.183	77.2	dd	3.2, 1.3
4'Mn H-2	4.104	77.1	dd	3.4, 1.3
3Mn H-6a	3.955	66.6	dd	11.2, 5.9
4,4'Mn H-6a	3.910	62.4	m	1.8
5,5'Gn H-6a	3.910	61.4	m	-
4,4'Mn H-3	3.889	70.2	m	
2Gn H-6a	3.874	60.7	dd	12.0, 1.3
1Gn H-2	3.843	54.5	dd	10.0, 10.0
1Gn H-6a	3.825	60.6	dd	12.0, 1.6
3Mn H-4	3.793	66.3	dd	10.0, 10.0
2Gn H-2	3.792	55.6	m	-
3Mn H-3	3.764	81.2		
3Mn H-6b	3.762	66.7		
1Gn H-3	3.756	73.0		
5,5'Gn H-6b	3.754	61.3		
2Gn H-3	3.744	74.2		
2Gn H-6b	3.741	60.7		
4'Mn H-5	3.731	66.4		
2Gn H-4	3.727	80.3	dd	8.9, 8.9
5 or 5'Gn H-2	3.705	56.1	dd	4.5, 8.5
5 or 5'Gn H-2	3.690	56.1	dd	4.5, 8.5
1Gn H-4	3.655	79.4	dd	8.7, 8.7
1Gn H-6b	3.639	60.6	dd	5.6, 12.2

## EXPERIMENTAL PROCEDURES

Assignment	$\delta[\text{ppm}]^1\text{H}$	$\delta[\text{ppm}]^{13}\text{C}$	Multiplicity	$J_{\text{H,H}}[\text{Hz}]$
4,4'Mn H-6b	3.614	62.4	m	-
4Mn, 3Mn H-5	3.610	75.1		
2Gn H-5	3.608	73.6		
1Gn H-5	3.571	77.0	ddd	9.6, 4.7, 1.8
5 or 5'Gn H-3	3.552	74.1	dd	9.4, 10.6
5 or 5'Gn H-3	3.535	74.1	dd	8.7, 10.4
4 or 4'Mn H-4	3.501	68.1	dd	9.7, 9.7
4 or 4'Mn H-4	3.489	68.1	dd	9.4, 9.4
5,5'Gn H-4	3.463	70.6	dd	9.4, 9.4
5,5'Gn H-4	3.454	70.6	m	-
5,5'Gn H-5	3.422	76.5	ddd	9.7, 5.7, 2.3
2Gn Ac	2.072	23.0	s	-
5Gn Ac	2.048	23.0	s	-
5'Gn Ac	2.047	23.1	s	-
1Gn 1 or 2Ac	2.001	23.1	s	-
1Gn 1 or 2Ac	2.000	22.5	s	-

### 7.2.5.3 GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4(Fuc $\alpha$ 1-6)GlcNAc (2a)

Table 7-3: Complete  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts of structural reporter groups of compound **2a** relative to acetone (2.22 ppm) at 700 MHz,  $\text{D}_2\text{O}$ , 300 K. [123]  $\alpha:\beta = 60:40$ .

Assignment	$\delta[\text{ppm}]^1\text{H}$	$\delta[\text{ppm}]^{13}\text{C}$	Multiplicity	$J_{\text{H,H}}[\text{Hz}]$
1Gn H-1( $\alpha$ )	5.180	91.0	d	3.4
4Mn H-1	5.116	100.3	s (b)	
4'Mn H-1	4.916	97.7	s (b)	1.3
Fuc H-1( $\beta$ )	4.895	100.3	d	4.4
Fuc H-1( $\alpha$ )	4.887	100.3	d	4.4
1Gn H-1( $\beta$ )	4.694	95.7	d	8
3Mn H-1	4.769	101.2	s (b)	
2Gn H-1( $\beta$ )	4.669	101.8	d	8.4
2Gn H-1( $\alpha$ )	4.664	101.8	d	8.1
5,5'Gn H-1	4.555	100.3	d	8.4
3Mn H-2	4.248	70.8	d	2.3
4Mn H-2	4.187	77.1	d	2.8
Fuc H-5( $\beta$ )	4.131	67.4	q	6.7
4'Mn H-2	4.105	76.9	d	3.8
Fuc H-5( $\alpha$ )	4.096	67.4	q	6.7
2Gn Ac ( $\alpha$ )	2.095	23.0	s	
2Gn Ac ( $\beta$ )	2.092	23.0	s	
5Gn Ac	2.052	23.0	s	
5'Gn Ac	2.052	23.0	s	
2Gn Ac	2.039	22.2	s	
Fuc H-6( $\beta$ )	1.219	15.9	d	6.7
Fuc H-6( $\alpha$ )	1.209	15.9	d	6.7

### 7.2.5.4 GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4(Fuc $\alpha$ 1-6)GlcNAc- $\beta$ 1NAc (**2b**)

Table 7-4:  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts of compound **2b** relative to acetone (2.22 ppm) at 700 MHz,  $\text{D}_2\text{O}$ , 300 K.

Assignment	$\delta[\text{ppm}] ^1\text{H}$	$\delta[\text{ppm}] ^{13}\text{C}$	Multiplicity	$J_{\text{H,H}} [\text{Hz}]$		
4Mn H-1	5.111	100.2	d	1.6		
1Gn H-1	5.033	78.8	d	9.9		
4'Mn H-1	4.912	97.5	d	1.7		
Fuc H-1	4.868	100.0	d	3.9		
3Mn H-1	4.767	101.1	(overlap with HDO resonance)			
2Gn H-1	4.678	101.6	d	8.3		
5,5'Gn H-1	4.549	100.2	d	8.4		
3Mn H-2	4.246	70.7	d	2.8		
4Mn H-2	4.183	77.1	dd	3.4, 1.6		
Fuc H-5	4.132	67.4	q	6.6		
4'Mn H-2	4.103	76.8	dd	3.5, 1.7		
3Mn H-6a	3.957	66.5	dd	11.3, 5.7		
4or4'Mn H-6a	3.920	62.4	dd	12.3, 2.3		
5,5'Gn H-6a	9.910	61.2	dd	2.1, 12.1		
4or4'Mn H-6a	3.909	62.4	m	-		
4,4'Mn H-3	3.886	70.1				
1Gn H-6a	3.879	67.3				
2Gn H-6a	3.870	60.5				
1Gn H-2	3.851	54.2				
Fuc H-3, H-4	3.793	72.5	m			
2Gn H-2	3.787	55.4				
3Mn H-4	3.783	66.1				
Fuc H-2	3.781	68.8				
3Mn H-6b	3.769	66.5				
3Mn H-3	3.766	81.0				
1Gn H-4	3.758	78.9				
5,5'Gn H-6b	3.754	61.2				
1Gn H-3	3.752	73.0				
2Gn H-3	3.746	73.9				
2Gn H-6b	3.743	60.4				
4'Mn H-5	3.733	66.3				
2Gn H-4	3.722	80.4				
5,5'Gn H-2	3.700	55.9			dd	10.4, 8.3
5,5'Gn H-2	3.694	55.9			dd	10.5, 8.3
1Gn H-6b	3.689	67.4	m			
1Gn H-5	3.681	75.8				
3Mn H-5	3.632	75.0	m			
4,4'Mn H-6b	3.614	62.4				
2Gn H-5	3.608	73.3				
4Mn H-5	3.608	75.1				
5,5'Gn H-3	3.551	73.9	dd	11.5, 8.5		

## EXPERIMENTAL PROCEDURES

Assignment	$\delta[\text{ppm}]^1\text{H}$	$\delta[\text{ppm}]^{13}\text{C}$	Multiplicity	$J_{\text{H,H}}[\text{Hz}]$
5,5'Gn H-3	3.536	73.9	dd	11.5, 8.5
4or4'Mn H-4	3.500	67.9	dd	9.9, 9.9
4or4'Mn H-4	3.490	67.9	dd	9.5, 9.5
5or5'Gn H-4	3.477-3.442	70.6	m	
5or5'Gn H-5	3.429	76.5	ddd	9.9, 5.6, 2.1
2Gn Ac	2.086	23.2	s	
5Gn Ac	2.048	23.1	s	
5'Gn Ac	2.046	23.1	s	
1Gn Ac	2.001	22.5	s	
1Gn Ac	1.997	22.5	s	
Fuc H-6	1.195	15.6	d	6.6

### 7.2.5.5 GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4(Fuc $\alpha$ 1-3)GlcNAc (3a)

Table 7-5:  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts of structural reporter groups of compound **3a** relative to acetone (2.22 ppm) at 700 MHz,  $\text{D}_2\text{O}$ , 300 K.  $\alpha:\beta = 55:45$ .

Assignment	$\delta[\text{ppm}]^1\text{H}$	$\delta[\text{ppm}]^{13}\text{C}$	Multiplicity	$J_{\text{H,H}}[\text{Hz}]$
Fuc H-1( $\alpha$ )	5.125	98.8	d	
4Mn H-1	5.115	100.2	m	1.2
Fuc H-1( $\beta$ )	5.115	98.8	m	
1Gn H-1a	5.079	100	d	4.0
4'Mn H-1	4.909	100.2	s (b)	
3Mn H-1	4.744	101.1	Overlap with HDO resonance	
Fuc H-5	4.718	67.1	Overlap with HDO resonance	
1Gn H-1b	4.69		Overlap with HDO resonance	
2Gn H-1( $\alpha$ )	4.559	101.3	d	8.4
5'Gn H-1	4.554	100.2	d	9.0
5Gn H-1	4.546	100.2	d	9.0
2Gn H-1( $\beta$ )	4.541	101.3	d	8.4
3Mn H-2	4.243	70.6	d	3.5
4Mn H-2	4.186	76.7	dd	3.8, 1.6
1Gn H-2( $\alpha$ )	4.134			
4'Mn H-2	4.105	76.7	dd	3.8, 1.6
5Gn Ac	2.052	22.5	s	
5'Gn Ac	2.051	22.5	s	
2Gn Ac	2.045	22.6	s	
1Gn Ac( $\alpha$ )	2.027	22.6	s	
1Gn Ac( $\beta$ )	2.021	22.6	s	
Fuc H-6	1.27	15.9	d	6.7

### 7.2.5.6 GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4(Fuc $\alpha$ 1-3)GlcNAc- $\beta$ 1NAc (**3b**)

Table 7-6:  $^1\text{H}$  chemical shifts of compound **3b** relative to acetone (2.22 ppm) at 700 MHz,  $\text{D}_2\text{O}$ , 300 K.

Assignment	$\delta[\text{ppm}] \text{ } ^1\text{H}$	Multiplicity	$J_{\text{H,H}}$ [Hz]
Fuc H-1	5.134	m	
4Mn H-1	5.114	s	
1Gn H-1	5.060	d	10.3
4'Mn H-1	4.908	s	
3Mn H-1	4.744	Overlap with HDO resonance	
Fuc H-5	4.701	Overlap with HDO resonance	
2Gn H-1	4.484	d	8.8
5Gn H-1	4.554	d	9.0
5'Gn H-1	4.546	d	9.0
3Mn H-2	4.142	d	2.5
4Mn H-2	4.185	d	3.2
4'Mn H-2	4.104	d	3.8
1Gn H-2	3.893	m	
2Gn Ac	2.051	s	
5Gn Ac	2.050	s	
5Gn Ac	2.045	s	
1Gn Ac	2.007	s	
1Gn Ac	2.002	s	
Fuc H-6	1.275	d	6.7

### 7.2.5.7 GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4(Fuc $\alpha$ 1-3)(Fuc $\alpha$ 1-6)GlcNAc (**4a**)

Table 7-7:  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts of structural reporter groups of compound **4a** relative to acetone (2.22 ppm) at 700 MHz,  $\text{D}_2\text{O}$ , 300 K.  $\alpha:\beta = 65:35$ . See also [124].

Assignment	$\delta[\text{ppm}] \text{ } ^1\text{H}$	$\delta[\text{ppm}] \text{ } ^{13}\text{C}$	Multiplicity	$J_{\text{H,H}}$ [Hz]
3Fuc H-1	5.125	98.6	m	
4Mn H-1	5.118	100.3	d	1.2
1Gn H-1a	5.069		d	3.7
6Fuc H-1	4.925		d	4.4
4'Mn H-1	4.909	97.6	d	1.3
3Mn H-1	4.748	101.0	Overlap with HDO resonance	
1Gn H-1b			Overlap with HDO resonance	
3Fuc H-5	4.719	67.2	d	8.0
2Gn H-1( $\beta$ )	4.689		d	8.7
2Gn H-1( $\alpha$ )	4.667		d	8.7
5Gn H-1	4.557	100.3	d	9.0
5'Gn H-1	4.549	100.3	d	9.0
3Mn H-2	4.247	70.7	d	3.2
4Mn H-2	4.188	76.8	dd	3.5, 1.7
6Fuc H-5( $\beta$ )	4.160		q	7.0

## EXPERIMENTAL PROCEDURES

Assignment	$\delta$ [ppm] $^1\text{H}$	$\delta$ [ppm] $^{13}\text{C}$	Multiplicity	$J_{\text{H,H}}$ [Hz]
1Gn H-2( $\alpha$ )	4.154		dd	10.3, 3.7
6Fuc H-5( $\alpha$ )	4.122		q	7
2Gn H-2( $\beta$ )	4.106	76.5	dd	3.5, 1.6
2Gn Ac( $\beta$ )	2.066	23.0	s	
2Gn Ac( $\alpha$ )	2.062	23.1	s	
5Gn Ac	2.055	23.1	s	
5'Gn Ac	2.049		s	
1Gn Ac( $\beta$ )	2.030	22.6	s	
1Gn Ac( $\alpha$ )	2.025	22.6	s	
3Fuc H-6( $\alpha$ )	1.282	16.3	d	6.6
6Fuc H-6( $\alpha$ )	1.221	16.3	d	6.9
6Fuc H-6( $\beta$ )	1.210	16.3	d	6.9

### 7.2.5.8 GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4(Fuc $\alpha$ 1-3)(Fuc $\alpha$ 1-6)GlcNAc (4b)

 Table 7-8:  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts of compound **4b** relative to acetone (2.22 ppm) at 700 MHz,  $\text{D}_2\text{O}$ , 300 K.

Assignment	$\delta$ [ppm] $^1\text{H}$	$\delta$ [ppm] $^{13}\text{C}$	Multiplicity	$J_{\text{H,H}}$ [Hz]
3Fuc H-1	5.129	98.1	d	3.8
4Man H-1	5.111	99.2	d	< 1
1Gn H-1	5.051	77.9	d	9.5
4'Mn H-1	4.902	96.5	d	< 1
6Fuc H-1	4.901	98.8	d	4.1
3Mn H-1	4.739	100.1	Overlap with HDO resonance	
3Fuc H-5	4.718	66.5	q	6.0
2Gn H-1	4.688	99.8	d	8.5
5Gn H-1	4.549	99.2	d	8.4
5'Gn H-1	4.543	99.2	d	8.3
3Mn H-2	4.245	69.8	dd	2.9; < 1
4Mn H-2	4.183	76.1	dd	3.4; < 1
6Fuc H-5	4.162	66.4	q	6.5
4'Mn H-2	4.101	75.9	dd	3.6; < 1
1Gn H-2	4.023	53.9	dd	9.6; 9.6
2Gn H-4	4.007	73.0	dd	10.0; 10.0
3Fuc H-2	3.982	68.9	dd	10.5; 3.6
6Fuc H-2	3.963	69.4	m	
1Gn H-6a	3.961	65.6		
4Man H-6a	3.919	61.2	m	
5'Gn H-6a	3.919	60.2		
1Gn H-6b	3.917	65.7		
1Gn H-3	3.913	74.7		
4'Man H-6a	3.911	61.0		
5Gn H-6a	3.906	60.4		
3Mn H-6a	3.906	65.9		
4'Mn H-3	3.889	69.2		
2Gn H-6a	3.887	60.7		

## EXPERIMENTAL PROCEDURES

Assignment	$\delta[\text{ppm}]^1\text{H}$	$\delta[\text{ppm}]^{13}\text{C}$	Multiplicity	$J_{\text{H,H}}$ [Hz]		
6Fuc H-5	3.878	70.0				
4Mn H-3	3.876	69.5	m			
3Fuc H-4	3.813	71.6				
3Mn H-4	3.792	65.0				
6Fuc H-4	3.791	71.1				
6Fuc H-3	3.788	67.7				
2Gn H-2	3.772	54.5				
5'Gn H-6b	3.765	60.3				
3Mn H-6b	3.762	65.9				
3Mn H-3	3.758	80.1				
5Gn H-6b	3.755	60.4				
2Gn H-3	3.742	73.2				
3Fuc H-3	3.718	67.4			m	
5'Gn H-3	3.716	71.8				
5Gn H-2	3.704	55.0				
5'Gn H-2	3.697	55.0				
1Gn H-5	3.670	75.6	m			
2Gn H-6b	3.643	60.8				
4'Man H-6b	3.617	61.1				
4Man H-6b	3.613	61.4				
3Mn H-5	3.607	74.1				
4'Mn H-5	3.605	72.5				
5Gn H-3	3.550	74.0	m			
5'Gn H-3	3.542	72.9				
4'Mn H-4	3.503	66.9	m			
4Mn H-4	3.496	67.0				
5'Gn H-4	3.460	69.7	m			
5Gn H-4	3.453	69.6				
1Gn H-4	3.444	80.7				
5'Gn H-5	3.437	75.5				
5'Gn H-5	3.428	75.5				
2Gn Ac	2.063	22.2	s			
5Gn Ac	2.050	22.2	s			
5'Gn Ac	2.043	22.2	s			
1Gn Ac1	1.992	21.6	s			
1Gn Ac2	1.984	21.6	s			
3Fuc H-6	1.278	15.0	d	6.7		
6Fuc H-6	1.191	14.6	d	6.7		

**7.2.5.9 Man $\alpha$ 1-3(GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc (5)**

Table 7-9:  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts of structural reporter groups of compound **5** relative to acetone (2.22 ppm) at 700 MHz,  $\text{D}_2\text{O}$ , 300 K.

Assignment	$\delta[\text{ppm}]^1\text{H}$	$\delta[\text{ppm}]^{13}\text{C}$	Multiplicity	$J_{\text{H,H}}$ [Hz]
1Gn H-1a	5.191	100	d	2.9
4Mn H-1	5.102	100.2	d	1.2

## EXPERIMENTAL PROCEDURES

Assignment	$\delta[\text{ppm}] \text{ } ^1\text{H}$	$\delta[\text{ppm}] \text{ } ^{13}\text{C}$	Multiplicity	$J_{\text{H,H}}$ [Hz]
4'Mn H-1	4.924	97.6	d	1.3
3Mn H-1	4.772	101	Overlap with HDO resonance	
1Gn H-1b	4.700	96.3	d	8
2Gn H-1( $\alpha$ )	4.614	102	d	8.5
2Gn H-1( $\beta$ )	4.606	102	d	8.3
5'Gn H-1	4.554	100.2	d	8.4
3Mn H-2	4.254	70.9	dd	1.8, 0.8
4Mn H-2	4.069	70.8	dd	3.4, 1.6
4'Mn H-2	4.109	77.0	dd	3.6, 1.7
2Gn Ac	2.081	23.0	s	
5'Gn Ac	2.051	23.1	s	
2Gn Ac	2.040	22.6	s	

**7.2.5.10 GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc (6)**

Table 7-10:  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts of structural reporter groups of compound **6** relative to acetone (2.22 ppm) at 700 MHz,  $\text{D}_2\text{O}$ , 300 K. [125]

Assignment	$\delta[\text{ppm}] \text{ } ^1\text{H}$	$\delta[\text{ppm}] \text{ } ^{13}\text{C}$	Multiplicity	$J_{\text{H,H}}$ [Hz]
1Gn H-1a	5.186	100.0	d	2.9
4Mn H-1	5.122	100.2	d	1.2
4'Mn H-1	4.924	97.6	d	1.3
3Mn H-1	4.772	101.0	Overlap with HDO resonance	
1Gn H-1b	4.637	95.4	d	7.7
2Gn H-1( $\alpha$ )	4.614	102.0	d	8.5
2Gn H-1( $\beta$ )	4.606	102.0	d	8.3
5Gn H-1	4.556	100.2	d	9.0
3Mn H-2	4.247	70.9	dd	1.8, 0.8
4Mn H-2	4.187	77.1	dd	3.4, 1.6
4'Mn H-2	3.977	70.3	dd	3.7, 1.4
2Gn Ac	2.079	23.0	s	
5Gn Ac	2.053	23.1	s	
1Gn Ac	2.038	22.6	s	

**7.2.5.11 GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4(Fuc $\alpha$ 1-3)GlcNAc (7)**

Table 7-11:  $^1\text{H}$  chemical shifts of structural reporter groups of compound **7** relative to acetone (2.22 ppm) at 700 MHz,  $\text{D}_2\text{O}$ , 300 K.

Assignment	$\delta[\text{ppm}] \text{ } ^1\text{H}$	Multiplicity	$J_{\text{H,H}}$ [Hz]
Fuc H-1	5.134	m	
4Mn H-1	5.114	m	
1Gn H-1a	5.083	d	4.0
1Gn H-1	5.060	d	
4'Mn H-1	4.914	s	
3Mn H-1	4.744	Overlap with HDO resonance	
Fuc H-5	4.710	Overlap with HDO resonance	
1Gn H-1b	4.688	d	8.7

## EXPERIMENTAL PROCEDURES

Assignment	$\delta$ [ppm] $^1\text{H}$	Multiplicity	$J_{\text{H,H}}$ [Hz]
2Gn H-1a	4.557	d	8.4
2Gn H-1b	4.545	d	8.4
5Gn H-1	4.553	d	9
3Mn H-2	4.249	d	2.5
4Mn H-2	4.185	d	3.2
1Gn H-2	4.138	dd	10.6, 3.9
4'Mn H-2	3.965	d	3.8
5Gn Ac	2.054	s	
2Gn Ac	2.054	s	
1Gn Ac( $\alpha$ )	2.031	s	
1Gn Ac( $\beta$ )	2.025	s	
Fuc H-6	1.275	d	6.7

### 7.3 Synthesis of chitotriose derivatives

#### 7.3.1 Deacetylation of peracetylated chitotriose (8a)

Under nitrogen atmosphere, 520 mg (540  $\mu\text{mol}$ ) peracetylated chitotriose **9** were dissolved in 35 mL abs. methanol. Sodium methanolate solution (prepared by dissolving 100 mg (4.3 mmol) sodium in 5.0 mL abs. methanol) was added until pH 10 had been reached (1.5 mL). Reaction progress was tracked by TLC. After 3.0 h, another 1.0 mL of sodium methanolate solution was added. The reaction was neutralized after 24 h by adding dry ice and the solvent was evaporated. The residue was not desalted, as minor impurities from sodium bicarbonate do not interfere with the following reaction. Therefore, no yield could be determined. Completeness of the reaction was verified by NMR and HPLC analysis.

TLC (silica, MeOH/DCM 1:1, detection: anisaldehyde in aq.  $\text{H}_2\text{SO}_4/\text{AcOH}$ ):  $R_f = 0.0$  (reactant **9**: 0.8).

PGC HPLC (Hypercarb 150  $\times$  4.6 mm; particle size 3  $\mu\text{m}$ ), flow rate 0.8 mL/min, eluent A: ultrapure water, 0.1% FA; eluent B: acetonitrile, 0.1% FA; gradient (share of eluent B): 0–15 min: 5–70%, UV-detection at 215 nm:  $R_t = 9.9$  min ( $\beta$ ) and 10.1 min ( $\alpha$ ).

HR-ESI MS:  $m/z = 650.2357$  [ $\text{M}+\text{Na}$ ] $^+$  (exact mass: 627.2487).

$^1\text{H}$  NMR (400 MHz, 298 K,  $\text{D}_2\text{O}$ ):  $\delta$  [ppm] = 5.223 (d,  $^3J_{\text{H,H}} = 1.8$  Hz, 1 H, H-1( $\alpha$ )); 4.729 (d,  $^3J_{\text{H,H}} = 7.8$  Hz, 1 H, H-1( $\beta$ )); 3.851 (d,  $^3J_{\text{H,H}} = 8.5$ , 1 H, H-1'); 3.851 (d,  $^3J_{\text{H,H}} = 8.5$ , 1 H, H-1''); 3.956 (dd,  $^2J_{\text{H,H}} = 12.5$  Hz,  $^3J_{\text{H,H}} = 1.7$  Hz, 1 H, H-6a); 3.937–3.519 (m, 16 H, H-2, H-2', H-2'', H-3, H-3', H-3'', H-4, H-4''; H-5, H-5', H-5''); 3.500 (dd,  $^3J_{\text{H,H}} = 9.5$  Hz,  $^3J_{\text{H,H}} = 9.5$  Hz, 1 H, H-4''); 2.101 (m, 6 H, 2 Ac), 2.073 (s, 3 H, Ac).

### 7.3.2 Synthesis of 1- $\beta$ -*N*-acetyl chitotriose **8b**

390 mg (630  $\mu$ mol) chitotriose were aminated with 30 mL saturated ammonium bicarbonate and acetylated with 2 x 1.5 mL (66 mmol) acetic anhydride in 30 mL saturated sodium bicarbonate solution as described in 7.2.3. The product was desalted by means of size exclusion chromatography (500 mL Biogel P-2, particle size 45-90  $\mu$ M, MWCO 1800 Da, Biorad, eluent: ultrapure water). The reaction yielded 230 mg (344  $\mu$ mol) product **8b** (55%).

PGC HPLC (Hypercarb 150  $\times$  4.6 mm; particle size 3  $\mu$ m), flow rate 0.8 mL/min, eluent A: ultrapure water, 0.1% FA; eluent B: acetonitrile, 0.1% FA ; gradient (share of eluent B): 0–15 min: 5-70%, UV-detection at 215 nm:  $R_t$  = 8.2 min

HR-ESI MS:  $m/z$  = 669.2819 [M+H]<sup>+</sup>; 691.2638 [M+Na]<sup>+</sup> (exact mass: 668.2752).

Table 7-12: <sup>1</sup>H and <sup>13</sup>C chemical shifts compound **8b** relative to TMS (0.0 ppm) at 500 MHz, D<sub>2</sub>O, 300 K. In addition to the usual experiments, chemical shifts were assigned by means of a series of TOCSY spectra with increasing spin-lock time.

Assignment	$\delta$ [ppm] <sup>1</sup> H	$\delta$ [ppm] <sup>13</sup> C	Multiplicity	$J_{H,H}$ [Hz]
1Gn H-1	5.051	78.0	d	9.7
2Gn H-1	4.601	101.2	d	9.3
3Gn H-1	4.593	101.2	d	9.5
1Gn H-6a	3.929	60.2	dd	12.3, 2.0
1Gn H-2	3.861	53.6	dd	9.8, 9.8
2 o. 3Gn H-6a	3.844	59.8	dd	12.0, 2.0
2 o. 3Gn H-6a	3.835	59.8	dd	12.2, 1.9
2Gn H-2	3.787	54.9	dd	10.5, 10.5
1Gn H-6b	3.753	60.2	m	
1Gn H-3	3.752	72.4		
3Gn H-2	3.750	55.3		
2Gn H-4	3.722	71.9		
2,3Gn H-6b	3.661	59.8	m	m
1Gn H-4	3.660	78.7	ddd	10.0, 4.2, 2.0
2Gn H-5	3.649	78.9		
2Gn H-3	3.576	76.0	m	
3Gn H-3	3.570	73.2		
1Gn H-5	3.562	74.4	ddd	9.9, 4.9, 2.0
3Gn H-5	3.498	75.7	ddd	9.9, 5.1, 1.9
3Gn H-4	3.472	69.6	dd	9.8, 8.4

### 7.3.3 1- $\beta$ -*N*-acetyl *N*-actetylglycosamine **19b**

1.0 g (4.5 mmol) GlcNAc (Sigma) were aminated with 50 mL saturated ammonium bicarbonate and acetylated with 4 mL (42 mmol) acetic anhydride in 40 mL saturated sodium bicarbonate solution as described in 7.2.3.

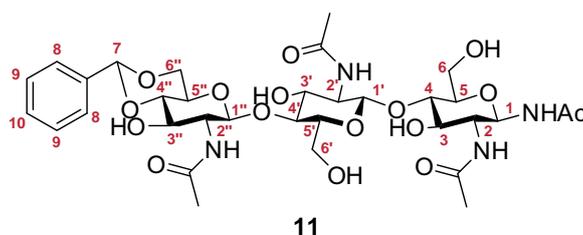
HR-ESI MS:  $m/z = 285.1060$   $[M+Na]^+$  (exact mass: 262.1165).

$^1H$  NMR (400 MHz, 298 K,  $D_2O$ ):  $\delta$  [ppm] = 5.024 (d,  $^3J_{H,H} = 9.8$  Hz, 1 H, H-1); 3.851 (dd,  $^2J_{H,H} = 12.3$  Hz,  $^3J_{H,H} = 1.8$  Hz, 1 H, H-6a); 3.789 (dd,  $^3J_{H,H} = 9.8$  Hz,  $^3J_{H,H} = 9.8$  Hz, 1 H, H-2); 3.720 (dd,  $^2J_{H,H} = 12.3$  Hz,  $^3J_{H,H} = 4.7$  Hz, 1 H, H-6b); 3.576 (dd,  $^3J_{H,H} = 9.8$  Hz,  $^3J_{H,H} = 8.5$  Hz, 1 H, H-3); 3.497 (ddd,  $^3J_{H,H} = 9.8$  Hz,  $^3J_{H,H} = 4.7$  Hz,  $^3J_{H,H} = 1.8$  Hz, 1 H, H-5); 3.449 (dd,  $^3J_{H,H} = 9.8$ ;  $^3J_{H,H} = 8.5$ , 1 H, H-4); 1.981 (s, 3 H, Ac-2), 1.978 (s, 3 H, Ac-1).

$^{13}C$  NMR (100 MHz, 298 K,  $D_2O$ ):  $\delta$  [ppm] = 175.0 (Ac-1 CO); 174.8 (Ac-2 CO); 78.4 (C-1); 77.6 (C-5), 74.2 (C-3), 69.4 (C-4), 60.5 (C-6), 54.2 (C-2), 22.0 (Ac-1  $CH_3$ ), 21.9 (Ac-2  $CH_3$ ).

### 7.3.4 Synthesis of 1- $\beta$ -*N*-acetyl 4'',6''-*O*-benzylidene chitotriose **11**

Under nitrogen atmosphere, 320 mg (480  $\mu$ mol) 1- $\beta$ -*N*-acetyl *N*-actetylglucosamine **19b** were dissolved into 280 mL dry DMF. 540 mg (2.7 mmol) *p*-toluenesulfonic acid and 36 mL (240 mmol)  $\alpha,\alpha$ -dimethoxytoluene were added. The reaction solution was stirred for 10 h at 40 °C. DMF was evaporated and the residue was successively coevaporated with methanol, toluene and DCM in order to remove the excess of  $\alpha,\alpha$ -dimethoxytoluene. The residue was dissolved in methanol and a mixed-bed ion exchange resin with color indicator (Amberlite MB-6113, Merck) was added in order to remove *p*-toluenesulfonic acid. The resin was removed by filtration and washed with methanol and water. The solvent was evaporated. The reaction yielded 140 mg (190  $\mu$ mol, 40%) of a colorless solid identified as **11**.



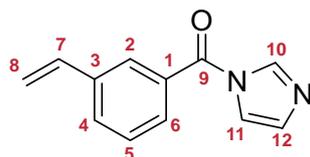
HR-ESI MS:  $m/z = 757.3137$   $[M+H]^+$ ; 779.2985  $[M+Na]^+$  (exact mass: 756.3065).

$^1H$  NMR (500 MHz, 300 K,  $D_2O$ ):  $\delta$  [ppm] = 7.581-7.563 (m, 2 H, H-8); 7.510-7.498 (m, 3 H, H-9, H-10); 5.796 (s, 1H, H-7); 5.067 (d,  $^3J_{H,H} = 9.7$  Hz, 1 H, H-1); 4.618 (d,  $^3J_{H,H} = 8.3$  Hz, 1 H, H-1'); 4.727 (overlap with HDO, H-1''); 4.618 (d,  $^3J_{H,H} = 8.2$  Hz; 1 H, H-1'); 4.408 (dd,  $^2J_{H,H} = 10.5$  Hz,  $^3J_{H,H} = 5.1$ , 1 H, H-6''a); 3.970-3.901 (m, 2 H, H-6''b, H-3''), 3.891-3.799 (m, 6 H, H-2, H-2'', H-3'', H-2', H-6'a, H-6a); 3.798-3.759 (m, 1 H, H-4''), 3.747-3.632 (m, 6 H, H-6'b H-3, H-4, H-5'', H-4', H-6b); 3.616-3.451 (m, 2 H, H-5, H-5'), 2.101 (s, 3 H, Ac); 2.086 (s, 3 H, Ac); 2.028 (s, 3 H, Ac); 2.026 (s, 3 H, Ac).

$^{13}\text{C}$  NMR (126 MHz, 300 K,  $\text{D}_2\text{O}$ ):  $\delta$  [ppm] = 125.7 (C-8); 128.8 (C-9); 128.2 (C-10); 101.4 (C-7); 101.2 (C-1''); 100.9 (C-1'); 79.9 (C-4''); 78.6 (C-4/4'); 77.9 (C-4/4'); 75.9 (C-5); 74.0 (C-5'); 71.7 (C-3); 69.9 (C-3'); 67.6 (C-6''); 67.0 (C-3''); 65.5 (C-5''); 59.7 (C-6, C-6'); 55.8 (C-2); 53.4 (C-2, C-2'), 21.8 (Ac  $\text{CH}_3$ ), 21.8 (Ac  $\text{CH}_3$ ), 21.6 (Ac  $\text{CH}_3$ ), 21.6 (Ac  $\text{CH}_3$ ).

### 7.3.5 Synthesis of 3-vinyl benzoic acid *N*-imidazolyl amide **18**

Under nitrogen atmosphere, 2.0 g (14 mmol) 3-vinylbenzoic acid (Sigma) were dissolved into 20 mL dry THF. 2.2 g (14 mmol) *N,N*-carbonyldiimidazole were added in portions. The solution was stirred at r.t. until 1 h after the evolution of gas had stopped (2 h total reaction time). The solvent was evaporated and the residue was extracted with 5 x 5 mL of hot cyclohexane. The solution was dried with sodium sulfate. The reaction yielded 2.5 g (13 mmol, 91%) of colorless syrup identified as compound **18**.



**18**

HR-ESI MS:  $m/z$  = 199.0851  $[\text{M}+\text{H}]^+$  (exact mass: 198.0793).

$^1\text{H}$  NMR (400 MHz, 298 K,  $\text{CDCl}_3$ ):  $\delta$  [ppm] = 8.085-8.080 (m, 1 H, H-10); 7.807-7.797 (m, 1 H, H-2); 7.718-7.693 (m, 1 H, H-4); 7.666-7.640 (m, 1 H, H-6); 7.550 (t,  $^3J_{\text{H,H}} = 1.5$  Hz, 1 H, H-11); 7.515 (t,  $^3J_{\text{H,H}} = 7.7$  Hz, 1 H, H-5); 7.714 (dd,  $^3J_{\text{H,H}} = 1.6$  Hz,  $^3J_{\text{H,H}} = 0.7$  Hz, 1 H, H-12); 6.760 (dd,  $^3J_{\text{H,H}} = 17.5$  Hz,  $^3J_{\text{H,H}} = 10.9$  Hz, 1 H, H-7); 5.849 (d,  $^3J_{\text{H,H}} = 17.5$  Hz, 1 H, H-8<sub>trans</sub>); 5.400 (d,  $^3J_{\text{H,H}} = 10.9$  Hz, 1 H, H-8<sub>cis</sub>)

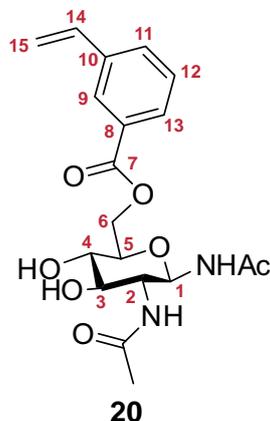
$^{13}\text{C}$  NMR (100 MHz, 298 K,  $\text{CDCl}_3$ ):  $\delta$  [ppm] = 166.2 (C-9); 138.7 (C-1); 135.3 (C-7); 132.4 (C-3); 131.1 (C-4); 131.0 (C-12); 129.2 (C-2); 128.8 (C-6); 127.5 (C-2); 118.1 (C-11); 116.5 (C-8).

### 7.3.6 Synthesis of 1- $\beta$ -*N*-acetyl (3-vinyl-)benzoyl *N*-actetylglycosamines

Under nitrogen atmosphere, 100 mg (380  $\mu\text{mol}$ ) 1- $\beta$ -*N*-acetyl *N*-actetylglycosamine **19b** were dissolved into 45 mL dry DMF. 1.0 mL of a sodium methanolate solution in dry DMF (4 mM) were added to give 1 mol% methanolate in the reaction solution. The reaction was followed with TLC over 7 d. After no further conversion of **19b** could be detected despite adding methanolate or heating to 70  $^\circ\text{C}$ , DMF was evaporated. The residue was coevaporated with methanol several times and subjected to separation by liquid chromatography (silica

20 mL, eluent methanol/chloroform 1:4). The reaction yielded the products **20**, **21** and **22** in the ratio 8:1:1.

### 7.3.6.1 1- $\beta$ -*N*-acetyl 6-(3-vinyl-)benzoyl *N*-actetylglycosamine

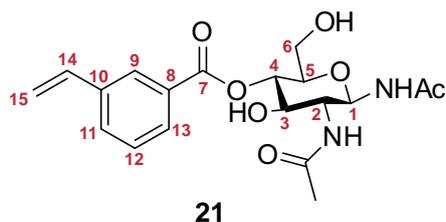


TLC (silica, MeOH/CHCl<sub>3</sub> 1:4, detection: anisaldehyde in aq. H<sub>2</sub>SO<sub>4</sub>/AcOH): *R<sub>f</sub>* = 0.8

<sup>1</sup>H NMR (500 MHz, 300 K, MeOD-*d*<sub>4</sub>):  $\delta$  [ppm] = 8.022-8.014 (m, 1 H, H-9); 7.877-7.839 (m, 1 H, H-11); 7.637-7.615 (m, 1 H, H-13); 7.387 (dd, <sup>3</sup>*J*<sub>H,H</sub> = 7.8 Hz, <sup>3</sup>*J*<sub>H,H</sub> = 7.8 Hz, 1 H, H-12); 6.741 (dd, <sup>3</sup>*J*<sub>H,H</sub> = 11.1 Hz, <sup>3</sup>*J*<sub>H,H</sub> = 17.7 Hz, 1 H, H-14); 5.819 (d, <sup>3</sup>*J*<sub>H,H</sub> = 17.7 Hz, 1 H, H-15<sub>trans</sub>); 5.279 (d, <sup>3</sup>*J*<sub>H,H</sub> = 11.1 Hz, 1 H, H-15<sub>cis</sub>); 5.024 (d, <sup>3</sup>*J*<sub>H,H</sub> = 9.8 Hz, 1 H, H-1); 4.618 (dd, <sup>2</sup>*J*<sub>H,H</sub> = 11.9 Hz, <sup>3</sup>*J*<sub>H,H</sub> = 2.0 Hz, 1 H, H-6a); 4.414 (dd, <sup>2</sup>*J*<sub>H,H</sub> = 11.9 Hz, <sup>3</sup>*J*<sub>H,H</sub> = 5.0, 1 H, H-6b); 3.790 (dd, <sup>3</sup>*J*<sub>H,H</sub> = 9.7 Hz, <sup>3</sup>*J*<sub>H,H</sub> = 9.7 Hz, 1H, H-2); 3.635 (ddd, <sup>3</sup>*J*<sub>H,H</sub> = 9.4 Hz, <sup>3</sup>*J*<sub>H,H</sub> = 5.0 Hz, <sup>3</sup>*J*<sub>H,H</sub> = 2.0 Hz, 1 H, H-5); 3.541-3.466 (m, 2 H, H-3, H-4); 1.933 (s, 3 H, Ac); 1.884 (s, 3 H, Ac).

<sup>13</sup>C NMR (126 MHz, 300 K, MeOD-*d*<sub>4</sub>):  $\delta$  [ppm] = 174.3 (CO-Ac); 173.8 (CO-Ac); 167.8 (C-7); 139.5 (C-8); 127.3 (C-9); 129.1 (C-11); 131.7 (C-13); 129.0 (C-12); 137.1 (C-14); 115.7 (C-15); 80.3 (C-1); 65.1 (C-6); 56.1 (C-2); 77.1 (C-5); 75.5 (C-3); 71.9 (C-4); 22.8 (CH<sub>3</sub>), 22.8 (CH<sub>3</sub>).

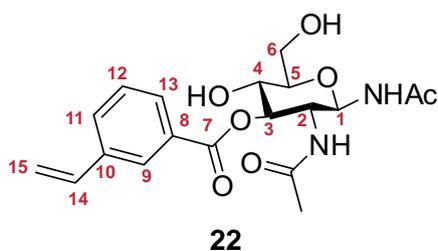
### 7.3.6.2 1- $\beta$ -*N*-acetyl 4-(3-vinyl-)benzoyl *N*-actetylglycosamine



TLC (silica, MeOH/CHCl<sub>3</sub> 1:4, detection: anisaldehyde in aq. H<sub>2</sub>SO<sub>4</sub>/AcOH): *R<sub>f</sub>* = 0.7

$^1\text{H}$  NMR (500 MHz, 300 K, MeOD- $d_4$ ):  $\delta$  [ppm] = 8.099-8.092 (m, 1 H, H-9); 7.953-7.933 (m, 1 H, H-11); 7.715-7.694 (m, 1 H, H-13); 7.461 (dd,  $^3J_{\text{H,H}} = 7.8$  Hz,  $^3J_{\text{H,H}} = 7.8$  Hz, 1 H, H-12); 6.807 (dd,  $^3J_{\text{H,H}} = 17.6$  Hz,  $^3J_{\text{H,H}} = 10.9$  Hz, 1 H, H-14); 5.873 (d,  $^3J_{\text{H,H}} = 17.6$  Hz, 1 H, H-15 $_{\text{trans}}$ ); 5.334 (d,  $^3J_{\text{H,H}} = 10.9$  Hz, 1 H, H-15 $_{\text{cis}}$ ); 5.110 (d,  $^3J_{\text{H,H}} = 9.5$  Hz, 1 H, H-1); 5.092 (dd,  $^3J_{\text{H,H}} = 9.5$  Hz,  $^3J_{\text{H,H}} = 9.3$  Hz, 1 H, H-4); 3.903 (dd,  $^3J_{\text{H,H}} = 9.5$  Hz,  $^3J_{\text{H,H}} = 10.0$  Hz, H-2); 3.875 (dd,  $^2J_{\text{H,H}} = 12.4$  Hz,  $^3J_{\text{H,H}} = 2.0$  Hz, 1 H, H-6a); 3.857 (dd,  $^3J_{\text{H,H}} = 10.0$  Hz,  $^3J_{\text{H,H}} = 9.5$  Hz, 1 H, H-3); 3.743 (dd,  $^2J_{\text{H,H}} = 12.4$  Hz,  $^3J_{\text{H,H}} = 5.4$  Hz, 1 H, H-6b); 3.711 (ddd,  $^3J_{\text{H,H}} = 9.3$  Hz,  $^3J_{\text{H,H}} = 5.4$  Hz,  $^3J_{\text{H,H}} = 2.2$  Hz, 1 H, H-5); 2.035 (s, 3 H, CH<sub>3</sub> Ac), 1.970 (s, 3 H, CH<sub>3</sub> Ac).

$^{13}\text{C}$  NMR (126 MHz, 300 K, MeOD- $d_4$ ):  $\delta$  [ppm] = 173.7 (CO Ac); 167.2 (C-7); 139.5 (C-8); 137.2 (C-14); 131.9 (C-13); 131.7 (C-10); 130.0 (C-11); 129.9 (C-12); 128.4 (C-9); 115.6 (C-15); 80.3 (C-1); 77.7 (C-5); 74.1 (C-3); 73.4 (C-4); 62.4 (C-6); 56.5 (C-2); 22.8 (Ac1); 22.6 (Ac2).



### 7.3.6.3 1-β-N-acetyl 3-(3-vinyl)-benzoyl N-acetylglycosamine

TLC (silica, MeOH/CHCl<sub>3</sub> 1:4, detection: anisaldehyde in aq. H<sub>2</sub>SO<sub>4</sub>/AcOH):  $R_f = 0.7$

$^1\text{H}$  NMR (500 MHz, 300 K, MeOD- $d_4$ ):  $\delta$  [ppm] = 8.081-8.076 (m, 1 H, H-9); 7.925-7.904 (m, 1 H, H-11); 7.690-7.669 (m, 1 H, H-13); 7.441 (dd,  $^3J_{\text{H,H}} = 7.8$  Hz,  $^3J_{\text{H,H}} = 7.8$  Hz, 1 H, H-12); 6.799 (dd,  $^3J_{\text{H,H}} = 17.5$  Hz,  $^3J_{\text{H,H}} = 10.9$  Hz, 1 H, H-14); 5.876 (d,  $^3J_{\text{H,H}} = 17.5$  Hz, 1 H, H-15 $_{\text{trans}}$ ); 5.326 (d,  $^3J_{\text{H,H}} = 10.9$  Hz, 1 H, H-15 $_{\text{cis}}$ ); 5.261 (dd,  $^3J_{\text{H,H}} = 10.2$  Hz,  $^3J_{\text{H,H}} = 9.5$  Hz, 1 H, H-3); 5.206 (d,  $^3J_{\text{H,H}} = 9.8$  Hz, 1 H, H-1); 4.111 (dd,  $^3J_{\text{H,H}} = 10.2$  Hz,  $^3J_{\text{H,H}} = 9.8$  Hz, H-2); 3.715 (dd,  $^3J_{\text{H,H}} = 9.7$  Hz,  $^3J_{\text{H,H}} = 9.4$  Hz, 1 H, H-4); 3.654 (dd,  $^2J_{\text{H,H}} = 12.2$  Hz,  $^3J_{\text{H,H}} = 2.0$  Hz, 1 H, H-6a); 3.566 (dd,  $^2J_{\text{H,H}} = 12.2$  Hz,  $^3J_{\text{H,H}} = 5.4$  Hz, 1 H, H-6b); 3.495 (ddd,  $^3J_{\text{H,H}} = 9.7$  Hz,  $^3J_{\text{H,H}} = 5.4$  Hz,  $^3J_{\text{H,H}} = 2.0$  Hz, 1 H, H-5); 1.981 (s, 3 H, CH<sub>3</sub> Ac), 1.957 (s, 3 H, CH<sub>3</sub> Ac).

$^{13}\text{C}$  NMR (126 MHz, 300 K, MeOD- $d_4$ ):  $\delta$  [ppm] = 173.9 (CO Ac); 167.8 (C-7); 139.4 (C-8); 137.2 (C-14); 131.8 (C-13); 131.5 (C-10); 130.0 (C-11); 129.8 (C-12); 128.4 (C-9); 115.6 (C-15); 79.7 (C-1); 79.7 (C-5); 78.4 (C-3); 69.6 (C-4); 62.4 (C-6); 54.3 (C-2); 22.8 (Ac1); 22.8 (Ac2).

## 7.4 SPR Binding Studies

All SPR experiments were performed at a Biacore T100 (GE Healthcare). Both FUT8 and FucTA were immobilized using the standard amide coupling method.

### 7.4.1 SPR affinity assays for ligands of FucTA

The SPR chip surface (coated with carboxymethylated dextrane, CM5 (GE Healthcare)) was activated by injecting a solution containing 50 mM of *N*-hydroxysuccinimide and 50 mM of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide at a flow rate of 10  $\mu$ L/min for 15 minutes on both flow cells. Subsequently, FucTA (30  $\mu$ g/mL in 25 mM BisTris/10 mM sodium acetate buffer, pH 6.0) was injected into one flow cell at a flow rate of 10  $\mu$ L/min for 30 minutes yielding 4000-6000 RU (70-100 fmol) immobilized protein. Finally, unreacted carboxy functions in both flow cells were capped by injecting 1 M ethanolamine for 7 minutes at a flow rate of 30  $\mu$ L/min. Immobilized FucTA was found to be stable for a couple of weeks.

Sterile filtered 50 mM BisTris, pH 6.7, containing 100 mM sodium chloride and 5 mM magnesium chloride was used as buffer system throughout the experiments. All ligand samples were injected for 1 minute (contact time) followed by 3 minutes injection of buffer (dissociation time) at a flow rate of 30  $\mu$ L/min into both flow cells. In order to determine thermodynamic binding constants, each ligand concentration was assayed twice. Donor substrate and related ligands (GDP-Fuc (Jennewein Biotechnology), GDP (Sigma), guanosine (Fluka)) were dissolved in buffer and diluted to the indicated concentrations (typically 0.25–500  $\mu$ M). For acceptor substrates **1a** and **2a**, concentrations between 0.5  $\mu$ M and 10 mM were employed.  $\beta$ - $\gamma$ -*N*-asparaginyl chitobiose (prepared as described [126]), was analyzed in concentrations of 2 mM and 10 mM and  $\alpha$ -methylfucoside in concentrations of 10 mM and 100 mM. The sensorgrams with the highest ligand concentrations possible were recorded and no SPR response could be detected. Therefore, no further experiments were conducted.

### 7.4.2 SPR affinity assays for ligands of FUT8

The SPR chip surface (coated with carboxymethylated dextrane, CM5 (GE Healthcare)) was activated by injecting a solution containing 50 mM of *N*-hydroxysuccinimide and 50 mM of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide at a flow rate of 10  $\mu$ L/min for 15 min on both flow cells. Subsequently, FUT8 (30  $\mu$ g/mL in 25 mM PBS/25 mM sodium acetate buffer, pH 5.5) was injected into one flow cell at a flow rate of 10  $\mu$ L/min for 60 s yielding 7300 RU (120 fmol) immobilized protein. Finally, unreacted carboxy functions in both flow

cells were capped by injecting 1 M ethanolamine for 5 min a flow rate of 30  $\mu\text{L}/\text{min}$ . Immobilized FUT8 was found to be stable for one day at 25  $^{\circ}\text{C}$ .

Sterile filtered 50 mM Mes–NaOH, pH 7.0, was used as buffer system throughout the experiments. All ligand samples were injected for 40 s (contact time) followed by 1 min injection of buffer (dissociation time) at a flow rate of 30  $\mu\text{L}/\text{min}$  into both flow cells. In order to determine thermodynamic binding constants, each ligand concentration was assayed twice. Donor substrate and related ligands (GDP-Fuc (Jennewein Biotechnology), GDP (Sigma), GMP (Sigma) and guanosine (Fluka)) were dissolved in buffer and diluted to the indicated concentrations (typically 0.25–500  $\mu\text{M}$ ).

## 7.5 STD NMR Experiments

STD NMR spectra were obtained at 300 K using a Bruker Avance 700 MHz NMR spectrometer equipped with a TXI triple resonance cryoprobe head incorporating gradients in the  $z$ -axis unless indicated otherwise.

### 7.5.1 STD NMR experiments of FucTA

Samples (150  $\mu\text{L}$  in 3 mm sample tubes) contained 10  $\mu\text{M}$  protein and 0.1–1.5 mM ligand. Deuterium oxide buffered with 50 mM BisTris- $d_{19}$  (Euroisotope Laboratories), pH 6.7 and containing 100 mM sodium chloride and 5 mM magnesium chloride was used as solvent. Data was acquired using a Bruker standard sequence (stddiffesgp.3) incorporating an excitation sculpting sequence with a  $180^{\circ}$  soft pulse of 8 ms duration for selective suppression of the HDO resonance. On resonance irradiation was applied at  $-0.9$  ppm or at 0 ppm and off resonance irradiation at 40 ppm. Saturation was achieved by a cascade of 40 Gaussian pulses with duration of 50 ms (field strength 120 Hz or 100 Hz) to give a total saturation time of 2 s. STD spectra and reference spectra were acquired with 32,768 data points and a total of 4048 transients. FIDs were multiplied with an exponential function (line broadening 3) before Fourier transformation. All ligands were assigned using  $^1\text{H}$ – $^1\text{H}$  total correlation and  $^1\text{H}$ – $^{13}\text{C}$  heteronuclear single quantum correlation spectra.

To determine the size of STD effects, resonances in the STD spectrum were integrated with respect to the reference spectrum. Therefore, an STD effect of 100% results if the resonances in both spectra show the same intensity. For the competition experiment, GDP was added in concentrations of 130  $\mu\text{M}$ , 450  $\mu\text{M}$ , 740  $\mu\text{M}$ , 1.0 mM, 1.4 mM, 1.7 mM and 2.1 mM to a sample containing 10  $\mu\text{M}$  FucTA and 100  $\mu\text{M}$  GDP-Fuc. STD effects of GDP-Fuc and GDP

resonances were multiplied with the excess of the respective ligand over FucTA to obtain the STD amplification factor.

## 7.5.2 STD NMR experiments of FUT8

After purification of the enzyme, the samples in 50 mM PBS buffer pH 7.0 were exchanged to deuterium oxide buffered with 50 mM Mes- $d_{13}$  (Cambridge Isotope Laboratories), pH 7.0 and concentrated via centrifugal filtering using Amicon Ultra filter devices (Millipore, MWCO 10,000 Da). Protein concentration was determined by absorption at 280 nm using a nanodrop UV-spectrometer. All NMR samples were prepared in 160  $\mu$ L of buffer solution and measured using 3-mm NMR tubes.

### 7.5.2.1 STD NMR experiments of FUT8 with GDP-Fuc

Spectra were obtained at 285 K. Samples contained 10  $\mu$ M protein and 1.0 mM ligand. Data was acquired using a Bruker standard sequence (stddiffesgp) incorporating an excitation sculpting sequence with a 180 degree soft pulse of 2-8 ms duration for selective suppression of the HDO resonance. On resonance irradiation was applied at -0.9 ppm and off resonance irradiation at 40 ppm. Saturation was achieved by a cascade of 40 Gaussian pulses with duration of 50 ms (field strength 120 Hz) to give a total saturation time of 2 s. STD spectra and reference spectra were acquired with 32,768 data points and a total of 4048 transients. FIDs were multiplied with an exponential function (line broadening 3) before Fourier transformation.

To determine the size of STD effects, resonances in the STD spectrum were integrated with respect to the reference spectrum. Therefore, a STD effect of 100% results if the resonances in both spectra show the same intensity. The resulting relative STD percentages were subsequently corrected by their  $T_1$  relaxation constants as determined by means of inversion recovery experiments. [109]

Table 7-13: Absolute and relative STD values for protons of GDP-Fuc in presence of FUT8. The values were corrected by the longitudinal relaxation times in order to obtain the final epitope. \*The value for the fucose anomeric proton was obtained from a spectrum recorded at 300K due to overlap of the signal with the HDO signal at 285 K.

Proton	H-8	H-1'	H-1''*	H-2'	H-3'	H-4'	H-5'	H-5''	H-4''	H-3''	H-2''	H-6''
Abs. STD%	2.48	2.84	0.91	2.38	2.12	1.83	0.92	0.49	0.81	0.78	0.83	1.01
Rel. STD%	87%	100%	32%	84%	75%	64%	32%	17%	29%	27%	29%	36%

## EXPERIMENTAL PROCEDURES

Proton	H-8	H-1'	H-1''*	H-2'	H-3'	H-4'	H-5'	H-5''	H-4''	H-3''	H-2''	H-6''
$T_1$ [ms]	560	2020		1630	1000	1000	360	540	840	870	2180	430
Abs. STD% <sub>corr</sub>	4.43	1.41		1.46	2.12	1.83	2.56	0.91	0.96	0.90	0.38	2.35
Rel. STD% <sub>corr</sub>	100%	32%		33%	48%	41%	58%	20%	22%	20%	9%	53%

**7.5.2.2 STD NMR experiments of FUT8 with the heptasaccharide 1b**

The spectrometer was equipped with a TXI triple resonance probe with  $z$ -gradients. Samples contained 10  $\mu$ M protein and 1.0 mM heptasaccharide **1b**. Data was acquired using a Bruker standard sequence (stdiffesgp) incorporating an excitation sculpting sequence with a 180 degree soft pulse of 8 ms duration for selective suppression of the HDO resonance. On resonance irradiation was applied at 0.2 ppm and off resonance irradiation at 40 ppm. Saturation was achieved by a cascade of 40 Gaussian pulses with duration of 50 ms (field strength 120 Hz) to give a total saturation time of 2 s. STD spectra and reference spectra were acquired with 32,768 data points and a total of 4048 transients. FIDs were multiplied with an exponential function (line broadening 3) before Fourier transformation. STD spectra from a sample without enzyme were acquired to determine artifact STD effects. This artifact spectrum was subtracted from the actual STD spectrum to give artifact adjusted STD effects. To determine the size of STD effects, resonances in the artifact adjusted STD spectrum were integrated with respect to the reference spectrum. Therefore, an absolute STD effect of 100% results if the resonances in both spectra show the same intensity. Artifacts are between 0.2 and 0.6%.

 Table 7-14: Absolute and relative STD effects for protons of the heptasaccharide **1b** with FUT8.

	4Mn H-1	1Gn H-1	2Gn H-1	5,5'Gn H-1	3Mn H-2	4Mn H-2	4'Mn H-2	3Mn H6a	4,4'Mn H6a	5,5'Gn H-6a	2Gn H6a	1Gn H-2	1Gn H6a	2Gn H-4	5'Gn H-2
STD % abs.	0.46	0.89	0.53	0.32	0.38	0.54	0.29	0.46	0.43	0.37	0.45	0.73	0.30	0.26	0.08
STD % rel.	52	100	60	36	43	61	32	52	49	41	51	82	33	30	9

## EXPERIMENTAL PROCEDURES

	5Gn H-2	1Gn H-4	1Gn H6b	3,4Mn H-5	1Gn H-5	5,5'Gn H-3	4,4'Mn H-4	5'Gn H-4	5Gn H-4	Gn5,5' H-5	2Gn Ac	3Gn Ac	6Gn Ac	1Gn Ac2	1Gn Ac1
STD % abs.	0.21	0.31	0.38	0.22	0.10	0.43	0.37	0.48	0.42	0.27	0.49	0.51	0.27	0.60	0.28
STD % rel.	23	35	43	25	12	49	41	53	47	31	55	57	30	68	31

### 7.5.3 STD NMR experiments with anti-CCD IgG

The spectrometer was equipped with a TXI triple resonance probe with  $z$ -gradients. Samples contained 1.3 mg/mL IgG (9.0  $\mu$ M) and 1.0 mM bifucosylated glycan **4b** PBS-D<sub>2</sub>O containing 100 mM NaCl at pH 7.0. Data was acquired using a Bruker standard sequence (stddiffesgp) incorporating an excitation sculpting sequence with a 180 degree soft pulse of 8 ms duration for selective suppression of the HDO resonance. On resonance irradiation was applied at  $-0.5$  ppm and off resonance irradiation at 40 ppm. Saturation was achieved by a cascade of 80 Gaussian pulses with duration of 50 ms (field strength 120 Hz) to give a total saturation time of 4 s. STD spectra and reference spectra were acquired with 32,768 data points and a total of 4048 transients. FIDs were multiplied with an exponential function (line broadening 1) before Fourier transformation. STD spectra from a sample with the same concentration of IgG with different specificity were acquired as a negative control.

Table 7-15: Absolute and relative STD effects for protons of the bifucosylated oligosaccharide **4b** with polyclonal anti-bromelain IgG.

	3Fuc H-1	4Man H-1	1Gn H-1	2Gn H-1	5Gn H-1	5'Gn H-1	3Mn H-2	4Mn H-2	6Fuc H-5	4'Mn H-2	1Gn H-2	1Gn H-4	3Fuc H-6	6Fuc H-6
STD% abs.	0.62	0.59	0.95	0.85	0.31	0.62	0.49	0.41	0.59	0.41	1.3	0.61	1.13	0.72
STD% rel.	48 %	45 %	73 %	65 %	24 %	48 %	38 %	32 %	45 %	32 %	100 %	47 %	87 %	55 %

Table 7-16: Absolute and relative STD effects for protons of the *N*-acetates of the bifucosylated oligosaccharide **4b** with polyclonal anti-bromelain IgG.

	2Gn Ac	5Gn Ac	5'Gn Ac	1Gn Ac	1Gn Ac
STD% abs.	1.10	0.96	0.96	1.38	1.38
STD% rel.	79.7%	69.6%	69.6%	100.0%	100.0%

## 7.6 TrNOESY experiments

Spectra were recorded at 310 K. Samples (160  $\mu$ L) contained 10-20  $\mu$ M FUT8 in PBS-D<sub>2</sub>O at pH 7.0. GDP-Fuc excesses over FUT8 were 1:2, 1:5, 1:6, 1:10 and 1:20, respectively. Mixing times for noesyegpph (2048 k data points, 256 increments, 192 transients) were chosen between 50 and 200 ms. In addition, spectra of samples at pH 8.5 were acquired in order to increase the binding constant (and thus  $k_{off}$ ). In all cases, no signals arising from specific trNOE of GDP-Fuc could be detected.

## 7.7 Enzyme kinetic studies

### 7.7.1 Enzyme kinetic studies by progress curve analysis of FucTA

For analysis of enzyme kinetics of FucTA, samples were prepared as described above. They contained 1.5  $\mu$ M FucTA, 10 U alkaline phosphatase, sequencing grade, from calf intestine (Roche) and 1.0 mM heptasaccharide **1a** or 1.6 mM octasaccharide **2a**, respectively. Spectra were recorded at 300 K. Prior to adding donor substrate, a 1D proton spectrum employing an excitation sculpting sequence as described above (zgesgp) was acquired with 32 transients (2 min experiment duration). Resonances from this spectrum yielded the intensities for  $t = 0$ . Thereafter, GDP-Fuc was added to the sample to give a concentration of 2 mM and the sample solution was mixed quickly. The sample was then reinserted into the magnet and after applying a short shimming routine, 1H NMR spectra using identical parameters as mentioned above were recorded at different time points for 18 h. FIDs were multiplied with an exponential function (line broadening 0.5) prior to Fourier transformation. All well dispersed resonances were integrated with respect to the sum integral of the H-1' resonances of GDP and GDP-Fuc, which corresponds to a concentration of 2 mM. From the integrals, substrate and product concentrations were calculated at each time point.

### 7.7.2 Enzyme kinetic studies by progress curve analysis of FUT8

For analysis of enzyme kinetics of FUT8, samples contained 2.0  $\mu$ M FUT8, 10 U alkaline phosphatase, sequencing grade, from calf intestine (Roche) and 0.9 mM GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(GlcNAc $\beta$ 1-2Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $\beta$ 1-4GlcNAc $\beta$ 1-N-Ac (heptasaccharide **1b**), respectively. BSA (Fluka, >98%) was added to give a concentration of 1.0 mg/mL in the sample to prevent adhesion and subsequent degeneration of FUT8 to the glass surface of the NMR tube at such low concentrations. Spectra were recorded at 310 K. Prior to adding donor substrate, a 1D proton spectrum employing an excitation sculpting sequence as described above (zgesgp) was acquired with 64 transients (4 min experiment duration). Resonances

from this spectrum yielded the intensities for  $t = 0$ . Thereafter, GDP-Fuc was added to the sample to give a concentration of 3.0 mM and the sample solution was mixed quickly. The sample was then reinserted into the magnet and after applying a short shimming routine,  $^1\text{H}$  NMR spectra using identical parameters as mentioned above were recorded at different time points for 30 min. FIDs were multiplied with an exponential function (line broadening 0.5) prior to Fourier transformation. All well dispersed resonances were integrated with respect to the sum integral of the H-1' resonances of GDP and GDP-Fuc, which corresponds to a concentration of 3.0 mM. From the integrals, substrate and product concentrations were calculated at each time point.

For analysis of the reaction with 1- $\beta$ -*N*-Acetyl chitotriose as acceptor substrate, samples contained 2.4  $\mu\text{M}$  FUT8, 10 U alkaline phosphatase, sequencing grade, from calf intestine (Roche) and 2.0 mM 1- $\beta$ -*N*-Acetyl chitotriose, 3.0 mM GDP-Fuc and 1 mg/mL BSA. Spectra were acquired using the zgesgp pulse sequence employing a selective  $180^\circ$  pulse with 2 ms duration and 32 transients over a total reaction time of 60 h. Integrals of the resonances were normalized with respect to spectra of the single compounds in order to eliminate contributions of  $T_1$  relaxation and solvent suppression. For resonances of fucose, the anomeric ratio ( $\alpha$ : $\beta$  30:70) was taken into account. The concentration of each compound was calculated with respect to the sum integral of all methyl resonances of fucosyl residues in the spectrum which corresponds to a concentration of 3.0 mM throughout the reaction. Concentration of components was calculated from the average of several resonances of the respective compound: GDP-Fuc (H-8, H-1', H-4', H-6''); guanosine (H-8, H-1', H-3'); 1- $\beta$ -*N*-Acetyl chitotriose (H-1, NAc  $\text{CH}_3$ ), 1- $\beta$ -*N*-Acetyl  $\alpha$ 1,6-fucosyl chitotriose (H-1, Fuc H-4, Fuc H-6), fucose (H-1( $\alpha$ ), H-6( $\beta$ )).

## 7.8 Molecular modeling

All simulations were performed using the Desmond Molecular Dynamics System, version 2.2, D 2009 (E. Shaw Research, New York, NY) as implemented in Maestro-Desmond Interoperability Tools, version 2.2 2009 (Schrödinger, New York, NY). [127] MD simulations are based on the crystal structure of human FUT8 [PDB ID: 2DE0] [44]. All hydrogen atoms were added and ionizable side chains as well as the C- and N-terminus were converted into their default ionization state and the system was minimized using a steepest descend algorithm.

### 7.8.1 Determination of the GDP-Fuc binding site

The donor binding site was determined by using the structure of *ce*POFUT in complex with GDP-Fuc [PDB ID: 3ZY6]. Peptide segments in both proteins that possess high structural similarity were identified: In the first step, the proteins were aligned along backbones of the beta sheets from Lys302 to Ser308 (*ce*POFUT) and from Lys402 to Thr408 (FUT8) in order to find further regions with high structural similarity. Four regions were identified to have high structural similarity (cf. Figure 4-15). In a second step, the segments Arg40-Leu58, Pro233-Arg240, Leu348-Asn352 and Val354-Gly369 of *ce*POFUT were aligned with Gly219-Thr237, Pro358-Arg365, Phe462-Thr466 and Ser468-His483 of FUT8, respectively. Further, GDP-Fuc from the *ce*POFUT structure was positioned at the putative binding site in the FUT8 crystal structure. After alignment of the side chain conformation of Arg365 of FUT8 to that of Arg240 in *ce*POFUT, the complex of FUT8 and GDP-Fuc was minimized.

### 7.8.2 Molecular dynamics simulations

For all MD simulations, the system was simulated at 310 K and at a pressure of 1.01325 bar over 1.5 ns with recording intervals of 0.1 ps for energy and 0.5 ps for the trajectory. The OPLS 2005 Force Field [128] was used in combination with the SPC model for water [129]. The method employed for electrostatic interactions was the Particle mesh Ewald (PME) [130] with a real space cut-off of 9.0 Å. The system was coupled to a Nose-Hoover-Chain thermostat [131] with a relaxation time of 1.0 ps and to an isotropic Martyna-Tobias-Klein barostat [132] with a relaxation time of 2.0 ps. The integration time step was set to 2.0 fs. Data analysis was conducted using Desmond Simulation Event Analysis, and all molecular images were made with PyMol (The PyMOL Molecular Graphics System, Schrödinger, LLC, 2011).

#### 7.8.2.1 FUT8-GDP-Fuc complex

The system was immersed into an orthorhombic water box containing 0.2 M sodium chloride and 16,847 solvent molecules in total. The box dimensions were calculated with the buffer method in such a way that the minimal distance between the periodic images of the box was 10 Å in each dimension and the box volume was 631,032 Å<sup>3</sup>. The total number of atoms during MD simulations was 59,073.

EXPERIMENTAL PROCEDURES

Table 7-17: Results of the simulation quality analysis.

<b>Block length for averaging</b>	10.0 ps
<b>Duration</b>	1.5 ns
<b>Time</b>	1.5 ns
<b>Degrees of freedom</b>	120,390
<b>Particles</b>	16,849 (1 protein, 1 ligand, 63 sodium ions, 63 chloride ions and 16,721 water molecules)
<b>Atoms</b>	59,073
<b>Target temp</b>	310 K
<b>Ensemble type</b>	Hoover-Noise chain thermostat, constant particle number, constant volume and constant temperature (HN_NVT)

Table 7-18: Average properties and their standard deviations of the MD simulation.

Properties	Average	Std. Dev.	Slope (ps <sup>-1</sup> )
<b>Total energy [kcal/mol]</b>	-158,468	9.7	0.014
<b>Potential energy [kcal/mol]</b>	-196,039	108.6	-0.089
<b>Temperature [K]</b>	308.620	0.4	0.0
<b>Pressure [bar]</b>	-1107	14.1	0.002
<b>Volume [Å<sup>3</sup>]</b>	631,032	0.0	0.0

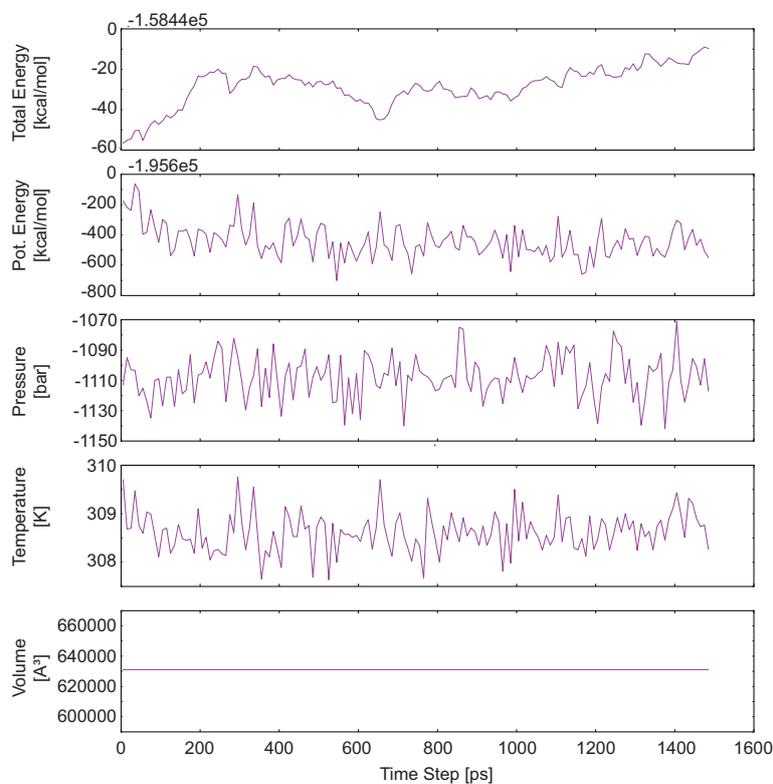


Figure 7-1: Behavior of properties during the MD simulation of the FUT8–GDP-Fuc complex over time.

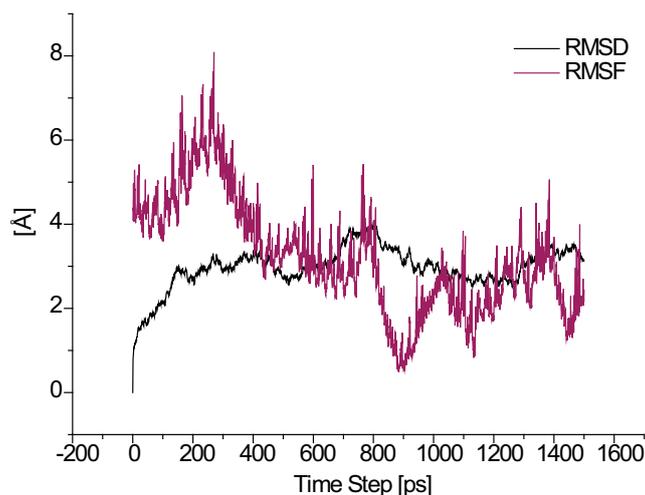


Figure 7-2: Root mean square deviation (RMSD, black) and Root mean square fluctuation (RMSF, purple) of all residues in the MD simulation of the FUT8–GDP-Fuc complex.

### 7.8.2.2 FUT8 ternary complex

MD simulations are based on the model the complex of FUT8 and GDP-Fuc [108]. As a starting point, last frame of the MD simulation of the FUT8–GDP-Fuc complex was used. A model of the heptasaccharide **1b** was manually docked into the putative binding pocket as described in Section 4.3.3. After minimization of the ternary complex, the system was immersed into an orthorhombic water box containing 0.2 M sodium chloride and 58,697 solvent molecules in total. The box dimensions were calculated with the buffer method in such a way that the minimal distance between the periodic images of the box was 10 Å in each dimension and the box volume was 713,764 Å<sup>3</sup>. The total number of atoms during MD simulations was 66,429. The MD simulation was performed with the parameters described above.

Table 7-19: Results of the simulation quality analysis for the MD of the ternary complex.

<b>Block length for averaging</b>	10.0 ps
<b>Duration</b>	1.5 ns
<b>Time</b>	1.5 ns
<b>Degrees of freedom</b>	136,929
<b>Particles</b>	58,700 (1 protein, 2 ligands, 76 sodium ions, 76 chloride ions and 58,545 water molecules)
<b>Atoms</b>	66,429
<b>Target temp</b>	310 K
<b>Ensemble type</b>	Hoover-Noise chain thermostat, constant particle number, constant volume and constant temperature (HN_NVT)

## EXPERIMENTAL PROCEDURES

Table 7-20: Average properties and their standard deviations of the MD simulation.

Properties	Average	Std. Dev.	Slope ( $\text{ps}^{-1}$ )
<b>Total energy (kcal/mol)</b>	-18,1387	57	0.13
<b>Potential energy (kcal/mol)</b>	-224,260	112	-0.044
<b>Temperature (K)</b>	308.65	0.4	0.0
<b>Pressure (bar)</b>	-959.8	14.3	0.0
<b>Volume (<math>\text{\AA}^3</math>)</b>	713,764.01	0.0	0.0

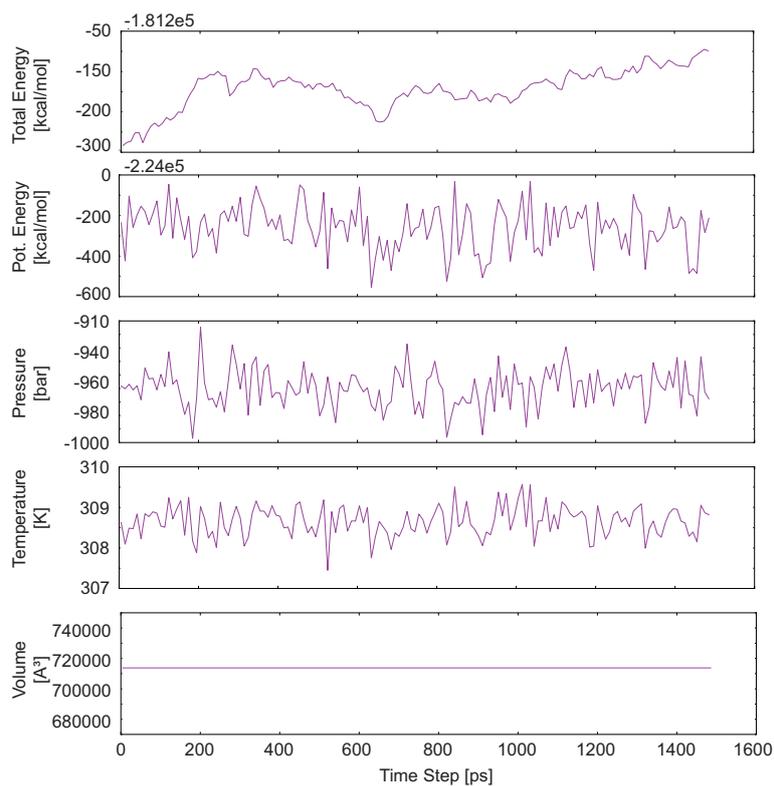


Figure 7-3: Behavior of properties during the MD simulation of the ternary complex over time.

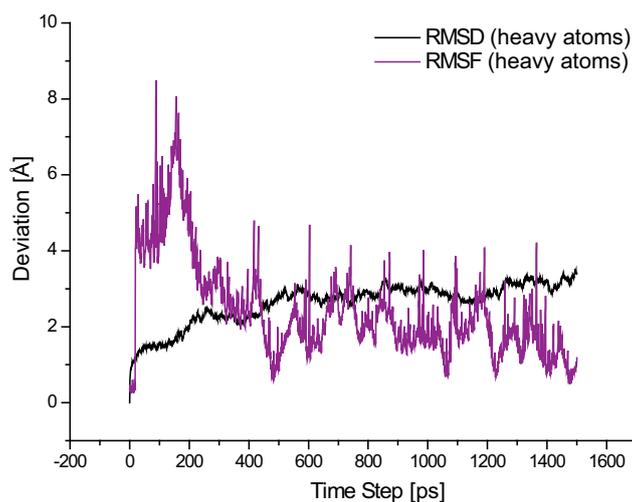


Figure 7-4: Root mean square deviation (RMSD, black) and Root mean square fluctuation (RMSF, purple) of all heavy atoms in the MD simulation. As the potential energy is stable throughout the simulation (cf. Figure 7-3), the higher RMSF and the increase in RMSD are explained by rearrangement of water molecules during the first 200 ps.

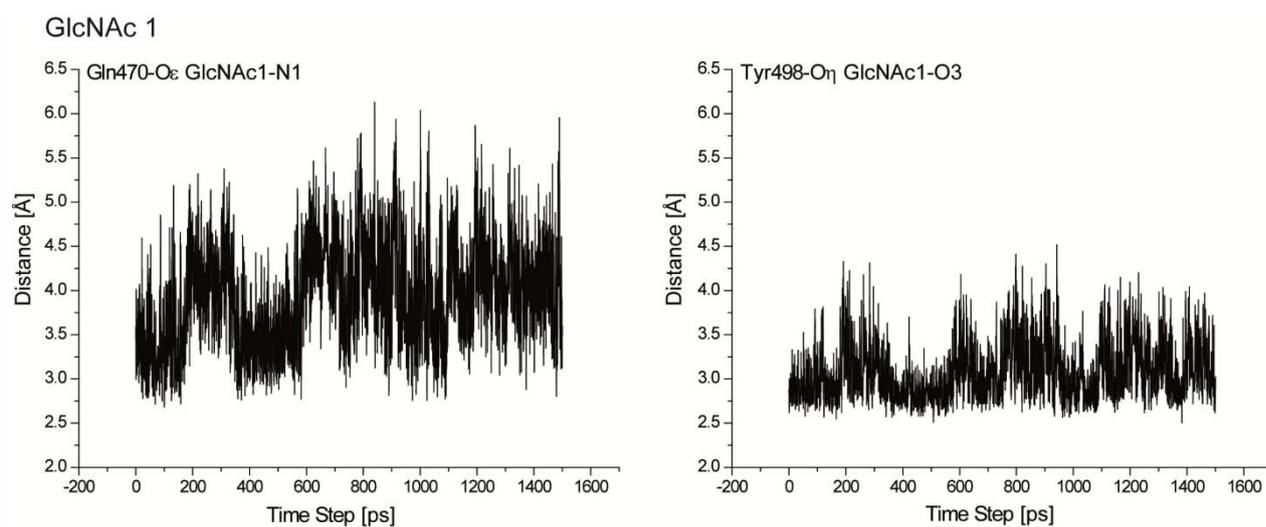


Figure 7-5: Trajectories of distances of atoms of FUT8 to atoms of the GlcNAc-1 of the acceptor substrate.

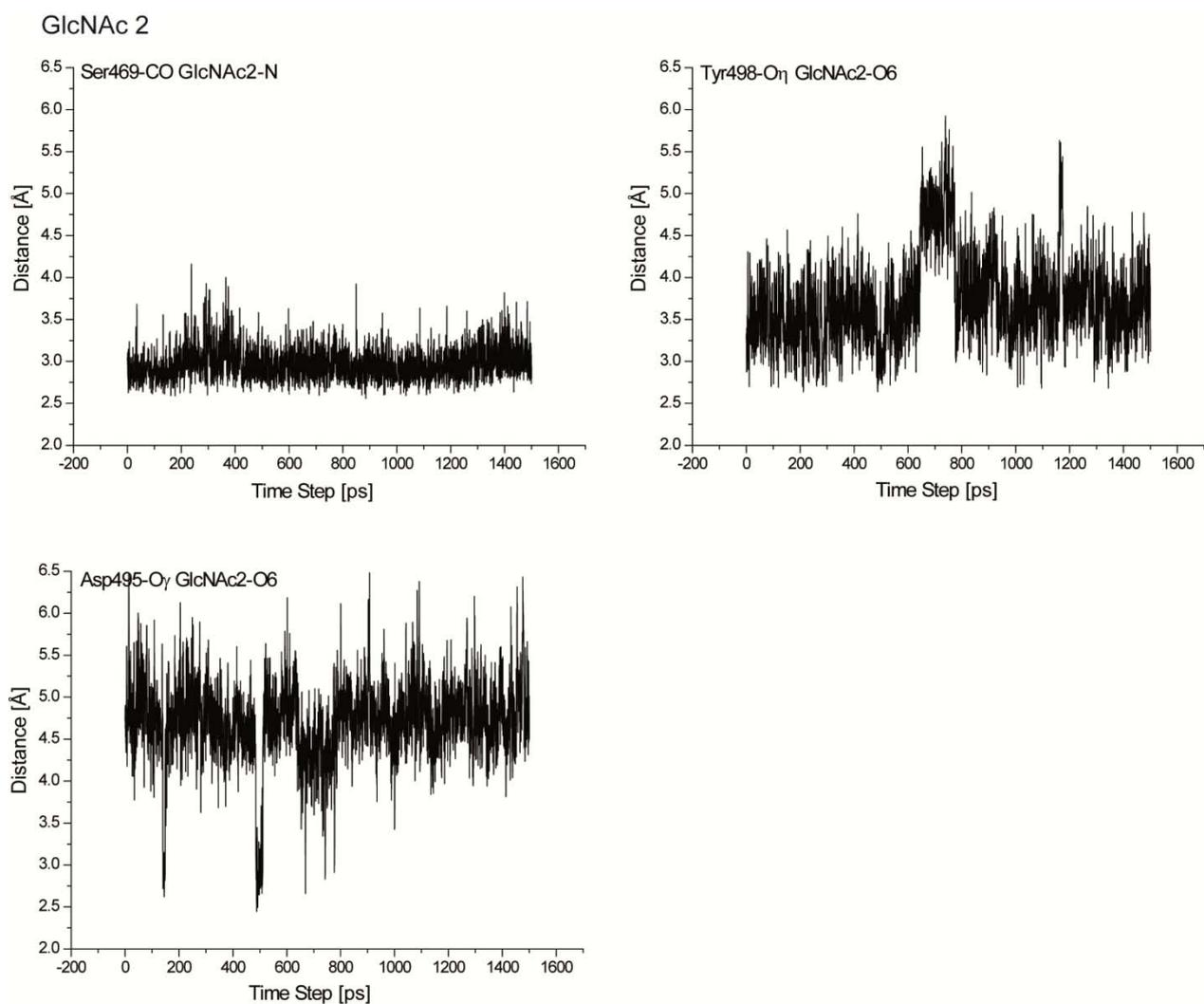


Figure 7-6: Trajectories of distances of atoms of FUT8 to atoms of the GlcNAc-2 of the acceptor substrate.

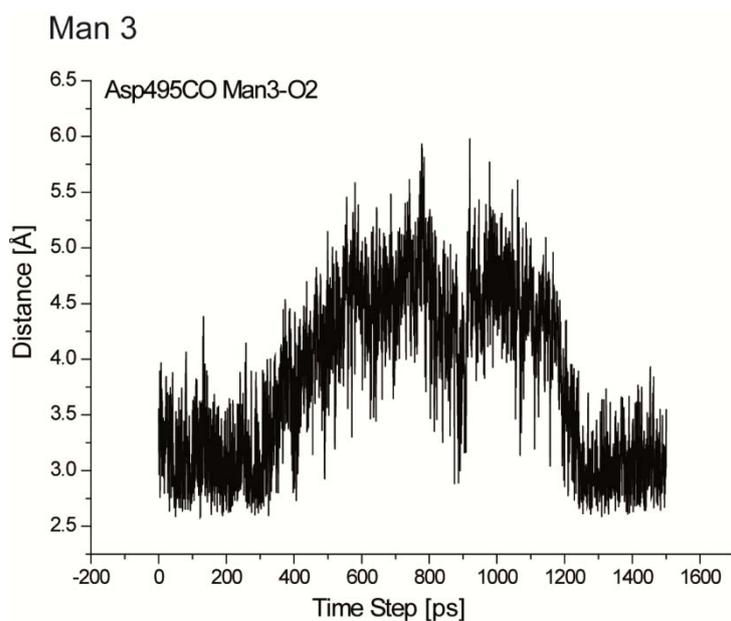


Figure 7-7: Trajectory of distances of atoms of FUT8 to atoms of the Man-3 of the acceptor substrate.

## Man 4

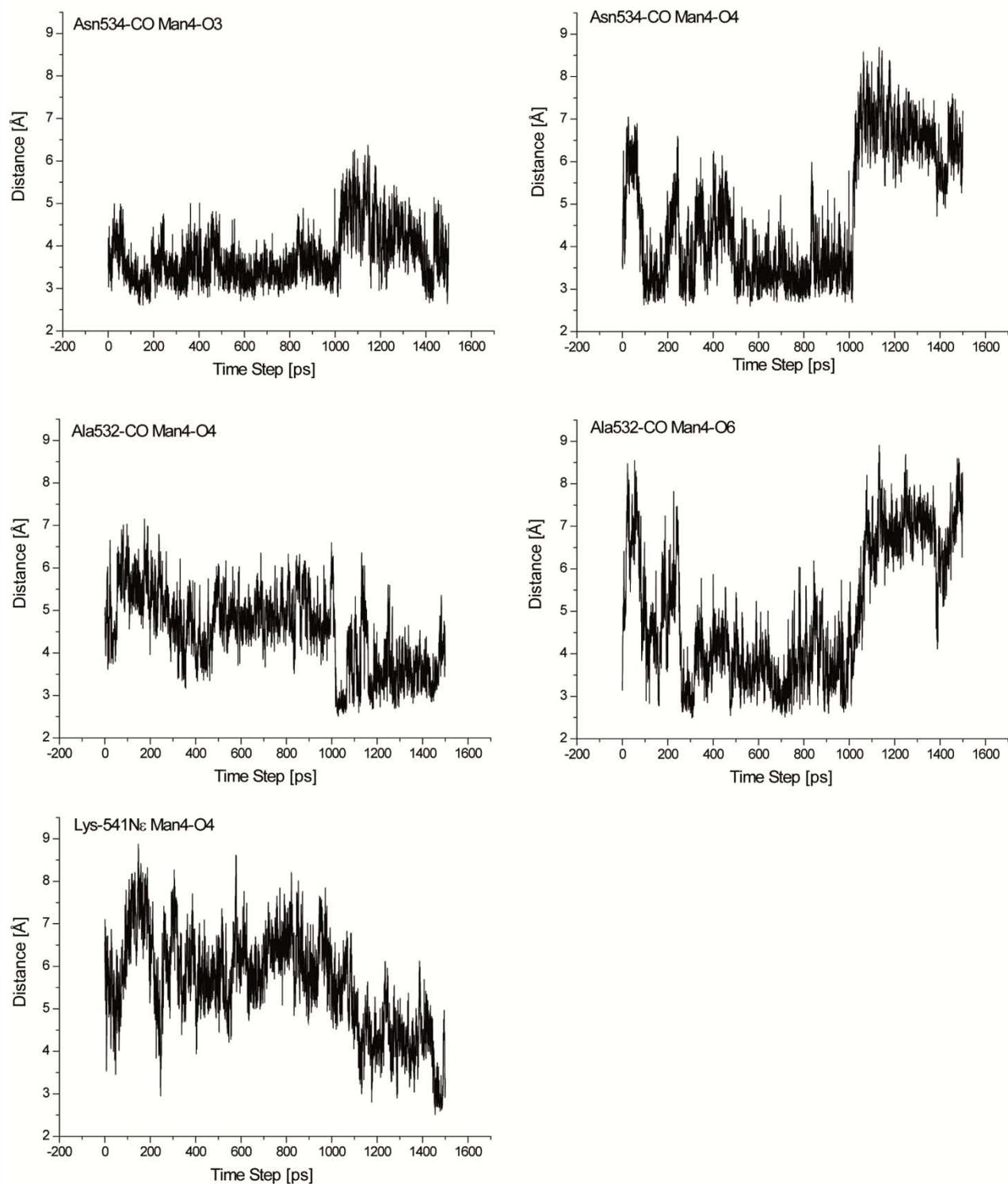


Figure 7-8: Trajectories of distances of atoms of FUT8 to atoms of the Man-4 of the acceptor substrate.

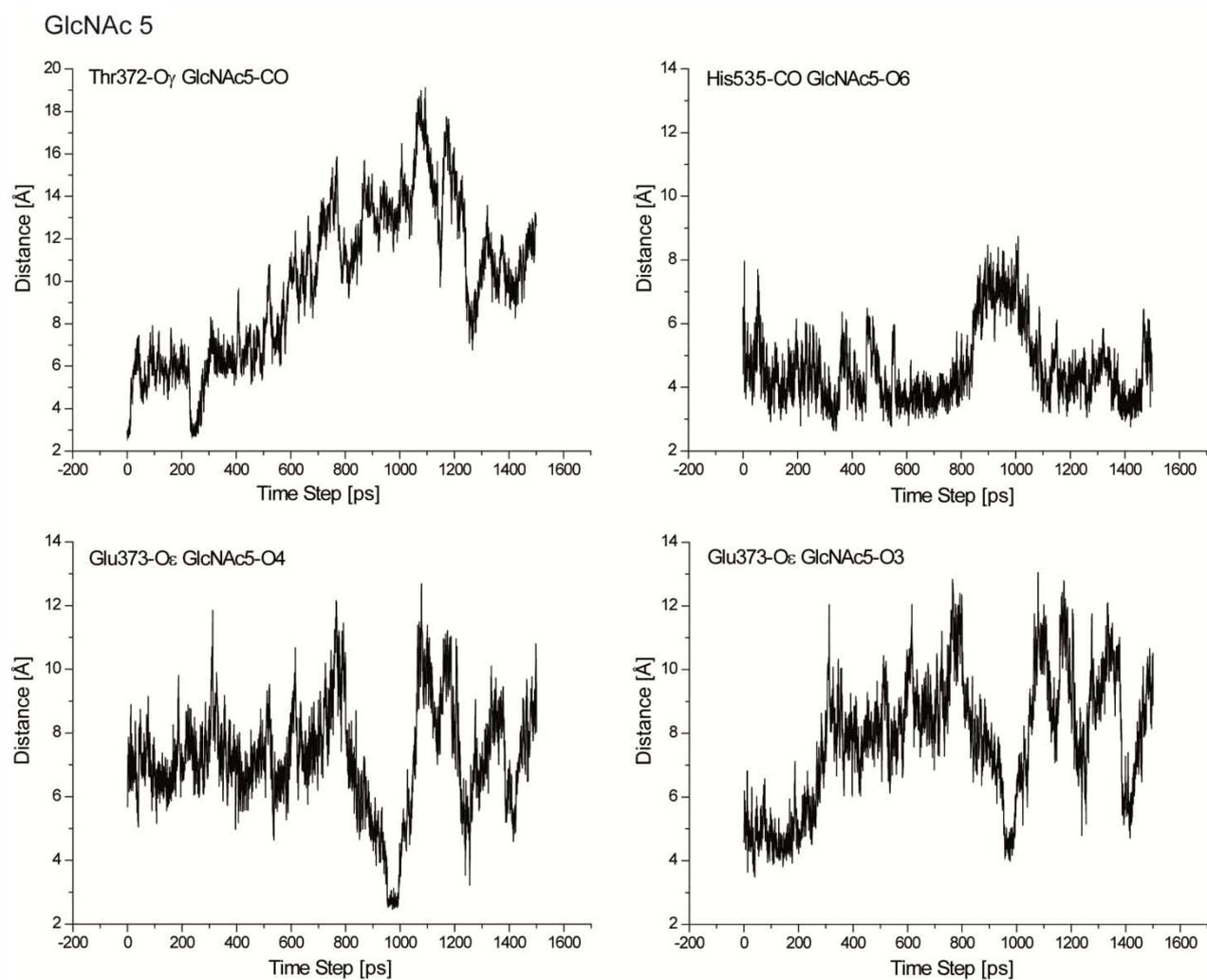


Figure 7-9: Trajectories of distances of atoms of FUT8 to atoms of the GlcNAc5 (3-mannose branch) of the acceptor substrate.

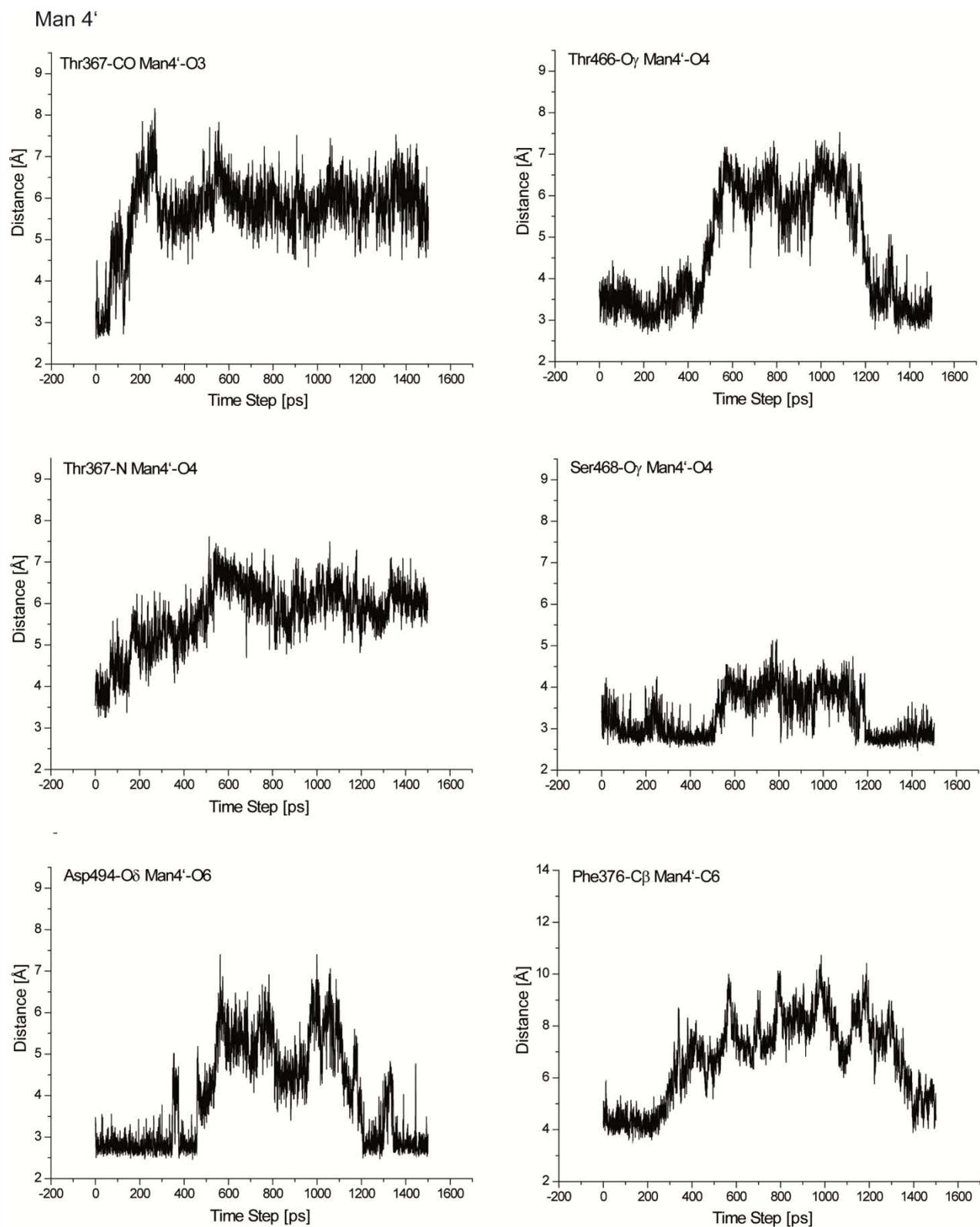


Figure 7-10: Trajectories of distances of atoms of FUT8 to the atoms of Man-4' (6-mannose branch) of the acceptor substrate.

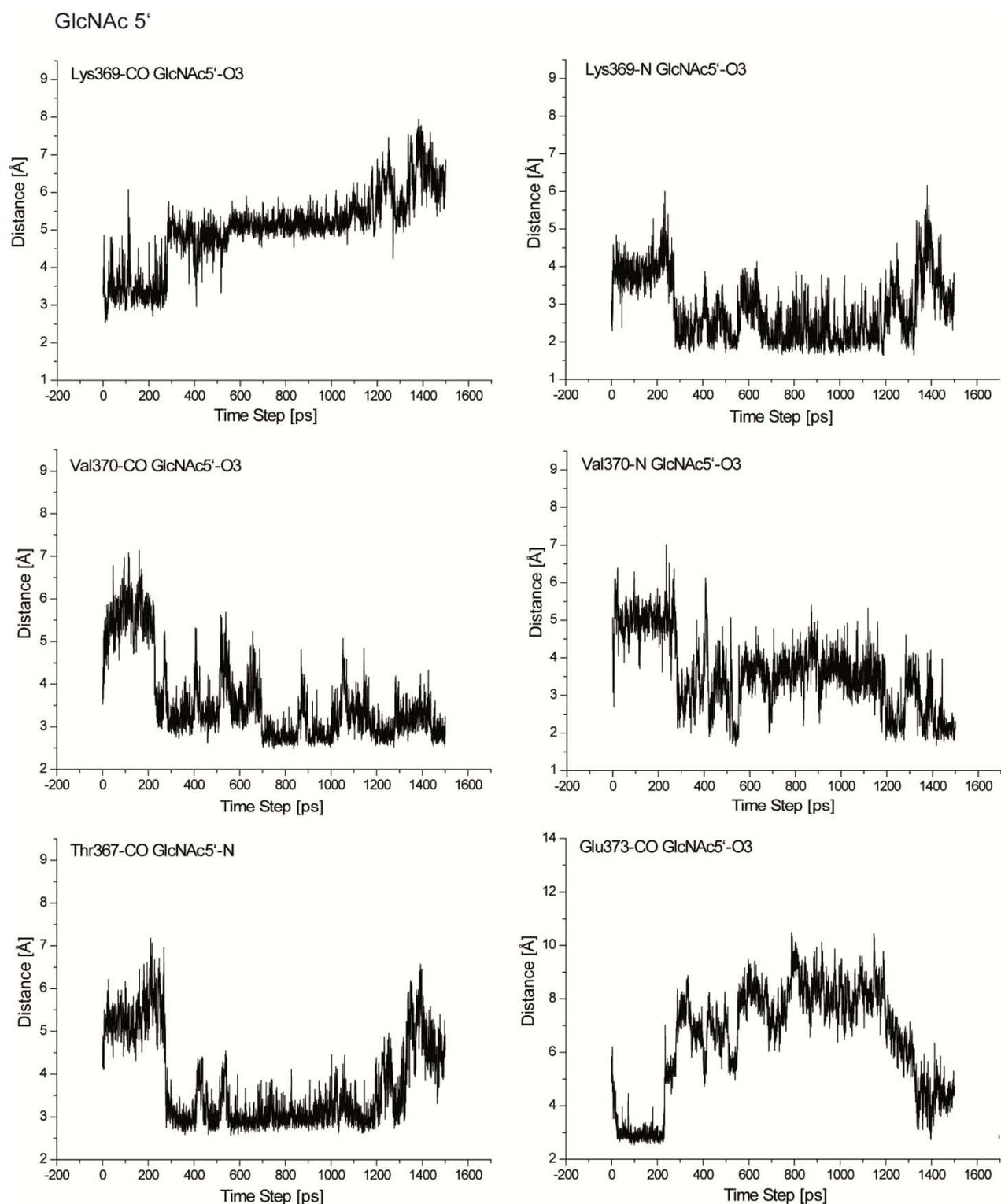


Figure 7-11: Trajectories of distances of atoms of FUT8 to the atoms (oxygen-3 and nitrogen) of GlcNAc5' (6-mannose branch) of the acceptor substrate.

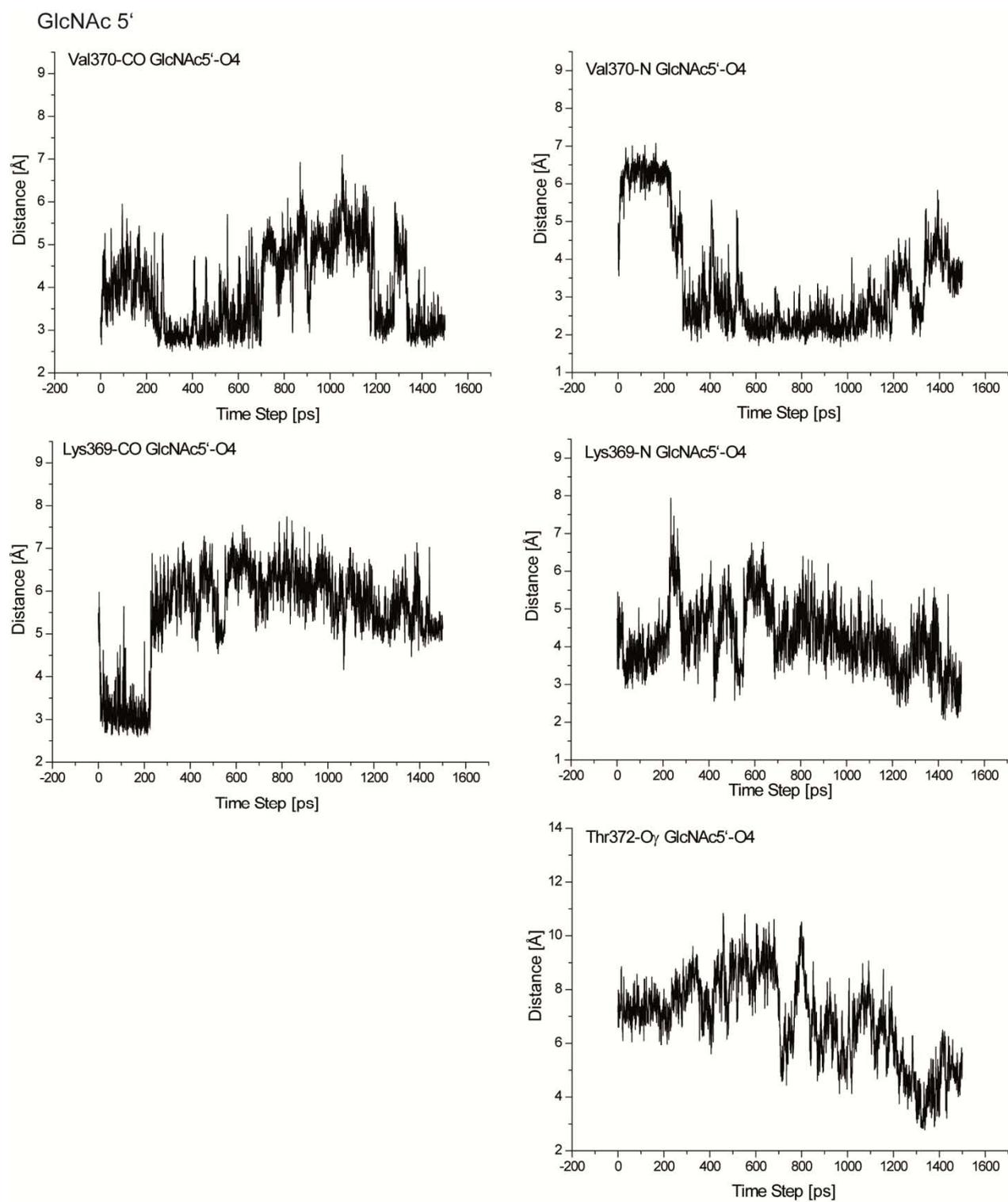


Figure 7-12: Trajectories of distances of atoms of FUT8 to the oxygen-4 atom of GlcNAc5' (6-mannose branch) of the acceptor substrate.

## EXPERIMENTAL PROCEDURES

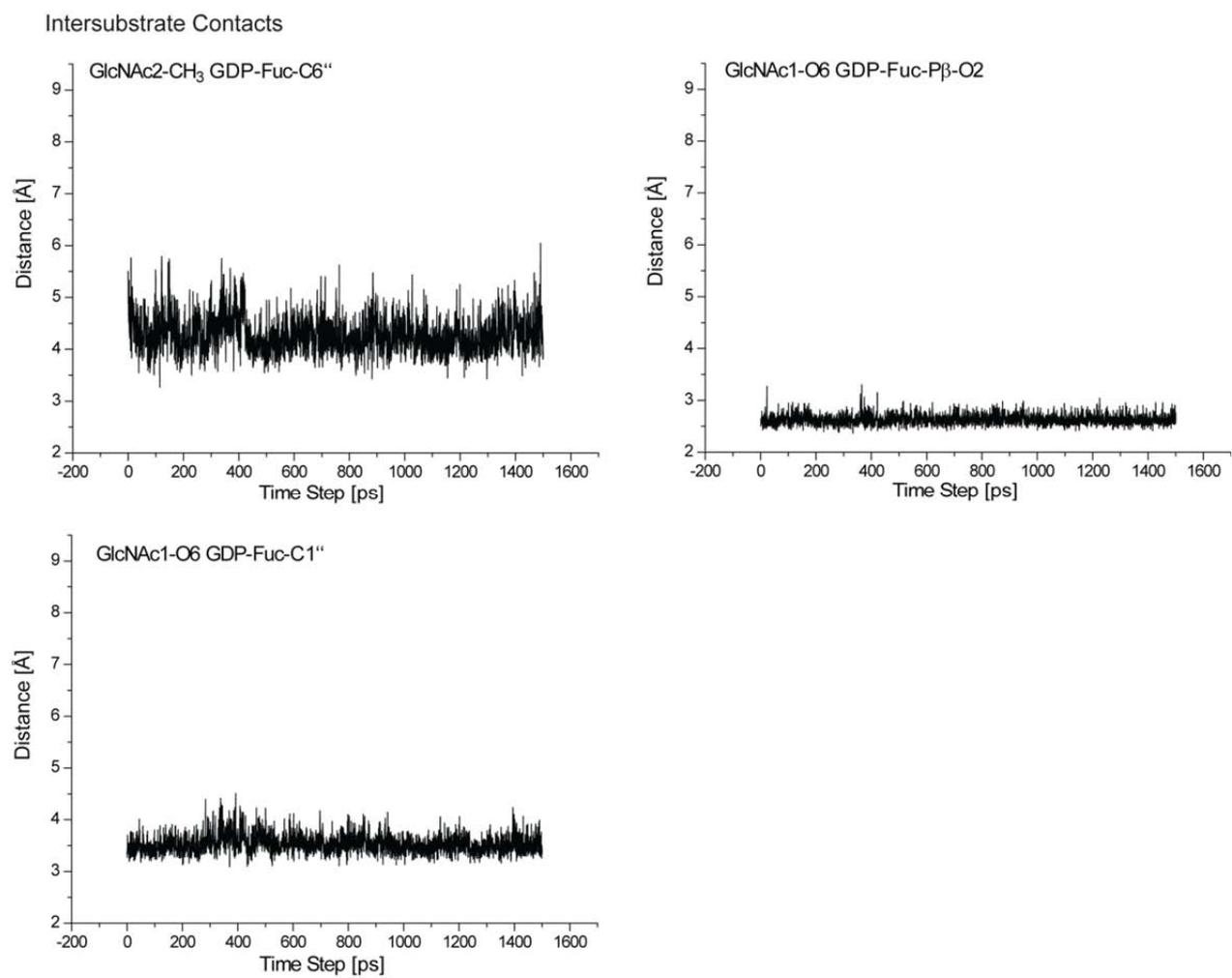


Figure 7-13: Trajectories of distances of atoms the donor substrate and of the acceptor substrate.

### 7.8.2.3 FUT8 ternary complex with chitotriose as acceptor substrate

MD simulations are based on the model the complex of FUT8 and GDP-Fuc [108]. As a starting point, last frame of the MD simulation of the FUT8 ternary complex was used. A model of the 1- $\beta$ -*N*-Acetyl chitotriose was manually docked into the putative binding pocket by means of fitting to the core trisaccharide of the heptasaccharide. The heptasaccharide was subsequently deleted and the system was minimized. The system was immersed into an orthorhombic water box containing 0.2 M sodium chloride and 51,862 solvent molecules in total. The box dimensions were calculated with the buffer method in such a way that the minimal distance between the periodic images of the box was 10 Å in each dimension and the box volume was 639,787 Å<sup>3</sup>. The total number of atoms during MD simulations was 59,504. The MD simulation was performed with the parameters described above.

Table 7-21: Results of the simulation quality analysis for the MD of the ternary complex with chitotriose.

<b>Block length for averaging</b>	10.0 ps
<b>Duration</b>	1.5 ns
<b>Time</b>	1.5 ns
<b>Degrees of freedom</b>	123,016
<b>Particles</b>	51,865 (1 protein, 2 ligands, 68 sodium ions, 68 chloride ions and 51,726 water molecules)
<b>Atoms</b>	59,504
<b>Target temp</b>	310 K
<b>Ensemble type</b>	Hoover-Noise chain thermostat, constant particle number, constant volume and constant temperature (HN_NVT)

Table 7-22: Average properties and their standard deviations of the MD simulation.

<b>Properties</b>	<b>Average</b>	<b>Std. Dev.</b>	<b>Slope (ps<sup>-1</sup>)</b>
<b>Total energy (kcal/mol)</b>	-162,197	31	0.068
<b>Potential energy (kcal/mol)</b>	-200,723	126	-0.14
<b>Temperature (K)</b>	308.62	0.44	0.0
<b>Pressure (bar)</b>	-1000	16	0.0
<b>Volume (Å<sup>3</sup>)</b>	639,786.58	0.0	0.0

## EXPERIMENTAL PROCEDURES

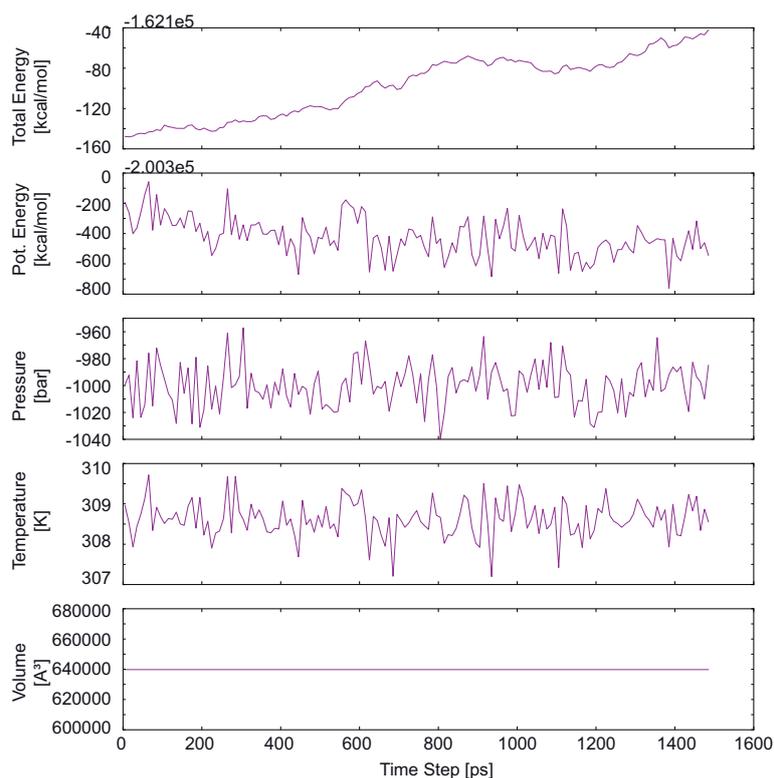


Figure 7-14: Behavior of properties during the MD simulation over time.

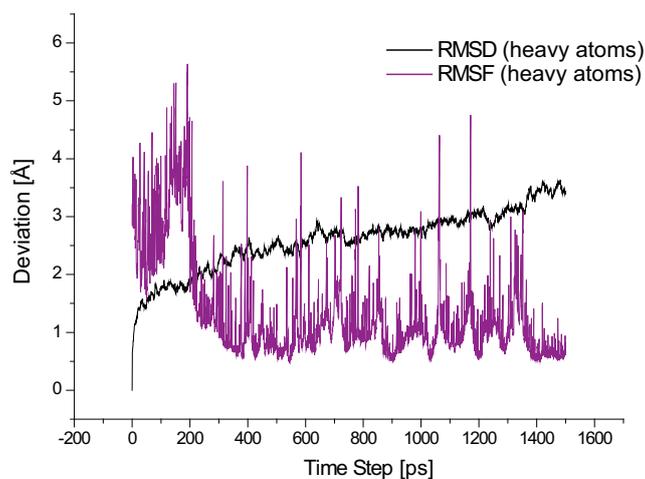


Figure 7-15: Root mean square deviation (RMSD, black) and Root mean square fluctuation (RMSF, purple) of all heavy atoms in the MD simulation. As the potential energy is stable throughout the simulation (cf. Figure 7-3), the higher RMSF and the increase in RMSD are explained by rearrangement of water molecules during the first 200 ps.

## 8 Hazards

Compound	Pictogram	Hazard statements	Precautionary statements
1-ethyl-3-(3-dimethylaminopropyl)carbodiimide		314	280-305 + 351 + 338-310
2-(2-aminoethoxy) ethanol		312, 314	280, 305 + 351 + 338, 310
2-( <i>N</i> -morpholino)-ethanesulfonic acid		315- 319- 335	261-305 + 351 + 338
3-vinylbenzoic acid		315- 319- 335	261-305 + 351 + 338
acetic acid		226, 314	280, 305 + 351 + 338, 310
acetic anhydride		226, 302, 314, 332	280, 305 + 351 + 338, 310
acetone		225, 319, 336, EUH066	210, 233, 305 + 351 + 338
acetonitrile		225, 302 + 312 + 332, 319	210, 305 + 351 + 338, 403 + 235
ammonia (aq., 25%)		314, 335, 400	280, 273, 301 + 330 + 331, 305 + 351 + 338, 309, 310
ammonium acetate		315- 319- 335	261-305 + 351 + 338
ammonium bicarbonate		302	
bis(2-hydroxyethyl)-amino-tris(hydroxymethyl)-methane		315, 319, 335	261, 305 + 351 + 338
chloroform		302, 315, 351, 373	281
cyclohexane		225- 304- 315- 336- 410	210-261-273-301 + 310- 331-501
dichloromethane		351	281
dithiothreitol		302, 315, 319, 335	261, 305 + 351 + 338
ethanol		225	210
ethyl acetate		225, 319, 336	210, 261, 305 + 351 + 338

## HAZARDS

Compound	Pictogram	Hazard statements	Precautionary statements
formic acid		226, 314	280, 305 + 351 + 338, 310
hydrochloric acid (37%)	 	314, 335	261, 280, 305 + 351 + 338, 310
iodoacetamide	 	301- 317- 334- 413	261-280-301 + 310-342 + 311
light petroleum (50-70)	   	225, 304, 340, 350, 361f, 373, 412	201, 210, 273, 281, 301 + 310, 308 + 313
methanol	  	225, 301, 311, 331, 370	210, 260, 280, 301 + 310, 311
<i>N,N</i> -carbonyldiimidazole	 	302- 314	280-305 + 351 + 338-310
<i>N,N</i> -dimethylformamide	 	226, 312, 319, 332, 360D	201, 280, 305 + 351 + 338, 308 + 313
<i>p</i> -toluenesulfonic acid monohydrate		315- 319- 335	261-305 + 351 + 338
pyridine	 	225, 302, 312, 332	210, 280
sodium	 	260, 314	223, 231 + 232, 280, 305 + 351 + 338, 370 + 378, 422
sodium azide	 	300- 410	264-273-301 + 310-501
sodium hydroxide		314	280, 305 + 351 + 338, 310
sulfuric acid		314	280, 305 + 351 + 338, 310
tetrahydrofurane	 	225, 319, 335	210, 261, 305 + 351 + 338
toluene	  	225, 304, 315, 336, 361d-373	210, 261, 281, 301 + 310, 331
triethylamine	  	225, 302, 312, 314, 332	210, 280, 305 + 351 + 338, 310
trifluoro acetic acid	 	314, 332, 412	273, 280, 305 + 351 + 338, 310
tris(hydroxymethyl)amino methane		315- 319- 335	261-305 + 351 + 338
trypan blue		350	201-308+313
$\alpha,\alpha$ -dimethoxy toluene		302	

## 9 References

- [1] E. Miyoshi, K. Noda, Y. Yamaguchi, S. Inoue, Y. Ikeda, W. Wang, J.H. Ko, N. Uozumi, W. Li, N. Taniguchi, The  $\alpha$ 1,6-fucosyltransferase gene and its biological significance, *Biochim. Biophys. Acta*, 1473 (1999) 9-20.
- [2] E. Staudacher, F. Altmann, I.B. Wilson, L. März, Fucose in N-glycans: from plant to man, *Biochim. Biophys. Acta*, 1473 (1999) 216-236.
- [3] D.J. Becker, J.B. Lowe, Fucose: biosynthesis and biological function in mammals, *Glycobiology*, 13 (2003) 41R-53R.
- [4] H. Behnken, M. Fellenberg, M. Koetzler, R. Jirmann, T. Nagel, B. Meyer, Resolving the problem of chromatographic overlap by 3D cross correlation (3DCC) processing of LC, MS and NMR data for characterization of complex glycan mixtures, *Anal. Bioanal. Chem.*, 404 (2012) 1427–1437.
- [5] M. Fellenberg, A. Coksezen, B. Meyer, Characterization of picomole amounts of oligosaccharides from glycoproteins by <sup>1</sup>H NMR spectroscopy, *Angew. Chem. Int. Ed. Engl.*, 49 (2010) 2630-2633.
- [6] A. Varki, R. Cummings, J. Esko, *Essentials of Glycobiology*, in, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), 2009.
- [7] R.J. Kelly, S. Rouquier, D. Giorgi, G.G. Lennon, J.B. Lowe, Sequence and expression of a candidate for the human Secretor blood group alpha(1,2)fucosyltransferase gene (FUT2). Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype, *J. Biol. Chem.*, 270 (1995) 4640-4649.
- [8] R.D. Larsen, L.K. Ernst, R.P. Nair, J.B. Lowe, Molecular cloning, sequence, and expression of a human GDP-L-fucose:beta-D-galactoside 2-alpha-L-fucosyltransferase cDNA that can form the H blood group antigen, *Proc. Natl. Acad. Sci. U S A*, 87 (1990) 6674-6678.
- [9] M. Kaneko, T. Kudo, H. Iwasaki, Y. Ikehara, S. Nishihara, S. Nakagawa, K. Sasaki, T. Shiina, H. Inoko, N. Saitou, H. Narimatsu, Alpha1,3-fucosyltransferase IX (Fuc-TIX) is very highly conserved between human and mouse; molecular cloning, characterization and tissue distribution of human Fuc-TIX, *FEBS Lett.*, 452 (1999) 237-242.
- [10] S. Natsuka, K.M. Gersten, K. Zenita, R. Kannagi, J.B. Lowe, Molecular cloning of a cDNA encoding a novel human leukocyte alpha-1,3-fucosyltransferase capable of synthesizing the sialyl Lewis x determinant, *J. Biol. Chem.*, 269 (1994) 20806.
- [11] B. Ma, J.L. Simala-Grant, D.E. Taylor, Fucosylation in prokaryotes and eukaryotes, *Glycobiology*, 16 (2006) 158R-184R.
- [12] P. Malý, A.D. Thall, B. Petryniak, C.E. Rogers, P.L. Smith, R.M. Marks, R.J. Kelly, K.M. Gersten, G. Cheng, T.L. Saunders, S.A. Camper, R.T. Camphausen, F.X. Sullivan, Y. Isogai, O. Hindsgaul, U.H. von Andrian, J.B. Lowe, The  $\alpha$ (1,3)Fucosyltransferase Fuc-TVII Controls Leukocyte Trafficking through an Essential Role in L-, E-, and P-selectin Ligand Biosynthesis, *Cell*, 86 (1996) 643-653.
- [13] J.W. Homeister, A. Daugherty, J.B. Lowe,  $\alpha$ (1,3)Fucosyltransferases FucT-IV and FucT-VII Control Susceptibility to Atherosclerosis in Apolipoprotein E<sup>-/-</sup> Mice, *Arteriosclerosis, Thrombosis, and Vascular Biology*, 24 (2004) 1897-1903.
- [14] X. Wang, S. Inoue, J. Gu, E. Miyoshi, K. Noda, W. Li, Y. Mizuno-Horikawa, M. Nakano, M. Asahi, M. Takahashi, N. Uozumi, S. Ihara, S.H. Lee, Y. Ikeda, Y. Yamaguchi, Y. Aze, Y. Tomiyama, J. Fujii, K. Suzuki, A. Kondo, S.D. Shapiro, C. Lopez-Otin, T. Kuwaki, M. Okabe, K. Honke, N. Taniguchi,

## REFERENCES

---

- Dysregulation of TGF- $\beta$ 1 receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucose-deficient mice, *Proc. Natl. Acad. Sci. U S A*, 102 (2005) 15791-15796.
- [15] K. Schaefer, J. Albers, N. Sindhuwinata, T. Peters, B. Meyer, A new concept for glycosyltransferase inhibitors: Nonionic mimics of the nucleotide donor of the human blood group B galactosyltransferase, *ChemBioChem*, 13 (2012) 443-450.
- [16] R.R. Schmidt, K.-H. Jung, Glycosyltransferase inhibitors, in: C.-H. Wong (Ed.) *Carbohydrate-based Drug Discovery*, vol. 2, Wiley-VCH, Weinheim, 2003, pp. 609-659.
- [17] J.R. Wilson, D. Williams, H. Schachter, Control of glycoprotein synthesis - N-acetylglucosamine linkage to a mannose residue as a signal for attachment of L-fucose to asparagine-linked N-acetylglucosamine residue of glycopeptide from  $\alpha$ 1-acid glycoprotein, *Biochem. Biophys. Res. Commun.*, 72 (1976) 909-916.
- [18] M. Takahashi, Y. Kuroki, K. Ohtsubo, N. Taniguchi, Core fucose and bisecting GlcNAc, the direct modifiers of the N-glycan core: their functions and target proteins, *Carbohydr. Res.*, 344 (2009) 1387-1390.
- [19] X. Wang, J. Gu, H. Ihara, E. Miyoshi, K. Honke, N. Taniguchi, Core fucosylation regulates epidermal growth factor receptor-mediated intracellular signaling, *J. Biol. Chem.*, 281 (2006) 2572-2577.
- [20] X. Wang, T. Fukuda, W. Li, C.X. Gao, A. Kondo, A. Matsumoto, E. Miyoshi, N. Taniguchi, J. Gu, Requirement of Fut8 for the expression of vascular endothelial growth factor receptor-2: a new mechanism for the emphysema-like changes observed in Fut8-deficient mice, *J. Biochem.*, 145 (2009) 643-651.
- [21] D. Osumi, M. Takahashi, E. Miyoshi, S. Yokoe, S.H. Lee, K. Noda, S. Nakamori, J. Gu, Y. Ikeda, Y. Kuroki, K. Sengoku, M. Ishikawa, N. Taniguchi, Core fucosylation of E-cadherin enhances cell-cell adhesion in human colon carcinoma WiDr cells, *Cancer Sci.*, 100 (2009) 888-895.
- [22] Y. Zhao, S. Itoh, X. Wang, T. Isaji, E. Miyoshi, Y. Kariya, K. Miyazaki, N. Kawasaki, N. Taniguchi, J. Gu, Deletion of core fucosylation on  $\alpha$ 3 $\beta$ 1 integrin down-regulates its functions, *J. Biol. Chem.*, 281 (2006) 38343-38350.
- [23] K. Noda, E. Miyoshi, N. Uozumi, C.X. Gao, K. Suzuki, N. Hayashi, M. Hori, N. Taniguchi, High expression of  $\alpha$ 1,6-fucosyltransferase during rat hepatocarcinogenesis, *Int. J. Cancer.*, 75 (1998) 444-450.
- [24] K. Noda, E. Miyoshi, N. Uozumi, S. Yanagidani, Y. Ikeda, C. Gao, K. Suzuki, H. Yoshihara, K. Yoshikawa, K. Kawano, N. Hayashi, M. Hori, N. Taniguchi, Gene expression of  $\alpha$ 1,6-fucosyltransferase in human hepatoma tissues: a possible implication for increased fucosylation of alpha-fetoprotein, *Hepatology*, 28 (1998) 944-952.
- [25] K. Moriwaki, K. Noda, T. Nakagawa, M. Asahi, H. Yoshihara, N. Taniguchi, N. Hayashi, E. Miyoshi, A high expression of GDP-fucose transporter in hepatocellular carcinoma is a key factor for increases in fucosylation, *Glycobiology*, 17 (2007) 1311-1320.
- [26] R.L. Shields, J. Lai, R. Keck, L.Y. O'Connell, K. Hong, Y.G. Meng, S.H. Weikert, L.G. Presta, Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fc $\gamma$ 3 and antibody-dependent cellular toxicity, *J. Biol. Chem.*, 277 (2002) 26733-26740.
- [27] T. Shinkawa, K. Nakamura, N. Yamane, E. Shoji-Hosaka, Y. Kanda, M. Sakurada, K. Uchida, H. Anazawa, M. Satoh, M. Yamasaki, N. Hanai, K. Shitara, The absence of fucose but not the presence of galactose

## REFERENCES

---

- or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity, *J. Biol. Chem.*, 278 (2003) 3466-3473.
- [28] S. André, T. Kožár, S. Kojima, C. Unverzagt, H.-J. Gabius, From structural to functional glycomics: core substitutions as molecular switches for shape and lectin affinity of N-glycans, in: *Biol. Chem.*, vol. 390, 2009, pp. 557.
- [29] W. Nishima, N. Miyashita, Y. Yamaguchi, Y. Sugita, S. Re, Effect of bisecting GlcNAc and core fucosylation on conformational properties of biantennary complex-type N-glycans in solution, *J. Phys. Chem. B*, 116 (2012) 8504-8512.
- [30] N. Uozumi, S. Yanagidani, E. Miyoshi, Y. Ihara, T. Sakuma, C.X. Gao, T. Teshima, S. Fujii, T. Shiba, N. Taniguchi, Purification and cDNA cloning of porcine brain GDP-L-Fuc:N-acetyl- $\beta$ -D-glucosaminide  $\alpha$ 1,6-fucosyltransferase, *J. Biol. Chem.*, 271 (1996) 27810-27817.
- [31] S. Yanagidani, N. Uozumi, Y. Ihara, E. Miyoshi, N. Yamaguchi, N. Taniguchi, Purification and cDNA cloning of GDP-L-Fuc:N-acetyl- $\beta$ -D-glucosaminide: $\alpha$ 1,6-fucosyltransferase ( $\alpha$ 1,6 FucT) from human gastric cancer MKN45 cells, *J. Biochem.*, 121 (1997) 626-632.
- [32] J.A. Voynow, R.S. Kaiser, T.F. Scanlin, M.C. Glick, Purification and characterization of GDP-L-fucose-N-acetyl  $\beta$ -D-glucosaminide  $\alpha$ 1,6-fucosyltransferase from cultured human skin fibroblasts. Requirement of a specific biantennary oligosaccharide as substrate, *J. Biol. Chem.*, 266 (1991) 21572-21577.
- [33] M. Costache, P.-A. Apoil, A. Cailleau, A. Elmgren, G. Larson, S. Henry, A. Blancher, D. Iordachescu, R. Oriol, R. Mollicone, Evolution of fucosyltransferase genes in vertebrates, *J. Biol. Chem.*, 272: (1997) 29721-29728.
- [34] Y. Yamaguchi, Y. Ikeda, T. Takahashi, H. Ihara, T. Tanaka, C. Sasho, N. Uozumi, S. Yanagidani, S. Inoue, J. Fujii, N. Taniguchi, Genomic structure and promoter analysis of the human  $\alpha$ 1,6-fucosyltransferase gene (FUT8), *Glycobiology*, 10 (2000) 637-643.
- [35] I. Martinez-Duncker, R. Mollicone, J.J. Candelier, C. Breton, R. Oriol, A new superfamily of protein-O-fucosyltransferases,  $\alpha$ 2-fucosyltransferases, and  $\alpha$ 6-fucosyltransferases: phylogeny and identification of conserved peptide motifs, *Glycobiology*, 13 (2003) 1C-5C.
- [36] C. Breton, R. Oriol, A. Imberty, Conserved structural features in eukaryotic and prokaryotic fucosyltransferases, *Glycobiology*, 8 (1998) 87-94.
- [37] R. Oriol, R. Mollicone, A. Cailleau, L. Balanzino, C. Breton, Divergent evolution of fucosyltransferase genes from vertebrates, invertebrates, and bacteria, *Glycobiology*, 9 (1999) 323-334.
- [38] T. Takahashi, Y. Ikeda, A. Tateishi, Y. Yamaguchi, M. Ishikawa, N. Taniguchi, A sequence motif involved in the donor substrate binding by  $\alpha$ 1,6-fucosyltransferase: the role of the conserved arginine residues, *Glycobiology*, 10 (2000) 503-510.
- [39] H. Ihara, Y. Ikeda, N. Taniguchi, Reaction mechanism and substrate specificity for nucleotide sugar of mammalian  $\alpha$ 1,6-fucosyltransferase - a large-scale preparation and characterization of recombinant human FUT8, *Glycobiology*, 16 (2006) 333-342.
- [40] J. Kaminska, M.C. Glick, J. Koscielak, Purification and characterization of GDP-L-Fuc: N-acetyl  $\beta$ -D-glucosaminide  $\alpha$ 1,6-fucosyltransferase from human blood platelets, *Glycoconj. J.*, 15 (1998) 783-788.
- [41] G.D. Longmore, H. Schachter, Product-identification and substrate-specificity studies of the GDP-L-fucose:2-acetamido-2-deoxy- $\beta$ -D-glucoside (FUC goes to Asn-linked GlcNAc) 6- $\alpha$ -L-fucosyltransferase in a Golgi-rich fraction from porcine liver, *Carbohydr. Res.*, 100 (1982) 365-392.

## REFERENCES

---

- [42] K. Paschinger, E. Staudacher, U. Stemmer, G. Fabini, I.B. Wilson, Fucosyltransferase substrate specificity and the order of fucosylation in invertebrates, *Glycobiology*, 15 (2005) 463-474.
- [43] M.C. Shao, C.W. Sokolik, F. Wold, Specificity studies of the GDP-L-fucose - 2-acetamido-2-deoxy- $\beta$ -D-glucoside (Fuc-Asn-linked GlcNAc) 6- $\alpha$ -fucosyl-transferase from rat-liver golgi membranes, *Carbohydr. Res.*, 251 (1994) 163-173.
- [44] H. Ihara, Y. Ikeda, S. Toma, X. Wang, T. Suzuki, J. Gu, E. Miyoshi, T. Tsukihara, K. Honke, A. Matsumoto, A. Nakagawa, N. Taniguchi, Crystal structure of mammalian  $\alpha$ 1,6-fucosyltransferase, FUT8, *Glycobiology*, 17 (2007) 455-466.
- [45] K. Brzezinski, Z. Dauter, M. Jaskolski, Structures of NodZ  $\alpha$ 1,6-fucosyltransferase in complex with GDP and GDP-fucose, *Acta Cryst. D*, 68 (2012) 160-168.
- [46] K. Brzezinski, T. Stepkowski, S. Panjikar, G. Bujacz, M. Jaskolski, High-resolution structure of NodZ fucosyltransferase involved in the biosynthesis of the nodulation factor, *Acta. Biochim. Pol.*, 54 (2007) 537-549.
- [47] H. Ihara, S. Hanashima, T. Okada, R. Ito, Y. Yamaguchi, N. Taniguchi, Y. Ikeda, Fucosylation of chitooligosaccharides by human  $\alpha$ 1,6-fucosyltransferase requires a nonreducing terminal chitotriose unit as a minimal structure, in: *Glycobiology*, vol. 20, 2010, pp. 1021-1033.
- [48] P.K. Qasba, B. Ramakrishnan, E. Boeggeman, Substrate-induced conformational changes in glycosyltransferases, *Trends Biochem. Sci.*, 30 (2005) 53-62.
- [49] J. Angulo, B. Langpap, A. Blume, T. Biet, B. Meyer, N.R. Krishna, H. Peters, M.M. Palcic, T. Peters, Blood Group B Galactosyltransferase: Insights into substrate binding from NMR Experiments, *J. Am. Chem. Soc.*, 128 (2006) 13529-13538.
- [50] H.Y. Sun, S.W. Lin, T.P. Ko, J.F. Pan, C.L. Liu, C.N. Lin, A.H. Wang, C.H. Lin, Structure and mechanism of *Helicobacter pylori* fucosyltransferase. A basis for lipopolysaccharide variation and inhibitor design, *J. Biol. Chem.*, 282 (2007) 9973-9982.
- [51] E. Lira-Navarrete, J. Valero-González, R. Villanueva, M. Martínez-Júlvez, T. Tejero, P. Merino, S. Panjikar, R. Hurtado-Guerrero, Structural insights into the mechanism of protein O-fucosylation, *PLoS One*, 6 (2011) e25365.
- [52] M.P. Kötzler, S. Blank, H.N. Behnken, D. Alpers, F.I. Bantleon, E. Spillner, B. Meyer, Formation of the immunogenic  $\alpha$ 1,3-fucose epitope: Elucidation of substrate specificity and of enzyme mechanism of core fucosyltransferase A, *Insect Biochem. Mol. Biol.*, 42 (2012) 116-125.
- [53] G.K. Shoemaker, N. Soya, M.M. Palcic, J.S. Klassen, Temperature-dependent cooperativity in donor-acceptor substrate binding to the human blood group glycosyltransferases, *Glycobiology*, 18 (2008) 587-592.
- [54] I.B. Wilson, F. Altmann, Structural analysis of N-glycans from allergenic grass, ragweed and tree pollens: core  $\alpha$ 1,3-linked fucose and xylose present in all pollens examined, *Glycoconj. J.*, 15 (1998) 1055-1070.
- [55] R.C. Aalberse, J. Akkerdaas, R. van Ree, Cross-reactivity of IgE antibodies to allergens, *Allergy*, 56 (2001) 478-490.
- [56] W. Hemmer, M. Focke, D. Kolarich, I. Dalik, M. Götz, R. Jarisch, Identification by immunoblot of venom glycoproteins displaying immunoglobulin E-binding N-glycans as cross-reactive allergens in honeybee and yellow jacket venom, *Clin. Exp. Allergy*, 34 (2004) 460-469.

## REFERENCES

---

- [57] U. Jappe, M. Raulf-Heimsoth, M. Hoffmann, G. Burow, C. Hubsch-Muller, A. Enk, In vitro hymenoptera venom allergy diagnosis: improved by screening for cross-reactive carbohydrate determinants and reciprocal inhibition, *Allergy*, 61 (2006) 1220-1229.
- [58] F. Altmann, E. Staudacher, I.B. Wilson, L. Marz, Insect cells as hosts for the expression of recombinant glycoproteins, *Glycoconj. J.*, 16 (1999) 109-123.
- [59] D. Rendic, J. Kludiny, U. Stemmer, J. Schmidt, K. Paschinger, I.B. Wilson, Towards abolition of immunogenic structures in insect cells: characterization of a honey-bee (*Apis mellifera*) multi-gene family reveals both an allergy-related core alpha1,3-fucosyltransferase and the first insect Lewis-histo-blood-group-related antigen-synthesizing enzyme, *Biochem. J.*, 402 (2007) 105-115.
- [60] K. Nguyen, I. van Die, K.M. Grundahl, Z.S. Kowar, R.D. Cummings, Molecular cloning and characterization of the *Caenorhabditis elegans* alpha1,3-fucosyltransferase family, *Glycobiology*, 17 (2007) 586-599.
- [61] H. Leiter, J. Mucha, E. Staudacher, R. Grimm, J. Glossl, F. Altmann, Purification, cDNA cloning, and expression of GDP-L-Fuc:Asn-linked GlcNAc alpha1,3-fucosyltransferase from mung beans, *J. Biol. Chem.*, 274 (1999) 21830-21839.
- [62] G. Fabini, A. Freilinger, F. Altmann, I.B.H. Wilson, Identification of core alpha1,3-fucosylated glycans and cloning of the requisite fucosyltransferase cDNA from *Drosophila melanogaster*, *J. Biol. Chem.*, 276 (2001) 28058-28067.
- [63] V. Kubelka, F. Altmann, L. Marz, The asparagine-linked carbohydrate of honeybee venom hyaluronidase, *Glycoconj. J.*, 12 (1995) 77-83.
- [64] D. Rendic, A. Linder, K. Paschinger, N. Borth, I.B. Wilson, G. Fabini, Modulation of neural carbohydrate epitope expression in *Drosophila melanogaster* cells, *J. Biol. Chem.*, 281 (2006) 3343-3353.
- [65] F. Altmann, G. Kornfeld, T. Dalik, E. Staudacher, J. Glossl, Processing of asparagine-linked oligosaccharides in insect cells. N-acetylglucosaminyltransferase I and II activities in cultured lepidopteran cells, *Glycobiology*, 3 (1993) 619-625.
- [66] E. Staudacher, L. Marz, Strict order of (Fuc to Asn-linked GlcNAc) fucosyltransferases forming core-difucosylated structures, *Glycoconj. J.*, 15 (1998) 355-360.
- [67] I. Jabbal, H. Schachter, Pork liver guanosine diphosphate-L-fucose glycoprotein fucosyltransferases, *J. Biol. Chem.*, 246 (1971) 5154-5161.
- [68] E. Struppe, E. Staudacher, Occurrence of GDP-L-fucose: beta-N-acetylglucosamine (Fuc to Asn-linked GlcNAc) alpha 1,6-fucosyltransferases in porcine, sheep, bovine, rabbit and chicken tissues, *Biochim. Biophys. Acta*, 1475 (2000) 360-368.
- [69] E. Staudacher, F. Altmann, J. Glossl, L. Marz, H. Schachter, J.P. Kamerling, K. Hard, J.F. Vliegthart, GDP-fucose: beta-N-acetylglucosamine (Fuc to (Fuc alpha1,6GlcNAc)-Asn-peptide)alpha1,3-fucosyltransferase activity in honeybee (*Apis mellifica*) venom glands. The difucosylation of asparagine-bound N-acetylglucosamine, *Eur. J. Biochem.*, 199 (1991) 745-751.
- [70] I. Martinez-Duncker, J.C. Michalski, C. Bauvy, J.J. Candelier, B. Mennesson, P. Codogno, R. Oriol, R. Mollicone, Activity and tissue distribution of splice variants of alpha6-fucosyltransferase in human embryogenesis, *Glycobiology*, 14 (2004) 13-25.
- [71] M. Bruggert, T. Rehm, S. Shanker, J. Georgescu, T.A. Holak, A novel medium for expression of proteins selectively labeled with <sup>15</sup>N-amino acids in *Spodoptera frugiperda* (Sf9) insect cells, *J. Biomol. NMR*, 25 (2003) 335-348.

## REFERENCES

---

- [72] A. Strauss, F. Bitsch, B. Cutting, G. Fendrich, P. Graff, J. Liebetanz, M. Zurini, W. Jahnke, Amino-acid-type selective isotope labeling of proteins expressed in Baculovirus-infected insect cells useful for NMR studies, *J. Biomol. NMR*, 26 (2003) 367-372.
- [73] A. Strauss, F. Bitsch, G. Fendrich, P. Graff, R. Knecht, B. Meyhack, W. Jahnke, Efficient uniform isotope labeling of Abl kinase expressed in Baculovirus-infected insect cells, *J. Biomol. NMR*, 31 (2005) 343-349.
- [74] H. Seismann, S. Blank, L. Cifuentes, I. Braren, R. Bredehorst, T. Grunwald, M. Ollert, E. Spillner, Recombinant phospholipase A1 (Ves v 1) from yellow jacket venom for improved diagnosis of hymenoptera venom hypersensitivity, *Clin. Mol. Allergy*, 8 (2010) 7.
- [75] S. Blank, Y. Michel, H. Seismann, M. Plum, K. Greunke, T. Grunwald, R. Bredehorst, M. Ollert, I. Braren, E. Spillner, Evaluation of different glycoforms of honeybee venom major allergen phospholipase A2 (Api m 1) produced in insect cells, *Prot. Pept. Lett.*, 18 (2011) 415-422.
- [76] I.B. Wilson, J.E. Harthill, N.P. Mullin, D.A. Ashford, F. Altmann, Core alpha1,3-fucose is a key part of the epitope recognized by antibodies reacting against plant N-linked oligosaccharides and is present in a wide variety of plant extracts, *Glycobiology*, 8 (1998) 651-661.
- [77] M. Bencurova, W. Hemmer, M. Focke-Tejkl, I.B.H. Wilson, F. Altmann, Specificity of IgG and IgE antibodies against plant and insect glycoprotein glycans determined with artificial glycoforms of human transferrin, *Glycobiology*, 14 (2004) 457-466.
- [78] A. Kurosaka, A. Yano, N. Itoh, Y. Kuroda, T. Nakagawa, T. Kawasaki, The Structure of a Neural Specific Carbohydrate Epitope of Horseradish-Peroxidase Recognized by Anti-Horseradish Peroxidase Antiserum, *J. Biol. Chem.*, 266 (1991) 4168-4172.
- [79] V. Kubelka, F. Altmann, E. Staudacher, V. Tretter, L. Marz, K. Hard, J.P. Kamerling, J.F. Vliegthart, Primary structures of the N-linked carbohydrate chains from honeybee venom phospholipase A2, *Eur. J. Biochem.*, 213 (1993) 1193-1204.
- [80] J.A. Voynow, T.F. Scanlin, M.C. Glick, A quantitative method for GDP-Fuc:N-acetyl-beta-glucosaminide alpha1,6fucosyltransferase activity with lectin affinity chromatography, *Anal. Biochem.*, 168 (1988) 367-373.
- [81] A. Roitinger, H. Leiter, E. Staudacher, F. Altmann, HPLC method for the determination of Fuc to Asn-linked GlcNAc fucosyltransferases, *Glycoconj. J.*, 15 (1998) 89-91.
- [82] N. Uozumi, T. Teshima, T. Yamamoto, A. Nishikawa, Y.E. Gao, E. Miyoshi, C.X. Gao, K. Noda, K.N. Islam, Y. Ihara, S. Fujii, T. Shiba, N. Taniguchi, A fluorescent assay method for GDP-L-Fuc:N-acetyl-beta-D-glucosaminide alpha 1-6fucosyltransferase activity, involving high performance liquid chromatography, *J. Biochem.*, 120 (1996) 385-392.
- [83] T. Tamura, M.S. Wadhwa, K.G. Rice, Reducing-end modification of N-linked oligosaccharides with tyrosine, *Anal. Biochem.*, 216 (1994) 335-344.
- [84] S. Yazawa, N. Kochibe, T. Nishimura, C. Shima, I. Takai, M. Adachi, T. Asao, T. Hada, Y. Enoki, L.R. Juneja, A novel method for determination of alpha1,6fucosyltransferase activity using a reducing oligosaccharide from egg yolk as a specific acceptor, *Glycoconj. J.*, 15 (1998) 863-871.
- [85] J.F.G. Vliegthart, H. Vanhalbeek, L. Dorland, The applicability of 500-MHz high-resolution <sup>1</sup>H NMR spectroscopy for the structure determination of carbohydrates derived from glycoproteins, *Pure Appl. Chem.*, 53 (1981) 45-77.
- [86] T. Turbadar, Complete Absorption of Light by Thin Metal Films, *P Phys Soc Lond*, 73 (1959) 40-44.

## REFERENCES

---

- [87] R.G. Duggleby, Analysis of enzyme progress curves by nonlinear regression, *Methods Enzymol.*, 249 (1995) 61-90.
- [88] R.G. Duggleby, R.B. Clarke, Experimental designs for estimating the parameters of the Michaelis-Menten equation from progress curves of enzyme-catalyzed reactions, *Biochim. Biophys. Acta*, 1080 (1991) 231-236.
- [89] B.A. Orsi, K.F. Tipton, L.P. Daniel, [8] Kinetic analysis of progress curves, in: *Methods Enzymol.*, vol. Volume 63, Academic Press, 1979, pp. 159-183.
- [90] C.T. Goudar, S.K. Harris, M.J. McInerney, J.M. Suflita, Progress curve analysis for enzyme and microbial kinetic reactions using explicit solutions based on the Lambert W function, *J. Microbiol. Methods*, 59 (2004) 317-326.
- [91] M. Helfgott, E. Seier, Some mathematical and statistical aspects of enzyme kinetics, in: *J. Online Math. Appl.*, vol. 7, The Mathematical Association of America, 2007, pp. 1-34.
- [92] B. Meyer, T. Peters, NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors, *Angew. Chem. Int. Ed.*, 42 (2003) 864-890.
- [93] M. Mayer, B. Meyer, Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy, *Angew. Chem. Int. Ed.*, 38 (1999) 1784-1788.
- [94] M. Mayer, B. Meyer, Group Epitope Mapping by Saturation Transfer Difference NMR To Identify Segments of a Ligand in Direct Contact with a Protein Receptor, *J. Am. Chem. Soc.*, 123 (2001) 6108-6117.
- [95] Y. Cheng, W.H. Prusoff, Relationship between the inhibition constant ( $K_1$ ) and the concentration of inhibitor which causes 50 per cent inhibition ( $I_{50}$ ) of an enzymatic reaction, *Biochem. Pharmacol.*, 22 (1973) 3099-3108.
- [96] A.R. Leach, *Molecular Modelling: Principles and Applications*, Prentice Hall, Harlow 2007.
- [97] D.C. Rapaport, *The Art of Molecular Dynamics Simulation*, Cambridge Univ. Press, Cambridge, 2004.
- [98] D.B. Kitchen, H. Decornez, J.R. Furr, J. Bajorath, Docking and scoring in virtual screening for drug discovery: methods and applications, *Nat. Rev. Drug Discov.*, 3 (2004) 935-949.
- [99] K.G. Rice, M.L. Corradi Da Silva, Preparative purification of tyrosinamide N-linked oligosaccharides, *J. Chrom. A*, 720 (1996) 235-249.
- [100] M. Pabst, J.S. Bondili, J. Stadlmann, L. Mach, F. Altmann, Mass + retention Time = structure: A strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to fibrin N-glycans, *Anal. Chem.*, 79 (2007) 5051-5057.
- [101] K. Koizumi, High-performance liquid chromatographic separation of carbohydrates on graphitized carbon columns, *J. Chrom. A*, 720 (1996) 119-126.
- [102] I.D. Manger, T.W. Rademacher, R.A. Dwek, 1-N-glycyl  $\beta$ -oligosaccharide derivatives as stable intermediates for the formation of glycoconjugate probes, *Biochemistry*, 31 (1992) 10724-10732.
- [103] A. Blume, J. Angulo, T. Biet, H. Peters, A.J. Benie, M. Palcic, T. Peters, Fragment-based screening of the donor substrate specificity of human blood group B galactosyltransferase using saturation transfer difference NMR, *J. Biol. Chem.*, 281 (2006) 32728-32740.
- [104] T. de Vries, R.M.A. Knegt, E.H. Holmes, B.A. Macher, Fucosyltransferases: structure/function studies, *Glycobiology*, 11 (2001) 119R-128.

## REFERENCES

---

- [105] F. Exnowitz, B. Meyer, T. Hackl, NMR for direct determination of  $K_m$  and  $V_{max}$  of enzyme reactions based on the Lambert W function-analysis of progress curves, *Biochim. Biophys. Acta*, 1824 (2012) 443-449.
- [106] M.L. Sinnott, Catalytic mechanism of enzymic glycosyl transfer, *Chem. Rev.*, 90 (1990) 1171-1202.
- [107] H. Bisswanger, Multiple Equilibria, in: *Enzyme Kinetics*, Wiley-VCH Verlag GmbH & Co. KGaA, 2008, pp. 7-58.
- [108] M.P. Kötzler, S. Blank, F.I. Bantleon, E. Spillner, B. Meyer, Donor substrate binding and enzymatic mechanism of human core  $\alpha$ 1,6-fucosyltransferase (FUT8), *Biochim. Biophys. Acta*, (2012) 1915–1925.
- [109] S. Kemper, M.K. Patel, J.C. Errey, B.G. Davis, J.A. Jones, T.D. Claridge, Group epitope mapping considering relaxation of the ligand (GEM-CRL): including longitudinal relaxation rates in the analysis of saturation transfer difference (STD) experiments, *J. Magn. Res.*, 203 (2010) 1-10.
- [110] T. Biet, T. Peters, Molecular recognition of UDP-Gal by beta-1,4-galactosyltransferase T1, *Angew. Chem. Int. Ed.*, 40 (2001) 4189-4192.
- [111] M.A. Macnaughtan, M. Kamar, G. Alvarez-Manilla, A. Venot, J. Glushka, J.M. Pierce, J.H. Prestegard, NMR Structural Characterization of Substrates Bound to N-Acetylglucosaminyltransferase V, *J. Mol. Biol.*, 366 (2007) 1266-1281.
- [112] S. Liu, L. Meng, K.W. Moremen, J.H. Prestegard, Nuclear Magnetic Resonance Structural Characterization of Substrates Bound to the  $\alpha$ 2,6-Sialyltransferase, ST6Gal-I, *Biochemistry*, 48 (2009) 11211-11219.
- [113] P. Petrova, C. Monteiro, C.H. de Penhoat, J. Koca, A. Imberty, Conformational behavior of nucleotide-sugar in solution: Molecular dynamics and NMR study of solvated uridine diphosphate-glucose in the presence of monovalent cations, *Biopolymers*, 58 (2001) 617-635.
- [114] L.L. Lairson, B. Henrissat, G.J. Davies, S.G. Withers, Glycosyltransferases: structures, functions, and mechanisms, *Ann. Rev. Biochem.*, 77 (2008) 521-555.
- [115] D.H. Williams, E. Stephens, D.P. O'Brien, M. Zhou, Understanding Noncovalent Interactions: Ligand Binding Energy and Catalytic Efficiency from Ligand-Induced Reductions in Motion within Receptors and Enzymes, *Angew. Chem. Int. Ed.*, 43 (2004) 6596-6616.
- [116] H. Paulsen, T. Peters, V. Sinnwell, R. Lebuhn, B. Meyer, Conformational analysis, XXIV. Determination of the conformations of tri- and tetrasaccharide sequences of N-glycoproteins. The problem of the (1  $\rightarrow$  6)-glycosidic bond, *Liebigs Ann. Chem.*, 1984 (1984) 951-976.
- [117] C. Breton, L. Šnajdrová, C. Jeanneau, J. Koča, A. Imberty, Structures and mechanisms of glycosyltransferases, *Glycobiology*, 16 (2006) 29R-37R.
- [118] B.W. Murray, V. Wittmann, M.D. Burkart, S.C. Hung, C.H. Wong, Mechanism of human  $\alpha$ 1,3-fucosyltransferase V: glycosidic cleavage occurs prior to nucleophilic attack, *Biochemistry*, 36 (1997) 823-831.
- [119] H.A. Staab, M. Lüking, F.H. Dürr, Darstellung von Imidazoliden. Synthese von Amidinen, Hydraziden und Hydroxamsäuren nach der Imidazolidmethode, *Chem. Ber.*, 95 (1962) 1275-1283.
- [120] H.A. Staab, Neuere Methoden der präparativen organischen Chemie IV Synthesen mit heterocyclischen Amidinen (Azoliden), *Angew. Chem.*, 74 (1962) 407-423.
- [121] F.M. Ausubel, in: *Current Protocols in Molecular Biology*, Wiley Interscience, 1996.

## REFERENCES

---

- [122] B. Bendiak, J. Orr, I. Brockhausen, G. Vella, C. Phoebe, Separation of neutral reducing oligosaccharides derived from glycoproteins by HPLC on a hydroxylated polymeric support, *Anal Biochem*, 175 (1988) 96-105.
- [123] A.A. Bergwerff, C.J. Stroop, B. Murray, A.P. Holtorf, G. Pluschke, J. Van Oostrum, J.P. Kamerling, J.F. Vliegthart, Variation in N-linked carbohydrate chains in different batches of two chimeric monoclonal IgG1 antibodies produced by different murine SP2/0 transfectoma cell subclones, *Glycoconj J*, 12 (1995) 318-330.
- [124] E. Staudacher, F. Altmann, L. März, K. Hård, J. Kamerling, J. Vliegthart,  $\alpha$ 1-6( $\alpha$ 1-3)-Difucosylation of the asparagine-bound N-acetylglucosamine in honeybee venom phospholipase A2, *Glycoconj. J.*, 9 (1992) 82-85.
- [125] K. Koles, P.H.C. van Berkel, F.R. Pieper, J.H. Nuijens, M.L.M. Mannesse, J.F.G. Vliegthart, J.P. Kamerling, N- and O-glycans of recombinant human C1 inhibitor expressed in the milk of transgenic rabbits, *Glycobiology*, 14 (2004) 51-64.
- [126] S.T. Cohen-Anisfeld, P.T. Lansbury, A practical, convergent method for glycopeptide synthesis, *J. Am. Chem. Soc.*, 115 (1993) 10531-10537.
- [127] K.J. Bowers, E. Chow, X. Huafeng, R.O. Dror, M.P. Eastwood, B.A. Gregersen, J.L. Klepeis, I. Kolossvary, M.A. Moraes, F.D. Sacerdoti, J.K. Salmon, S. Yibing, D.E. Shaw, Scalable algorithms for molecular dynamics simulations on commodity clusters, in: *SC 2006 Conference, Proceedings of the ACM/IEEE*, 2006, pp. 43-43.
- [128] W.L. Jorgensen, J. Tirado-Rives, The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin, *J. Am. Chem. Soc.*, 110 (1988) 1657-1666.
- [129] H.J.C. Berendsen, W.F. Postma, W.F. van Gunsteren, J. Hermans, *Intermolecular forces*, Reidel, Dordrecht, 1981.
- [130] T. Darden, D. York, L. Pedersen, Particle Mesh Ewald - an N.Log(N) method for Ewald sums in large systems, *J. Chem. Phys.*, 98 (1993) 10089-10092.
- [131] D.J. Evans, B.L. Holian, The Nose-Hoover thermostat, *J. Chem. Phys.*, 83 (1985) 4069-4074.
- [132] G.J. Martyna, D.J. Tobias, M.L. Klein, Constant-pressure molecular-dynamics algorithms, *J. Chem. Phys.*, 101 (1994) 4177-4189.

## Acknowledgements

### *Mein Dank gilt...*

Meinen Eltern für ihre liebevolle Unterstützung in allen Lebenslagen, für ihr Interesse an meiner Arbeit und hilfreiche Ratschläge.

Meinen beiden Schwestern, meinen Freunden und allen, die mir nahe stehen oder nahe standen für das Leben außerhalb des Labors, das für das Gelingen dieser Arbeit so wichtig war.

Der Arbeitsgruppe Meyer

Meike, für alles, was wir über NMR gelernt haben (egal ob diskutierend oder schraubend) und für deinen Sherlock-Holmes-Blick auf das eine oder andere meiner Spektren.

Henning, für deine Expertise bei der Glycananalytik mit HPLC und MS und für deine mitreißende Art, die Dinge anzupacken.

Katrin, für deine Tipps Rund um Ionentauscher, die Merck-HPLC und Synthesefragen.

Bea, für deine Hilfe beim Aufsetzen der ersten MD-Simulationen.

Moritz, für deinen aufopfernden Einsatz an den HPLCs und für Tipps bei meinen Trennproblemen.

Kathrin, für dein Wissen rund um alles, was mit Chito... anfängt und für die Motivation zu Aufräumaktionen im Labor.

Martin, für deine spontane Hilfe bei Modellingproblemen.

Anna, für deine Einweisung am Biacore und Patrizia, für deine Hilfe bei Problemen am Biacore.

Alex und Karsten, für die prompte Installation aller möglichen dringend benötigten Programme auf meinem Laborrechner.

Allen anderen Mitgliedern für Hilfsbereitschaft bei den alltäglichen Laborproblemchen und die gute Stimmung innerhalb und außerhalb der Arbeitszeiten.

## ACKNOWLEDGEMENTS

---

Den Arbeitsgruppen von Prof. Dr. Bredehorst und PD Dr. Spillner, insbesondere Frank und Simon, für die großartige Zusammenarbeit und für das Anlernen in Proteinexpression und –reinigung.

Dr. Thomas Hackl, für Beantwortung von Fragen bei der Progresskurvenanalytik und Diskussionen um Enzymkinetik.

Dr. Maria Trusch, für die Aufnahme von MS-Spektren von FUT8 und der MS-Serviceabteilung für die vielen gemessenen Glycanfraktionen meiner Chromatographieläufe.

Meinem Bachelorstudenten Dirk sowie meinen Praktikanten Oliver, Katja, Dénes, Sven und Matthias, für eure Mitarbeit im Labor.

Der *Hamburg School of Structure and Dynamics in Infection (SDI)* für die Finanzierung dieser Arbeit und die Möglichkeiten, über mein Arbeitsgebiet hinaus zu lernen. Meinen *SDI-fellows* für die vielen bereichernden Diskussionen, insbesondere Julia für die Hilfe mit PyMol und das Entdecken des *cePOFUT*-Papers, Sophie für das Korrekturlesen der Einleitung und Raphael für Hilfe bei der Proteinreinigung.

**Curriculum vitae**

*(Not applicable for reasons of protection of personal data)*

## CURRICULUM VITAE

---

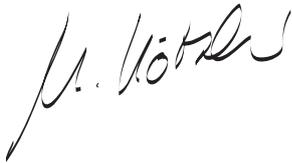
*(Not applicable for reasons of protection of personal data)*

## **Affidavits**

### **Erklärung**

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbstständig angefertigt und keine anderen als die von mir angegebenen Hilfsmittel und Quellen verwendet habe. Ich versichere weiterhin, dass die vorliegende Dissertation weder in gleicher noch in veränderter Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Hamburg, den 15.10.2012

A handwritten signature in black ink, appearing to read 'M. Kötzler', written in a cursive style.

Miriam Kötzler