Recombinant glycosyltransferases:  
Expression, production and use for the synthesis 
of O-linked selectin ligands

A dissertation 
submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY

ZÜRICH

For the degree of 
Doctor of Technical Sciences

Presented by:

Steffen Zeng 

Dipl. Ing. 
Studienrichtung Biotechnologie 
Technical University of Berlin 

born January 13, 1970 
citizen of Germany 

Accepted on the recommendation of:

James E. Bailey, examiner 
Eric G. Berger, co-examiner 

1999
Table of contents

1. Zusammenfassung 6

2. Summary 8

3. Introduction 10
   3.1 Glycosylation 10
   3.2 Glycosyltransferases 11
   3.3 Glycoproteins 14
      3.3.1 N-linked glycosylation 14
      3.3.2 O-linked glycosylation 17
   3.4 Inflammation 21
   3.5 Lectins 24
      3.5.1 Selectin structure 25
         3.5.1.1 E-Selectin 26
         3.5.1.2 L-Selectin 27
         3.5.1.3 P-Selectin 28
      3.5.2 Selectin ligands 29
         3.5.2.1 E-Selectin ligand 1 29
         3.5.2.2 CD34 29
         3.5.2.3 GlyCAM-1 30
         3.5.2.4 P-Selectin glycoprotein ligand 1 30
   3.6 Medical targets 33
   3.7 Antiadhesive compounds 34

4. Aim of the thesis 37

5. Results 38

6. Discussion and Outlook 43
7. Manuscript I: 48
   Pilot scale expression and purification of soluble Protein A tagged beta
   1,6N-acetylglucosaminyltransferase in CHO cells
   [Biochim Biophys Res Commun 237: 653-8.]

8. Manuscript II: 63
   Improved product formation in high density Chinese hamster ovary cell
   cultures transfected at confluency
   [Biotechnol Bioeng 1999, in revision]

9. Manuscript III: 83
   Complete enzymic synthesis of the mucin-type sialyl Lewis X epitope,
   involved in the interaction between PSGL-1 and P-selectin
   [manuscript in preparation]

10. References 107

11. Abbreviations 131

12. Curriculum vitae 133
To Gudrun and Jan Lukas
Acknowledgments

The idea on which my thesis is based was suggested by Professor Dr. Eric G. Berger who supervised my work in his group. I like to thank him for giving me the opportunity to work on an interesting project, for introducing me to the field of glycobiology as well as for the numerous discussions and hints.

I am honestly thankful to Professor James E. Bailey who was my supervisor at the ETH. My colleague Andre Dinter, with whom I really enjoyed working, I thank for explaining all the little miracles of molecular biology to me and especially for the stimulating discussions and his permanent interest.

To Martine Malissard and Lubor Borsig I am grateful for their contributions to my work as well as for their interesting scientific discussions and hints.

For their helpful hands in the lab I like to thank Marianne Farah, Claudia Ruedin and Bea Berger, who additionally to that contributed a lot in making my stay so pleasant.

Peter Burger, Timo Imbach, Maud Ramuz, Sylvia Schnyder, Andreas Stutz and Dapeng Zhou as well as all the other co-workers in the Institute of Physiology with whom I did not work personally, I am grateful for their individual contribution to the friendly atmosphere.

I also greatly appreciated the fruitful collaboration with Detlef Eisenkrätzer, Gregor Dudziak, Manfred Biselli and Christian Wandrey from the Research center Jülich. Hans Kamerling, Ricardo Gutierrez-Gallego from the Bijvoet Center of the Utrecht University I thank for doing all the tedious analysis and making our collaboration a real exchange of competence.

To Adriana Rosa Guerini I am grateful for being my point of contact on the Institute of Biotechnology at the ETH Hoenggerberg during all these three years.

I also like to thank Frank Bootz from the Institute of Labortierkunde of the University of Zürich for providing us with fresh rat liver as well as Markus Streiff from Novartis Pharma for supplying us with some sialyltransferase.

PD Dr. Kaltner from the Ludwig-Maximilians University of Munich I thank for the friendly permission to reprint two of his illustrations.

Dr. Thierry Hennet I thank for the fruitful scientific discussions and the receptivity for contributions from people with another educational background.
Parts of this thesis have been published in:


1. Zusammenfassung

Bei der Entzündungsreaktion, in deren Verlauf Leukocyten den Blutstrom verlassen und in das Gewebe einwandern, wird der erste Kontakt zwischen Endothel und Leukocyt über das spezifische Tetrasaccharid sialyl Lewis X (Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc) gebildet. Die vorliegende Dissertation beschreibt die Präparation von fünf verschiedenen Glykosyltransferasen sowie deren Verwendung für die erste voll enzymatische Synthese von O-glykosidisch gebundenem sialyl Lewis X.

In einem ersten Schritt wurde β3-Galaktosyltransferase aus Ratten Leber extrahiert und für die Verlängerung des N-Acetylglukosaminsamin verwendet. Die so erhaltene Galaktose β1-3 N-Acetylglukosamin (core 1) Struktur diente als Substrat für die Core 2 β6-GlcNAcT.

In einem zweiten Schritt wurde eine "Chinese Hamster Ovary" (CHO) Zelllinie etabliert, die Core 2 β6-N-Acetylglukosaminsaminyltransferase (β6-GlcNAcT) stabil exprimierte. Nach der Fermentation wurde das Enzym aus dem Überstand über eine Affinitätschromatographie aufgereinigt.

Während des dritten Schrittes wurde das eingebaute N-Acetylglucosamin mit der β4-Galaktosyltransferase, die über Affinitätschromatographie aus menschlicher Milch gewonnen wurde, verlängert.

Mit rekombinanter Ratten α3-Sialyltransferase, welche in Sf9 Zellen hergestellt und aufgereinigt wurde, konnte das erhaltene Tetra- zum Pentasaccharid umgewandelt werden.

Schließlich wurde lösliche humane α3-Fukosyltransferase 6 in der Hefe Pichia pastoris exprimiert und aus dem Überstand mit einer Affinitätschromatographie gereinigt. Mit dem so hergestellten rekombinanten Enzym wurde der letzte Syntheseschritt katalysiert. Das Endprodukt wurde über Massenspektrometrie und 1H Magnetische Kernresonanz-Spektroskopie als Core 2 basiertes sialyl Lewis X identifiziert. Der Gesamtertrag von 32 % über alle Syntheseschritte zeigt, dass die hochspezifischen Glykosyltransferasen in der Tat für die Synthese von komplexen Glykanen geeignet sind.
2. Summary

In the inflammatory process, during which leukocytes leave the blood stream and migrate into the tissue, the initial contact between the endothelia and the leukocyte is established via the specific tetrasaccharide \textit{sialyl Lewis X} (Neu5Ac(\alpha2-3)Gal(\beta1-4)[Fuc(\alpha1-3)]GlcNAc). This thesis describes the preparation of five different glycosyltransferases as well as their use for the first full enzymic synthesis of the O-glycosidically linked \textit{sialyl Lewis X}.

In a first step \(\beta3\)-galactosyltransferase was extracted from rat liver and used for the elongation of \(\beta4\)-acetylgalactosamine. The product galactose \(\beta1-3\) \(\beta4\)-acetylgalactosamine (core 1) structure could afterwards be branched with the Core 2 \(\beta6\)-GlcNAcT.

In a second step a Chinese Hamster Ovary (CHO) cell line was established that stably expressed Core 2 \(\beta6\)-N-acetylgalactosaminyltransferase (\(\beta6\)-GlcNAcT). After fermentation the enzyme was purified from the supernatant by affinity chromatography.

During the third step the incorporated \(\beta6\)-acetylgalactosamine was elongated with \(\beta4\)-galactosyltransferase, obtained through affinity chromatography from human milk.

Using recombinant rat \(\alpha3\)-sialyltransferase, which was produced with Sf9 cells and purified, the obtained tetrasaccharide could be converted to the pentasaccharide.

Finally soluble human \(\alpha3\)-fucosyltransferase 6 was expressed in the yeast \textit{Pichia pastoris} and purified from the supernatant by affinity chromatography. This recombinant enzyme catalyzed the last step of the synthesis. The final product was identified by Mass Spectrometry and \(^1\text{H}\) Nuclear Magnetic Resonance as Core 2 based \textit{sialyl Lewis X}. The overall yield of 32 \% obtained for the whole synthesis demonstrates that the highly specific glycosyltransferases are suited for the synthesis of complex glycans.

As a side project we showed that the productivity of our Core 2 \(\beta6\)-GlcNAcT CHO clone continuously decreased during cultivation in the fluidized bed reactor after a steady state cell concentration was reached. Cells taken from the reactor after 400 hours reached again maximum productivity in a batch fermentation. The analysis of the cell cycle distribution revealed that the cells conventionally transfected after 24 hours formed 5 times more product in S- than in G0/G1-phase under synchronized conditions.
Evaluating the transfection procedure revealed that at this time point the majority of the cells are in the S-phase. Transfection in S-phase thus led to high productivity in S-phase. Thereafter, cells were transfected, when the majority of the cells were in G0/G1-phase, by shifting the time point of transfection. Through this modification a changed expression pattern could be achieved for two independent reporter proteins. Transfection in G0/G1-phase, therefore, led to significantly higher production in G0/G1-phase. These results indicate that for an optimal production the interplay of transfection and expression with regard to the choice of the cultivation system should not be neglected.
Also the longest journey starts with a first step.

(Marco Polo)

3. Introduction

3.1 Glycosylation

The attachment of sugar molecules to proteins and lipids is a frequent event in the living cell. Over the past decades the field of Glycobiology has emerged, reflected by an ever increasing wealth of data. In many cellular activities complex carbohydrates are involved, as for example topogenesis, signaling, cell-cell interaction, cell growth, differentiation and apoptosis. These chapters cover important biological functions and explain the rising interest in the field of Glycobiology.

Among posttranslational modifications, glycosylation which is defined as enzyme catalyzed transfer of sugars, is unique in its complexity. Five different parameters affect the structural diversity of the glycans. First, composition of unit monosaccharides; to date nine different (Gal, GalNAc, Glc, GlcA, GlcNAc, Fuc, Man, Neu5Ac and Xyl) main carbohydrates have been found in glycans of higher eukaryotes. Second, the length of the glycan chain; chain length can vary between one to over several hundred monosaccharide units. Third, the linkage type of glycosidic bond; carbohydrates can be connected to five different hydroxyl groups of the acceptor pyranose sugar by condensation. Fourth, anomic configuration; the corresponding groups can occur in the two different α and β anomic forms. Fifth, the branching; up to four carbohydrates can be connected to one monosaccharide. In addition to this main variability, glycans can further be diversified by the attachment of small substituents such as phosphate, acetate or sulfate (Varki, 1996; Brockhausen and Kuhns, 1997). However, it is essential to add that the theoretical number of combinations has by far not been unearthed in
nature. This is probably due to the fact that the corresponding high number of glycosyltransferases, which are catalyzing the formation of glycans, is not available in the cell (Whitfield and Douglas, 1996). The actual diversity of heteroglycans is therefore limited by the expressed highly stereo- and regiospecific glycosyltransferases, which are located in the endoplasmic reticulum (ER) and the Golgi apparatus (GA).

Many of the cellular biomolecules can be decorated with glycans. These so called glycoconjugates can further be distinguished in glycoproteins, glycolipids and proteoglycans. The glycolipids are important constituents of biomembranes and can carry rarely more than a dozen monosaccharides long carbohydrates. By presenting their glycan to the outer environment they thus constitute the gly cocalyx and play an important role in the nervous system (for pathway and review see Reuter and Gabius, 1999). In glycoproteins the glycan chains are covalently linked to suitable functional groups of amino acid side chains within a protein or peptide. Depending on the chemical function on which the glycans are transferred three main types of glycosylation can be distinguished:

1. **O-glycosylation** where the carbohydrates are bound to the hydroxyl group of Ser/Thr residues, when the protein / peptide is transported from the *cis* Golgi to the *trans* Golgi network. Usually the synthesis of O-glycans starts with a GalNAc; however, less commonly GlcNAc (Hart et al., 1996), Fuc (Moloney et al., 1997), Xyl (Schwartz, 1995), Gal (Clark CC and Kefalides, 1976) and Man (Krusius et al., 1987) have been found directly attached to the protein in higher eukaryotes.

2. **N-glycosylation**, where the starting GlcNAc sugar is linked to the carboxy amide group of an asparagine and which occurs in both endoplasmic reticulum and Golgi apparatus.

3. **C-glycosylation**, where a mannose is attached to the carbon 2 of tryptophan, discovered on RNase 2 (Loffler et al., 1996).

### 3.2 Glycosyltransferases

The glycosyltransferases catalyze the transfer of glycosyl residues from nucleotide-activated sugars to other carbohydrates or to aglycons according to the following general equation:
Glycosyltransferases are grouped according to the donor substrates. Further classification can include the used acceptor substrate, the type of linkage, the sequence of their action and their first description. The high donor and acceptor specificity of this proteins has led to the so called "central dogma of glycobiology" (Kleene and Berger, 1993) that for each carbohydrate linkage a specific glycosyltransferase gene has to be provided. This general rule has been proposed by Hagopian and Eylar (1968) and is known as the "one enzyme-one linkage" concept. Only a few exceptions to this rule have been described. For example β4-GalT specificity can be switched from the acceptor GlcNAc to Glc in the presence of α-lactalbumin; the Lewis type FucTIII is able to catalyze two different linkages, e.g. Gal β1-3/4 [Fuc α1-3/4] GlcNAc. Interestingly, several redundant genes are present in the cell which code for different glycosyltransferases, all forming the same linkage (e.g. α3-FucT, pp-GalNAcT, β4-GalT). The biological relevance of these families are not fully understood yet; it has been demonstrated that individual members of a family can exhibit distinct substrate specificities and are regulated differently. The expression of these enzymes may therefore be regulated in a tissue and developmental stage specific manner, thereby explaining the observed differences in glycan structure of glycoproteins synthesized in different cell types and at different stages of differentiation and development. Besides, the observation that a knock-out of one polypeptide-GalNAc-transferase produced no significantly changed phenotype (Hennet et al., 1995), suggests that the family members can also function as back-ups. Within these families of enzymes forming the same linkage and using the same acceptor a certain homology has been observed which was the starting point for the successful cloning of other family members by BLAST search on EST database. Some of these motifs, as for example the L and the S-sialylmotif of the sialyltransferases, have been shown to be involved in the donor and acceptor binding (Wen et al., 1992; Datta and Paulson, 1997).

At least two main groups of glycosyltransferases can be distinguished: First, the endoplasmic reticulum enzymes, which contain several possible transmembrane domains and are involved in the early formation of N-glycans. Second, the type II Golgi membrane enzymes which are responsible for O-glycan synthesis and N-glycan
elongation and termination. This group shares a common domain structure but has no overall sequence homology. The domain structure of glycosyltransferases consists of a short amino-terminal cytoplasmic tail, a membrane anchored region, an extended stem region and a large carboxy-terminal catalytic domain (Paulson and Colley, 1989) which is oriented to the lumen of the Golgi apparatus (Figure 1). Interestingly, most glycosyltransferases are themselves glycosylated and in some cases, this seems of critical importance for activity (Nagai et al., 1997; Toki et al., 1996; Toki et al., 1997).

![Figure 1: Schematic representation of the domain structure of Golgi glycosyltransferases](image)

The cytoplasmic portion, the stem region and the transmembrane domain are thought to be involved in Golgi localization. Two models have been formulated, the bilayer thickness model (Bretscher and Munro, 1993), in which the length of the transmembrane domain is believed to be responsible for the site of retention and the kin recognition / oligomerization model (Nilsson et al., 1993). The latter explains the requirement for non-transmembrane sequences or multiple domains in the Golgi retention process. Collecting the published data, Colley (1997) concluded, that neither model alone may explain all findings and it appears that two independent additive mechanisms would be more appropriate. Exchanging the transmembrane domain from a late (trans Golgi) to an early (cis Golgi) acting glycosyltransferase and vice versa, was shown to remarkably influence the glycosylation pattern (Skrincosky et al., 1997).
this aspect, engineering of the transmembrane domain and the domains involved in the
kin recognition may be an interesting approach to modify the glycosylation pattern. In
addition, replacing cytoplasmic tail and transmembrane region by a cleavable signal
sequence resulted in efficient transport of the otherwise membrane bound
glycosyltransferase into the extracellular space as an active soluble enzyme (Weinstein
et al., 1987; US patent to Paulson et al., 1991; US patent to Paulson et al., 1996; US
patent to Paulson et al., 1998). This technique has been used extensively in the past
years to efficiently produce soluble glycosyltransferases as enzymic catalysts (Gallet et
al., 1998; Borsig et al., 1997).

However, soluble glycosyltransferases have also been reported to naturally occur
in different body fluids (as for example β4-GalT1 in human milk). Some of these
enzymes derive from proteolytic release from the membrane-bound forms (Lammers
and Jamieson, 1989).

3.3 Glycoproteins

Besides their role in cell-cell interaction and signaling, glycans can have a
remarkable influence on proteins, by mediating protein half-life \textit{in vivo}, by stabilizing
the protein structure, by shielding the protein from proteases (Baynes and Wold, 1976;
US patent to Bergh et al., 1990), by orienting the protein on the cell surface, and by
assisting protein folding (Hebert et al., 1997). In addition glycosylation can also
participate in protein regulation by competing, for example, with phosphorylation (Hart
et al., 1996). These important functions are mediated by N- and O-linked glycans.

3.3.1 N-linked glycosylation

Many membrane-bound and soluble glycoproteins contain Asn-linked
oligosaccharides often in addition to O-glycans. It has been recognized that for N-
glycosidic linkage a peptide sequence of Asn-X-Ser/Thr-Y is required. In this the
variable amino acid X is important for the efficiency of recognition whereas the amino
acid Y has a notable impact on the glycosylation efficiency. In some proteins the
unusual sequon Asn-X-Cys has also been identified as being the starting point of N-
glycosylation (Vance et al., 1997).
In contrast to O-glycans the initial step of N-glycan synthesis starts in the endoplasmic reticulum, where N-glycans are pre-assembled on a lipid-based anchor, the dolicholphosphate. First, on the cytoplasmic side of the endoplasmic membrane, two GlcNAc and five mannose residues are added in a well defined manner. The resulting structure is translocated across the ER membrane to the luminal side. Here four mannose and three glucose residues are added as shown in Figure 2A. The resulting dolichol linked Glc\(_5\)Man\(_6\)GlcNAc\(_2\) structure is later transferred en bloc by the oligosaccharyltransferase to the appropriate Asn of the polypeptide on the luminal side. The three glucose residues and one mannose from the Man\(_\alpha1\)-6 arm are subsequently cleaved by the action of two distinct glucosidases and a \(\alpha\)-mannosidase within the ER. Three additional mannose residues are removed by mannosidase I which is localized in the cis-Golgi or in pre-Golgi transitional compartments. The resulting Man\(_5\)GlcNAc\(_2\) structure has first to be elongated by the branching GlcNAcT I, which initiates the synthesis of antennae of complex or hybrid N-glycans (Figure 2B). Hybrid chains have Man residues on the Man\(_\alpha1\)-6 arm as well as GlcNAc and complex-type structure on the Man\(_\alpha1\)-3 arm. By the action of GnT III a bisecting GlcNAc can be introduced, which blocks the action of Golgi mannosidase II. This mannosidase, which is also inhibited by swainsonine, releases two additional mannose units leading to the formation of complex glycans. By the action of GnT III the bisected hybrid chains are therefore committed to the hybrid structure.

Subsequently other antennae are added by GnT IV and GnT V. The antennae may grow by the attachment of linear or branched poly \(N\)-acetyllactosamine units and terminal structures (as \(\alpha1,6\)Fuc or \(\alpha2,3\) or \(\alpha2,6\)Neu5Ac). Many of these reactions are similar to those involved in the extension of O-glycans.

Besides outer chain elongation, also the intermediate core residues are a subject of change. Namely L-Fucose can be attached to the GlcNAc connected to Asn. In mammals it is bound to the 6' position whereas in plants the 3' position is used. Interestingly, in insect cells 6' and 3' bifucosylated GlcNAc could be found. The antigenicity displayed by the \(\alpha1,3\) linked Fuc is one reason for concern to employ insect or plant cells for the production of therapeutic glycoproteins.
Figure 2: Pathway of N-glycosylation
3.3.2 O-linked glycosylation

O-glycans, which are, by contrast to N-glycans, more compact structures, occur on soluble secreted and membrane-bound glycoproteins. Mucins as well as mucin-like glycoproteins are rich in O-glycan chains and may contain 50-80% of carbohydrates.

In contrast to the well defined sequon for the attachment of N-glycans, which is recognized by a single enzyme complex, i.e. the oligosaccharyltransferase, more than 10 different UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferases (pp-GalNAcT) acting on the peptide backbone have been identified (Hagen et al., 1993; Homa et al., 1993; White et al., 1995; Bennet et al., 1996; Clausen and Bennett, 1996; Hagen et al., 1997; Bennet et al., 1998; Ten Hagen et al., 1998; Hagen and Nehrke, 1998). Thus, no specific peptide sequence requirements have been described so far for classical mucin-type O-glycosylation. Only a few algorithms identifying potential O-glycosylation are available which do not specify the pp-GalNAcT involved in a particular O-glycosylation reaction. The Ser/Thr residues are preferentially surrounded by Ser, Thr, Pro, Ala and Gly. Additionally, studies with purified enzyme from natural sources revealed that the start of O-glycosylation is a selective process dependent on both the amino acid sequence and also the prior glycosylation of adjacent sites of the peptide backbone (Brockhausen et al., 1996). Supporting these findings the cloned pp-GalNAcT family members have been demonstrated to act in a different order on adjacent Ser/Thr residues in the MUC1 mucin (Hanisch et al., 1999; Bennet et al., 1998; Wandall et al., 1997).

In contrast to N-glycans O-glycan synthesis starts in the cis Golgi in a stepwise manner directly on the protein. Processing by glycosidases, as found in the synthesis of N-glycans has not been reported to occur in O-glycan synthesis.

The pathways leading to the different core structures are shown in Figure 3A. As the first step GalNAc is linked via reaction a to the corresponding serine/threonine residue. This reaction is catalyzed by a whole family of pp-GalNAcT (Clausen and Bennet, 1996), having distinct acceptor specificity (Wandall et al., 1998). Once the GalNAc is bound it can be sialylated by reaction o formed by the ST6Gal1 (Kurosawa et al., 1994) or further elongated to the different core structures. Eight of these core structures have been found to naturally occur in mammals (Table 1).
Table 1: O-glycan core structures and other common carbohydrates

<table>
<thead>
<tr>
<th>Core</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Gal(β1-3)GalNAc(α1-Ser/Thr)</td>
</tr>
<tr>
<td>2</td>
<td>GlcNAc(β1-6)[Gal β1-3]GalNAc(α1-Ser/Thr)</td>
</tr>
<tr>
<td>3</td>
<td>GlcNAc(β1-3)GalNAc(α1-Ser/Thr)</td>
</tr>
<tr>
<td>4</td>
<td>GlcNAc(β1-6)[GlcNAc(β1-3)]GalNAc(α1-Ser/Thr)</td>
</tr>
<tr>
<td>5</td>
<td>GalNAc(α1-3)GalNAc(α1-Ser/Thr)</td>
</tr>
<tr>
<td>6</td>
<td>GlcNAc(β1-6)GalNAc(α1-Ser/Thr)</td>
</tr>
<tr>
<td>7</td>
<td>GalNAc(α1-6)GalNAc(α1-Ser/Thr)</td>
</tr>
<tr>
<td>8</td>
<td>Gal(α1-3)GalNAc(α1-Ser/Thr)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>sLex</td>
<td>Neu5Ac(α2-3)Gal[β1-4][Fuc(α1-3)]GlcNAc((β1-3)Gal-)</td>
</tr>
<tr>
<td>sLeα</td>
<td>Neu5Ac(α2-3)Gal[β1-3][Fuc(α1-4)]GlcNAc((β1-3)Gal-)</td>
</tr>
<tr>
<td>Tn</td>
<td>GalNAc(α1-Ser/Thr)</td>
</tr>
<tr>
<td>T</td>
<td>Gal(β1-3)GalNAc(α1-Ser/Thr)</td>
</tr>
<tr>
<td>Sialyl TF</td>
<td>Neu5Ac(α2-3)Gal[(β1-3)[Neu5Ac(α2-6)]GalNAc(α1-Ser/Thr)</td>
</tr>
<tr>
<td>type I</td>
<td>Gal(β1-3)GlcNAc</td>
</tr>
<tr>
<td>type II</td>
<td>Gal(β1-4)GlcNAc</td>
</tr>
<tr>
<td>type III</td>
<td>Gal(β1-3)GalNAc</td>
</tr>
</tbody>
</table>

They can further be elongated and branched by a variety of enzymes leading to highly complex glycans. The complexity of O-glycans has impressively been demonstrated by discovering more than hundred different O-glycans in human lung tissue extracts (Klein et al., 1992).

In the context of this thesis, I therefore restrict the description to those O-glycans relevant for selectin ligands. For a recent review of all enzymes involved N- and O-glycan synthesis the reader is referred to Brockhausen et al. (1998).

As shown in Figure 3A, bound GalNAc can be branched with a GlcNAc in 6’ position following the pathway b,c as well as e,f or d. Reactions c and f are catalyzed by two distinct β6-GlcNAcTs, the Core 2 β6-GlcNAcT and the Core 2/4 β6-GlcNAcT. The cDNA of the first enzyme has been cloned in 1992 (Bierhuizen and Fukuda, 1992).
whereas the latter became available only recently (Schwientek et al., 1999). Specificity studies of the enzymes forming reactions c and f have early been investigated by Brockhausen and others using purified preparations from natural sources (Brockhausen et al., 1986; Kuhns et al., 1993). However, due to the naturally occurring core 6 structure (Wu et al., 1984; Hounsell et al., 1985; van Halbeek et al., 1985), it remained unclear whether the substitution in 3’ position formed by β3-GalT (reaction b) or β3-GlcNAcT (reaction e), which have not been cloned yet, is absolutely necessary for the action of β6-GlcNAcTs. At least three pathways leading to this structure are conceivable. First, a separate β6-GlcNAcT, which catalyzes reaction d directly without substitution in 3’position may exist. Second, it was unknown whether one of the characterized β6-GlcNAcTs has at least a side activity on this acceptor. Third, the core 6 structure can also be the result of a β-galactosidase degradation reaction of core 2.

Nevertheless core 2, core 4 and core 6 structure can further be elongated by β3-GalT (Hennet et al., 1998; Amado et al., 1998; Kolbinger et al., 1998) by reaction j (Figure 3B). The resulting structures can subsequently be sialylated (reaction k) by ST3GalIII (Kitagawa and Paulson, 1994; Wen et al., 1992) and fucosylated (reaction l) by α3/4-FucTIII (Kukowska-Latallo et al., 1990). This pathway finally ends up with the formation of the O-linked core 2 based tetrasaccharide sialyl Lewis a, which is composed of Neu5Ac, Gal, GlcNAc and Fuc.

Elongation of core 2, core 4 or core 6 with a family of β4-GalTs (Narimatsu et al., 1986; Almeida et al., 1997) following reaction g constitutes a different pathway. After sialylation with ST3GalIV (Sasaki et al., 1993; Kitagawa and Paulson, 1993) and fucosylation by a family of α3/4-FucT (Kukowska-Latallo et al., 1990; Goelz et al., 1990; Weston et al., 1992; Koszdin and Bowen, 1992; Sasaki et al., 1994; Natsuka et al., 1994) it ends up with the formation of sialyl Lewis X (sLeX). This tetrasaccharide has the same composition as sialyl Lewis a but two different linkages. It is present on different selectin ligands.

Prior to the addition of sialic acid, β3-GlcNAcT (Zhou et al., 1999; Sasaki et al., 1997) can elongate the Gal linked to the branched GlcNAc. Together with the β4-GalTs it can extend the chain to poly-N-acetyllactosaminoglycans, which can again be further fucosylated and sialylated (see pathway on the left side of Figure 3B). Such polylactosamino glycans bearing multiple Fuc and terminal sialyl Lewis X also have
been found to naturally occur on selectin ligands (Wilkins et al., 1996; Aeed et al., 1999).

Figure 3:  A) synthesis of O-glycan core structures

B) elongation pathway of O-glycan core structures to sLe^x and sLe^a
3.4 Inflammation

During inflammation, leukocytes which are normally patrolling the body in search for foreign antigens and invaders, leave the blood stream and migrate through the endothelial cells into the inflamed tissue. The extravasation itself is a multistep process mediated by a variety of cell adhesion molecules. Figure 4 shows the three main steps occurring during inflammation: Tethering and rolling, arrest and firm adhesion and finally migration into the underlying tissue also known as diapedesis.

![Diagram of blood vessel with leukocyte extravasation](image)

**Figure 4: Inflammation - Extravasation of leukocytes**

The inflammatory response is initiated by the action of a variety of biological response modifiers released after injury of the underlying tissue. These molecules, as for example interleukin-8, cause venule dilatation, leading to a decelerated blood flow. Leukocyte extravasation subsequently starts with a reversible transient adhesive contact of leukocytes with the endothelium, which is mediated by three different proteins. They are named P-, E- and L-selectin and belong to the family of lectins, which are by definition carbohydrate binding proteins (see section 3.5). Interactions of endothelial and leukocyte selectins with carbohydrate motifs displayed as oligosaccharide capping groups on membrane constituents result in transient tethering and rolling of leukocytes along the endothelial surface. Rolling of leukocytes ensues from rapid binding and
dissociation of selectin ligands, with selectins. Finally, leukocytes are arrested and adhere firmly to integrins (For review, see Kaltner and Stierstorfer, 1998).

Interestingly, all three of the selectins can mediate rolling on stimulated endothelium; however, it appears that they have to work in concert. The very first contact between leukocyte and endothelia is thought to be mediated by P-selectin, which is pre-made and stored in Weibel-Palade bodies of the endothelial cells. 5-20 minutes following activation, these storage bodies fuse with the cell membrane. As a result, P-selectin is expressed on the endothelia surface as shown in Figure 5 (Stenberg et al., 1985) where it can bind to its ligands. One of them is the P-selectin glycoprotein ligand-1 (PSGL-1) which is expressed on the microvilli of the leukocytes. The rapid transient presentation has led to the suggestion that P-selectin functions mainly during the early inflammatory response (Lawrence and Springer, 1991).

A second selectin, e.g. E-selectin, is synthesized de novo. It becomes maximally expressed on the endothelium 3-4 hours after activation (Mulligan et al., 1991) and forms also transient tethers with its ligands. The main ligand of E-selectin on leukocytes is the E-selectin ligand 1 (ESL-1).

Figure 5: Molecular interactions mediating tethering and rolling
At later time points rolling is largely L-selectin dependent. This third selectin is, in contrast to E- and P-selectins, constitutively expressed. Unlike them, L-selectin is located on the microvilli of leukocytes in concentrated patches and has its corresponding ligands on the endothelia. L-selectin binds to the glycosylation dependent cell adhesion molecule-1 (GlyCAM-1, the mucosal adressin cell adhesion molecule-1 (MAdCAM-1) and CD34. However, the specific functions of these interactions have to be elucidated. It seems that L-selectin is also involved in other adhesive processes as for example cell trafficking to specific tissues / organs such as lymphocyte homing (Arbones et al., 1994).

These processes of leukocyte arrest and firm adhesion are mainly mediated by the integrins, through protein-protein interactions. This family of adhesion molecules consists of heterodimeric glycoproteins composed of two non-covalently associated α and β subunits. To date 22 heterodimers originated of 16 different α and 8 different β subunits can be distinguished.

On top of the layer of arrested leukocytes additional leukocytes from the bloodstream can again start tethering and rolling. These interactions between leukocytes are thought to be mediated by L-selectin / PSGL-1 tethers, which are both present on leukocytes.

Finally, diapedesis starts, at sites where the cells migrate across the endothelia into the underlying tissue. Stable adhesion and transmural migration predominantly involve both integrins (LFA-1, etc.) and members of the immunoglobulin superfamily such as the intercellular adhesion molecules (ICAM-1, 2) and the vascular cell-adhesion molecule (V-CAM). It should be noted, that only the interplay of selectins, chemotactants, integrins, ICAMs and V-CAM makes it possible that leukocytes can transmigrate to infectious loci in underlying tissues.

The entire process of lymphocyte sticking to endothelia (rolling, activation, arrest) takes only a few seconds (Bargatze et al., 1994), while transendothelial migration through the endothelial basement membrane occurs in approximately ten minutes (Smith and Ford, 1983). The mechanisms of leukocyte-endothelial interaction are markedly similar in various organs under both physiological and pathophysiological conditions.

Several substances including humanized antibodies (Co et al., 1999) and several other substances have been tested in vitro for their selectin binding capability. One of
them, binding to all three selectins, however with a low affinity, was the simple tetrasaccharide sLe\(^a\) (Imai et al., 1992; Nelson et al., 1993b). In vivo experiments with anti-adhesive substances (Nelson et al., 1993a; Winn et al., 1998; Hayashi et al., 1999) show that blocking the leukocyte-endothelial interaction reduces the cellular inflammatory infiltrate.

### 3.5 Lectins

To decode the structural information comprised in the diverse sugar structures, molecules differentially recognizing these various glycans are required. A huge class of proteins specifically binding glycans and thereby displaying the mentioned function, has long been known to occur in plants. They have been extensively used as a popular laboratory tool for structural analysis as well as purification in the past years (Cummings, 1997).

**Table 2: Main families of animal lectins (taken from Gabius, 1997).**

<table>
<thead>
<tr>
<th>Family</th>
<th>Structural motif</th>
<th>Carbohydrate ligand</th>
<th>Modular arrangement</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-type</td>
<td>Conserved CBD</td>
<td>variable (mannose, galactose, fucose, heparin tetrasaccharide)</td>
<td>yes</td>
</tr>
<tr>
<td>I-type</td>
<td>Immunoglobulin-like CBD</td>
<td>variable (Man(_6)GlcNAc(_2), HNK-1 epitope, hyaluronic acid, (\alpha2,3: \alpha2,6)-sialyllactose)</td>
<td>yes</td>
</tr>
<tr>
<td>Galectins (S-type)</td>
<td>Conserved CBD</td>
<td>(\beta)-galactosides</td>
<td>variable</td>
</tr>
<tr>
<td>Pentraxins</td>
<td>Pentameric subunit arrangement</td>
<td>4,6-cyclic acetal of (\beta)-galactose, galactose, sulfated and phosphorylated monosaccharides</td>
<td>yes</td>
</tr>
<tr>
<td>P-type</td>
<td>Conserved CBD</td>
<td>mannose-6-phosphate-containing glycoproteins</td>
<td>yes</td>
</tr>
</tbody>
</table>

CBD = carbohydrate binding domain
These carbohydrate binding proteins, which are neither enzymes nor immunoglobulins, have been referred to as lectins (Barondes, 1988). Several years ago they have been discovered to also occur in mammals (for review see Gabius, 1997). Some of them have been demonstrated to be involved in several important biological functions as for example proper protein folding (Trombetta and Helenius, 1998).

As compiled in table 2 five main groups of animal lectins can be distinguished based on structural alignments (Drickamer, 1995).

The dependence of sugar binding on the presence of $\text{Ca}^{2+}$ ions and the preservation of a common sequence motif of 14 invariable and 18 highly conserved amino acid residues have been the prerequisite for defining C-type lectins. This family is also divided into several subgroups (for a more extended review see Gabius, 1997), one of them are the selectins.

### 3.5.1 Selectin structure

The three proteins that comprise the class of selectins are referred to as E-, P- and L-selectin. They have been discovered in 1989 (Bevilacqua et al., 1989; Johnston et al., 1989; Siegelman et al., 1989), and are highly related by their primary structure, as can be seen in Figure 6. Selectins are all linear type I transmembrane multi-domain proteins. The lectin domain, which is involved in the $\text{Ca}^{2+}$ dependent carbohydrate binding, is located on the amino terminus. This carbohydrate binding domain (CBD) is followed by the epidermal growth factor like domain (EGF), which may be important for stabilizing the structure of the protein and partially for its specificity (Kansas et al., 1994; Siegelman et al., 1990). Several consensus repeats (CR) are located adjacent to the EGF domain which are also important for the specificity (Tu et al., 1996). The number of CRs is dependent on cell and selectin type. They are followed by a short transmembrane domain and a cytoplasmic tail on the carboxy terminus.

All three selectins bind sialylated and fucosylated oligosaccharides such as the tetrasaccharide sialyl Lewis X as well as sialyl Lewis a (Varki et al., 1994), a terminal component of glycans attached to glycoproteins and glycolipids, found on most leukocytes and some endothelial cells. However, E-, L- and P-Selectin each displays distinct carbohydrate ligand preferences (Brandley et al., 1993).
3.5.1.1 E-Selectin

The inducible cell adhesion molecule E-Selectin (also known as ELAM-1, LECAM-2, CD62E) is a 95-115 kDa glycoprotein expressed on the surface of endothelial cells and contains 4 to 6 consensus repeats. Following activation it is *de novo* synthesized and the peak of transcriptional activation is reached after 4-6 h. 24-28 h post activation the E-selectin mRNA again reaches basal levels (Tedder *et al.*, 1995; Kansas, 1996). It has been shown to occur in capillary endothelial cells in several disease states such as rheumatoid arthritis (Koch *et al.*, 1991), sepsis (Engelberts *et al.*, 1992), skin inflammation (Picker *et al.*, 1991; Rohde *et al.*, 1992), organ transplant rejection (Briscoe *et al.*, 1991).

E-selectin binds both P-selectin glycoprotein ligand-1 as well as E-selectin ligand-1 and is responsible for slow granulocyte rolling (Kunkel and Ley, 1996).

In support of the participation of E-selectin in acute inflammatory models, E-selectin deficient mice showed a more rapid leukocyte rolling than wild-type (Milstone...
et al., 1998; Ley et al., 1998). However, they show a normal phenotype, but host defense was found to be impaired (Munoz et al., 1997). In addition it was concluded from both E- and P-selectin deficient mice, that migration of leukocytes requires an endothelial selectin, whereas E-selectin and P-selectin may be functionally redundant to some extent (Homeister et al., 1998; Labow et al., 1994).

3.5.1.2 L-Selectin

L-Selectin (previously known as MEL-14, LAM-1, gp90, LECAM-1, CD62L), which is constitutively expressed on the surface of leukocytes, preferentially at the tips of the microvilli, contains two consensus repeats. This distinct position facilitates L-selectin ligand interactions with approaching cells despite its short stretch of consensus repeats.

As shown by Kahn et al. (1994) L-selectin can be cleaved by a cell-associated protease at a specific site close to its transmembrane domain. Interestingly, plasma L-selectin may thereby reach levels that partially inhibit leukocyte adhesion (Schleiffenbaum et al., 1992). On the other hand this protease action seemed to be involved in the rolling process. Walcheck et al. (1996) demonstrated that L-selectin is proteolytically shed from leukocytes after cell-cell interaction and activation. They concluded that L-selectin is cleaved in seconds during the process of rolling under hydrodynamic flow, and that shedding of L-selectin may contribute significantly to the velocity of leukocyte rolling. Furthermore, L-selectin shedding during rolling interactions may be physiologically important for limiting leukocyte aggregation and accumulation at sites of inflammation.

Experiments with L-selectin deficient mice revealed that L-selectin is indeed involved in lymphocyte migration to sites of inflammation (Arbonnes et al., 1994; Steeber et al., 1996; Tang et al., 1997). However, leukocytes appear to roll faster on L- than on P- and E-selectin (Alon et al., 1997). L-selectin alone can mediate rolling in E- and P-selectin deficient mice (Jung et al., 1998). The results also confirmed that leukocyte migration plays an important role in the initiation of humoral responses and demonstrated complementarity of the selectins.

The known ligands for L-selectin are PSGL-1, CD34, GlyCAM1 and MAdCAM-1.
Besides being involved in the extravasation process, L-selectin plays an important role in the homing of T-cells. This interaction enables the entry of peripheral lymphocytes into the lymphatic system for immune surveillance (Gallatin et al., 1983).

### 3.5.1.3 P-Selectin

P-Selectin (previously known as PADGEM, GMP140, GP140, LECAM-3, CD62P) is an integral 140 kDa membrane glycoprotein which is transiently expressed on the surface of endothelial cells and platelets up on activation. It is stored in α granules of platelets and Weibel-Palade bodies of endothelial cells and is rapidly mobilized to the cell surface within 10 minutes after stimulation by agents such as oxygen-derived radicals, thrombin (platelets and endothelial cells) or histamine, phorbol ester (endothelial cells) (McEver et al., 1989; Bonfanti et al., 1989).

P-selectin contains 8 - 9 consensus repeats and thus spans the longest distance from the cell surface. Therefore, it has been concluded that it must project its lectin domain well above the membrane to mediate optimum attachment of neutrophils under shear forces (Patel et al., 1995). The consensus repeats of P-selectin therefore may facilitate interactions with its major ligand PSGL-1 on flowing leukocytes, and increase the intermembrane distance where specific bonds form, minimizing contacts between the glycocalyces that result in cell-cell repulsion. Supporting this theory the flowing neutrophils roll much less efficiently on transfected cells expressing P-selectin molecules that are shortened by deletion of some of the CR (Patel et al., 1995).

Mice deficient in P-selectin are defective in leukocyte rolling and emigration in acute inflammatory models (Mayadas et al., 1993). E- and P-selectin deficiency severely affects leukocyte homeostasis and indicates that these two selectins are as important for normal leukocyte function as are the leukocyte β2 integrins (Frenette et al., 1996). The low extravasation at sites of inflammation makes these animals susceptible to opportunistic bacterial infections to which they succumb.

After activation P-selectin may be cleaved off as it has been demonstrated for L-selectin. This assumption is substantiated by the observation, that a soluble form of P-selectin is present in plasma, which may reach higher levels in patients with rheumatoid arthritis (Ertenli et al., 1998).
3.5.2 Selectin ligands

Figure 7: Scheme of selectin ligands
(reprinted with the friendly permission from PD Dr. Kaltner)

3.5.2.1 E-Selectin-ligand-1
E-Selectin-ligand-1 (ESL-1), which is thought to be the major ligand of E-selectin, has been affinity isolated in 1993 (Levinovitz et al., 1993) and cloned in 1995 by Steegmaier. In contrast to PSGL-1 and the L-selectin ligands, it is not a highly O-glycosylated sialomucin. Instead, it carries up to five N-glycans but no demonstrated O-glycan (Steegmaier et al., 1995). Therefore the N-glycans are thought to carry the carbohydrate chain, being responsible for its ligand activity.

3.5.2.2 CD34
Three molecules binding to L-selectin have been identified yet. CD34 (formerly named Sgp90, or leukosialin) is one of them (Puri et al., 1995). It is a mucin like glycoprotein expressed on high endothelial venules in peripheral lymph nodes and in
several other tissues. CD34 contains a single membrane spanning domain and may present highly clustered O-linked oligosaccharides. It has been found to contain sulfated and fucosylated glycans. Sulfation of CD34 (Hemmerich et al., 1994) has been shown to be required for L-selectin recognition. Furthermore, binding was sensitive to neuraminidase and resistant to treatment by N-glycanase, demonstrating that the epitope responsible for the recognition by L-selectin is located on an O-glycan.

3.5.2.3 GlyCAM-1

The soluble glycosylation dependent cell adhesion molecule-1 (GlyCAM-1, formerly named Sgp^59) was discovered by Baumhueter et al in 1993 and is the best characterized ligand for L-selectin. It has no transmembrane domain and is secreted from high endothelial venules of peripheral lymph nodes (Rosen et al., 1997). Because of this it is more likely that it may inhibit further interactions by blocking the binding sites and/or transduce signals into leukocytes rather than to support leukocyte adhesion (Giblin et al., 1997; Hwang et al., 1996). To bind L-selectin, GlyCAM-1 must be modified with branched core 2 O-glycans that are sialylated, fucosylated and sulfated (Imai et al., 1993). Sulfate ester is found to be attached to the C6 position of Gal and GlcNAc residues on the GlyCAM-1 O-glycan (Hemmerich et al., 1994a).

3.5.2.4 PSGL-1

P-selectin glycoprotein-ligand 1 (PSGL-1) has been cloned in 1993 (Sako et al., 1993) and is thought to be the major ligand for P-selectin. It is a type I transmembrane protein which occurs as a disulfide bonded homodimeric sialomucin with two 120-kDa subunits (Moore et al., 1992). Translation starts with a 18 amino acid signal sequence followed by a propeptide spanning residues 19-41. These residues are removed in processed PSGL-1 in leukocytes (Li et al., 1996) by the enzyme called PACE (paired basic amino acid converting enzyme; Rehemtulla and Kaufman, 1992). The extracellular domain of the processed mature protein extends from residue 42-318 and contains between 15 (HL-60) to 16 (leukocytes) decameric repeats (Sako et al., 1993). Besides many serines and threonines only three N-linked glycans are present, which are not required for P-selectin binding (Norgard et al., 1993). The extracellular domain is followed by a 25 residue transmembrane domain and a 69-residue cytoplasmic tail.
It has been demonstrated that the mAb PL1 blocks binding of purified PSGL-1 to P-selectin in vitro (Moore et al., 1995; Patel et al., 1995). This antibody has its epitope on the amino terminus of PSGL-1 where the tyrosine sulfation sites and the Thr-57 carrying the critical O-glycan are located. It also dramatically reduces in vivo tethering and rolling of neutrophils on HL60 cells (Norman et al., 1995). Thus interaction between P-selectin and PSGL-1 are essential for adhesion under flow. PSGL-1 is constitutively expressed on leukocytes and concentrated, like L-selectin, on the tips of the microvilli (Moore et al., 1995).

PSGL-1 binds to all three selectins (McEver, 1997; Guyer et al., 1996). Interestingly, P- and E-selectin recognize a related N-terminal binding site. However, P-selectin binds with higher affinity (Moore et al., 1994; Patel et al., 1995). For binding to P-selectin, PSGL-1 has to be sulfated whereas sulfation is not required for binding to E-selectin (Li et al., 1996). In contrast to GlyCAM-1 the O-glycans on PSGL-1 are found not to be sulfated (Wilkins et al., 1995). Binding of PSGL-1 to P-selectin requires a fucosylated and sialylated core 2 branched O-glycan attached to threonine-57 (Liu et al., 1998). In addition, Liu et al. demonstrated that binding is enhanced when at least one of the three tyrosine residues at positions 46, 48 and 51, respectively, on the N-terminus of the processed molecule are sulfated. As predicted from structural analysis of PSGL-1 released from HL-60 (Wilkins et al., 1996) only a minority of the glycans on PSGL-1 is fucosylated. These structures occur as two major species both containing the sialyl Lewis X epitope (Figure 8).

Despite the presence of PSGL-1 on all lymphocytes only a minority of these cells binds P-selectin. These findings suggests that the glycosylation may be highly regulated (Moore et al., 1995).
Structure A

\[
\begin{align*}
\text{Fuc} & \rightarrow 3 \\
\text{Neu5Aca}_2 & \rightarrow 3\text{Gal}\beta_1 & \rightarrow 4\text{GlcNAc}\beta_1 \\
\downarrow & & \\
6 & \\
\text{Neu5Aca}_2 & \rightarrow 3\text{Gal}\beta_1 & \rightarrow 3\text{GalNAc} (\alpha_1 - \text{Thr})
\end{align*}
\]

Structure B

\[
\begin{align*}
\text{Fuc} & \rightarrow 3 \\
\text{Neu5Aca}_2 & \rightarrow 3\text{Gal}\beta_1 & \rightarrow 4\text{GlcNAc}\beta_1 & \rightarrow 3\text{Gal}\beta_1 & \rightarrow 4\text{GlcNAc}\beta_1 & \rightarrow 3\text{Gal}\beta_1 & \rightarrow 4\text{GlcNAc}\beta_1 \\
\downarrow & & & & & & \\
6 & & & & & & \\
\text{Gal}\beta_1 & \rightarrow 3\text{GalNAc} (\alpha_1 - \text{Thr})
\end{align*}
\]

Figure 8: Two main structures naturally occurring on PSGL-1 from HL60 cells.
3.6 Medical targets

As mentioned in section 3.4 all three selectins are involved in the initial step of the inflammatory cascade. Blocking selectin carbohydrate interactions has been demonstrated to significantly reduce leukocyte infiltration into the underlying tissue.

Inflammation is usually a beneficial response of the immune system for the control of infection and injury. However, occasionally it also can result in acute systemic collapse or, more frequently, a chronic reaction such as for example in autoimmune diseases. Attenuated migration has been demonstrated to be advantageous for the treatment of several different indications as illustrated in the following examples:

In acute neutrophil dependent lung injury a diminished tissue accumulation of neutrophils was observed after infusion with sLe^X oligosaccharides, most probably due to blocking selectin binding sites (Mulligan et al., 1993).

Also in obliterated blood vessels, as for example the coronary artery in the case of infarction, expression of the selectins is up-regulated. During reperfusion, this leads to strong leukocyte tethering and rolling, which again can result in hindrance of the yet re-established blood flow and ensuing tissue damage. Blocking selectin carbohydrate interactions has been shown to reduce reperfusion damage in these cases (Buerke et al., 1994; Murohara et al., 1995).

Furthermore in inflammatory diseases, as for example rheumatoid arthritis (Ochi et al., 1993) and inflammatory bowel disease, a reduced influx of leukocytes may at least alleviate the effect of these chronic diseases (reviewed in Kansas, 1996; Varki, 1997).

Leukocyte infiltration is a hallmark of acute rejection in solid organ transplants such as cardiac allograft. It has been demonstrated that leukocyte extravasation to cardiac grafts undergoing rejection is largely due to interactions between leukocyte L-selectin and its sLe^X decorated ligands (Turunen et al., 1994). An inhibitor of L-selectin-dependent leukocyte adhesion to graft endothelium therefore is desirable to prevent transplant rejection.

Another potential indication is cancer and metastasis. The spread out of cancer cells over the body has been shown to be attenuated in P-selectin deficient mice (Kim et al., 1998). Therefore, blocking P-selectin binding may have a beneficial effect in...
preventing metastasis. In addition, an inhibitory effect of selectin blockers on angiogenesis has been reported (Nguyen et al., 1996).

Besides several other diseases selectin carbohydrate interactions are involved also in psoriasis (Wakita et al., 1994) and asthmatic diseases (reviewed in Bloemen et al., 1997).

3.7 Anti-adhesive compounds

The numerous medical targets make the development of selectin blockers, which display a new class of anti-adhesive compounds, a relevant task. Several strategies have been followed to reach this challenging goal. At least two main points of action can be distinguished, the selectin ligands, and the selectins themselves.

An interesting approach to inhibit selectin mediated interactions on the ligand side aims at blocking the biosynthetic pathway leading to the expression of sLe\(^x\). Most efforts focus on the sialyl- and the fucosyltransferases. Small molecules designed to specifically inhibit these enzymes are the targets of many groups (Chandrasekaran et al., 1995; Jefferies and Bowen, 1997). For cultured cells another interesting approach has been published by Prati et al. (1998). They demonstrated, by use of antisense approach to inhibit expression of fucosyltransferase, that enzyme activity and sLe\(^x\) synthesis could be decreased concomitant with reduced adhesion of endothelial cells. However, permanently down regulating fucosyltransferase activity in vivo may be problematic. It has been shown that in leukocyte adhesion deficiency (LAD) patients, which lack sialyl Lewis X ligands, a defect in the synthesis pathway of GDP-fucose led to recurrent episodes of bacterial infections (Karsan et al., 1998).

In addition a short eight amino acid long conserved peptide present in all three selectins, has been shown to bind Le\(^x\) as well as sLe\(^x\). It thus can be used to block the recognized carbohydrate structures present on the selectin ligands. By inhibiting leukocyte extravasation, this small peptide may be useful as an anti-inflammatory non antigenic compound (Briggs et al., 1996).
To inhibit leukocyte / endothelia interaction on the selectin side different classes of anti-adhesive compounds have been developed and tested. They include antibodies, peptides, sulfatides, mono- and polyvalent oligosaccharides as well as several sLe\(^x\) based analogs.

Notably, a short therapeutic antibody generated against the selectins, hinder leukocyte extravasation by shielding the selectins carbohydrate binding sites (Co et al., 1999; Hayashi et al., 1999). However, due to potential antigenicity life time in the circulation and recurrent usage over a longer period of time may be problematic. Another approach applied phage display to identify peptide ligands that bind to E-selectin and thereby inhibit leukocyte adhesion. Following this screening several peptides have been identified that prevent adhesion in the nanomolar range leading to reduced neutrophil transmigration in mice undergoing an acute inflammatory response (Martens et al., 1996).

Among the naturally occurring selectin ligands the sulfatide Gal(3-SO\(_4\))\(\beta\)1-1Cer has been identified. It inhibits P-selectin binding at concentrations of IC\(_{50}\) = 10-20 μM and binds to all three selectins (Watson et al., 1990; Aruffo et al., 1991; Nair et al., 1994).

According to this, beneficial effects have been obtained in vivo (Mulligan et al., 1995; Kajihara et al., 1995). Based on this sulfatide analogs have been developed and demonstrated to have enhanced in vivo anti-inflammatory properties with a half life time in the circulation of 7 h (Todderud et al., 1997). Potential antigenicity may again limit the usage of this compound to single treatment of acute diseases.

Another naturally occurring compound binding to the selectins is the major sialyl Lewis X structure. This simple oligosaccharide has therefore been synthesized and tested for its anti-inflammatory potential. The studies indeed revealed a beneficial effect of the oligosaccharide at high dosage only for its low affinity. (Polley et al., 1991; Nelson et al., 1993; Mulligan et al., 1992). To enhance binding, polyvalent structures bearing multiple sialyl Lewis X epitopes on different backbones, such as for example polylactosamine stretches, have been developed (Renkonen et al., 1997). Unfortunately, besides their low oral availability the unsubstituted saccharides are rapidly excreted.

After the first full chemical synthesis of sLe\(^x\) (Kameyama et al., 1991) another strategy has focussed on the development of several glyco-mimetics (reviewed in
Simanek et al., 1998). Some of these components have been demonstrated to bind the selectins with higher affinity than simple sLe\(^\alpha\) (Wittmann et al., 1998) and led to the identification of the important functions and groups recognized by the selectins. Unfortunately, the glyco-mimetics may have again a high antigenic potential. Therefore, a non antigenic component which binds with high affinity to the selectins is still missing.

As mentioned in section 3.4. P-selectin is thought to be involved in the very first contact between leukocytes and endothelia and that its major ligand is PSGL-1. It has been demonstrated that the selectins clearly bind with higher affinity to glycoproteins than to the free glycan ligands. Therefore, naturally occurring PSGL-1, which presents sLe\(^\alpha\) on a protein backbone could be an interesting, non antigenic candidate binding to P-selectin with high affinity. In fact, a short fully glycosylated and sulfated N-terminal stretch of 19 amino acid of PSGL-1 has been shown to bind with high affinity (Goetz et al., 1997). Besides blocking the very first interactions, PSGL-1 will also bind to L- as well as to E-selectin.
4. Aim of the thesis

The aim of this thesis was to construct an oligosaccharide scaffold for the attachment of the \textit{sialyl Lewis X} recognition structure on N-acetylgalactosamine-peptides using an enzymic approach. This goal should be accomplished by using the highly stereo- and regiospecific glycosyltransferases as enzymic catalysts \textit{in vitro}. As this work was embedded in an EU-project that aimed at the development of selectin blockers, part of the outlined project had to be implemented in collaboration with partner groups.

In a first step, a GlcNAc residue, which is the backbone of the \textit{sialyl Lewis X} structure, had to be transferred to the core GalNAc which links O-glycans to the peptide. Therefore, a reaction transferring GlcNAc to GalNAc had to be performed. Because the core 6 structure has been reported to naturally occur it was assumed that one of the \(\beta_6\)-GlcNAcT would exhibit a certain side activity towards the unsubstituted O-glycosidically linked GalNAc (Figure 3A, reaction d). The Core 2 \(\beta_6\)-GlcNAcT, which has been cloned (Bierhuiizen and Fukuda, 1992) and was known to branch Gal(\(\beta_1\)-3)GalNAc structures (Figure 3A, reaction c), therefore had to be produced as a recombinant enzymic catalyst in sufficient amounts and used for this reaction.

In a second step, after introducing the N-acetylglucosamine, the resulting core 6 should be elongated using a \(\beta_4\)-GalT following reaction g (Figure 3B), which could be purified from human milk (Gerber \textit{et al.}, 1979).

In a third step, a \(\alpha_3\)-SiaT should be used to sialylate the obtained type II structure (Figure 3B, reaction h). For this purpose one of the \(\alpha_3\)-SiaT has to be expressed as a recombinant enzyme, produced and purified, so that it could be used as an enzymic catalyst.

Finally, the lacking fucose had to be added in \(\alpha_1\)-3 position to the GlcNAc (Figure 3B, reaction i) to obtain the desired O-glycosidically linked core 6 based sLe\(^\alpha\). For this reaction, one of the known \(\alpha_3\)-FucT had to be expressed, produced and purified to enable its use for the synthesis.
5. Results

The first task was to make available in sufficient amounts the β6-GlcNAcT, which was required to link GlcNAc to the core GalNAc of O-glycans.

In paper 1, entitled "Pilot scale expression and purification of soluble Protein A tagged β1,6N-acetylglucosaminyltransferase in CHO cells" the development of a CHO cell line stably secreting the Core 2 β6-GlcNAcT and its cultivation in a fluidized bed reactor is described. After the purification the specificities of this enzyme were investigated.

To facilitate enrichment and purification my strategy aimed at producing a secretable Core 2 β6-GlcNAcT. Additionally a feature was desirable, that allows fast screening and affinity purification. To fulfill these criterias the pProtAC2GnT plasmid was chosen. It codes, under the control of the SV40 promoter, for a truncated version of Core 2 β6-GlcNAcT (aa 38-429), which lacks the transmembrane domain responsible for Golgi retention. In addition, the enzyme is expressed N-terminally fused to the Protein A tag, which specifically binds IgG. Together with the pSV2Neo, a vector conferring the neomycin resistance, CHO cells were co-transfected following the standard Lipofectamine procedure. Clones were isolated manually and the best performing clones were determined by monitoring Protein A concentration in the supernatant by ELISA using an antibody to Protein A. Stability of the selected clones was monitored over 8 weeks and finally confirmed by activity measurements.

For the cultivation of the established cell line a continuously operating fluidized bed reactor, supporting high cell density of up to 2.3 x 10⁸ cells/ml carrier, was chosen in collaboration with D. Eisenkrätzer (Research center of Jülich). In this bioreactor, which has formerly successfully been used for the cultivation of Hybridoma cells, a maximum activity of 0.95 U/l was obtained. Unfortunately, after about 220 h, when the highest titer and a steady state cell density were reached, enzyme activity decreased approximately 9 fold to 0.15 U/l after 400 h. This observation was further investigated as described in manuscript 2.

The obtained supernatant was collected and purified by a single step affinity chromatography on IgG-Sepharose. Eluting by shifting the pH to 3, followed by rapid neutralization, yielded a purification factor of 213 fold at a recovery of 32 %. The specific activity was 44 mU/mg protein and total activity 23,7 U/l. SDS PAGE and
western blot analysis revealed that the protein was glycosylated most probably not only the two potential N-glycosylation sites of the Core 2 β6-GlcNAcT but also at the Protein A tag. It was further shown that another protein co-eluted, which was neither recognised by the polyclonal antibody directed against Core 2 β6-GlcNAcT nor by the monoclonal antibody against Protein A.

We proceeded by investigating the acceptor specificities of the produced catalyst. Because no significant activity could be detected when using the core 3 (GlcNAc(β1-3)GalNAc) acceptor the enzyme was assigned to be of the L-type and not a Core 2/4 β6-GlcNAcT (M-type). Unfortunately, our Core 2 β6-GlcNAcT did not incorporate any GlcNAc on GalNAc(α1-Benzyl. Even by using a high excess of enzyme under optimised assay conditions and a peptide bound GalNAc as acceptor substrate, no incorporation was detectable. Thus, our initial strategy of obtaining the core 6 structure by following reaction d (Figure 3A) and synthesizing sLe³ based on this core therefore had to be reconsidered.

In manuscript 2, entitled "Improved product formation in high density Chinese hamster ovary cell cultures transfected at confluency" the improvement of a Lipofectamine mediated transfection procedure is described in collaboration with D. Eisenkrätzer (Research center of Jülich) leading to a changed expression pattern. As mentioned above, cultivation of conventionally transfected CHO cells expressing Core 2 β6-GlcNAcT in a fluidized bed reactor led to a decreasing productivity when the cell density reached a steady state. To investigate whether this decrease was due to genetic instability caused by cultivation without selection pressure, cells were taken from the reactor and cultivated again in batch mode. Interestingly, they again expressed maximum productivity.

Because some metabolites (e.g. antibiotics) are known to be expressed in a growth-phase dependent manner (Smiley et al., 1989; Suzuki and Ollis, 1990; Lloyd et al., 1999), cell cycle distribution was investigated using the propidium iodide method. We observed, by analyzing the relative distribution of the cells, a drop of the S/(G0/G1)-ratio during cultivation. Thus we concluded that the immobilization at high cell density led to a reduction of proliferation, concomitantly with a decrease of the proportion of cells in S-phase. In a next step cell specific productivity was monitored during a whole fermentation period in a spinner flask. Indeed, a strong correlation
between productivity and amount of cells in S-phase could be detected. To further substantiate this observation, cultures were synchronized with the non-cytotoxic DNA-polymerase inhibitor aphidicolin, arresting 95% of the cells in G0/G1-phase. After resuspending these cells in fresh aphidicolin free media activity was measured at different time points as a function of the cell cycle. Activity of recombinant Core 2 β6-GlcNAcT was detected to be 5 times higher in S-phase than in G0/G1-phase. Thus, Core 2 β6-GlcNAcT production was indeed coupled to cell growth. This observation correlates with the productivity pattern obtained in the fluidized bed reactor. While cells grow the highest enzyme productivity was detected whereas at steady state cell density, productivity decreased. We then assessed the cell cycle distribution during transfection at subconfluency, when conventional transfection is carried out. At this point the majority of the cells were in the S-phase. Taken together, transfection at subconfluency, where cells are mainly in S-phase, resulted in maximum expression in S-phase.

Because the use of a high cell density cultivation system is advantageous in terms of high space time yield and titer, we aimed at developing an appropriate cell line suitable for maximum expression in growth inhibited stage. To obtain a clone showing this changed expression pattern, cells were transfected 2 to 3 days after seeding at confluency instead of subconfluency. The best performing clones were detected by Protein A ELISA and their performance was confirmed by enzyme activity measurements. A suitable clone was cultured in spinner flasks and cell specific productivity was measured as a function of the cell cycle. In contrast to the pattern observed with clones transfected at subconfluency, the clones transfected at confluency expressed maximum productivity in GO/G1-phase.

To substantiate these observation, both transfection procedures, the conventional and the confluent stage transfection, were applied to another reporter protein. Based on the pPROTA vector, the pPROM plasmid was constructed by introducing a multi cloning site. A truncated version, lacking the transmembrane domain, of α3-fucosyltransferase III was cloned in this vector. After screening and isolation cell specific productivity was investigated and found to show a similar pattern as the one observed with Core 2 β6-GlcNAcT.

In summary, a clone obtained by transfection at confluency, where cells are mainly arrested in G-phase, exhibited maximum expression in G-phase.
In manuscript 3, entitled "Complete enzymic synthesis of the mucin-type sialyl Lewis X epitope, involved in the interaction between PSGL-1 and P-selectin" the stepwise synthesis of our target structure using the different glycosyltransferases and intermediate product purification is described. $^1$H-NMR and MS analysis by R. G. Gallego (Utrecht University) confirmed the identity as O-linked sLe$^X$.

Starting with the model compound GalNAc(α1-pNp, an enzyme transferring GlcNAc to this structure was needed. As mentioned in manuscript 1 neither unsubstituted GalNAc nor GlcNAc(β1-3)GalNAc are acceptors for the produced Core 2 β6-GlcNAcT. Therefore, an enzymic activity substituting GalNAc in 3’ position with galactose is a prerequisite for the full enzymic synthesis of O-glycosidically linked sLe$^X$ using the Core 2 β6-GlcNAcT. Unfortunately, the cDNA of the β3-GalT is not available yet despite intensive screening efforts of a whole family of β3-GalT (Hennet et al., 1998). However, the corresponding enzyme has been enriched from natural sources in sufficient quantities. Thus, a postnuclear fraction of a Triton extract of rat liver was prepared and used to elongate the pNp bound GalNAc (reaction b Figure 3A). Incubating the assay for six days and adding fresh enzyme as well as new donor in between finally yielded an incorporation of 52 %. After purification via hydrophobic chromatography the above mentioned Core 2 β6-GlcNAcT was used to branch the obtained Gal(β1-3)GalNAc(α1-pNp (reaction c Figure 3A). With 64 mU of this purified recombinant enzyme a yield of >85% was reached within 32 hours. In a next step a galactose had to be attached in 4’ position to the purified branched GlcNAc. This reaction (reaction g Figure 3B) was performed by β4-GalT, an enzyme which naturally also occurs as a soluble protein in different body fluids. Despite the fact that a recombinant form of this glycosyltransferase was available (Malissard et al., 1996; Zigova et al., 1999), we used the native enzyme which could easily be purified from human milk (Gerber et al., 1979). Incubation with 200 mU of this β4-GalT yielded an incorporation of >85% into the core 2 compound after 5 h. To elongate the purified structure an α3-SiaT was necessary (reaction h Figure 3B). We used purified recombinant rat ST3Gal3 which was a gift from M. Streiff (Novartis). Two sequential incubations with 170 mU and 72 mU of the ST3Gal3, respectively, were needed to obtain a yield of 87 %. Finally, the resulting structure has to be fucosylated (reaction i Figure 3B) to obtain the O-glycosidically linked sLe$^X$. A soluble form of human α3-FucT6 was therefore cloned under the control of the strong alcohol oxidase promoter.
(AOX) into *Pichia pastoris*. Following induction by methanol for 48 h, the enzyme was secreted into the media, from which it was enriched and affinity purified. By using 70 mU of this recombinant enzyme in a 6 h incubation, a complete conversion of the acceptor was observed on HPLC as well as by small scale parallel incubations, in which incorporation of sugars was monitored by use of radioactive donor substrates. Taken together, the overall yield was 32% of hexasaccharide product formation in a five step full enzymic synthesis using three recombinant enzymes, a purified preparation of a native enzyme and one extract.

In collaboration with R. Gutiérrez Gallego (Bijvoet Center, Utrecht University) MALDI TOF mass spectra, 1D and 2D TOCSY and ROESY $^1$H-NMR analysis was carried out and confirmed the obtained structure.
6. Discussion and perspectives

Reducing leukocyte extravasation by blocking selectin carbohydrate interactions has been demonstrated to be of beneficial effect for several indications. During the past years the mostly pursued efforts in the development of such antagonists have focused on the preparation of huge amounts of the minimal binding epitope, the glycan sLe\(^\text{x}\) as well as its mimetics. The developed carbohydrate based anti-adhesive components will most probably be confined to the treatment of often life-threatening acute diseases including for example organ transplantation and ischemia-reperfusion injury. In these circumstances the application of injected carbohydrates seems to be an appealing approach, although they are rapidly excreted and are low affinity ligands. Orally applicable blockers, having a longer life time in the circulation and binding with higher affinity are therefore highly desirable for the recurrent treatment of acute as well as more chronic inflammatory diseases. An interesting candidate fulfilling all these criteria is the natural P-selectin ligand PSGL-1. It has been shown that for high affinity binding the protein backbone of PSGL-1 is playing an important role and that PSGL-1 binds E- as well as L- selectin. Sulfation of at least one of the three tyrosine's on the N-terminus results in enhanced binding properties. Therefore, the development of such a fully glycosylated and sulfated PSGL-1 is a feasible task towards achieving the above mentioned goal.

In this thesis the first five steps to a full enzymic synthesis leading to a core 2 based sLe\(^\text{x}\) antagonist is described. Our method can be used for the in vitro glycosylation of a given protein backbone. Without further optimization, we reached a yield of 32% O-glycosidically linked core 2 type sLe\(^x\). Compared with the chemical synthesis using more than 20 steps and obtaining a yield below 10 %, this approach displays a real improvement (Parekh and Edge, 1994). In addition the chemical synthesis led to toxic byproducts like Hg(CN)\(_2\), HgBr\(_2\), ethanolamine etc. (Nicolaou et al., 1991) while the enzymic synthesis ends up with unproblematic waste. At present, the synthesis of our core 2 linked sLe\(^x\) based selectin blocker in terms of laboratory chemicals, not taking into account the costs of the enzymes, would lead to a price of 1200 SFr per mg under non optimized conditions. The IC\(_{50}\) of simple sLe\(^x\) for binding to P-, L- and E-selectin is ranging from 0,16 to 3 mM (Brandley et al., 1993;
Parekh and Edge, 1994; Koenig et al., 1997). Interestingly, these values obtained in vitro are at least two to three orders of magnitude higher than the concentration of 1 μM found in vivo to be efficient for the treatment of acute lung injury (Mulligan et al., 1993). The basis of this discrepancy has to be elucidated and can be due to several factors as for example measuring without flow or under dynamic shear-stress conditions. However, as a result of the ongoing clinical trials with carbohydrates the required concentration in man is expected to become clearer soon.

Assuming no renal clearance of our product will occur, due to the presence of a suitable protein backbone and that the binding properties of the protein bound sLe^x will at least be in the same range as simple sLe^x oligosaccharide, it would be possible to attenuate leukocyte migration with an amount of 5 μmol core 2 based sLe^x in human. A single dosage of 5 μmol glycosylated sLe^x with the described non optimized method would lead to a price of around 6500 SFr. Clearly our method has to be improved in order to reduce costs. A challenging strategy towards this goal is the implication of in situ donor recycling cycles as proposed by Ichikawa et al. (1992). Simanek et al. (1998) showed a 10 enzyme system for the production of sLe^x tetrasaccharide. Besides the three glycosyltransferases (β4-GalT, α3-SiaT, and α3-FucT) two kinases (Pyruvate kinase, myokinase) were needed to generate the sugar nucleotide triphosphates. Furthermore a synthase (CMP-Neu5Ac-synthase), two phosphorylases (UDP-Glc pyrophosphorylase GDP-Fuc pyrophosphorylase), an epimerase (UDP-Glc epimerase) and a pyrophosphatase to drive the reaction towards product formation were used. Another UDP-Gal recycling cycle using sucrose synthase has been published by Elling et al. (1993). For a full in vitro synthesis of core 2 based sLe^x a recycling cycle for UDP-GalNAc and UDP-GlcNAc would be needed in addition. UDP-GalNAc can be generated from UMP and sucrose via a 7 enzyme system (nucleoside monophosphate kinase, sucrose synthase, galactose-1-phosphate uridyltransferase, phosphoglucomutase, glucose-6-phosphate dehydrogenase, pyruvate kinase, lactate dehydrogenase) as published by Bültter et al. (1997). UDP-GlcNAc recycling cycles starting with inexpensive GlcNAc-1-P and using UDP-GlcNAc pyrophosphorylase have been established by Look et al., (1993) and de Luca et al. (1995).

Even if, all other steps can be scaled-up, the remaining bottle neck of the described method is the availability of β3-GalT. It can so far only be obtained in small
quantities from natural sources. Until this gene is not available other possibilities to achieve β3-GalT activity or circumvent its requirement have to be considered. They can be distinguished in *in vitro*, *in vivo* and combined *in vivo - in vitro* approaches.

One possible *in vitro* procedure to link GlcNAc to the core GalNAc could use the transglycosylation reaction catalyzed by a β-galactosidase. A three enzyme system using Core 2 β6-GlcNAcT and β-galactosidase has been demonstrated to form β6-GlcNAc branches with a yield of >90% (Dudziak et al., 1998). This method leads to the formation of a core 2 and core 6 mixture and can only be driven towards high core 6 yields. The galactose on the 3-arm however, which is lacking in the core 6 structure, seems to be important for binding. Jain *et al.* (1998) demonstrated that a sialic acid bound to this 3-galactose on core 2 based Le^x^ led to a 5-6 fold better inhibition of P- and L-selectin than sLe^x^ tetrasaccharide. Hence, for the development of a high affinity selectin blocker therefore this Gal residue will be important.

As soon as β3-GlcNAcT is cloned another approach to circumvent the need of β3-GalT could make use of the recently cloned Core 2/4 β6-GlcNAcT (Yeh *et al.*, 1999). This enzyme has a broader acceptor specificity than Core 2 β6-GlcNAcT but interestingly also has no demonstrated side activity towards unsubstituted GalNAc (Schwientek *et al.*, 1999).

At present a combined *in vivo in vitro* strategy seems to be promising. The production of the desired short N-terminal PSGL-1 protein backbone in CI10 cells, which lacks the Core 2 β6-GlcNAcT but contains β3-GalT, will most likely lead to a sulfated product carrying a sialylated core 1 structure (sialyl TF) at the corresponding threonine. It can be assumed that the *in vivo* glycosylated protein, after treatment by neuraminidase, can be branched *in vitro* and extended to a sLe^x^ presenting product by using the methods described here.

In addition, a full *in vivo* strategy can be designed. Considering that CHO cells are missing the Core 2 β6-GlcNAcT and the desired α3-FucT, glycosylation engineering will be necessary to co-express these two enzymes and the protein backbone of PSGL-1. For the simultaneous cloning of three independent genes an approach using multicistronic vectors allowing coupled expression of different genes is highly suitable as demonstrated by our group and others (Fussenegger *et al.*, 1998; Dinter *et al.*, 1999).
Meanwhile, the elongation to polylactosamine structures will be an interesting task in order to enhance affinity for all strategies. The analysis of PSGL-1 derived O-glycans from HL60 cells revealed that also multifucosylated structures on polylactosamine chains with terminal sLe\(^\text{X}\) may be present as shown in Figure 8B (Wilkins et al., 1996; Aeed et al., 1998). It has been demonstrated that structures containing several LacNAc repeat units bind more efficiently to E-selectin (Stroud et al., 1995). The enzymes responsible for the formation of these stretches known as β4-GalT4 (Ujita et al., 1998) and β3-GlcNAcT (Zhou et al., 1999) have been identified and cloned recently. Additionally, two distinct fucosyltransferases, α3-FucT4 and α3-FucT7 will be necessary (Niemelä et al., 1998).

In a side study we developed a new transfection method for CHO which led to a changed expression pattern for two independent reporter proteins. The use of a high density cultivation systems is of general interest in biotechnological process development for several reasons including higher space time yields and reduced equipment costs. These main advantages led to the development of several high density cultivation systems, as for example the fluidized bed reactors, which have been extensively used for the fermentation of hybridoma cells in the past years (Lüllau et al., 1992). Such high density systems, supporting up to 2 x 10\(^6\) cells ml\(^{-1}\), have not been reported for CHO cells. Since this cell line is becoming increasingly important as a universal production host for therapeutic proteins, we decided to investigate the drop of productivity in more detail. As shown in manuscript I the use of CHO cells in such a fluidized bed reactor resulted in a decreasing productivity when the cell density reached a steady state. Investigating this effect revealed that production was growth-associated for CHO cells transfected at subconfluency and was 5 times higher in S- than in G0/G1-phase in batch. Using these data for modeling the productivity pattern in the fluidized bed reactor substantiated our assumption that cells immobilized in the carriers of a continuously operating fluidized bed after 300 hours are mainly arrested in G0/G1-phase. In order to benefit from the advantages this bioreactor offers, a cell line mainly producing in G0/G1-phase was needed. By shifting the time point of transfection from subconfluency to confluency this goal was reached. Considering that proteases and hydrolases are thought to be predominantly formed and secreted in this phase, product quantity and quality have to be investigated during test cultivations in fluidized bed
reactors with the thus generated cell lines. However antibody production using this cultivation system, perhaps also due the short residence time, demonstrates impressively its feasibility.
7. Manuscript I:

Pilot scale expression and purification of soluble Protein A tagged β1,6N-acetylglucosaminyltransferase in CHO cells

*Biochem Biophys Res Commun, 1997 237: (3) 653-658
PILOT SCALE EXPRESSION AND PURIFICATION OF SOLUBLE PROTEIN A TAGGED \( \beta1,6N\)-ACETYLGOSAMINYL-TRANSFERASE IN CHO CELLS

Steffen Zeng\(^1\), Andre Dinter\(^1\), Detlef Eisenkrätzer\(^2\), Manfred Biselli\(^2\), Christian Wandrey\(^2\), and Eric. G. Berger\(^1\)*

\(^1\)Institute of Physiology, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland; \(^2\)Institute of Biotechnology, Forschungszentrum Jülich, D-52425 Jülich, Germany.

Received July 22, 1997

Expression of recombinant soluble protein A tagged mouse core 2 \( \beta6-N\)-acetylglucosaminyltransferase (E.C. 2.4.1.102) has been scaled-up in CHO cells using a continuously operating fluidized bed system yielding 0.3 U/day. A one step 213 fold purification by affinity chromatography on IgG-Sepharose yielded a stable enzyme preparation with a specific activity of 44 mU/mg. The enzyme was shown to belong to the L-type with a highly restricted specificity for the acceptor substrate \( \text{Gal}^1 \rightarrow 3 \text{GalNAc}^\alpha \rightarrow \text{R} \) (core 1). Only little activity towards \( \text{GlcNAc}^\beta \rightarrow 3 \text{GalNAc}^\alpha \rightarrow \text{R} \) (core 3) (<1%) and no incorporation on unsubstituted benzyl or peptide-bound GalNAc was detected. \( \text{Zn}^{2+} \) and to a lesser extent \( \text{Mn}^{2+} \) were found to be inhibitory whereas \( \text{Mg}^{2+} \) could activate the enzyme. The enzyme preparation proved suitable for in vitro application as a catalyst for the synthesis of core 2 structures\(^1\).

\(^1\)To whom correspondence should be addressed. Fax: ++41/1/364 0564
E-mail: egberger@physiol.unizh.ch

Core 1 structure: \( \text{Gal}^1 \rightarrow 3 \text{GalNAc}^\alpha \rightarrow \text{R} \); core 2 structure: \( \text{Gal}^\beta \rightarrow 3 [\text{GlcNAc}^\beta \rightarrow 6] \text{GalNAc}^\alpha \rightarrow \text{R} \)
core 3 structure: \( [\text{GlcNAc}^\beta \rightarrow 3 \text{GalNAc}^\alpha \rightarrow \text{R} \); core 4 structure: \( [\text{GlcNAc}^\beta \rightarrow 3 [\text{GlcNAc}^\beta \rightarrow 6] \text{GalNAc}^\alpha \rightarrow \text{R} \)

Blood group I: \( \text{GlcNAc}^\beta \rightarrow 3 [\text{GlcNAc}^\beta \rightarrow 6] \text{Gal}^1 \rightarrow \text{R} \)
Glycosyltransferases are highly specific enzymes involved in structuring complex carbohydrates. Recent efforts in many laboratories in cloning and expression of these enzymes make it possible to use recombinant forms of these enzymes also for in vitro synthesis of oligosaccharides and glycoconjugates (1). To render glycosyltransferases suitable for this purpose several prerequisites must be fulfilled: The enzyme should be available in high quantities as a soluble reagent, its specificity towards both donor and acceptor substrates should be known and the reagent should be stable for prolonged incubation times.

In the case of glycosyltransferases, several expression systems have successfully been applied for the production of soluble forms of these enzymes such as yeasts (2-4), fungi (5), Sf9 cells (6) and CHO cells (7).

While enzymes involved in N-glycosylation are relatively well known (8) those structuring O-glycans have attracted more attention only recently. Therefore, enzymatic synthesis of O-glycans has not been possible mainly because corresponding enzymes are not available. It is, for example, not easy to synthesize enzymatically the core 2 structure because the cDNA encoding the corresponding galactosyltransferase has not been reported yet and the core 2 β6-N-acetylglucosaminytransferase (C2GnT) (E.C. 2.4.1.102) has not been made available in large quantity while its cDNA has been cloned by Bierhuizen and Fukuda (9) and its activity characterized in lysates of different tissues (for review see 10). In fact, a preparation of this enzyme purified from bovine tracheal epithelium has been reported to be very unstable (11). As a step towards the goal of enzymatic synthesis of core 2 O-glycans as possible carriers of sialyl Lewis X determinants such as in PSGL-1 (12), we developed a CHO cell line stably producing a soluble form of C2GnT fused with protein A and report here production, purification and characterization of the enzyme.
Materials and Methods

Materials
The following chemicals and reagents were from: ECL reagent, \[^{14}\text{C}]\text{UDP-GlcNAc}\) (Amersham); molecular weight markers (Bio-Rad); Streptavidin-Peroxidase conjugate (Boehringer Mannheim); pSV2Neo (Clontech); DMEM, FCS, HAM’s F12, G418, Lipofectamin™, Optimem™ (Gibco); 1,2 Phenylendiamine (Merck); Peptide: N-glycosidase F (New England Biolabs); Scintillant Irga-Safe Plus (Packard); CNBr-activated Sepharose, IgG-Sepharose G Fast Flow (Pharmacia Biotech AB); Asialofetuin, Asialomucin, biotinylated goat anti-Protein A antibody, BSA, core 3-Benzyl\(^1\), GalNAc-Benzyl, UDP-GlcNAc, human IgG 1 Lambda, Ovalbumin, Protein A Peroxidase, recombinant IgG-binding fragment of Protein A, (Sigma); core 1-p-nitrophenyl (pNp)\(^1\), core 1-Benzyl\(^1\) (Toronto Research Chemicals); SepPak Cartridges (Waters). Human IgG was a gift from Dr. P. Späth, Central Laboratories of Swiss Red Cross, Berne. The construct pProtAC2GnT coding for the soluble mouse C2GnT (EC 2.4.1.102) was kindly provided by C.E.Warren and J.Dennis. A polyclonal antibody against human C2GnT used for immunoblotting was obtained from A.Datti.

Transfection and generation of stable clones
CHO K1 cells were cotransfected with pProtAC2GnT and pSV2Neo following the Lipofectamin procedure from Gibco. The pProtAC2GnT plasmid codes for a truncated soluble version (aa 38-429) of the mouse C2GnT (gb:u19265 sw:q09324) as a soluble Protein A fusion-protein. Selection pressure was applied by using 800 \(\mu\text{g/ml}\) G418 (Geneticin™) in HAM’s F12 with 10 \% FCS. After 14 days of growth single colonies were picked and cultivated further. Media was harvested and analyzed by ELISA and assay for C2GnT-activity.

C2GnT Assay
A) For low level expression (<100 mU/l) the enzyme was enriched from 1.5 ml supernatant by binding to 40 \(\mu\text{l}\) of a 50\% IgG-bead slurry overnight at 4°C. B) Media containing higher levels (>100 mU/l) of C2GnT were directly assayed using 20 \(\mu\text{l}\) of supernatant. The standard assay mixture contained the following ingredients in a total
volume of 40 µl (supernatant assay) or 80 µl (slurry assay): 100 mM GlcNAc, 100 mM MES Puffer pH 6.5, 1 mM core l-pNp or core l-benzyl, 50 µg/ml BSA, 5.5 mM AMP, 1 mM UDP-GlcNAc, 1 µl [14C]UDP-GlcNAc (= 50,000 cpm). AMP, GlcNAc and BSA were used for measurements of crude material (media) only and not for assays performed with the purified enzyme. Mixtures were incubated at 37°C for 1 to 3 h. Reactions were stopped by adding 0.5 - 0.75 ml of ice-cold water. For separation, the mixture was loaded on a SepPak Cartridge mounted on a vacuum chamber. To remove unreacted radiolabeled UDP-GlcNAc the cartridge was washed with 15 ml of water. The hydrophobic glycosylated acceptor was eluted with 5 ml of methanol. This eluate was mixed with 10 ml of scintillation cocktail and counted. Typically, cpm values of 1000 - 5000 were counted. High molecular weight acceptors (10 mg/ml) were assayed as described previously (13).

**Protein A ELISA**

The Protein A part of the fusion-protein was assayed by a Protein A ELISA (details of this procedure will be published elsewhere by Kolbinger F., Streiff, M.B. and Katopodis A.G., manuscript in preparation). Briefly, human-IgG coated microtiter plates were incubated either with 100 µl of sample or Protein A standard, washed, further incubated with Streptavidin-Peroxidase conjugate and developed using 1,2 phenylenediamine as substrate.

**Media for mass cultivation:**

Cells were grown in a 3:1 mixture of DMEM and Ham’s F12 (Gibco, Eggenstein, FRG). Due to the high consumption rates of the CHO cells both compounds have been used at a concentration of 120 % compared to the Gibco protocol (Osmolarity: 350 mOsmol/kg). The media has been supplemented with essential amino acids, trace elements, vitamins, 6 mg/l insulin, 6 mg/l transferrin, 0.1952 mg/l linoleic acid, 0.04636 mg/l thioctic acid and 1% (v/v) fetal calf serum. Neither selection marker nor antibiotics were used for the fermentations.

**Cultivation System:**

Batch cultivations were carried out in T-flasks, spinner flasks and stirred vessels. Continuous fermentations were performed in a fluidized bed reactor. The cells were
immobilized in open porous glass carriers with a diameter of 400-700 µm and a porosity of 50% (SIRAN®, Schott Mainz). Bubble free aeration was realized by diffusion of oxygen across a silicon tube integrated in the circulation loop. The pH was controlled at 7.0 to 7.2 and the oxygen saturation in the suspension was regulated to 10 to 20 % air saturation.

**Affinity chromatography on IgG Sepharose**

A 6 ml affinity chromatography column was prepared by coupling 5.2 mg/ml of human IgG to CNBr-activated Sepharose resin following the Pharmacia protocol. Prior to loading the supernatant, cells were separated and 4 liters of media were dialyzed for 3 days against 2 x 45 l PBS pH 7.2 at 4°C. 3.4 Liters of the dialysed media was loaded on a 6 ml column at a flow rate of 100 ml/h. After washing with 1 liter of PBS, bound proteins were eluted with 0.1 M Glycin-HCl pH 3.0 at a flow rate of 80 ml/h of. The eluate was rapidly neutralized with 1/10 volume of 0.3 M TRIS-Glycine pH 8.9. Fractions of 2.6 ml were collected.

**Electrophoresis and Immunoblotting**

Eluted neutralized fractions from IgG Sepharose were resolved on a 10 % SDS-polyacrylamide gel according to Laemmli (14). For Peptide: N-glycosidase F (PNGase F) treatment 20 µg of sample was denatured at 95 °C for 10 min in 0.5% SDS 1% β-mercaptoethanol, followed by 1 h of incubation in 50 mM Sodium Phosphate (pH 7.5) containing 1 % NP-40 and 100 U of PNGase F. Electrophoresis was followed by immunoblotting to nitrocellulose membrane as described previously (15).

A) Protein A immunoblot: Membranes were incubated for 60 minutes with the biotinylated monoclonal antibody to Protein A in a 400 fold dilution. After 3 times washing with blocking buffer Streptavidin-Peroxidase diluted 8000 fold was applied for 1 h and detected by ECL following the manufacturers protocol.

B) C2GnT immunoblot: Incubation with 400 fold diluted anti-human-C2GnT antiserum was followed by goat-anti rabbit horseradish peroxidase (1:5000) and developed as above.
Results

1. Establishment of C2GnT producing cell line and fermentation

CHO K1 cells were transfected with the pProtAC2GnT plasmid encoding soluble C2GnT directly fused to the O-terminus of Protein A as indicated above. Two weeks after transfection, 20 G418 resistant clones were picked and expanded. Best performing clones were selected by assessing C2GnT activity in the supernatant. To ascertain stability, the clones were further cultivated for 8 weeks under selection pressure and monitored by repeated activity measurements. The best clone, achieving a maximum enzyme activity of 520 mU/l during cultivation in a spinner flask was further cultivated. Since this system is not suitable for scale-up a fermentation in a continuously operating fluidized bed reactor was performed (fig. 1). After about 220 h of fermentation the enzyme activity in the supernatant reached a maximum value of 950 mU/l corresponding to a space time yield of 4560 mU/l of total reactor volume and day while the cell specific productivity was 84 µU/(10^6 cells*day). Later on the cell specific productivity decreased to 13.2 µU/(10^6 cells*day) and the space time yield to 720 mU/(l*d) due to the formation of thick cell layers on the carriers and reduction of cell viability. In summary, we produced about 18 Units C2GnT using a small scale fluidized bed reactor in a continuous fermentation of 1400 h corresponding to an average enzyme activity of 300 mU/l. Average space time yield [1440 mU/(l*d)] was 20 times higher than in the spinner [65 mU/(l*d)] even though the average cell specific productivity was low compared to the maximum productivity in the batch system.

2. Downstream processing and characterisation of C2GnT

The cultivation was followed by a rapid one step purification. An affinity chromatography on an IgG-Sepharose column was performed, by taking advantage of the Protein A part of the fusion-protein. Loading and elution were controlled by Protein A ELISA and could be correlated with activity measurements. The enzyme activity remained stable during 30 min of incubation in elution buffer. The capacity of the 6 ml column was not exceeded by loading 3.4 liters of dialyzed supernatant (not shown). Purification with this technique revealed a specific activity of 44 U/g protein at a recovery of 32 %. The SDS-PAGE analysis of the purification is shown in Fig. 2. The expressed protein was not detectable in the supernatant nor in the flow through. In the
purified fraction 3 major bands could be observed (panel A). The upper two bands were recognized by immunoblotting with an antiserum to human C2GnT (panel B) as well as with the monoclonal antibody to Protein A (panel C). Since the protein of 50 kD visible on the Coomassie blue stained PAGE was recognized neither by the antiserum to human C2GnT nor by the monoclonal antibody to Protein A we assumed that this protein was a contaminant. Sequence analysis of C2GnT predicted the presence of 6 potential N-glycosylation sites, of which 4 are assigned to the Protein A part and 2 to the soluble portion of C2GnT. Removing the N-glycosylation by Peptide: N-glycosidase F (PNGase F) treatment resulted in a shift of 8 to 10 kD (panel D), indicating that the protein was N-glycosylated also on the protein A moiety.

To characterize the purified soluble recombinant mouse enzyme several acceptors were tested. To distinguish between L- and M-type (see reference 16 and discussion) the core 3 structure was chosen. Purified recombinant C2GnT showed nearly no incorporation (<1%) of GlcNAc on core 3 (Table 1).

<table>
<thead>
<tr>
<th>Acceptor (1 mM)</th>
<th>Enzyme activity (U/l)</th>
<th>Enzyme activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gal β1-3 GalNAcα Benzyl</td>
<td>87.8</td>
<td>100</td>
</tr>
<tr>
<td>Gal β1-3 GalNAcα pNP</td>
<td>73.9</td>
<td>84</td>
</tr>
<tr>
<td>Gal β1-3 GalNAcα Octyl</td>
<td>71.0</td>
<td>81</td>
</tr>
<tr>
<td>GlcNAc β1-3 GalNAcα pNP</td>
<td>0.63</td>
<td>0.71</td>
</tr>
<tr>
<td>GalNAcα Benzyl</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
</tr>
</tbody>
</table>

Thus, this mouse enzyme was assigned to be of the L type. No activity on GalNAc-Benzyl as well as on a peptide substituted with one GalNAc residue (not shown) could be detected. These data confirm the previously found strict dependence of C2GnT (L-type) on GalNAc substituted in 3’ position with Gal. Furthermore, a variety of glycoprotein acceptor substrates were investigated (Table 2).
As expected, asialoglycophorin, containing 15 clustered O-glycans, proved to be an excellent acceptor. The low activity towards ovalbumin was ascribed to a contaminant since ovalbumin is not known to be O-glycosylated (17).

Interestingly, in lysates containing endogenously expressed β1,6GlcNAc-transferases unusual metal cofactors dependence was observed. We therefore investigated the influence of several metal cations on recombinant soluble mouse C2GnT. Enzyme activity was only slightly affected by 10 mM Cu²⁺, Co²⁺ and Ca²⁺ ions. The most pronounced effects were observed for Mg²⁺(10 mM) which stimulated the enzyme 1.5 fold and 10 mM Zn²⁺ which inhibited it to 15 % remaining activity. The influence of Zn²⁺ was strongly dose dependent and could be reversed by addition of EDTA.

Among various buffers including Mes, Cacodylate, Phosphate, Tris-maleate, Mops and Heps, Mes proved to be the best system. Testing purified soluble recombinant C2GnT at different pH revealed a broad pH - optimum between 6 and 7 (not shown).

**TABLE 2**

Specificity of the recombinant purified soluble mouse C2GnT for glycoprotein acceptors

<table>
<thead>
<tr>
<th>Acceptors</th>
<th>Content (mg)</th>
<th>Molecular weight (kDa)</th>
<th>O-glycans</th>
<th>N-glycans</th>
<th>Activity (U/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asialoglycophorin A</td>
<td>0.1</td>
<td>31</td>
<td>15</td>
<td>1</td>
<td>5.54</td>
</tr>
<tr>
<td>Asialofetuin</td>
<td>1</td>
<td>64</td>
<td>3</td>
<td>3</td>
<td>4.67</td>
</tr>
<tr>
<td>Asialofetuin</td>
<td>0.1</td>
<td>64</td>
<td>3</td>
<td>3</td>
<td>2.03</td>
</tr>
<tr>
<td>Asialomucin bovine</td>
<td>1</td>
<td>60</td>
<td>?</td>
<td>?</td>
<td>0.58</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>1</td>
<td>37</td>
<td>0</td>
<td>1</td>
<td>0.23</td>
</tr>
<tr>
<td>BSA</td>
<td>1</td>
<td>66</td>
<td></td>
<td></td>
<td>0.04</td>
</tr>
</tbody>
</table>
Discussion

A number of other glycosyltransferases has already been produced in different cell types (4,18) to make this important catalysts available in suitable amounts. Here we report the first large scale production of a glycosyltransferase in CHO cells. The use of many other expression systems can cause problems: in S. cerevisiae hyperglycosylation may occur (for review see 13,19), while the baculovirus expression system is inappropriate for continuous cultivation (20). Thus CHO cells, which are used to manufacture several pharmaceutical proteins (e.g. refs.21-23), seemed to be an interesting host for the production of functional mammalian glycosyltransferases.

C2GnT plays a key role in the synthesis of branched O-linked glycans, which are important for numerous biological functions such as fertilization, mucin synthesis, receptor functions, inflammation, etc. (10,24) In addition, increased C2GnT activity can be cancer-associated (25). Therefore, the aim of our investigation was to produce suitable amounts of this enzyme to use it as a catalyst for the enzymatic in vitro synthesis of the O-linked sLex structure which plays an important role in the inflammation process by conferring P- and E-selectin binding properties to PSGL-1 (26). Besides this the availability of recombinant C2GnT allows to perform studies of specific inhibitors (27).

According to the experiments performed with lysates (28) Core2, Core4 and the activity responsible for the synthesis of the I antigen can be found together in several cell lines (for example in pig gastric mucosal microsomes). From competition experiments using this crude material, it was concluded that the same β1-6GnT was responsible for the synthesis of core 2, core 4 as well as for the blood group I structure1. In acute myeloid leukemic cells only core 2 and no activity on core 4 and I GnT2 was found (16). Therefore two types of β1-6 Gnt have been distinguished. One type taken from leukocytes and thus called L-type was highly specific for the core 1 structure and did not act on core 31. Whereas the other type, found in mucin secreting tissues and therefore called M-type, has a broad specificity and was associated with the formation of core 4 and the I antigen1. To assign the purified soluble mouse enzyme to the L or M type respectively, the core 3 structure1 GlcNAc β1-3 GalNAc was used. No significant incorporation of GlcNAc could be observed. Thus the mouse C2GnT was assigned to the L-type.
Although the existence of GlcNAcβ1-6GalNAc has been observed (29-31) and an enzyme activity leading to this structure has been reported (32) we were not able to perform this reaction neither with GalNAc-Benzyl nor with peptide-bound GalNAc under optimized conditions even by excess of the purified CHO-derived enzyme. These findings can be explained by the theory proposed by Schachter et al. (33) that the occurrence of this structure may be due to degradation, or formed by another yet unknown β1-6 GnT. It appeared that the influence of the protein backbone of the acceptor substrate is of lesser importance for the C2GnT than the activities of β1-3 GaT, α2-3- and α2-6-sialyltransferases which seemed to be responsible for the quantity of branching points found on O-linked glycans.

The influence of metal ions on several glycosyltransferases has been recognized early (34). In the case of β1-6GnTs different enzyme sources have already been investigated (16,28,35-37) but the data appeared quite heterogeneous probably due to the use of crude enzyme preparations. Functionally pure recombinant enzyme, therefore, proves suitable to identify unusual kinetic properties such as sensitivity to Zn$^{2+}$ of C2GnT.

**Acknowledgments**

This work was supported by grant BIO4-CT95-0138 Engineering O-Glycosylation of the EU. The authors also like to thank Bea Berger for maintenance of the CHO cells.
References


FIG. 1. Time course of cell concentration and enzyme activity during a continuous fermentation in a fluidized bed reactor. The reactor was inoculated with $2.28 \times 10^8$ cells directly transferred into the circulation loop. Within 1 day the cells attached and the circulation medium was cleared. Feeding with fresh medium containing 100 mg/l BSA and 1% (v/v) FCS was started with a rate of 120 ml/d. Afterwards the feeding rate was increased according to the increasing glucose consumption rate, keeping the glucose concentration at 1.5 g/l. After 6 days a stable high cell density was reached and the flow rate was set to 1.2 liters per day. After 600 h the FCS and BSA content were successively reduced to zero without any influence on the enzyme production.
FIG. 2. SDS-PAGE and immunoblot of purified fraction. Panel A: Coomassie blue stained SDS-PAGE. Panel B: immunoblot with polyclonal Anti-human C2GnT antibody. Panel C: immunoblot with monoclonal Anti-Protein A antibody. Panel D: immunoblot with monoclonal Anti-Protein A Antibody. Lane M Marker, Lane 1 supernatant, Lane 2 purified fraction, Lane 3 flow through, Lane 4 PNGase F treated purified fraction.
8. Manuscript II*:

Improved product formation in high density Chinese hamster ovary cell cultures transfected at confluence

* submitted to Biotech Bioeng, in revision
Improved product formation in high density Chinese hamster ovary cell cultures transfected at confluency

Steffen Zeng1, Detlef Eisenkrätzer2, André Kiesewetter2, André Dinter1, Eric G. Berger1* and Manfred Biselli2

1Institute of Physiology, University of Zurich, Zurich, Switzerland
2Institute for Biotechnology, Research Center Jülich, Jülich, Germany
*Corresponding author

Address: Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
Tel: Int+ 41 1 635 50 70
Fax: Int+ 41 1 635 18 44

Summary

A reliable method to generate Chinese hamster ovary (CHO) cell clones that stably express heterologous proteins in high density cell cultures is not available. High production yield of CHO cells is usually restricted to the exponential growth phase. To improve productivity in batch cultured CHO cells that are predominantly in G0/G1 phase we compared CHO cells transfected by liposome-mediated gene transfer at confluency (G0/G1 phase) with cells that were conventionally transfected at subconfluency (S phase). The recombinant reporter glycoproteins investigated in this study were soluble proteinA fusion proteins of α1-3 fucosyltransferase III (ProtA-Fuc-TIII) and core2 β1-6 N-acetyl glucosaminyltransferase (ProtA-C2GnT), respectively. Cells that were transfected at subconfluence decreased their productivity at high density in G0/G1 phase. By contrast, cells that were transfected at confluency maintained their productivity under similar conditions. This difference was shown not to depend on the type of promoter or selection marker. Thus, continuous high expression rates of soluble glycosyltransferases at G0/G1-phase of CHO cell cultures was achieved by selection of production clones transfected at confluency.

Key-words: Cell cycle, fluidized bed reactor, fucosyltransferase III, N-acetylglucosaminyltransferase
Introduction

For cost effective production of recombinant glycoproteins long term cultures of CHO cells that maintain a high level of productivity are needed. As shown for hybridoma cell lines, specific productivity increased in G1 phase at high density (Al-Rubeai, et al., 1992). By contrast, under similar conditions, production in CHO cells decreased (Brightwell, et al., 1997; Gu, et al., 1994). Therefore, the main challenge is generation and selection of stable CHO cell lines with high productivity in G0/G1 phase. Several strategies have been followed on the basis of reactor design (Fenge, et al., 1993), generation of appropriate vectors (Lin, et al., 1994), supplementation of nutrients (Schmid, et al., 1991) or addition of xenobiotic agents as for example the use of Na butyrate to elevate gene expression of a randomly incorporated plasmid (Palermo, et al., 1991). A much more sophisticated approach was recently described to arrest CHO cells in G0/G1 phase that have been engineered to express cytostatic p27 in response to the addition of tetracycline (Fussenegger, et al., 1998).

Conventional non-viral transfection randomly incorporates the plasmid into the genome (for review see Koch-Brandt, 1993). Little is known on the access to chromatin of foreign plasmid DNA as a function of the cell cycle. Usually cells are transfected at the stage of subconfluency or in suspension since efficiency of transfer and transient expression appears higher than following transfection of contact-inhibited cells (Guide To Eukaryotic Transfections With Cationic Lipid Reagents. http://www2.lifetech.com:80/catalog/techline/molecular_biology/protocols/txnguide.html). We argued that cells transfected at confluency would exhibit higher productivity since plasmid DNA might incorporate at sites subjected to G0/G1 phase-dependent transcriptional regulation. In fact, experimental support for this assumption has recently been provided (Dean, 1997) along with the notion that chromosomes and chromatin are dynamic structures during the cell cycle (for review see Lamond and Earnshaw, 1998). We used recombinant glycosyltransferases as reporter molecules, e.g. β1,6N-acetylglucosaminyltransferase (so-called core 2 type GnT or C2GnT) and α1,3fucosyltransferase III (Fuc-TIII) both involved in the construction of O-glycosidic selectin ligands (Li, et al., 1996).
Materials and Methods

Plasmids and cloning

pProtA vector (Sanchez-Lopez, et al., 1988) harbouring soluble mouse C2GnT (amino acids 38 to 429 of the mouse C2GnT, accession# 19265), was kindly provided by Charles Warren (Toronto, Canada), (Figure 1A).

pPROM: An insertion of a multi cloning site into the unique EcoRI cloning site of pProtA resulted in a general cloning vector suitable for the production of other glycosyltransferases as fusion proteins with ProteinA. 2 µg of each 70mer synthetic, single strand DNAs were phosphorylated with polynucleotide kinase, combined and annealed in 20 mM Tris/Cl pH 7.6, 200 mM NaCl and 2 mM EDTA over night (starting at 60 °C to 25 °C). 12 ng were ligated to pProtA and cut with EcoRI without prior purification. The integrity of the newly generated plasmid pPROM was confirmed by sequencing the multicloning site from both sides (Figure 1B).

To clone a soluble form of Fuc-TIII into pPROM, the plasmid pcDNA3-Fuc-TIII, kindly provided by Dietmar Vestweber, (Münster, Germany) was digested with AvrII/XbaI and ligated into XbaI digested pPROM. The correct plasmid pPROM-Fuc-TIII, comprising aa 44-361 of the human Fuc-TIII (accession# X53578) was selected by restriction analysis followed by Fuc-TIII activity assays on transiently transfected CHO cells.

Vector pcDSA was a kind gift of Shuichi Tsuji, (Riken, Japan). A soluble construct was generated from full-length mouse C2GnT.

Vector pSecTag (Invitrogen) was used to express mouse C2GnT under a CMV promoter. A soluble construct was generated from full-length mouse C2GnT. In addition pSecTag was equipped with a neomycin resistancy by virtue of an IRES from pIRES1neo (Clontech). Schemes of the constructs are shown on fi. 7

Transfection procedure and selection

CHO cells were transfected by a LipofectAMINE™ mediated procedure following the manufacturer's protocol (Life Technologies). Briefly, 0.4 x 10^6 CHO cells were plated on a 3.5 cm dish and grown over night in Ham's/10% fetal calf serum for subconfluency and two to three days for confluency. 3 µg plasmid DNA and 10 µl LipofectAMINE™ were combined in OptiMem™ (Life Technologies) and added to the cells. Where necessary 1 µg of pGKneo or pPUR (puromycin resistance) were added to cotransfect the resistance markers. At this
stage the protocols of the transfection procedures differed: a) For C2GnT expression, subconfluent cells were transfected. Two days after start of transfection selection was started with 1000 μg/ml G418 with media change every two days. After 8-15 days single cell clones were picked from the plate by a manual procedure using a Gilson pipet. Individual clones were grown and tested by activity assays. In the case of transfection of confluent cells, G418 was added 6 days after start of transfection. Medium was changed every two days and plates were grown to confluency in 10-21 days. After detachment with trypsin/verseine cells were sorted in 500 single wells using a cell sorter (FACStar\textsuperscript{Plus}, Becton Dickinson). Due to the high clone number in this procedure, manual collection was not possible. After another 14 days, cell supernatants were measured for ProteinA expression by ELISA, selected and subsequently assayed for glycosyltransferase activity as described below. b) For Fuc-TIII expression, transfection in subconfluent and confluent cells were performed in exactly the same way. Either confluent or subconfluent cells were transfected. Selection was started 6 days after transfection with 1000 μg/ml G418 with media change every two days. Emerging clones were diluted by detachment followed by reseeding the whole population on the same plate. After reaching confluency cells were sorted as described before. Where necessary puromycin was applied in a concentration of 10 μg/ml and zeocin (Invitrogen) in a concentration of 800 μg/ml.

Media for mass cultivation were as described before (Zeng, et al., 1997).

**Synchronization of Cell Cycle**

Synchronization was achieved using the non-cytotoxic DNA-polymerase inhibitor aphidicolin (Fox and Bedford, 1987, Spadari, et al., 1985) (SIGMA) in spinner flasks in a volume of 100 ml. Cells grown for 3 days to a density of $1 \times 10^5$ cells/ml were pelleted at 200 g, resuspended in media containing 3 mg/l aphidicolin and incubated for 14 h. Then the cells were harvested and resuspended in fresh medium without aphidicolin. The spinner was placed in an incubator and henceforth every hour a 2 ml aliquot was taken from the culture. Enzyme activity was assayed as described below (assay A) and cell specific productivity was calculated based on the increase of the enzyme activity.

**Cultivation Systems**

Batch cultivations were carried out in T-flasks or spinner flasks, continuous cultivations in a fluidized bed reactor as described (Zeng, et al., 1997; Lüllau, et al., 1992).
ProteinA ELISA

Supernatants from clones in 96 well plates were subjected to a ProteinA ELISA (Zeng, et al., 1997; Kolbinger, et al., 1998).

Activity measurements of glycosyltransferases

C2GnT assay: (A) For low level expression (<100 mU/l) the C2GnT-proteinA fusion protein was enriched by binding 1.5 ml of the supernatant to 40 µl of a 50% IgG-Sepharose bead slurry in conditioned medium (0.05% Tween20, 5 mM Tris-HCl pH 7.6) overnight at 4 °C. The beads were washed, resuspended in an equal volume and assayed in 80 µl containing 40 µl of 50% bead slurry, 100 mM 2-[N-morpholino]ethanesulfonate buffer pH 6.5, 1 mM Gal(β1→3)GalNAca-α-pNp as acceptor, 1 mM UDP-GlcNAc as donor, 1 µl 14C-UDP-GlcNAc (50'000 cpm), and as stabilizers 100 mM GlcNAc, 5.5 mM AMP and 20 µg/ml bovine serum albumin. The assay was incubated gently shaking for 2-3 h at 37 °C to yield an incorporation between 1000 and 10000 cpm. The incubation was stopped with 1 ml cold water and loaded on a C18 Sep-Pack Cartridge (Waters). The column was washed twice with 5 ml H2O and eluted with 2 ml of methanol. (B) Media containing higher levels (>100 mU/l) of C2GnT were directly assayed using 20 µl of supernatant in a total volume of 40 µl. As inhibitory substances in the supernatant were not removed by this procedure this assay yielded lower activity values compared to (A).

Fuc-TIII assay: Fuc-TIII-protein A fusion protein was enriched as described above. The 80 µl assay contained 40 µl of 50% bead slurry, 40 mM cacodylate buffer pH 6.2, 5 mM LacNAc as acceptor, 0.1 mM GDP-Fucose as donor, 1 µl 14C-GDP-Fucose (50'000 cpm), 10 mM MnCl2, 5 mM ATP and 10 mM fucose. Incubation and purification was as above.

Determination of cell specific productivity in batch culture

Cell specific productivity was determined by taking cells either from tissue flasks or from spinner flasks and resuspending them in fresh medium at a cell density of (1x 10⁶ cells per mL). After 2 h of incubation enzyme activity was measured in the supernatant (assay A) and cell specific productivity was calculated based on the increase of enzyme activity.

Cell cycle phase determination

Cell cycle analysis was performed using a standard procedure (Ronot, 1996). Cells either from tissue flasks or from microcarriers were trypsinized, centrifuged (10 min, 200 g) resuspended, frozen in 70% ethanol and stored at -20 °C. Before measurement cells were
centrifuged and resuspended in PBS. An aliquot of 1-5 x 10^6 cells was incubated with propidium iodide (18 μg/mL/10^6 cells) and RNAse (40 μg/mL/10^6 cells) for 1 h at room temperature. Flow cytometric analysis was performed using a FACS-Calibur (Becton Dickinson; excitation wavelength at 488 nm, emission wavelength at 630 nm). DNA histograms were analyzed using the program ModFit.

Results

For the comparison of lipofection at subconfluency versus confluency we transfected a construct consisting of a vector harbouring the SV40 transcription promoter, the sequence encoding the transin signal peptide, linked in frame to a fusion construct of protein A with the cDNA encoding a truncated form of C2GnT (see Methods). This vector (designated pProtA-C2GnT) was designed to direct an active form of the enzyme to the supernatant where it can be assayed by activity and ELISA using an antibody to protein A (Zeng, et al., 1997; Kolbinger, et al., 1998).

For transfection, cells were plated at a density of 40,000/cm^2 and grown for 24 h. Subconfluent CHO cells (Figure 2 inset) shown to have a proportion of 45% in S phase (corresponding to a S/G0/G1 ratio of 0.8) (Figure 2) were co-transfected with pProtA-C2GnT and pGKneo, harbouring the neomycin resistance gene. Selection was started day 2 after transfection and clones were manually detached and grown to confluency. The best performing clone was expanded to spinner flasks and subsequently inoculated in a fluidized bed reactor where the cells are immobilized on macroporous carriers at high cell density (Lüllau, et al., 1992). As described previously, within 220 h of fermentation, recombinant ProtA-C2GnT reached 950 mU/l in the supernatant, then dropped to 150 mU/l after 400 h (Zeng, et al., 1997). At this stage cell density was 2.3 x 10^8 cells/mlcarrier. Thus cell specific productivity decreased from 84.0 to 13.2 μU/10^6cells/day. After 880 h of continuous fermentation without selection pressure, cells could be recovered by trypsinization and expressed again maximum productivity of recombinant ProtA-C2GnT in a batch fermentation (not shown). By analyzing the relative distribution of the cells in S versus G0/G1 phase we observed a drop of the S/(G0/G1)-ratio. Thus, we concluded that the immobilization at high cell density led to a reduction of proliferation, concomitant with a decrease of the proportion
of cells in S-phase, eventually resulting in a drop of product formation. By analyzing cell specific productivity along the whole fermentation period in spinner flasks we observed a strong correlation of productivity with the S-phase (Figure 3). To substantiate this possibility we synchronized the culture using 3 mg/l aphidicolin (Fox and Bedford, 1987) for 14 h resulting in over 95% of the cells in G0/G1 phase. Following removal of aphidicolin the cells remain synchronized for 5 h during S and G2 phase (data not shown). The cell specific productivity was then assessed as a function of the cell cycle phase by measuring the activity of recombinant C2GnT in the supernatant or by protein A ELISA (not shown) and found to be five times higher in S-phase than in G0/G1-phase (Figure 4). Thus, the high productivity in S phase seemed to be associated with transfection at subconfluency. Cell specific productivity decreased by using higher concentrations of aphidicolin (data not shown).

Next, we aimed at comparing productivity of cells transfected at confluency (Figure 2, inset) using the same vector harbouring a fusion construct of protein A with fucosyltransferase III as follows: CHO cells were plated at the same density as above and grown to confluency for 2 to 3 days. After lipofection cells were cultivated for 6 days in non-selective medium. Following addition of the selection marker, cells were again grown to confluency, clones were isolated, their performance quantitated by ELISA and confirmed by activity measurements. Batch cultures in spinner flasks were established from a suitable clone. Cell specific productivity was again assessed as a function of the S/(G0/G1)-phase ratio. In contrast to the results shown on Figure 4, expression was shown to be favored in G0/G1-phase. Therefore, cells transfected at confluency showed the highest expression level at minimal proliferation rate, i.e. low S/(G0/G1) ratio (Figure 5B). Conversely, cells transfected at subconfluency showed highest productivity during proliferation (Figure 5A) as already shown for production of C2GnT (Figure 4). Qualitatively similar experiments were obtained using ProtA-C2GnT as reporter enzyme (Figure 6 A and B).

To exclude effects unrelated to the cell cycle phase at transfection, a variety of controls were carried out by following four strategies (Figure 7): First, a vector was constructed in which the SV40 promoter, the transin signal sequence and the selection marker were replaced by the CMV promoter, the IgK signal sequence and zeocin. Second, pSecTag-C2GnT was equipped with the neomycin resistance marker directly coupled to the expression cassette via an internal ribosomal entry site (IRES), to exclude loss of the plasmid harbouring the resistance gene. Third, the pcDSA vector harbouring a SRα promoter, a SV40 promoter/enhancer cassette, a HTLV-1 LTR and a IgM signal peptide was investigated. Fourth, a double selection procedure was followed (Wirth, et al., 1988) for which the resistance marker genes
neomycin and puromycin were co-transfected on independent plasmids. In none of these instances did we detect better performing clones in dense CHO cell cultures.

**Discussion**

The data presented here show that the efficiency of CHO cells for the production of recombinant proteins depends on their cell cycle status at the time of transfection: When cells were transfected in a subconfluent state, production was most effective during cell growth whereas transfection at confluency generated stable CHO cells that produced most efficiently in a stationary phase. Improving efficiency of recombinant product formation in animal cells has been on the agenda since many years. A higher cell specific productivity yields a higher product concentration which is a prerequisite for the cost-effective synthesis of complex glycoproteins using mammalian cell cultures (Werner, 1994). Reactor design (Fenge, et al. 1993), control of temperature (Weidemann, et al. 1994; Hortacsu and Ryu, 1990), nutrient supply (Doverskog, et al. 1997; Hayter, et al. 1991; Schmid, et al. 1991), dependency on foetal calf serum (Leelavatcharamas, et al., 1994), promoter design (Yano, et al. 1994) have long been known to influence the production rate of recombinant proteins. In general, production in CHO cells was usually associated with high proliferation rate of the host cells (e.g. Hayter, et al., 1991; Deutschmann, 1996; this paper). Conversely, a reduction of proliferation and concomitant drop of product formation by immobilization of BHK cells in high cell density in a fluidized bed reactor has been described by Deutschmann (1996). A similar reduction in cell specific productivity was observed using CHO cells secreting a N-acetylgalactosaminyl transferase (our unpublished data).

Longevity of cells has been found to be higher in non-dividing cells when attached on microcarriers allowing for continuous cultures over several months (Ogata, et al., 1993). These cultures, however, consist of cells arrested in G0/G1 phase. Therefore, more recent attempts at improving the production rate focused on cell cycle control by aiming at developing cells that would exhibit a higher production rate in G0/G1 than in S phase. In special cases such as hybridoma cells, addition of interleukin-6 has been shown to suppress growth rate and to increase product formation five-fold (Makishima, et al., 1992). In CHO cells an analogous switch from S to G1 phase may be achieved by shifting the temperature
from 37 °C to 30 °C; this was accompanied by a 3.4 higher overall product yield (Kaufmann, et al., 1999). Another strategy was followed by Mazur, et al. (1998), who stably overexpressed the cyclin-dependent kinase inhibitor p27 in a tetracycline dependent manner; removal of this agent resulted in growth arrest and increased cell-specific production.

Surprisingly little is known concerning the dependence of transfection on the cell cycle phase. The difference between transfection at confluency vs subconfluency was apparently not due to the type of the viral promoter since the use of both CMV and SV40 promoters led to similar results. We also observed this phenomenon for two unrelated recombinant proteins, e.g. C2GnT and fuc-TIII rendering a specific effect of the recombinant product unlikely.

To our knowledge, a systematic study examining integration of chemically transfected (as opposed to virally infected) foreign DNA into the host genome as a function of the cell cycle is not available. The following aspects may be of relevance in an attempt to find an explanation for our results. The nucleus is regarded as a highly dynamic organelle along the cell cycle as shown by a GFP-tagged probe that binds specifically to foreign DNA stably transfected into CHO cells (Li, et al., 1998). Position effects have been described to contribute to the regulation of gene transcription (for review see Lamond and Earnshaw, 1998). It is tempting to speculate that insertion of chemically transfected foreign DNA does not occur randomly but predominantly at sites of active transcription. Some experimental support for this assumption has been provided by Dean (1997). Thus, foreign DNA integrated at sites that are actively transcribed during G0/G1 phase (at confluency, fig. 2) would perhaps also predominantly be expressed in G0/G1 phase. This explanation does not lack simplicity but is certainly worth providing experimental support as this may have also a bearing on gene or cell therapy beyond optimization for biotechnological production. Thus, further work is needed to investigate integration of foreign DNA as a function of the cell cycle.

Acknowledgements

The authors wish to thank Prof. Dr. Christian Wandrey for the excellent support of this work, Bea Berger and Eva Niederer for technical assistance. We thank Prof. Sandro Rusconi and Dr. Martin Fussenegger for critically reading this manuscript. This work was supported by grant BIO4-CT95-0138 of the EU framework 4.
References


Kaufmann, H., Mazur, X., Fussenegger, M., Bailey, J.E. 1999 Influence of low temperature on productivity, proteome and protein phosphorylation of CHO cells, in press.


Figure 1. A: Features of vector pProtA. SV40 transcription promoter, transin signal peptide sequence, proteinA domain, ampicillin resistance gene. A construct, lacking the transmembrane domain of C2GnT was generated by PCR comprising aa 38-429 and cloned into the single EcoRI site of pProtA. B: A construct, lacking the transmembrane domain of Fuc-TIII was generated by restriction cuts of the full-length form comprising aa 44-361 and cloned into XbaI site of pPROM.
Figure 2. Ratio of S/(G0/G1) in adherent CHO cell layers during transition from subconfluency to confluency prior to transfection. Cells were plated to 40'000/cm² and grown for the indicated time. Cells were detached, and processed for cell cycle phase determination as indicated in the Method section. Histograms generated by FACS were analyzed by ModFit (BectonDickinson). Insets show cell density at transfection time by interference contrast (bar = 120 μm).
Figure 3. Time course of cell specific productivity and $S/(G_0/G_1)$ ratio during batch fermentation in spinner flasks of CHO cells stably transfected with pProtA-C2GnT. Cell specific productivity was determined by harvesting cells at the indicated time point, followed by incubation in fresh medium and activity measurements. $S/(G_0/G_1)$ ratio was determined by the propidium iodide method. Each point refers to cells recovered from a single flask.
Figure 4. Cell specific productivity related to the cell cycle phase. Secretion of active C2GnT was measured depending on the cell cycle in spinner flasks. CHO cells were synchronized with 3 mg/l aphidicolin for 14 h, resulting in an arrest at G0/G1-phase. After media change expression was followed by activity measurements and correlated with the cell cycle phase.
Figure 5. Cell specific productivity related to the proportion of cells in S-phase vs G0/G1 phase. CHO cell clones stably expressing pPROM-Fuc-TIII were transfected at subconfluency (A) or confluency (B), respectively. The cells were cultured in parallel T-flasks up to 1 week. At different time points cells were harvested, an aliquot was analyzed for cell cycle phases and another aliquot incubated in fresh medium for determination of cell specific productivity.
Figure 6. Cell specific productivity related to the proportion of cells in S-phase vs G0/G1 phase. CHO cell clones stably expressing pProtA-C2GnT were transfected at subconfluency (A) or confluency (B), respectively, and assessed as described.
**Figure 7.** Different plasmids used in the study: CMV, SV40, SRα are different promoters; transin, IgK and IgM are different signal sequences. In addition two different selection markers zeo (zeoR) and neo (neoR) are used.
9. Manuscript III⁺:

Complete enzymic synthesis of the mucin-type sialyl Lewis x epitope, involved in
the interaction between PSGL-1 and P-selectin

⁺ Manuscript in preparation
Complete enzymic synthesis of the mucin-type sialyl Lewis x epitope, involved in the interaction between PSGL-1 and P-selectin

Steffen Zeng\textsuperscript{1}, Ricardo Gutiérrez Gallego\textsuperscript{2}, Andre Dinter\textsuperscript{1}, Martine Malissard\textsuperscript{1}, Johannis P. Kamerling\textsuperscript{2}, Johannes F.G. Vliegenthart\textsuperscript{2}, Eric G. Berger\textsuperscript{1*}

\textsuperscript{1}Physiologisches Institut, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
\textsuperscript{2}Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, P.O. Box 80 075, NL-3508 TB Utrecht, The Netherlands

Sialyl Lewis x is an established selectin ligand occurring on N- and O-linked glycans. Using a complete enzymic approach starting from p-nitrophenyl N-acetyl-\(\alpha\)-D-galactosaminide (GalNAc(\(\alpha\)-pNp) as core substrate, the sLe\(^x\)-oligosaccharide Neu5Ac(\(\alpha\)2-3)Gal(\(\beta\)1-4)Fuc(\(\alpha\)1-3)GlcNAc(\(\beta\)1-6)Gal(\(\beta\)1-3)GalNAc(\(\alpha\)-pNp, representing the O-linked form, has been synthesized in an overall yield of 32\%. In a first step, Gal(\(\beta\)1-3)GalNAc(\(\alpha\)-pNp was prepared in a yield of 52\% using UDP-Gal and an enriched preparation of \(\beta\)3-galactosyltransferase (EC 2.4.1.122) from rat liver. UDP-GlcNAc and a recombinant affinity-purified preparation of core 2 \(\beta\)6-N-acetylgalactosaminyltransferase (EC 2.4.1.102) fused to Protein A were used to branch the core 1 structure, affording GlcNAc(\(\beta\)1-6)[Gal(\(\beta\)1-3)]GalNAc(\(\alpha\)-pNp in a yield of >85\%. The core 2 structure was galactosylated using UDP-Gal and purified human milk \(\beta\)4-galactosyltransferase 1 (EC 2.4.1.38) (yield of >85\%), then sialylated using CMP-Neu5Ac and purified recombinant \(\alpha\)3-sialyltransferase 3 (EC 2.4.99.X) (yield of 87\%), and finally fucosylated using GDP-Fuc and recombinant human \(\alpha\)3-fucosyltransferase 6 (EC 2.4.1.152) produced in \textit{Pichia pastoris} (yield of 100\%). MALDI TOF mass spectra, and 1D and 2D TOCSY and ROESY \(^1\)H NMR analysis confirmed the obtained structure.

\textsuperscript{1}Both authors contributed equally.
\textsuperscript{*To whom correspondence should be addressed. Phone: ++41 1 635 5070; Fax: ++41 1 635 6814; E-mail: eberger@physiol.unizh.ch
Keywords: Sialyl Lewis x; enzymic synthesis; N-acetylglucosaminyltransferase; fucosyltransferase; galactosyltransferase; sialyltransferase; recombinant glycosyltransferases; P-selectin; PSGL-1.

Abbreviations: ß3-GalT, UDP-Gal:GalNAcα-R ß-1,3-galactosyltransferase; C2GnT, UDP-GlcNAc:Gal(ß1-3)GalNAc(α1-R (GlcNAc to GalNAc) ß-1,6-N-acetylglucosaminyltransferase; ß4-GalT1, UDP-Gal:GlcNAc ß-1,4-galactosyltransferase 1; ST3Gal3, CMP-Neu5Ac:Gal(ß1-4)GlcNAc α-2,3-sialyltransferase 3; α3-FucT6, GDP-Fuc:Gal(ß1-4)GlcNAc (Fuc to GlcNAc) α-1,3-fucosyltransferase 6; Caco, sodium cacodylate (Na(CH$_2$)$_2$ AsO$_2$) - HCl buffer; CIAP, Calf intestinal alkaline phosphatase; core 1, Gal(ß1-3)GalNAc; core 2, GlcNAc(ß1-6)[Gal(ß1-3)]GlcNAc; core 4, GlcNAc(ß1-6)[GlcNAc(ß1-3)]GlcNAc; core 6, GlcNAc(ß1-6)GalNAc; FPLC, Fast Protein Liquid Chromatography; HPLC, High Performance Liquid Chromatography; LacNAc, Gal(ß1-4)GlcNAc; lacto-N-biose, Gal(ß1-3)GlcNAc; MALDI TOF, Matrix Assisted Laser Desorption Ionisation Time Of Flight; MES, 2-(N-Morpholino)ethanesulfonic acid - NaOH buffer; Me$_2$SO, Dimethyl sulfoxide; MLEV, Malcolm Levit; NMR, Nuclear Magnetic Resonance; pNp, p-Nitrophenyl; PSGL-1, P-selectin glycoprotein ligand 1; ROESY, Rotating Frame Nuclear Overhauser Enhancement Spectroscopy; sLe$^x$, Sialyl Lewis x; TOCSY, Total Correlation Spectroscopy; WEFT, Water Eliminated Fourier Transform.
Introduction

Selectin-mediated cell adhesion via the sialyl Lewis x (sLe^x) epitope (Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc(β1-) [1,2] is playing an important role in many diseases, e.g. in cancer [3,4], inflammatory diseases (reviewed in [5,6]), and acute rejections of solid organ transplants [7]. Oligosaccharides containing the sLe^x structure have been demonstrated in vivo [8] and in vitro [9] to inhibit E-, L- and P-selectin-mediated adhesive interactions. Up till now these components as well as their mimetics have been tested and demonstrated to be beneficial in lung injury [10], myocardial ischemia and reperfusion injury [11,12] as well as in the inhibition of angiogenesis [13].

Blocking sLe^x-selectin binding, therefore, represents a highly relevant therapeutic target. To develop selectin blockers, different strategies have been followed. Since the first chemical synthesis of sLe^x in 1991 [14], several elegant organic synthetic routes for this epitope have been described [14-16]. Alternatively, enzymic as well as chemoenzymic approaches using glycosyltransferases and/or glycosidases have been investigated [17,18]. The quite low affinity of simple sLe^x oligosaccharides [19-21] as well as the short life time in the circulation both led to the development of multiple sLe^x-containing structures [22,23] and numerous mimetics (reviewed in [24]), showing enhanced binding properties compared to simple sLe^x [25,26]. While the analogues may be toxic or antigenic, the synthesis of complex glycoconjugates based on naturally occurring selectin ligands seems to be a possible alternative. A promising candidate is the hexasaccharide O-linked to Thr57 of the N-terminus of PSGL-1, one of the best characterized glycoproteins involved in P- and E-selectin binding [27].

Here, we report the first full enzymic synthesis of O-linked core 2 type sLe^x, Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNp, based on the sequential transfer of appropriate monosaccharides from nucleotide analogues using suitable glycosyltransferases as biocatalysts. The identity of this product is confirmed by MALDI TOF mass spectra, 1D and 2D TOCSY and ROESY ^1H NMR analysis.
Materials and methods

Chemicals

All reagents were commercially available and of highest purity unless otherwise noted. UDP-Gal, UDP-GlcNAc, CMP-Neu5Ac, GDP-Fuc, and GalNAc(α1-pNP were purchased from Sigma (Buchs, Switzerland). The corresponding ^14^C-labeled nucleotide sugars were obtained from Amersham International plc (Zürich, Switzerland). D$_2$O was purchased from Isotex (Veenendaal, The Netherlands), HPLC-grade acetonitrile from Rathburn (Walkerburn, Scotland), and ammonium bicarbonate and 6-aza-2-thiothymine from Sigma (Zwijndrecht, The Netherlands). Scintillant Irga-Safe Plus was from Packard (Zürich, Switzerland), and calf intestinal alkaline phosphatase (CIAP) from Boehringer Mannheim (Mannheim, Germany).

Preparation of glycosyltransferases

β3-Galactosyltransferase (β3-GalT)

Following the method of Schachter and Brockhausen [28,29], 40 g of fresh rat liver taken from 7 month old Long evans male were rinsed, minced with scissors and homogenized in 80 ml of 50 mM MES, pH 6.5, containing 250 mM sucrose and 20 mM MgCl$_2$, in a 1 l waring blender (two 20 s burst at high settings, with 40 s rest). After centrifugation at 680 g for 10 min, the pellet was homogenized using a Potter-Elvehjem glass homogenizer with a motor driven pestle by making three passes at 800 rpm. Then, the homogenate was centrifuged for 1 h at 10,000 g, and 22.5 ml of the pellet were extracted five times overnight with intermediate centrifugation at 100,000 g using equal volumes of 50 mM MES, pH 6.5, containing 5 mM MnCl$_2$, 0.02% NaN$_3$, 1% Triton X-100, and 0.1 M NaCl. The highest activity was found in the third extract after 3 days of incubation. All steps were performed at 4 °C.

α3-Fucosyltransferase 6 (α3-FucT6)

Human α3-FucT6 (GenBank Accession number M98825) was produced using the Pichia pastoris expression system from Invitrogen (Groningen, The Netherlands). The generation of the Pichia strain will be described elsewhere. Briefly, a strain was used that secreted 1 U/l of a soluble form of α3-FucT6 into the supernatant, as measured by using N-acetyllactosamine
(LacNAc) as acceptor. The enzyme was enriched by ultrafiltration and affinity-purified on GDP-hexanolamine-agarose.

Other glycosyltransferases
Recombinant mouse core 2 β6-N-acetylglucosaminyltransferase (C2GnT) (GenBank Accession number U19265) was expressed as a Protein A tagged soluble enzyme in CHO cells and purified as described previously [30]. β4-Galactosyltransferase 1 (β4-GalT1) was purified from human milk as described previously [31]. Purified recombinant rat α3-sialyltransferase 3 (ST3Gal3) (GenBank Accession number M97754), produced in Sf9 cells, was obtained as a gift from M. Streiff (Novartis Pharma, Basel, Switzerland).

Enzymic synthesis protocols

The experimental conditions for the various steps in the enzymic synthesis of Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNp are presented in Table 1. All incubations of saccharide acceptors and nucleotide sugar donors in the presence of glycosyltransferases were carried out at 37 °C in a water bath, except for the incubation related to β3-GalT (25 °C).

Yields were determined via parallel incubations with 14C-labeled nucleotide sugars, using 6-13% of the incubation mixtures; 20-μl samples of these mixtures were taken at different time intervals. Reactions were stopped by adding 0.5 ml of ice-cold water. Separations were carried out on a Sep-Pak C18 cartridge (Waters, Milford, USA) mounted on a vacuum chamber equilibrated with 10 ml of MeOH followed by 10 ml of H2O. To remove unreacted radiolabeled nucleotide sugar, the cartridge was washed with 15 ml of H2O. The non-radioactive acceptor and the radiolabeled product were eluted with 5 ml of MeOH, and the eluate was mixed with 10 ml of scintillant Irga-Safe Plus. Typically, cpm values of 1,000 – 5,000 were counted.

The non-radioactive incubation mixtures were separated as mentioned above, using three sequentially coupled Sep-Pak C18 cartridges. In each case, the MeOH phase was concentrated in a Speed vak, and the residues were used directly in the next glycosylation step, except for the fucosylation.
After the sialylation and fucosylation steps, the products were purified via gel filtration and HPLC. Reaction mixtures were desalted on Sephadex G-25 (FPLC system; HiTrap, 4x5 ml bedvolume; Pharmacia, Uppsala, Sweden) using 5 mM ammonium bicarbonate as eluent at a flow rate of 1.5 ml/min, and subsequent lyophilisation. Effluents were monitored by UV at 214 nm, and conductivity. Sephadex G-25 fractions were further purified by HPLC on a ChromSpher 5 C8 reversed phase column (10 x 250 mm, Chrompack, Bergen op Zoom, The Netherlands) at a flow rate of 2.0 ml/min using a Kratos SF 400 HPLC system (ABI Analytical, Kratos Division). The column was eluted isocratically with solvent A (aqueous 80% acetonitrile) during 5 min, followed by a gradient from 100% solvent A - 0% solvent B (aqueous 20% acetonitrile) to 65% solvent A - 35% solvent B in 18 min. The effluents were monitored at 280 nm using a 757 absorbance detector (ABI Analytical, Kratos Division). The collected fractions were immediately lyophilised for further analysis.

Mass Spectrometry

Negative-ion mode MALDI TOF mass spectrometric analysis of the products was performed on a Voyager-DE (PerSeptive Biosystems) instrument operating at an accelerating voltage of 22 kV (grid voltage 92%, ion guide wire voltage 0.1%) and equipped with a VSL-337ND-N2 laser. The sample was dissolved in bidistilled water (1 μg/μl) and subsequently mixed in the sample well with 6-aza-2-thiothymine (10 mg/ml in water:acetonitrile 1:1, v/v) at a ratio of 1:3. Linear mass scans were recorded over 1000 dalton using a pulse delay time of 90 ns. Recorded data were processed using GRAMS/386 software (v. 3.04, Galactic Industries Corporation).

NMR Spectroscopy

Prior to analysis the reaction products were repeatedly exchanged in 2H2O (99.9 atom % 2H) with intermediate lyophilisation and finally dissolved in 450 μl 2H2O (99.96 atom % 2H). Resolution-enhanced 1H 1D and 2D NMR spectra were recorded on Bruker AMX-500, DRX-500 or DRX-600 (Department of NMR Spectroscopy, Utrecht University) spectrometers, at probe temperatures of 300 K. Chemical shifts (δ) were expressed in parts per million relative to internal acetate (δ 1.908; acetone δ 2.225). HO2H signal suppression was achieved by
applying a WEFT pulse sequence [32] in 1D $^1$H experiments and by presaturation for 1 s in 2D experiments. 2D TOCSY spectra were recorded by using MLEV-17 mixing sequences [33] with effective spin-lock times between 20 and 100 ms. 2D ROESY [34] spectra were recorded with a mixing time of 250 ms. The spin-lock field strength corresponded to a 90° pulse of about 120 $\mu$s. A 2D 600 MHz off-resonance ROESY spectrum of the final product was recorded according to [35]. The spin-lock field strength corresponded to a 90° pulse of about 115 $\mu$s. $^1$H 1D and 2D spectra were processed on Silicon Graphics IRIS work stations (Indigo 2 and O2) using XINSF2 software (Bijvoet Center, Department of Bio-Organic Chemistry).
Results

Enzymic synthesis of the mucine-type sialyl Lewis x epitope

A survey of the strategy and the amounts of donors, acceptors, and glycosyltransferases used to build up the O-linked hexasaccharide, Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNp, bearing the sLe\(^x\) epitope, are presented in Table 1. Yields were determined via parallel incubations with radioactive nucleotide sugar donors (see Materials and methods). During preliminary studies, the individual glycosyltransfer reactions were optimized, and the products Gal(β1-3)GalNAc(α1-pNp (1), Gal(β1-4)GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNp (2), and Gal(β1-4)[Fuc(α1-3)]GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNp (3) were purified for characterization by mass spectrometry and \(^1\)H NMR spectroscopy (Table 2). Therefore, in the complete enzymic synthesis, as described below, the steps 1-4 were carried out without intermediate chromatographic purifications and detailed structural analysis.

GalNAc(α1-pNp was chosen as a starting core substrate. This compound was elongated to yield Gal(β1-3)GalNAc(α1-pNp using UDP-Gal and a rat liver β3-GalT (EC 2.4.1.122; 0.5 U/g) Triton X-100 extract. To increase yield, calf intestinal alkaline phosphatase (EC 3.1.3.1; 2500U/mg) was added [36]. Even though the activity at 25 °C was only 67% compared to an incubation at 37 °C, 25 °C was chosen because of a better stability of the transferase. Furthermore, it was found that at least 17% of UDP-Gal was degraded within 24 h. For these reasons, fresh UDP-Gal and enzyme preparation were added after 3 days of incubation. At this time (step 1a) the product yield was 29%; 3 days later (step 1b) an overall yield of 52% was reached.

For the branching of the core 1 structure Gal(β1-3)GalNAc(α1-pNp, yielding the core 2 structure, UDP-GlcNAc and purified mouse recombinant C2GnT (EC 2.4.1.102) were applied (step 2). Mouse C2GnT was expressed as a Protein A fusion in CHO cells [30]. GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNp was obtained in a yield of 85% after 32 h of incubation.

The core 2 structure was further elongated at O4 of GlcNAc using UDP-Gal and purified human milk β4-GalT1 (step 3). Under the applied conditions Gal(β1-4)GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNp was generated in a yield of >85% after 5 h of incubation.
To sialylate the tetrasaccharide, CMP-Neu5Ac and purified recombinant rat ST3Gal3 were used [37]. In two sequential incubations of 16 h and 5 h (step 4; for conditions, see Table 1) Neu5Ac(α2-3)Gal(β1-4)GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNp was synthesized in a total yield of 87%. For the purification and 1H NMR analysis (see below) use was made of HiTrap chromatography (Figure 1a), which yielded two carbohydrate-containing fractions. The two fractions, denominated F1 and FII, respectively, were further purified by reversed phase HPLC (Figure 1b) yielding the carbohydrate material in the void volume (fractions F1a (compound 5) and FIIa (compound 4)).

Finally, the pentasaccharide (4) was fucosylated with GDP-Fuc in the presence of recombinant enriched human α3-FucT6 (EC 2.4.1.152) (step 5), affording the sLex epitope-containing title compound 6 after 6 h of incubation in a yield of 100%, as observed by HPLC and supported by a radioactive parallel incubation. In contrast to insect cell supernatant [38] no exoglycosidase degrading the acceptor substrate for α3-FucT6 could be detected.

Taken together, the mucine-type sLex-containing hexasaccharide could be prepared in an overall yield of 32%.

Structural analysis of synthesized products 4 and 6

The negative-ion mode MALDI TOF mass spectrum (Figure 2a) of compound 4 showed one intense signal at m/z 1159.4, corresponding to the deprotonated pseudo-molecular ion of a pentasaccharide with the brutoformula Neu5AcHex2HexNAc2-pNp.

The 1D 1H NMR spectrum (Figure 3a) of compound 4 revealed three main resonances downfield of the HO2H signal (δ 4.766) at δ 8.307 (Jm,0 9.0 Hz), 7.269 (Jm,0 9.0 Hz) and 5.805 (Jm,1 4.0 Hz). The two most downfield signals were assigned to the m- and o-protons of the p-nitrophenyl aglycon [39], respectively, while the remaining resonance was attributed to the anomeric proton of GalNAc (A; pyranose ring form), α-glycosidically linked to the p-nitrophenyl aglycon. Upfield of the HO2H resonance three additional anomeric resonances at δ 4.538 (Jp,1 7.5 Hz), 4.467 (Jp,1 7.5 Hz) and 4.447 (Jp,1 7.5 Hz) were identified and assigned to β-1,3-linked Gal3 (B; pyranose ring form), β-1,6-linked GlcNAc (C; pyranose ring form), and β-1,4-linked Gal4 (D; pyranose ring form), respectively (compare with compounds 1-3 in Table 2 and literature data [39]). The three N-acetyl resonances representing three protons each at δ 2.030, 2.003 and 1.899 could be attributed to Neu5Ac, GalNAc and...
GlcNAc, respectively [39]. The presence of only one Neu5Ac residue (E) in this oligosaccharide was confirmed by comparing the intensities of the H3e and H3a resonances at δ 2.753 and 1.790, respectively, with those of the discrete signals of the p-nitrophenyl moiety.

By means of 2D TOCSY most of the resonances present in the 1D spectrum could be identified (Table 2, compound 4). In the TOCSY spectrum (100 ms, not shown) the anomeric track of the Gal4 residue D revealed three cross-peaks at δ 3.540 (H2), 4.097 (H3), and 3.951 (H4). When compared to the corresponding signals of precursor 2, these resonances showed downfield shifts of 0.028 ppm for Gal4 H2, 0.447 ppm for Gal4 H3, and 0.033 ppm for Gal4 H4, indicating that Neu5Ac is linked at O3 of Gal4. The anomeric track of the Gal5 residue B revealed three cross-peaks at δ 3.553 (H2), 3.654 (H3), and 3.928 (H4), in agreement with a terminal position for this residue (c.f. compound 2, Table 2), thereby excluding a possible sialylation of the Gal5 residue (compare with compound 2). The combined MS and NMR results justify the conclusion that oligosaccharide 4 has the structure Neu5Ac(α2-3)Gal(β1-4)GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNp.

In the 1D 1H NMR spectrum (Figure 3b) of fraction Fla, the carbohydrate originated resonances were virtually identical to those of fraction FlIa, with the exception that integration of the Neu5Ac-H3a, -H3e, and N-acetyl resonances demonstrated the presence of 2 Neu5Ac residues. The second Neu5Ac residue (F) was presumably α-2,3-linked to the β-1,3-linked galactose residue B. The amount of disialylated oligosaccharide was determined by UV (280 nm) absorption to be 12.76 % (with respect to the total amount of oligosaccharide recovered). This observation is in agreement with Kono et al. [40] who demonstrated a ST3Gal3-side activity on the core 1 galactose.

The negative-ion mode MALDI TOF mass spectrum (Figure 2b) of compound 6 showed one major peak at m/z 1305.0 corresponding to the deprotonated pseudo-molecular ion of a hexasaccharide with the brutiformula Neu5AcHex2dGalHexGalNAc-pNp.

The 1D 1H NMR spectrum (Figure 3c) of compound 6 revealed five anomeric signals at δ 5.807 (J1,2 4.0 Hz), 5.030 (J1,2 4.0 Hz), 4.545 (J1,2 7.5 Hz), 4.482 (J1,2 7.5 Hz), and 4.428 (J1,2 7.5 Hz) which could be identified on guidance of the NMR data of compounds 1-4 (Table 2). The additional anomeric resonance at δ 5.030 in compound 6 belongs to the incorporated α-Fuc residue (G; pyranose ring form). 2D TOCSY NMR spectroscopy allowed the identification of most of the signals (Table 2). The introduction of Fuc at O3 of GlcNAc
(C) resulted in distinct downfield shifts of GlcNAc H1 (0.015 ppm), H2 (0.181 ppm), H3 (0.272 ppm), and H4 (0.082 ppm), a feature which is also observed when the NMR data of compounds 2 and 3 are compared (Table 2).

In order to confirm the various glycosidic linkages in compound 6 a 2D off-resonance ROESY (Figure 4) experiment was performed. The interresidual cross-peaks between Gal3 H1 and GalNAc H3 (δ 4.545/4.297), GalNAc H6a and GlcNAc H1 (δ 3.763/4.482), Fuc H1 and GlcNAc H3 (δ 5.030/?????), and between Gal4 H1 and GlcNAc H4 (δ 4.428/3.745) proved the presence of the Gal(β1-3)GalNAc, GlcNAc(β1-6)GalNAc, Fuc(α1-3)GlcNAc, and Gal(β1-4)GlcNAc linkages, respectively. It should be noted that the interresidual cross-peak between Neu5Ac H3e and Gal4 H3 as reported by Ball et al. [41] was not observed in our study. The combined MS and NMR data indicate that oligosaccharide 6 has the structure Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc(β1-6)[Gal(β1-3)]GlcNAc(α1-pNp.
Discussion

The production of selectin blockers seems to be a rewarding task with respect to several therapeutic targets, such as inflammatory diseases, transplant rejection and metastasis. This new class of anti-adhesive compounds can be enzymically generated by taking advantage of the high stereo- and regioselectivity of glycosyltransferases. This approach which ascertains an easy access to the naturally occurring PSGL-1 type sLe\(^\alpha\) epitope and the disialylated equivalent, in relatively high yields, opens the possibility for the synthesis of glycopeptides by making use of the four cloned animal polypeptide GalNAc-transferases [42-46]. The structural identity with the natural ligand would not only render these compounds highly compatible in the competition with PSGL-1, but also assure a prolonged life-time and reduced antigenicity when compared with the non-natural ligands.

To demonstrate the suitability of glycosyltransferases for the synthesis of complex oligosaccharides, we used GalNAc(\(\alpha1\)-pNp as the starting substrate for the elongation of O-linked glycans. As shown previously [30] the branching C2GnT has a strict requirement for a 3-substitution of GalNAc. Even though a \(\beta3\)-GalT, involved in the GM1/GD1 synthesis, transferring Gal to lipid-linked GalNAc [47], has been cloned, yet no recombinant enzyme is available elongating peptide-bound GalNAc. Recently, an elegant and efficient method (yield >90%) to attach GlcNAc to O-linked GalNAc using \(\beta\)-galactosidase from bovine testis, thereby circumventing the need of \(\beta3\)-GalT, has been published [48]. Since this synthesis ended up with a mixture of core 2 and core 6 (GlcNAc(\(\beta1\)-6)GalNAc) structures, and can only be driven towards high yields of the core 6 structure, we preferred the use of a crude \(\beta3\)-GalT preparation. The rapid degradation of UDP-Gal in our system correlates well with the observation [49] that UDP-Gal decomposes rapidly in the presence of Mn\(^{2+}\), a metal ion needed to maintain \(\beta3\)-GalT activity. Therefore, a two step incubation was carried out, resulting in a final yield of 52% for Gal(\(\beta1\)-3)GalNAc(\(\alpha1\)-pNp (core 1 structure).

To date three different C2GnTs, of use for the branching of the core 1 structure, have been cloned [50-52]. Two of the three transferases [51,52] are also capable of synthesizing the core 4 structure (GlcNAc(\(\beta1\)-6)(GlcNAc(\(\beta1\)-3))GalNAc). As shown recently, the core 1 disaccharide can alternatively be branched with the \(\beta\)-N-acetyl-D-hexosaminidase from Nocardia orientalis [53]; unfortunately, the resulting yield is only around 6%. Applying a crude mouse kidney C2GnT preparation, the branching of a core 1 structure was realised in a
yield of 74% [54]. By the use of a recombinant purified mouse C2GnT we achieved a yield of >85% for GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNp (core 2 structure).

The core 2 structure can be galactosylated at O4 of GlcNAc by various β4-GalTs. By now, a whole family of β4-GalTs has been identified and cloned (reviewed in [55]). Although β4-GalT1 has been reported to be inefficient for the elongation of the core 2 structure [56], the purified human milk enzyme, known as β4-GalT1, was successfully applied to extend the core 2 structure, yielding Gal(β1-4)GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNp. In spite of the earlier reported data that a strong inhibition of β4-GalT1 should exist for core 2 acceptor concentrations of 5 mM [54], we were able to reach a yield of >85% with 200 mU of enzyme at an acceptor concentration of 10 mM. It should be noted that within 5 h of incubation we did not achieve the earlier reported 100% yield with the enzyme available from Sigma corporation [54].

The tetrasaccharide as generated above, was further extended by sialic acid using ST3Gal3. Out of the four to date cloned enzymes only ST3Gal3 and ST3Gal4 have been demonstrated to sialylate N-acetyllactosamine (LacNAc) [40]. In our studies we used purified recombinant rat ST3Gal3, despite the preference of this enzyme for Gal(β1-3)GlcNAc (lacto-N-biose). Both ST3Gal3 and ST3Gal4 have a side activity in sialylating the Gal residue in a core 1 structure [40]. However, Gal(β1-4)GlcNAc(β1-6)[Neu5Ac(α2-3)Gal(β1-3)]GalNAcα1-pNp was not found in our incubation mixture. This indicates that the side activity is strictly confined to the absence of a Gal(β1-4)X and the presence of a Gal(β1-3)X acceptor. Using a similar substrate in an earlier study, a yield of 68% was obtained with the partially purified recombinant rat ST3Gal3 [54]. Here, we obtained Neu5Ac(α2-3)Gal(β1-4)GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNp in a yield of 87% via a two step incubation using 170 and 72 mU of the purified enzyme.

To obtain the sLe^x epitope, the foregoing structure was fucosylated at O3 of GlcNAc. At present five cloned α3-FucTs (α3-FucT3 - α3-FucT7), displaying different acceptor specificities, are known. α3-FucT4 nearly exclusively fucosylates LacNAc, whereas α3-FucT7 only acts on (α2-3)-sialylated LacNAc [57,58]. α3-FucT6, used in this study, fucosylates both sialylated and non-sialylated LacNAc. This enzyme, which has been shown to be highly active in vivo [59], is also of interest for the preparation of difucosyl sLe^x structures on poly-LacNAc chains [60]. In a 6 h incubation with 70 mU of recombinant α3-
FucT6, Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNp was synthesized in 100% yield.

In summary, we performed the first full enzymic synthesis of a core 2 type sLe^x- containing hexasaccharide in an overall yield of about 32%. Once recombinant β3-GalT forming the core 1 structure will become available, scaling-up of the procedure described here should be feasible.

Acknowledgments

This work was supported by grant BIO4-CT95-0138 ‘Engineering O-Glycosylation’ of the EU to EGB/JFGV and grant 5002-46084 (SPP-Biotech) of the Swiss National Science Foundation to EGB. The authors also thank Markus Streiff (Novartis Pharma, Basel, Switzerland) and Frank Bootz (University of Zürich, Switzerland) for providing us with ST3Gal3 and fresh rat liver, respectively.
References


Figure 1
Purification of the reaction mixture after Neu5Ac incorporation: 1a represents the HiTrap desalting profile, 1b represents the HPLC (C8) profile of HiTrap fraction 2 (FII), and 1c represents the HPLC (C8) profile of HiTrap fraction 1 (FI).
Figure 2
MALDI TOF mass spectra (negative-ion mode) of (a) Neu5AcHex3HexNAc2-pNp, and (b) Neu5AcHex3dHexHexNAc2-pNp.
Figure 3

One-dimensional 500 MHz or 600 MHz $^1$H NMR spectra of (a) Neu5Ac($\alpha$2-3)Gal($\beta$1-4)GlcNAc($\beta$1-6)[Gal($\beta$1-3)]GalNAc($\alpha$1-pNp 4, (b) Neu5Ac($\alpha$2-3)Gal($\beta$1-4)GlcNAc($\beta$1-6)[Neu5Ac($\alpha$2-3)Gal($\beta$1-3)]GalNAc($\alpha$1-pNp, and (c) Neu5Ac($\alpha$2-3)Gal($\beta$1-4)[Fuc($\alpha$1-3)]GlcNAc($\beta$1-6)[Gal($\beta$1-3)]GalNAc($\alpha$1-pNp 6 in D$_2$O at 300 K (referenced to internal acetone ($\delta$ 2.225)).
Figure 4
Two-dimensional 600 MHz ROESY spectrum (mixing time 200 ms) of Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc(β1-6)[Gal(β1-3)]GalNAc(α1-pNP in D2O at 300 K (referenced to internal acetone (δ 2.225).
Table 2. 500 and 600-MHz $^1$H-NMR chemical shifts of pNp-oligosaccharides recorded at 300K referenced to internal acetate $\delta$ 1.908 (acetone $\delta$ 2.225). $\triangleright$=GalNAc, $\bullet$=GlcNAc, $\blacksquare$=4-Gal, $\blacklozenge$=3-Gal, $\box$=Fuc, and $\triangle$=Neu5Ac.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>pD</td>
<td>8.272 (9.0) 8.306 (9.0) 8.304 (9.0) 8.307 (9.0) 8.307 (9.0) 8.310 (9.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pD</td>
<td>7.281 (9.0) 7.270 (9.0) 7.265 (9.0) 7.269 (9.0) 7.271 (9.0) 7.270 (9.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>5.831 (4.0) 5.810 (4.0) 5.805 (4.0) 5.805 (4.0) 5.807 (4.0) 5.807 (4.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-6a</td>
<td>3.738      3.771      3.730      3.776      n.d.      3.763</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-6b</td>
<td>3.762      3.999      3.989      3.990      3.990      3.994</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ac</td>
<td>2.013      2.004      2.002      2.003      2.004      2.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>4.472 (7.5) 4.485 (7.5) 4.467 (0.010) 4.468 (0.010) 4.482 (0.012)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-6b</td>
<td>3.904      3.917      3.971      3.971      3.971      3.971</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Ac</td>
<td>1.901      1.886      1.899      1.900      1.890      1.890</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-j</td>
<td>4.386 (7.5) 4.370 (7.5) 4.447 (7.5) 4.449 (7.5) 4.428 (7.5) 4.428 (7.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td>3.512      3.468      3.540      3.539      3.504      3.504</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-3</td>
<td>3.650      3.629      4.097      4.098      4.069      4.069</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-4</td>
<td>3.918      3.889      3.951      3.954      3.935      3.935</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-5</td>
<td>n.d.      3.697      n.d.      3.717      n.d.      3.717</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j</td>
<td>4.548 (7.5) 5.541 (7.5) 4.539 (7.5) 4.538 (7.5) 4.540 (7.5) 4.545 (7.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-3</td>
<td>3.657      3.654      3.639      3.654      3.646      3.646</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-4</td>
<td>3.932      3.933      3.927      3.928      3.937      3.937</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-5</td>
<td>n.d.      n.d.      n.d.      3.70?      n.d.      3.70?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-6a</td>
<td>n.d.      n.d.      n.d.      3.717      n.d.      3.717</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>j</td>
<td>-         -         5.034 (4.0)  -         -         5.030 (4.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td>-         -         3.651      -         -         3.645</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-3</td>
<td>-         -         3.771      -         -         3.763</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-4</td>
<td>-         -         3.849      -         -         3.849</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-5</td>
<td>-         -         4.776      -         -         4.777</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-6</td>
<td>-         -         1.151 (6.5) -         -         1.147 (6.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3a</td>
<td>-         -         -         1.790      1.790      1.789</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-3c</td>
<td>-         -         -         2.753      2.753      2.763</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-4</td>
<td>-         -         -         3.670      3.667      3.667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-5</td>
<td>-         -         -         3.834      3.848      3.848</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-6</td>
<td>-         -         -         3.836      3.836      3.836</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-7</td>
<td>-         -         -         n.d.      3.685R²  3.685R²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-8</td>
<td>-         -         -         n.d.      3.889R²  3.889R²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-9a</td>
<td>-         -         -         n.d.      3.865R²  3.865R²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-9b</td>
<td>-         -         -         n.d.      3.649R²  3.649R²</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ac</td>
<td>-         -         -         2.030      2.030      2.036</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1. Assay conditions

<table>
<thead>
<tr>
<th>Step</th>
<th>Acceptor</th>
<th>Donor</th>
<th>Transferase</th>
<th>Temperature (°C)</th>
<th>Assay components</th>
<th>Time (h)</th>
<th>Step Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>15 mM (= 10.3 mg) GalNAc-α-pNp</td>
<td>20 mM (= 24.4 mg) UDP-Gal</td>
<td>enriched rat liver β3-GalT (5.6 mU)</td>
<td>25</td>
<td>1% Triton X-100 10% Me₂SO 0.1 M MES pH 6.5 10 mM MnCl₂ 60 U CIAP</td>
<td>72</td>
<td>29 %*</td>
</tr>
<tr>
<td>1b</td>
<td>10 mM (= 10.3 mg) GalNAc-α-pNp</td>
<td>20 mM (= 24.4 mg) UDP-Gal</td>
<td>enriched rat liver β3-GalT (3 mU)</td>
<td>25</td>
<td>1% Triton X-100 10% Me₂SO 0.1 M MES pH 6.5 10 mM MnCl₂ 60 U CIAP</td>
<td>72</td>
<td>23 %*</td>
</tr>
<tr>
<td>2</td>
<td>20 mM (= 7.8 mg) Core 1-α-pNp</td>
<td>25 mM (= 12.6 mg) UDP-GlcNAc</td>
<td>recombinant mouse C2GnT (64 mU)</td>
<td>37</td>
<td>0.1 M MES pH 6.5 10 mM MgCl₂</td>
<td>32</td>
<td>85 %*</td>
</tr>
<tr>
<td>3</td>
<td>10 mM (= 8.7 mg) Core 2-α-pNp</td>
<td>15 mM (= 11.3 mg) UDP-Gal</td>
<td>human milk β4-GalT1 (200 mU)</td>
<td>37</td>
<td>0.1 M Tris-HCl pH 7.4 10 mM MgCl₂</td>
<td>5</td>
<td>85 %*</td>
</tr>
<tr>
<td>4a</td>
<td>10 mM (= 8.3 mg) Galβ1-4Core 2-α-pNp</td>
<td>15 mM (= 8.8 mg) CMP-Neu5Ac</td>
<td>recombinant rat ST3Gal3 (170 mU)</td>
<td>37</td>
<td>0.1 M Caco pH 6.8 10 mM MgCl₂ 2 mM CaCl₂ 20 U CIAP</td>
<td>16</td>
<td>27 %*</td>
</tr>
<tr>
<td>4b</td>
<td>10 mM (= 7.7 mg) Galβ1-4Core 2-α-pNp</td>
<td>15 mM (= 8 mg) CMP-Neu5Ac</td>
<td>recombinant rat ST3Gal3 (72 mU)</td>
<td>37</td>
<td>0.1 M Caco pH 6.8 10 mM MgCl₂ 2 mM CaCl₂ 20 U CIAP</td>
<td>5</td>
<td>60 %*</td>
</tr>
<tr>
<td>5</td>
<td>3.5 mM (= 2 mg) Neu5Acα2-3Galβ1-4Core 2-α-pNp</td>
<td>20 mM (= 6 mg) GDP-Fuc</td>
<td>recombinant human α3-FucT6 (70 mU)</td>
<td>37</td>
<td>0.1 M Caco pH 6.8 10 mM MgCl₂ 2 mM CaCl₂ 20 U CIAP</td>
<td>6</td>
<td>100 %*</td>
</tr>
</tbody>
</table>

*calculated by radioactive labeled parallel incubations. Yields are given on the basis of acceptor used for the specific reaction. †calculated from the HPLC profile
10. References


11. Abbreviations

Ala Alanin
Asn Asparagine
BLAST Basic Local Alignment Search Tool
cbd carbohydrate binding domain
CD cluster of differentiation
CHO Chinese Hamster Ovary
core1-8 see table 1 page 11
CR consensus repeats
Cys Cysteine
e.g. example given
EGF epidermal growth factor like domain
ELISA enzyme linked immunosorbent assay
ER Endoplasmic reticulum
ESL-1 E-selectin ligand-1
EST expressed sequence tags
Fuc Fucose
FucT Fucosyltransferase
G0-phase gap phase 0, resting phase G0 can be entered from G1
G1-phase gap phase 1, between M-phase and S-phase
G2-phase gap phase 2, between S-phase and M-phase
GA Golgi apparatus
Gal Galactose
GalT Galactosyltransferase
GalNAc N-acetylgalactosamine
Glc Glucose
GlcA Glucuronsäure
GlcNAc N-acetylglucosamine
GlcNAcT N-acetylglucosaminyltransferase
Gly Glycine
GlyCAM-1 glycosylation dependent cell adhesion molecule-1
GnT N-acetylgalcosaminyltransferase
h hours
HPLC High Performance Liquid Chromatography
i.e. id est
MAdCAM-1 mucosal adresin cell adhesion molecule-1
MALDI TOF Matrix Assisted Laser Desorption Ionisation Time Of Flight
Man Mannose
MS Mass spectrometry
Neu5Ac Sialic acid
NMR Nuclear magnetic resonance
pNp para-Nitro-phenyl
pp-GalNAcT UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase
Pro Proline
PSGL-1 P-selectin glycoprotein ligand-1
ROESY Rotating Frame Nuclear Overhauser Effect Spectroscopy
Ser Serine
SiaT Sialyltransferase
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>sLeα</td>
<td><em>sialyl Lewis a</em> for the structure see table 1 page 11</td>
</tr>
<tr>
<td>sLeX</td>
<td><em>sialyl Lewis X</em> for the structure see table 1 page 11</td>
</tr>
<tr>
<td>S-phase</td>
<td>DNA synthesis phase</td>
</tr>
<tr>
<td>ST3GalX</td>
<td>α3-sialyltransferase No. X incorporating Neu5Ac to Gal</td>
</tr>
<tr>
<td>Thr</td>
<td>Threonine</td>
</tr>
<tr>
<td>TOCSY</td>
<td>Total Correlation Spectroscopy</td>
</tr>
<tr>
<td>type I - III</td>
<td>see table 1 page 11</td>
</tr>
<tr>
<td>Xyl</td>
<td>Xylose</td>
</tr>
</tbody>
</table>
12. Curriculum vitae

Name: Steffen Zeng
Date of Birth: 13. January 1970
Place of Birth: Hamburg, Germany
Citizenship: German
Marital status: married, one son

School education

1976 - 1980 Primary school in Hamburg
1980 - 1989 Gymnasium in Hamburg

University

1989 – 1995 Studies in Biotechnology at the Technical University of Berlin, finishing with a degree in technical science
1996 - 1999 Doctoral thesis at the Institute of Physiology, University of Zürich under the supervision of Prof. Dr. J.E. Bailey, Institute of Biotechnology, Swiss Federal Institute of Technology (ETH) and Prof. Dr. E.G. Berger Institute of Physiology, University of Zurich.