

## Review

# Hans Paulsen: Contributions to the Investigations of Glycoprotein Biosynthesis

Inka Brockhausen

Department of Biomedical and Molecular Sciences, Queen's University, Kingston, ON K7L 3N6, Canada;  
brockhau@queensu.ca

## Abstract

Hans Paulsen was one of the first scientists who believed that chemistry should be applied to biology and medicine. His interest in natural products and their roles solidified in the 1970s. He passed on his knowledge to hundreds of students and coworkers and advanced science with many national and international collaborators. No matter where he was, at home or travelling, he was always curious and keen to learn, from chemistry to enzymes, their roles in diseases, and the possible applications of synthetic compounds. His creative chemistry and synthesis of novel compounds made essential contributions to elucidating the mechanisms and pathways of glycoprotein biosynthesis. This review describes the biosynthetic pathways of the O- and N-glycans of glycoproteins and studies of novel substrates and inhibitors developed by Hans Paulsen's group.

**Keywords:** glycosyltransferases; glycoproteins; N-glycosylation; mucins; O-glycosylation; substrates; inhibitors; glycopeptides

## 1. Introduction

Hans Paulsen was a highly respected and successful carbohydrate chemist, valued and admired by his many coworkers, students, and collaborators. He was extremely knowledgeable in chemistry and science, but also in other cultures, politics, and history. Hans Paulsen was kind and compassionate and always interested until his passing in 2024 at the age of 102. His major strengths included an incredible memory and his ability to give informative and humorous lectures. When he developed a research plan, his excitement was contagious, and his brilliance stimulated his students to do well. He was one of the first scientists who believed that chemistry should be applied to biology and medicine. His interest in natural products and their roles solidified in the 1970s. He passed on his knowledge to hundreds of students and coworkers, and advanced science with many national and international collaborators. Paulsen visited many countries and studied their cultures and history. He enjoyed visits to Toronto, Amherst Island, and the Northern wilderness of Canada. The natural environment was conducive to making plans and discussing the value of current and future experiments and approaches to solving biochemical problems. The wilderness did not distract him from learning about enzymes, their roles in diseases, and their possible applications in solving problems through biochemistry. Always curious and full of humour, politeness, and modesty, he was a real pleasure to collaborate with.

Knowledge of the enzymes and mechanisms involved in transferring sugar residues to proteins was minimal in the 1980s. Hans Paulsen's tremendous efforts and many contributions were critical for the advancement in the field of glycoprotein biosynthesis.



Academic Editor: Zhongwu Guo

Received: 25 July 2025

Revised: 8 September 2025

Accepted: 9 September 2025

Published: 14 September 2025

**Citation:** Brockhausen, I. Hans Paulsen: Contributions to the Investigations of Glycoprotein Biosynthesis. *Molecules* **2025**, *30*, 3735. <https://doi.org/10.3390/molecules30183735>

**Copyright:** © 2025 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

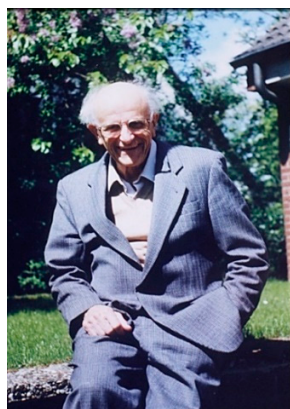
Accounts of Hans Paulsen's long and successful life as a carbohydrate chemist professor and emeritus have been published in journals and symposia. His main interests were carbohydrate chemistry, synthesis of unusual carbohydrate residues and linkages, selective oxidation, and specific modifications and substitutions of carbohydrate derivatives. This led to many glycan derivatives that proved to be valuable in determining enzyme specificities for their carbohydrate substrates and biosynthetic pathways.

Multiple novel chemical reactions were developed by Paulsen that established the groundwork for further chemistry of carbohydrates. While working with Kurt Heyns, Paulsen accomplished the catalytic oxidation of unusual carbohydrates [1]. He studied the role of sugar configuration and conformation in catalysis using platinum contact [2]. A major effort was made in the synthesis of carbohydrates containing nitrogen in the ring and their conformational analyses by NMR [3,4]. Many of these compounds were later shown to be potent glycosidase inhibitors [5]. Conformational analyses by NMR were developed and became standard analyses for carbohydrates [6]. He contributed to an understanding of the exo-anomeric effect and the dynamics of carbohydrate conformations. The chemistry of acyloxonium ion rearrangements was applied to many different compounds [7,8], including phosphate-containing carbohydrates [9]. These studies were extended to the synthesis and analysis of glycoprotein fragments [10]. Carbohydrates were analyzed in solution and in crystal form [11,12], contributing to our knowledge of carbohydrate structures and their biological roles.

Paulsen's efforts to advance the analyses of carbohydrate conformations were helpful in delineating the substrate binding of glycosyltransferases (GTs) and in designing specific substrate analog inhibitors. He also performed analyses of natural products, which turned out to be glucosidase and mannosidase inhibitors. These compounds could be valuable for therapy development in diseases associated with alterations of specific glycans or for controlling the immune system that relies on glycan recognition. The advancements in the structural analysis of glycans were crucial to the success of synthetic chemistry and its application in glycobiology and medicine.

The intensive efforts of the Paulsen lab succeeded in the synthesis of specific glycan derivatives for studies of glycobiology and the role of glycoproteins in cancer and immunology. Many other results of Paulsen's research led to knowledge of bacterial lipopolysaccharides and the production of specific antibodies and glycoprotein antigens. Thus, Paulsen's ideas and efforts in chemical synthesis and refined methods of structural analysis advanced our knowledge, especially in cancer cell biology. The wealth of synthetic compounds led to new knowledge of glycosyltransferase (GT) specificity, pathways, and glycoprotein functions. Compounds representing specific GT substrates were synthesized [13], allowing assays of enzyme preparation from cells or tissues expressing multiple GTs [14].

This review describes the biosynthesis of N-glycans and mucin type O-glycans by human enzymes and summarizes the work done by the Paulsen group to use synthetic chemistry for the acquisition of knowledge of biosynthetic enzymes and insight into the role of glycoproteins in health and disease.



## 2. Glycosyltransferases

The enzymes that build complex sugar chains of glycoproteins are glycosyltransferases (GTs), which have been classified into 138 families by the CAZy database (as of July 2025), based on their amino acid sequence, predicted 3-dimensional fold, mechanism, and known activity. Inverting GTs invert the anomeric configuration of the sugar in the donor substrate to form the opposite linkage in the reaction product, while retaining GTs retain this linkage. The nucleotide sugar donor substrates for glycoproteins are UDP- $\alpha$ GlcNAc, UDP- $\alpha$ Gal, UDP- $\alpha$ GalNAc, GDP- $\alpha$ Man, GDP- $\beta$ Fuc and CMP- $\beta$ Sialic acid. Dolichol (Dol) is used for membrane-bound intermediates in the endoplasmic reticulum (ER)—in particular, Dol-P-Man $\beta$  and Dol-P-Glc $\beta$ . The GT protein folds are GT-A, GT-B, and GT-C (CAZy), which have been confirmed with a limited number of crystal structures (Table 1). Table 1 lists the major GTs involved in the biosynthesis of N-glycans and mucin-type O-glycans.

**Table 1.** List of major human glycosyltransferases and glycosidases involved in glycoprotein biosynthesis. The GT names and accession numbers are derived from the Uniprot database. Included here are also the major glycoside hydrolases (GHs glucosidases and mannosidases) that are essential for the maturing of N-glycans in the ER and Golgi. \*GTs with known crystal structure from human or animal sources; GALNT1,2,3,4,7,10,12 have been crystallized. CAZy GT, Carbohydrate Active Enzyme database classification; Dol, dolichol; GnT, GlcNAc-transferase; P, phosphate; Sia, sialyl; T, transferase. GT Mechanisms and Folds, and CAZy GT classification: Inverting (11,14,16,17,18,29,31,54), Inverting GT-A (2,7,12,13), Inverting GT-B (1,10,23,33), Inverting GT-C (22,57,58,59), Retaining (32), Retaining GT-A (6,27), Retaining GT-B (4).

| Enzyme Name               | Uniprot | CAZy GT | Donor      | Product                                       |
|---------------------------|---------|---------|------------|-----------------------------------------------|
| ER: N-glycan biosynthesis |         |         |            |                                               |
| *GlcNAc-P-T DPAGT         | Q9H3H5  | -       | UDP-GlcNAc | GlcNAc-PP-Dol                                 |
| $\beta$ 4GlcNAcT ALG13/14 | Q9NP73  | 1       | UDP-GlcNAc | GlcNAc <sub>2</sub> -PP-Dol                   |
|                           | Q96F25  |         |            |                                               |
| DPM1                      | O60762  | 2       | GDP-Man    | Man $\beta$ -P-Dol                            |
| $\beta$ 4ManT ALG1        | Q9BT22  | 33      | GDP-Man    | Man-GlcNAc <sub>2</sub> -PP-Dol               |
| $\alpha$ 3/6ManT ALG2     | Q9H553  | 4       | GDP-Man    | Man <sub>3</sub> -GlcNAc <sub>2</sub> -PP-Dol |
| $\alpha$ 2ManT ALG11      | Q2TAA5  | 4       | GDP-Man    | Man <sub>5</sub> -GlcNAc <sub>2</sub> -PP-Dol |
| $\alpha$ 3ManT ALG3       | Q92685  | 58      | Dol-P-Man  | Man <sub>7</sub> -GlcNAc <sub>2</sub> -PP-Dol |
| $\alpha$ 2ManT ALG9       | Q9H6U8  | 22      | Dol-P-Man  | Man <sub>9</sub> -GlcNAc <sub>2</sub> -PP-Dol |
| $\alpha$ 2ManT ALG12      | Q9BV10  | 22      | Dol-P-Man  | Man <sub>9</sub> -GlcNAc <sub>2</sub> -PP-Dol |

Table 1. Cont.

| Enzyme Name                  | Uniprot | CAZy GT | Donor      | Product                                                              |
|------------------------------|---------|---------|------------|----------------------------------------------------------------------|
| DolPGlc synthase ALG5        | Q9Y673  | 2       | UDP-Glc    | Glc $\beta$ -P-Dol                                                   |
| $\alpha$ 3GlcT ALG6          | Q9Y672  | 57      | Dol-P-Glc  | Glc-Man <sub>9</sub> -GlcNAc <sub>2</sub> -Dol                       |
| $\alpha$ 3GlcT ALG8          | Q9BVK2  | 57      | Dol-P-Glc  | Glc <sub>2</sub> -Man <sub>9</sub> -GlcNAc <sub>2</sub> -Dol         |
| $\alpha$ 2GlcT ALG10         | Q5BKT4  | 59      | Dol-P-Glc  | Glc <sub>3</sub> -Man <sub>9</sub> -GlcNAc <sub>2</sub> -Dol         |
| OST complex                  | -       | -       | Glycan-Dol | Glycan-Asn                                                           |
| $\alpha$ 2Glc MOGS           | Q13724  | GH63    |            | Glc <sub>2</sub> -Man <sub>9</sub> -GlcNAc <sub>2</sub> -Asn         |
| $\alpha$ 3Glc GANAB          | Q14697  | GH31    |            | Man <sub>9</sub> -GlcNAc <sub>2</sub> -Asn                           |
| $\alpha$ 2Man MAN1B1         | Q9UKM7  | GH47    |            | Man <sub>8</sub> -GlcNAc <sub>2</sub> -Asn                           |
| Golgi: N-glycan processing   |         |         |            |                                                                      |
| $\alpha$ 2Man MAN1A1         | P33908  | GH47    |            | Man <sub>5</sub> GlcNAc <sub>2</sub> -Asn                            |
| * $\beta$ 2GnT I MGAT1       | P26572  | 13      | UDP-GlcNAc | GlcNAc $\beta$ 2Man <sub>5</sub> GlcNAc <sub>2</sub> -               |
| $\alpha$ 3,6Man MAN2A1       | Q16706  | GH38    |            | GlcNAc $\beta$ 2-Man <sub>3</sub> GlcNAc <sub>2</sub> -              |
| * $\beta$ 2GnT II MGAT2      | Q10469  | 16      | UDP-GlcNAc | GlcNAc $\beta$ 2 <sub>2</sub> Man <sub>3</sub> GlcNAc <sub>2</sub> - |
| $\beta$ 4GnT III MGAT3       | Q09327  | 17      | UDP-GlcNAc | bisecting GlcNAc $\beta$ 4Man $\beta$ -                              |
| $\beta$ 4GnT IV MGAT4A       | Q9UM21  | 54      | UDP-GlcNAc | GlcNAc $\beta$ 3Man <sub>3</sub> GlcNAc <sub>2</sub> -               |
| * $\beta$ 6GnT V MGAT5       | Q09328  | 18      | UDP-GlcNAc | GlcNAc $\beta$ 4Man <sub>3</sub> GlcNAc <sub>2</sub> -               |
| *GNPTAB                      | Q3T906  | -       | UDP-GlcNAc | GlcNAc-6P-Man                                                        |
| GNPTG complex                | Q9UJJ9  | -       | UDP-GlcNAc | GlcNAc-6P-Man                                                        |
| * $\alpha$ 6FucT FUT8        | Q9BYC5  | 23      | GDP-Fuc    | N-glycan core Fuc                                                    |
| * $\alpha$ 6SiaT ST6Gal1     | P15907  | 29      | CMP-Sia    | complex N-glycans                                                    |
| $\alpha$ 8SiaT ST8SIA2       | Q92186  | 29      | CMP-Sia    | Sia $\alpha$ 2-8Sia-                                                 |
| * $\alpha$ 8SiaT ST8SIA3     | O43173  | 29      | CMP-Sia    | Sia $\alpha$ 2-8Sia-                                                 |
| $\alpha$ 8SiaT ST8SIA4       | Q92187  | 29      | CMP-Sia    | Sia $\alpha$ 2-8Sia-                                                 |
| Golgi: O-glycan biosynthesis |         |         |            |                                                                      |
| *GALNT1                      | Q10472  | 27      | UDP-GalNAc | O-glycan initiation                                                  |
| * $\beta$ 3GalT C1GALT1      | Q9NS00  | 31      | UDP-Gal    | core 1                                                               |
| * $\beta$ 6GnT GCNT1         | Q02742  | 14      | UDP-GlcNAc | core 2                                                               |
| $\beta$ 6GnT GCNT3           | O95395  | 14      | UDP-GlcNAc | core 2/4                                                             |
| $\beta$ 6GnT GCNT4           | Q9P109  | 14      | UDP-GlcNAc | core2                                                                |
| $\beta$ 3GnT B3GNT3          | Q9Y2A9  | 14      | UDP-GlcNAc | core 1,2 elongation                                                  |
| $\beta$ 3GnT B3GNT6          | Q6ZMB0  | 31      | UDP-GlcNAc | core 3                                                               |
| $\alpha$ 4GnT A4GNT          | Q9UNA3  | 32      | UDP-GlcNAc | GlcNAc $\alpha$ 4-Gal of core 2                                      |
| * $\alpha$ 3SiaT ST3Gal1     | Q11201  | 29      | CMP-Sia    | Sia-core 1                                                           |
| $\alpha$ 6SiaT ST6GalNAc1    | K7EMB6  | 29      | CMP-Sia    | Sia-GalNAc-Thr                                                       |
| * $\alpha$ 6SiaT ST6GalNAc2  | Q9UJ37  | 29      | CMP-Sia    | Sia-GalNAc/Gal-GalNAc-Thr                                            |
| $\alpha$ 6SiaT ST6GalNAc3    | Q8NDV1  | 29      | CMP-Sia    | Sia-Gal-(Sia)-GalNAc                                                 |
| $\alpha$ 6SiaT ST6GalNAc4    | Q9H4F1  | 29      | CMP-Sia    | Sia-Gal (Sia)-GalNAc                                                 |

Table 1. Cont.

| Enzyme Name                                            | Uniprot | CAZy GT | Donor      | Product                       |
|--------------------------------------------------------|---------|---------|------------|-------------------------------|
| Golgi: N- and O-glycan chain extension and termination |         |         |            |                               |
| *i $\beta$ 3GnT B3GNT2                                 | Q9NY97  | 31      | UDP-GlcNAc | i antigen                     |
| $\beta$ 3GnT B3GN8                                     | Q7Z7M8  | 31      | UDP-GlcNAc | i antigen                     |
| $\beta$ 6GnT GCNT2                                     | Q8N0V5  | 14      | UDP-GlcNAc | I antigen                     |
| $\beta$ 4GalNAcT B4GALNT3                              | Q6L9W6  | 7       | UDP-GalNAc | Lac-diNAc                     |
| * $\alpha$ 3GalNAcT, GTA, BGAT                         | P16442  | 6       | UDP-GalNAc | Blood group A                 |
|                                                        | V5ZDP0  |         |            |                               |
| $\beta$ 4GalNAcT B4GALNT2                              | Q8NHY0  | 12      | UDP-GalNAc | Sda antigen                   |
| * $\alpha$ 3GalT GTB                                   | V9GWR7  | 6       | UDP-Gal    | Blood group B                 |
| * $\beta$ 4GalT B4GALT1                                | P15291  | 7       | UDP-Gal    | GlcNAc extension              |
| $\beta$ 4GalT B4GALT2                                  | O60909  | 7       | UDP-Gal    | GlcNAc extension              |
| $\beta$ 4GalT B4GALT3                                  | O60512  | 7       | UDP-Gal    | GlcNAc extension              |
| * $\beta$ 3GalT B3GALT5                                | Q9Y2C3  | 31      | UDP-Gal    | GlcNAc extension              |
| $\alpha$ 2FucT FUT1                                    | P19526  | 11      | GDP-Fuc    | Blood group H/O               |
| $\alpha$ 2FucT FUT2                                    | Q10981  | 11      | GDP-Fuc    | Blood group H/O               |
| $\alpha$ 3/4FucT FUT3                                  | P21217  | 10      | GDP-Fuc    | Lewis antigens                |
| $\alpha$ 3FucT FUT4                                    | P22083  | 10      | GDP-Fuc    | Lewis antigens, type 2 chains |
| $\alpha$ 3/4FucT FUT5                                  | Q11128  | 10      | GDP-Fuc    | Lewis antigens                |
| $\alpha$ 3FucT FUT7                                    | Q11130  | 10      | GDP-Fuc    | sialyl-Lewis                  |
| * $\alpha$ 3FucT FUT9                                  | Q9Y231  | 10      | GDP-Fuc    | Lewis antigens                |
| $\alpha$ 3SiaT ST3Gal3                                 | Q11203  | 29      | CMP-Sia    | Sia-Gal-GlcNAc-               |
| $\alpha$ 3SiaT ST3Gal4                                 | Q11206  | 29      | CMP-Sia    | Sia-core 1/Gal-GlcNAc         |
| $\alpha$ 3SiaT ST3Gal6                                 | Q9Y274  | 29      | CMP-Sia    | Sia-Gal-GlcNAc                |
| $\alpha$ 8SiaT ST8SIA6                                 | P61647  | 29      | CMP-Sia    | Sia $\alpha$ 2-8Sia           |

Some of these GTs have broad specificities and thus can extend a number of glycoconjugates, including glycolipids and/or proteoglycans. Many GTs occur as natural variants that may have slightly different properties and specificities. Only one GT (C1GALT1) requires the coexpression of a chaperone (COSMC). Often, GTs form dimers or multi-enzyme complexes that support activities. GTs requiring Dol intermediates are localized to the membranes of the ER. The initial N-glycan structures up to Man<sub>5</sub>GlcNAc<sub>2</sub>- are made on the cytoplasmic side of the ER membrane, and after being flipped to the luminal side, N-glycans are completed and transferred from Dol to nascent peptide. Further N-glycan processing and extension reactions occur after the transport of the glycoprotein to the Golgi.

GTA-folded GTs have one major fold with a central catalytic domain. GT-B folded GTs have 2 typical Rossman-like (nucleotide-binding) folds and the active site located in between them. GT-C folded GTs are uncommon, and Dol-binding GTs often have multiple membrane-spanning domains that are typical of GT-C folded enzymes [15]. The enzyme mechanisms of a few GTs have been suggested by their crystal structures, analysis of mutants, and enzyme kinetics. Many GTs are N-linked glycoproteins and can also be O-glycosylated.

In contrast to GTs, glycoside hydrolases are classified into 194 families (as of July 2025), with many subfamilies. These enzymes hydrolyze a glycosidic bond using a catalytic acid as proton donor and nucleophilic catalytic base. Like GTs, hydrolases act with an inverting or retaining mechanism and have a limited number of protein folds.

Most inverting GTs appear to follow a single displacement mechanism where a specific hydroxyl of the acceptor substrate forms a nucleophile that attacks carbon-1 of the sugar linked to phosphate of the nucleotide sugar donor substrate. This often involves a catalytic Asp residue in a DxD sequence [16]. The mechanisms of retaining GTs where sugar linkages in donor and product are the same remain to be clarified further. In the Golgi, GTs are commonly type II membrane proteins with a short N-terminal cytoplasmic domain followed by a transmembrane <sup>TM</sup> domain, and the folded GT domain in the lumen of the Golgi.

GlcNAc-transferases (GnT, MGAT) that synthesize N-glycan antennae are highly specific for the antennae they synthesize [17]. This specificity is likely a result of the large internal substrate-binding cavity that accommodates more than the terminal glycan. This is in contrast to certain Gal-transferases (GalTs) that can act on a single sugar residue. For example, the  $\beta$ 1,4-GalT B4GALT1 only requires GlcNAc as a substrate, although some branch specificity for N-glycan acceptors has been observed, which may be due to steric factors [18,19]. Acceptor substrates have been designed that have all of the structures required for recognition by a specific enzyme but are as small as possible, conveniently synthesized, and handled by enzyme assays. GTs are usually highly specific for their donor substrates, with small modifications of the sugar moiety being tolerated and allowing a modified sugar to be transferred.

### 3. Measurements of Glycosyltransferase Activities

The general GT reaction is Acceptor + Donor  $\rightarrow$  Product + Nucleotide. The reaction often requires a divalent metal ion to complex the nucleotide (e.g.,  $Mn^{2+}$  or  $Mg^{2+}$ ), along with a buffer and detergent if a membrane-bound version of GT is used. If the enzyme source has contaminating hydrolases, specific inhibitors can increase the product yield. The GT activities can be measured accurately by tracking the transfer of a radioactive sugar residue to form the reaction product. When neutral acceptor substrates are used, the radioactive nucleotide sugar and sugar-phosphate breakdown products can be separated after incubation by anion exchange chromatography. Scintillation counting of the reaction product in the flow-through fractions will indicate the exact amount of sugar transferred, allowing the specific enzyme activity and kinetics to be determined. Most of Paulsen's compounds have been examined using the sensitive radioactive method.

If charged acceptors are used, then other assay methods could include hydrophobic chromatography, HPLC, electrophoresis, or paper chromatography. Assays using non-radioactive donor substrates require HPLC separation of assay components and mass spectrometry (MS) for estimating the product yield. TLC can also be used with a panel of standard compounds. Another method is to detect the reaction product UDP by bioluminescence (UDP-Glo <sup>TM</sup>). In order to determine the linkage synthesized, specific hydrolases, antibodies, or lectins could be employed. If sufficient amounts of reaction product are available, NMR is the best method for determining anomeric configuration and sugar linkage in the product.

For accurate and efficient measurements of GT activities, it is crucial to design a suitable acceptor substrate and select the best method to determine the amount and the new linkage in the reaction products. Paulsen's N-glycan derivatives with an octyl aglycone group were ideal for isolating reaction products by columns of silica-linked hydrophobic chains of 18 carbons. By omitting specific sugar residues, the minimum size of the active

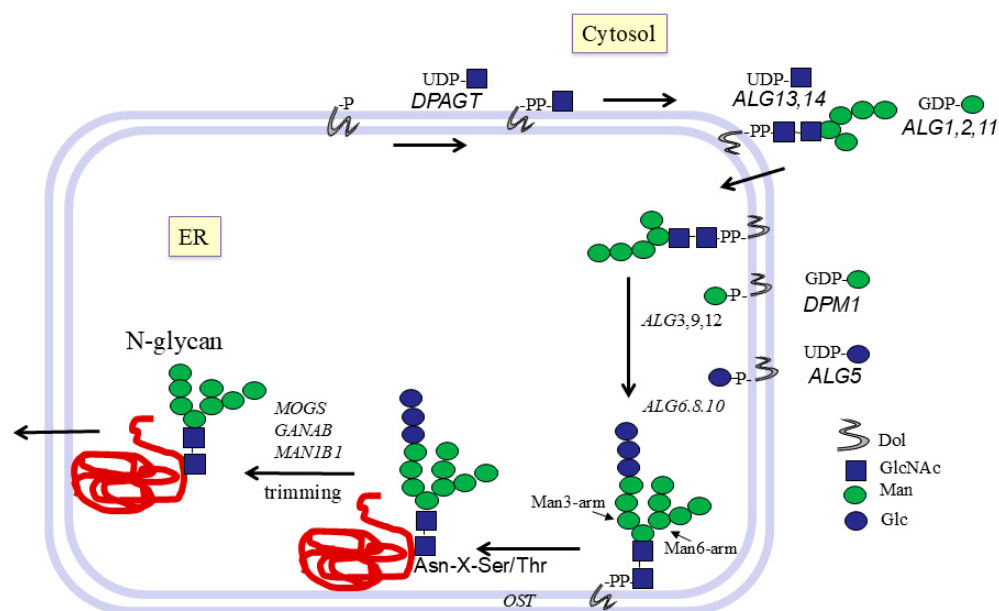
acceptor could be assessed. For GTs acting on O-glycans, aryl aglycone groups have been successful in both anion exchange and hydrophobic chromatography assays. Glycopeptide acceptors were used mainly by anion exchange chromatography and HPLC.

#### 4. Biological Significance of Glycoproteins

Most eukaryotic proteins are post-translationally modified with glycans and are found inside cells, on cell surfaces, and in secretions. The Asn-linked N-glycan and Ser/Thr-linked mucin-type O-glycan structures in human glycoproteins contain the same basic sugar residues, except that these O-glycans do not have Glc and Man residues. Many of the internal and terminal carbohydrate epitopes of extended glycan structures are similar. The N- and O-glycans of glycoproteins serve to provide protein stability, conformation, folding, intracellular targeting, resistance to degradation, and dehydration. Protein glycosylation is involved in many biological functions, in the immune system, during development and fertilization, and as lubricant. The glycan moieties can interact with microbes and protect against the invasion of pathogens. Lectins and antibodies bind to glycans, which play important roles in cell–cell and cell surface receptors interactions.

#### 5. Glycoprotein Biosynthesis

The initial pathways of N-glycan synthesis involve Dol lipid carriers at the cytoplasmic and luminal sides of the ER membrane. Preformed oligosaccharides are transferred from Dol to peptides in the ER (Figure 1). In contrast, O-glycans are all added to the protein in the Golgi and do not involve Dol. During N-glycan processing, Glc and Man are removed by specific hydrolases, a process not known for O-glycan synthesis. Sulfotransferases can add sulfate esters in the Golgi to Gal or GlcNAc residues of both N- and O-glycans by transferring sulfate from 3'-phosphoadenosine-5'-phosphosulfate.



**Figure 1.** Early pathways of N-glycan biosynthesis in the ER.

Paulsen's work has contributed immensely to our knowledge of specific enzymes and pathways. The substrate binding and catalytic mechanisms of the enzymes identified in this work may differ between homologues from different species and tissues, and many enzyme variants occur. Nature has provided human mutations that unfortunately result in disease and suffering but has helped the biochemist to identify enzyme functions and the

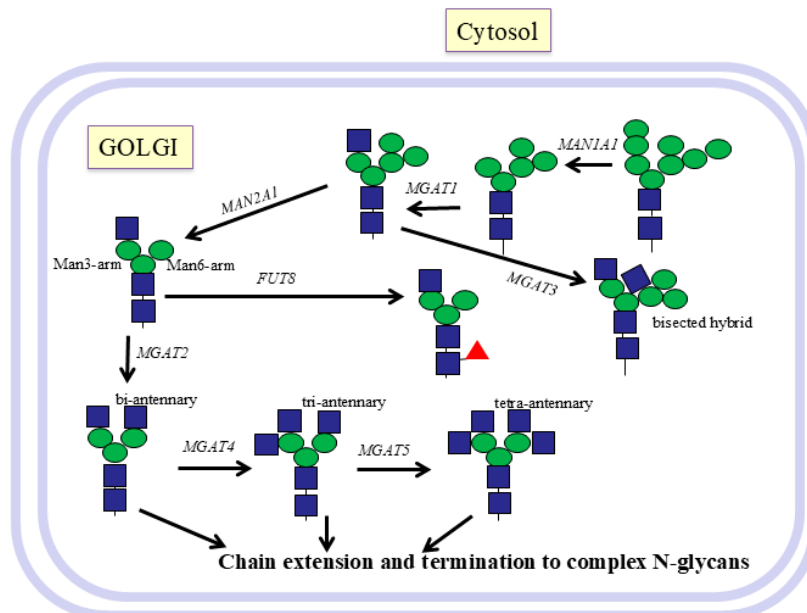
necessity for protein glycosylation. Thus, chemistry also played an important role in the discovery of the significance of glycosylation in disease.

Mucin glycoproteins have a high content of mucin-type GalNAc $\alpha$ -Ser/Thr-based O-glycans. Although many alterations of O-glycans in disease have been identified, there are no human mutants that lack mucin O-glycans, and 20 or more polypeptide GalNAc-transferases (GALNTs) cooperate in humans to assemble these O-glycans, indicating that they are essential for survival. The O-glycosylation of non-mucin proteins depends on the protein folding and exposure of Ser/Thr residues as potential O-glycosylation sites.

There are a number of other O-glycans that are important, e.g., O- $\alpha$ Man glycans of dystroglycans that may be altered in muscular dystrophies. O- $\alpha$ Fuc is found in signalling proteins, and O- $\beta$ Gal in collagens. O-GlcNAc $\beta$ -Ser/Thr (O-GlcNAc), along with phosphate modifications, is common for cytoplasmic and nuclear proteins and functions in a wide range of biological activities, from apoptosis to oxidative stress responses. O-GlcNAc is not extended by additional sugars in the Golgi. However, it can be modified *in vitro* by  $\beta$ 1,4-GalT. Alterations of O-GlcNAc have been implicated in protein aggregation, neurological diseases, cancer, and diabetes.

## 6. N-Glycans

N-glycans are found in all animals, as well as in many other species, and even in archaea and certain bacteria. Viruses usually carry glycoproteins with glycans that are characteristic of the cell type the virus originates from. High-mannose N-glycans have only Man residues extending the Man<sub>3</sub>GlcNAc<sub>2</sub> core and are the least processed chains. Hybrid N-glycans have several Man residues on the Man<sub>6</sub> arm and a GlcNAc $\beta$ 1-2 antenna on the Man residue of the Man<sub>3</sub> arm (Figure 2).



**Figure 2.** Golgi pathways of N-glycosylation. N-glycans are processed to hybrid or complex N-glycans and extension of the antennae in the Golgi. Glycans are trimmed by mannosidase, followed by the transfer of GlcNAc in  $\beta$ 1-2 linkage to Man $\alpha$ 1-3 of the Man<sub>3</sub>-arm by MGAT1. This facilitates further synthesis by  $\alpha$ 6-fucosylation of core GlcNAc attached to Asn, trimming of Man residues on the Man<sub>6</sub>-arm, and synthesis of further antennae by MGAT2,4,5 and bisecting structures by MGAT3. The antennae are then extended by a wealth of different enzymes that introduce repeating Gal-GlcNAc units and many possible branched and linear epitopes, Lewis antigens, blood groups, and tissue-specific structures. Glycoproteins are then transported to their preferred locations.

The antennae of complex N-glycans often consist of repeating Gal $\beta$ 1-4GlcNAc $\beta$ 1-3- (LacNAc) sequences that can carry many possible internal and terminal sugars and epitopes, such as GlcNAc, Gal, GalNAc, Fuc, sialic acid (Sia), blood groups, i and I antigens, and LacdiNAc and Lewis determinants (Table 2). Many of these are shared with complex O-glycans and glycosphingolipids. The aberrant synthesis of N-glycans and O-glycans, leading to altered amounts of epitopes in cancer, is controlled by many enzymes. The enzymatic basis for these changes is extremely complex and usually involves the altered expression of a number of enzymes involved in the glycosylation pathways.

**Table 2.** List of mucin core structures and epitopes of both O- and N-glycans. Human glycoproteins have O-glycan structures initiated with GalNAc or core structures 1 to 4. Other core structures (5 to 8) are rare. Based on the variable tandem repeat sequences of MUC1, MUC2, MUC3, and MUC4, Paulsen's group synthesized hundreds of glycopeptides carrying one or more O-glycans with core 1 to 7. The occurrences of most of these structures have been found to be altered (increased or decreased) in cancer [20–31].

| Epitope            | Structure                                                                              | Alteration in Cancer |
|--------------------|----------------------------------------------------------------------------------------|----------------------|
| Tn antigen         | GalNAc $\alpha$ -Ser/Thr-                                                              | Increased            |
| Sialyl-Tn antigen  | Sia $\alpha$ 2-6GalNAc $\alpha$ -                                                      | Increased            |
| Core 1, T antigen  | Gal $\beta$ 1-3GalNAc $\alpha$ -                                                       | Increased            |
| Sialyl-T antigen   | Sia $\alpha$ 2-3Gal $\beta$ 1-3GalNAc $\alpha$ -                                       | Increased            |
| Core 2             | GlcNAc $\beta$ 1-6(Gal $\beta$ 1-3)GalNAc $\alpha$ -                                   | Variable             |
| Core 3             | GlcNAc $\beta$ 1-3GalNAc $\alpha$ -                                                    | Decreased            |
| Core 4             | GlcNAc $\beta$ 1-6(GlcNAc $\beta$ 1-3)GalNAc $\alpha$ -                                | Decreased            |
| Core 5             | GalNAc $\alpha$ 1-3GalNAc $\alpha$ -                                                   | Increased            |
| Core 6             | GlcNAc $\beta$ 1-6GalNAc $\alpha$ -                                                    |                      |
| Core 7             | GalNAc $\alpha$ 1-6GalNAc-                                                             |                      |
| Core 8             | Gal $\alpha$ 1-3GalNAc-                                                                |                      |
| Sialyl-Lewis x     | Sia $\alpha$ 2-3Gal $\beta$ 1-4(Fuc $\alpha$ 1-3)GlcNAc $\beta$ -                      | Increased            |
| Sda/Cad            | GalNAc $\beta$ 1-4(Sia $\alpha$ 2-3)Gal $\beta$ -                                      | Decreased            |
| i antigen          | Gal $\beta$ 1-4GlcNAc $\beta$ 1-3Gal $\beta$ 1-                                        | Variable             |
| I antigen          | Gal $\beta$ 1-4GlcNAc $\beta$ 1-6<br>(Gal $\beta$ 1-4GlcNAc $\beta$ 1-3)Gal $\beta$ 1- |                      |
| LacdiNAc           | GalNAc $\beta$ 1-4GlcNAc $\beta$ 1-3-                                                  |                      |
| Sialyl-LacNAc      | Sia $\alpha$ 2,6Gal $\beta$ 1-4GlcNAc                                                  | Increased            |
| Blood group O(H)   | Fuc $\alpha$ 1-2Gal $\beta$ -                                                          |                      |
| Blood group A      | GalNAc $\alpha$ 1-3(Fuc $\alpha$ 1-2)Gal $\beta$ -                                     |                      |
| Blood group B      | Gal $\alpha$ 1-3(Fuc $\alpha$ 1-2)Gal $\beta$ -                                        |                      |
| Complex N-glycans  | Highly branched N-glycans                                                              | Increased GnT V      |
| Sialylated termini | Sialylated N-glycans                                                                   | Increased ST6Gal1    |

N-glycans have gained considerable interest due to their multiple functions in protein structure, biology, and medicine. The functionally important N-glycan antennae cover a significant amount of space around glycoproteins and can be recognized by antibodies or lectins and carbohydrate-binding proteins of the immune system. N-glycan structures are often altered in cancer and in patients with congenital disorders of glycosylation (CDG),

which can affect the properties and essential functions of glycoproteins [32]. For example, in cancer cells, GnT V (Table 1, Figure 2) was shown to increase in activity, thus expanding the size and complexity of branched tetra-antennary N-glycans. Thus, our focus was on designing inhibitors for N-glycan branching enzymes and controlling the functions of N-glycans.

The pathways of N-glycan synthesis are highly complex, involving many highly conserved enzymes in the ER, Golgi, cytoplasmic, and nuclear compartments and various mechanisms of transport and localization. In the early 1980s, Hans Paulsen started a fruitful collaboration with Harry Schachter (University of Toronto) to elucidate the substrate specificities and control of N-glycosylation pathways. For this purpose, Paulsen synthesized a series of N-glycans with systematic modifications of hydroxyl groups in order to explore their importance in GnT activities. Compounds included deoxy derivatives or those with specific reactive or non-reactive modifications and linkers. N-glycans were synthesized as reducing sugars or with a hydrophobic aglycone group that was useful for enzymatic reactions and isolation of enzyme products by HPLC. These derivatives helped determine the detailed recognition of substrates and enzyme specificities. Depending on the enzyme, specific hydroxyl groups were found to be critically required, and modifications in the stereochemistry, bulkiness, and hydrophobicity of substituents had variable effects on activity. These studies were complemented by our understanding of protein structures, their proposed mechanism, and the amino acids controlling the active site [33].

## 7. Early Pathways of N-Glycosylation at ER Membranes

N-glycosylation involves both nucleotide sugars and Dol-P-sugars. The early pathways to the  $\text{Man}_5\text{GlcNAc}_2\text{-PP-dolichol}$  structure are catalyzed by enzymes in the ER membrane. Reactions require UDP-GlcNAc and GDP-Man nucleotide sugars present in the cytosol. The N-glycan is then flipped across the membrane for further processing. In order to synthesize N-glycans, the cell must provide nucleotide sugar donor substrates and acceptor substrates that are Dol-P intermediates for the initial pathways (Figure 1) localized to the ER membrane. These early reactions have been studied mainly in yeast [34]. The biochemical importance of several of these enzymes was identified in humans with gene mutations leading to the CDG syndrome. Due to the reduction or absence of specific activities involved in the assembly of N-glycan structures, glycoproteins lack N-glycans and exhibit different electrophoretic mobility. This defect can be used as a diagnostic marker for CDG.

The initial cytoplasmic reaction is the transfer of  $\text{GlcNAc}\alpha\text{-P}$  from UDP-GlcNAc to Dol-P by GlcNAc-1-P-transferase, named ALG7, in yeast, to form Dol-PP- $\alpha\text{GlcNAc}$  (Figure 1). UDP-GlcNAc: Dol-P GlcNAc phosphotransferase DPAGT1 binds to the UDP-GlcNAc analog antibiotic tunicamycin [35]. Tunicamycin also binds to GlcNAc-1-P transferases from other species, including bacteria.

An oligosaccharide consisting of 2 GlcNAc and 5 Man residues is assembled on Dol-P from UDP- $\alpha\text{GlcNAc}$  and GDP- $\alpha\text{Man}$  on the cytoplasmic side of the ER membrane. The GnT complex ALG13/ALG14 first synthesizes the di-N-acetyl-chitobiose structure  $\text{GlcNAc}\beta 1\text{-4GlcNAc-}$  [36]. Subsequently, the  $\text{Man}\beta 1\text{-4}$  linkage to GlcNAc is synthesized by  $\beta 1,4\text{-Man-transferase (ManT) ALG1}$  (Table 1). ALG2 transfers  $\text{Man}\alpha 3$  as well as  $\text{Man}\alpha 6$  (forming the Man3- and Man6-arms) from GDP-Man to the  $\text{Man}\beta 1\text{-4}$  residue. Two Man residues are then added in  $\alpha 1\text{-2}$  linkage to the Man6 arm by  $\alpha 2\text{ManT ALG11}$ . The Dol-PP-linked  $\text{Man}_5\text{GlcNAc}_2$  oligosaccharide is then transferred to the ER lumen for the further addition of Man and Glc residues. Dol is a large membrane lipid that has a significant effect on the structure and leakiness of the membrane and may be involved in the translocation

process of the  $\text{Man}_5\text{-GlcNAc}_2$  intermediate. The enzyme that transfers Man to Dol-P has been shown to require membrane phospholipids for its activity [37].

For the reactions taking place at the luminal side of the ER membrane, both acceptor and donor substrates utilize membrane-bound Dol-P sugars that flip across the ER membrane. The donor substrate Dol-P- $\beta$ Man is synthesized at the cytoplasmic side from Dol-P and GDP-Man by Dol-P-Man synthase DPM1. Dol-P- $\beta$ Glc is synthesized by Dol-P-Glc synthase ALG5 using UDP- $\alpha$ Glc. The Glc and Man residues of these membrane-bound donor substrates are then translocated to the lumen of the ER, where they are used by GTs to transfer Man and Glc to  $\text{Man}_5\text{-GlcNAc}_2\text{-PP-Dol}$  to synthesize the large  $\text{Glc}_3\text{-Man}_9\text{-GlcNAc}_2$  structure (Figure 2). The transfer of Man residues involves  $\alpha 3\text{ManT}$  ALG3 and  $\alpha 2\text{ManTs}$  ALG9 and ALG12. The final Glc residues are transferred to the Man3-arm by  $\alpha 3\text{Glc-transferase (GlcT)}$  ALG6 and ALG8, and the terminal Glc residue is transferred by  $\alpha 2\text{GlcT}$  ALG10 (Table 1).

## 8. Transfer of N-Glycans from Dolichol to Protein

After the large oligosaccharide has been assembled, it is transferred from Dol-PP to the nitrogen side chain of Asn of the nascent glycoprotein in the ER by the oligosaccharyltransferase complex OST. This inverts the  $\alpha$ -linkage of the first GlcNAc to a  $\beta$ -linkage. The Asn-x-Ser/Thr sequon is absolutely required for protein N-glycosylation, and this requirement has been maintained in all eukaryotic species throughout evolution, including in archaea and several bacteria. A Pro residue between Asn and Ser/Thr or following Ser/Thr does not allow N-glycosylation, likely by preventing the formation of an essential hydrogen bond between Asn and Ser/Thr.

Processing reactions start in the ER lumen after the attachment of N-glycans to peptides. A number of chaperone and GlcT proteins are involved in protein folding and quality control. The terminal  $\text{Glc}\alpha 1\text{-2}$  residue is cleaved by glucosidase MOGS, followed by cleavage of two  $\text{Glc}\alpha 1\text{-3}$  residues by GANAB. Mannosidase MAN1B1 then cleaves  $\text{Man}\alpha 1\text{-2}$ , leaving a mixture of N-glycan structures from  $\text{Man}_8\text{-GlcNAc}_2$  to  $\text{Man}_5\text{-GlcNAc}_2$ , which are precursors of hybrid and complex N-glycans. These glycoside hydrolases are type II membrane proteins, but are classified into different CAZy GH families (Table 1). The glycoproteins are then transported from the ER to the Golgi, where further glycan processing and synthesis occurs.

## 9. Golgi Reactions

All GTs in the Golgi utilize nucleotide sugar donor substrates that need to be imported into the Golgi by specific transporters. The localization of Golgi proteins is due to the gradual transport of membranous vesicles through Golgi compartments. The relatively short tm domain of GTs does not appear to be important for activity, and recombinant GTs produced in vitro usually lack the tm domain to enhance solubility. However, it is possible that the short N-terminal cytoplasmic tail of GTs plays a role in the localization of GTs within various Golgi compartments. The conserved oligomeric Golgi protein complex (COG) has been shown to control GT localization as well as retrograde Golgi to ER transport. In addition to the ER proteins, the Golgi resident protein QSOX1 also catalyzes the disulfide bonding of specific GTs, e.g., sialyltransferases (SiaTs), as well as the intermolecular bonds of large mucins [38]. However, the complex mechanisms that localize GTs and re-localize them under pathological conditions remain to be further explored. Processed and extended glycoproteins are then delivered to their final destination on the cell surface, in secretions, or in intracellular membranes.

## 10. Synthesis of N-Glycan Antennae by GlcNAc-Transferases

After transport of the glycoprotein from the ER to the Golgi, further Man residues are cleaved from N-glycans by mannosidases that are type II membrane proteins. Mannosidase 1 MAN1A1 progressively cleaves Man $\alpha$ 1-2 residues. MGAT1 (GnT I) is the first enzyme to add an antenna and transfers GlcNAc in  $\beta$ 1-2 linkage from UDP-GlcNAc to the Man $\alpha$ 1-3 residue (Man3-arm) of Man<sub>5</sub>-GlcNAc<sub>2</sub>-Asn-R (Figure 2). The conversion of hybrid to complex N-glycans is then controlled by mannosidase 2, MAN2A1, which cleaves both the Man $\alpha$ 1-3 and Man $\alpha$ 1-6 residues from Man $\alpha$ 1-6 (Man6-arm), which then exposes the substrate for GnT II.

Harry Schachter, together with Pamela Stanley [39], discovered GnT I in mutant CHO cells. It was later shown that this enzyme is essential for further modifications of N-glycans. This includes the removal of 2 Man residues by Golgi mannosidase 2,  $\alpha$ 6-fucosylation of the core GlcNAc linked to Asn, and further branching by GnT II (MGAT2) and IV (MGAT4), which then allows the formation of the fourth antenna by GnT V (MGAT5) (Figure 2). MGAT1 is also required for introducing the bisecting GlcNAc by GnT III (MGAT3) [32].

## 11. GnT I-MGAT1

The characterization of GnT enzymes was greatly facilitated by collaboration with Hans Paulsen. Paulsen's group was the first to synthesize N-glycan oligosaccharides having the Man  $\beta$ 1-4GlcNAc linkage. The configuration and rotation of the N-acetyl group of GlcNAc represented a steric hindrance that favoured the synthesis of Man $\alpha$ -glycosides. Thus, the method for Man $\beta$ - synthesis was a breakthrough for the synthesis of GnT substrates of N-glycans up to 9 residues [40]. Glycan linkages and 3D conformations were extensively analyzed by NMR [41].

As an effective acceptor substrate for GnT I [39] that adds GlcNAc to the Man3-arm, Paulsen's group synthesized Man $\alpha$ 6(Man $\alpha$ 3)Man $\beta$ -octyl and many related compounds. It was shown that the minimum requirement for an acceptor substrate for purified rat liver GnT I is Man $\alpha$ 1-3Man $\beta$ 1-4GlcNAc, but optimal activity is observed when the Man $\alpha$ 1-6 residue is also present. Other groups, e.g., Ole Hindsgaul (University of Alberta), Jeremy Carver (University of Toronto), and Khushi Matta (Roswell Park Institute, Buffalo, NY), also participated in the synthesis of acceptor substrates for GnTs with hydrophobic groups replacing the di-N-acetylchitobiose of the N-glycan core and helped to successfully characterize properties of GnTs.

The crystal structure of rabbit GnT I suggested its catalytic mechanism as an inverting GT-A-folded enzyme that binds UDP-GlcNAc and Mn<sup>2+</sup> [33] (Table 1). The protein has a 211EDD<sup>213</sup> motif involved in forming hydrogen bonds and interacting with UDP-GlcNAc and Mn<sup>2+</sup>. D291 was proposed to be the catalytic base that can deprotonate the 2-hydroxyl of Man $\alpha$ 1-3 in the acceptor, which then becomes a nucleophile that attacks carbon-1 of GlcNAc in the donor UDP-GlcNAc. The proposed acceptor binding site was shown by Man $\alpha$ 1-3 modeled into the GnT I crystal structure, which is consistent with the specificity results using Paulsen's compounds.

Paulsen's group made significant efforts in replacing hydroxyls of the Man $\alpha$ 1-3(Man $\alpha$ 1-6)Man $\beta$ -octyl acceptor for GnT I with a number of small and large substituents, including O-methyl, O-propyl, O-pentyl, 4-pentenyl, 4-pentanolyl, 4-oxo-pentyl, 4,4-azopentyl (diazirino-pentyl), 4,5-epoxy-pentyl, 5-amino-pentyl, 5-pentanolyl, 5-iodoacetamidopentyl. In other derivatives, the oxygen was removed or replaced with NH<sub>3</sub> [42]. The deoxy compounds indicated the requirement of the hydroxyl group, possibly as a hydrogen bond donor, while loss of activity after replacement of hydroxyl groups with larger substituents may be a result of steric hindrance.

After this extended production of new compounds, GnT I assays showed that the 2,3 and 6-hydroxyls of the Man $\alpha$ 1-3 residue as well as the 2- and 4-hydroxyl of the Man $\beta$  residue were required for activity. The N-acetylchitobiose linked to Asn-peptide was not required and could be replaced by a number of different aglycones, including GlcNAc, hexyl, octyl, or other hydrophobic groups [14,17]. Although the Man $\alpha$ 1-6 residue was required for optimal activity, removal of its 2-hydroxyl or replacement of the 3-, 4-, or 6-hydroxyl with an O-methyl group still allowed GnT I activity at various levels. These compounds helped to explore the substrate specificity of GnT I and the enzymatic production of acceptor substrates for the next GnTs in the pathway. This approach was also part of the design of specific competitive or irreversible inhibitors. A number of these derivatives were inhibitors with  $K_i$  values between 0.76 and 10.2 mM. The most potent inhibitor was Man $\alpha$ 6(6-O-methyl-Man $\alpha$ 3)Man $\beta$ -octyl, where presumably the methyl group clashed with binding of the substrate to the active site of GnT I (Table 3).

**Table 3.** Acceptor substrate analog inhibitors for glycosyltransferases. Paulsen's group synthesized many acceptor substrates that are analogs of the natural N-glycans found in humans. Using these compounds, the activities of several glycosyltransferases were determined, and inhibitors were found. The compounds listed in this table showed significant inhibition of GT activity. GlcNAc-naphthyl derivatives were synthesized by W.A. Szarek, Queen's University. oct, octyl; pnp, *p*-nitrophenyl; -R, various hydrophobic groups.

| Enzyme             | Inhibitors                                                                       |
|--------------------|----------------------------------------------------------------------------------|
| GnT I              | Man $\alpha$ 6(Man $\alpha$ 3)4-O-methyl-Man $\beta$ 4GlcNAc                     |
|                    | Man $\alpha$ 6(6-O-methyl-Man $\alpha$ 3)Man $\beta$ -oct                        |
|                    | Man $\alpha$ 6(6-O-4,5-epoxy-pentyl-Man $\alpha$ 3)Man $\beta$ -oct              |
|                    | Man $\alpha$ 6(6-O-4,4-azo-pentyl-Man $\alpha$ 3)Man $\beta$ -oct                |
| GnT II             | 2-deoxy-Man $\alpha$ 6(GlcNAc $\beta$ 2Man $\alpha$ 3)Man $\beta$ -oct           |
|                    | GlcNAc $\beta$ 2Man $\alpha$ 3Man $\beta$ -oct                                   |
|                    | 3-O(4,4 azo)pentylMan $\alpha$ 6(GlcNAc $\beta$ 2Man $\alpha$ 3)Man $\beta$ -oct |
|                    | 2-deoxy-Man $\alpha$ 6(GlcNAc $\beta$ 2Man $\alpha$ 3)Man $\beta$ -oct           |
| GnT V              | GlcNAc $\beta$ 2(6-deoxy)Man $\alpha$ 6Glc $\beta$ -O-R                          |
|                    | GlcNAc $\beta$ 2(4-O-methyl)Man $\alpha$ 6Glc $\beta$ -O-R                       |
|                    | GlcNAc $\beta$ 2(6-deoxy, 4-O-methyl)Man $\alpha$ 6 Glc $\beta$ -O-R             |
|                    | GlcNAc $\beta$ 2(4,6-di-O-methyl)Man $\alpha$ 6Glc $\beta$ -pnp                  |
| O-glycan synthesis | GalNAc $\alpha$ -aryl                                                            |
| C1GALT1            | 6-O-(4,4-azo)pentyl-GalNAc $\alpha$ -Bn + UV                                     |
| GCNT1              | Gal $\beta$ 3GalNAc $\alpha$ -pnp + UV                                           |
|                    | Gal $\beta$ 3(6-deoxy) GalNAc $\alpha$ -Bn                                       |
| $\beta$ 4GalT      | GlcNAc $\beta$ -naphthyl derivatives                                             |
|                    | 1-thio-N-butyryl-GlcNAc $\beta$ -2-naphthyl                                      |

## 12. GlcNAc-Transferases (MGAT2-6)

After GnT I has acted, GnT II synthesizes the second antenna by transferring a GlcNAc residue in  $\beta$ 1-2 linkage to the Man $\alpha$ 1-6 residue (Figure 2). The serious human condition of CDG IIa is due to a mutation in GnT II. The synthesis of a large series of substrate analogs of the standard GnT II acceptor Man $\alpha$ 1-6(GlcNAc $\beta$ 1-2Man $\alpha$ 1-3)Man $\beta$ -octyl allowed determination of the specificity of partially purified rat liver GnT II [43,44]. GnT II requires the

3-hydroxyl of the GlcNAc added by GnT I on the Man3-arm as well as the 3-hydroxyl of the Man $\alpha$ 1-3 residue, which is essential for binding to GnT II. Interestingly, the trisaccharide GlcNAc $\beta$ 1-2Man $\alpha$ 1-3Man $\beta$ -octyl that lacks the Man $\alpha$ 1-6 residue is not a substrate but is a good inhibitor of GnT II with a  $K_i$  of 0.9 mM (Table 3).

Removal of the 2-hydroxyl of Man $\alpha$ 1-6, the site of GlcNAc addition, creates a competitive inhibitor for GnT II with a  $K_i$  of 0.13 mM, while replacement by a methyl group eliminates activity and does not inhibit, likely because of steric hindrance [45]. In contrast, the 3-, 4-, and 6-hydroxyls of Man $\alpha$ 1-6 are not required for activity. In contrast to GnT I, GnT II does not need to interact with the 4-hydroxyl of Man $\beta$ -. However, substitution of the 4-hydroxyl of Man $\beta$ - by a methyl group blocked GnT II activity. This may explain why GnT II cannot act after GnT III, which introduces a bisecting GlcNAc $\beta$ 1-4 residue linked to Man $\beta$  [43,44].

The structure of human MGAT2 was determined in complex with  $Mn^{2+}$ /UDP-GlcNAc or with the donor analog UDP and with the acceptor GlcNAc $\beta$ 1-2Man $\alpha$ 1-3(Man $\alpha$ 1-6)Man $\beta$ 1-4GlcNAc $_2$ -Asn [46]. This suggested that the inverting GnT II has an overall fold resembling GT-A and that D347 is the catalytic base residue.

After the second antenna has been established, GnT IV and V can add additional GlcNAc residues that initiate tri- and tetra-antennary N-glycans, respectively (Figure 2).

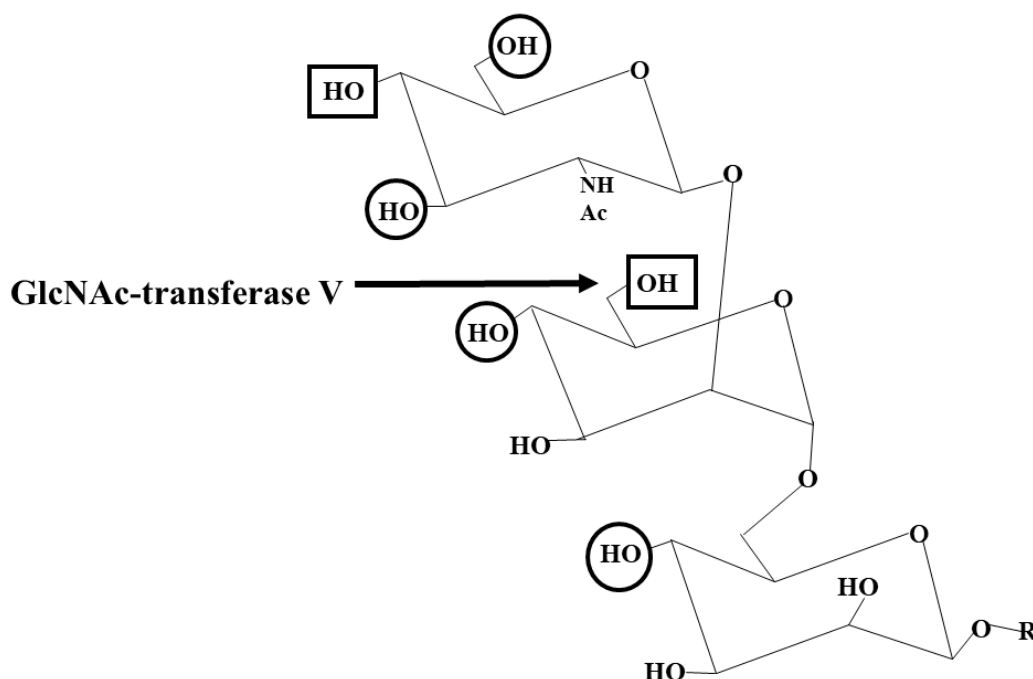
GnT IV from the hen oviduct was shown to be a  $\beta$ 1,4-GnT that can act on glycopeptides with the bi-antennary structure GlcNAc $\beta$ 1-2Man $\alpha$ 1-3 (GlcNAc $\beta$ 1-2Man $\alpha$ 1-6) Man $\beta$ 1-4GlcNAc1-4GlcNAc-Asn- where both Man3- and Man6-arms had to have GlcNAc residues (added by GnT I and II) for full activity. Extension of these GlcNAc residues by Gal yielded 5% activity, while sialylation completely blocked GnT IV activity.

The tetra-antennary N-glycans are synthesized by GnT V, which is involved in cell adhesion and recognition and in cancer development [47] (Table 2). Thus, GnT V inhibitors are of importance to obtain potential tools to fight cancer and metastasis (Table 3). In contrast to GnT II, GnT V from the hamster kidney, which also acts on the Man6-arm, does not require the Man3-arm in its acceptor substrate and can act on linear trisaccharides such as GlcNAc $\beta$ 1-2Man $\alpha$ 1-6Man $\beta$ -octyl [48,49]. Paulsen's group synthesized biantennary substrate analogs that showed how specific hydroxyls of both the Man3- and Man6-arms influence GnT V activity (Figure 3). However, while the 4-hydroxyl of GlcNAc $\beta$ 1-2 and the 4- and 6-hydroxyls of Man $\alpha$ 1-6 on the Man6-arm were absolutely essential for activity, replacement of the 3-hydroxyl of GlcNAc $\beta$ 1-2 with a larger pivaloyl group prevented activity, presumably due to steric interference [48]. Kanie et al. [50,51] also reported the critical role of the 3-, 4-, and 6-hydroxyls of the  $\beta$ 1-2-linked GlcNAc in enzyme recognition. This clearly established the pathway of biosynthesis, where GnT II must add a GlcNAc $\beta$ 1-2 residue to the Man6-arm before GnT V can act, and transfer of Gal $\beta$ 1-4 to GlcNAc $\beta$ 1-2 blocks GnT V action.

GnT V is an inverting GT, but does not appear to have a DxD motif. Instead, a Glu residue forms the catalytic base, and the activity is independent of divalent metal ions. The crystal structure of the luminal domain shows a GT-B-like fold with 2 Rossmann domains [52]. A pocket that accommodates the Man1-6 residue of the acceptor is lined by the aromatic rings of F380 and W401. GnT V requires a non-catalytic domain for activity towards glycoproteins [53,54].

GnT III requires the prior action of GnT I to add the bisecting GlcNAc $\beta$ 1-4 to the internal core Man $\beta$ 1-4 residue. The enzyme can also act on biantennary substrates and produce hybrid and biantennary bisected N-glycans, which causes a crowded configuration that may be the reason for the reduction of further N-glycan antennae synthesis and processing. Paulsen synthesized N-glycans with 9 sugar residues, including the bisecting

GlcNAc [40]. The steric effect of the bisecting GlcNAc was examined by NMR, which showed that the Man $\alpha$ 1-6 residue is present in a different (gg) preferred conformation [55].



**Figure 3.** Specificity of GlcNAc-transferase V. GnT V transfers a GlcNAc $\beta$  residue to the 6-hydroxyl of the Man $\alpha$ 1-6 residue of the N-glycan core. The acceptor substrate for purified GnT V from the hamster kidney was a derivative of the 6-arm, GlcNAc $\beta$ 1-2Man $\alpha$ 1-6Man $\beta$ -R. The 3-arm of the N-glycans is not essential for GnT V activity. R can be an oligosaccharide or a hydrophobic group such as octyl. The Man $\beta$ -residue can be replaced by Glc $\beta$  having full activity. To determine acceptor specificity, hydroxyls were deleted or replaced by a larger group. The rectangular boxes indicate an absolute requirement for the hydroxyl, while the circles indicate that the hydroxyl is important for the activity. Replacement of the 4- and 6-hydroxyls of the Man $\alpha$ 1-6 residue with an O-methyl group yielded the GnT V inhibitor GlcNAc $\beta$ 1-2(4,6-di-O-methyl)Man $\alpha$ 1-6Glc $\beta$ -pnp.

GnT VI synthesizes an additional antenna on the Man6-arm. This  $\beta$ 1,4-GnT activity was found in hen oviduct [56,57] and fish. GnT VI can act on the trisaccharide acceptor GlcNAc $\beta$ 1-2 (GlcNAc $\beta$ 1-4)Man $\alpha$ -methyl, which resembles the product of GnT II and GnT V.

### 13. Modifications of N-Glycan Antennae

Lysosomal hydrolases have N-glycans with a specific Man6-phosphate (Man6-P) marker that targets them to the lysosomes via a specific Man6-P receptor [58]. In Golgi, the GlcNAc-P-transferase GNPT ABG complex transfers GlcNAc-P from UDP-GlcNAc to the 6-hydroxyl of a Man residue on lysosomal enzymes. This is followed by hydrolysis of GlcNAc to expose the Man6-P marker that can bind to the Man6-P receptor in the Golgi, which reroutes the glycoprotein to acidic compartments and to the lysosomes.

Paulsen's compounds helped to define the association of glycoprotein biosynthesis in biological processes, such as apoptosis in porcine aortic endothelial cells [59]. GnTs I and II are usually the major GnT activities in tissues and cultured cells, with GnTs III, IV, and V activities being low in vitro. Further extension of N-glycan antennae includes the addition of Gal by  $\beta$ 1,4-GalT B4GalT1 [18]. The inhibition of Gal extension could reduce the addition of Sia and multiple other epitopes that may play a role in the immune system and in cancer [22] (Table 3). Sialic acids, which are often found in increased amounts in cancer, form selectin (Siglec) ligands and play a role in cancer cell metastasis, immune evasion,

and cancer cell survival. The expression of ST6Gal1 has been shown to be increased in cancer [24]. ST6Gal1 has been crystalized with CMP. The inverting enzyme has a GT-A-like fold and may act via an  $S_N2$  mechanism [60]. However, much of the knowledge about SiaT mechanisms comes from studies of bacterial enzymes with similar activities.

## 14. Inhibition of Glycosyltransferases

GTs bind their substrates by hydrogen bonding and hydrophobic interactions and prefer specific conformations and flexibilities. Inhibitors may compete with substrates or covalently bind to GTs to cause loss of activity. Paulsen's approach to inhibitor design was to synthesize many short and long analogs of natural acceptor substrates for GTs. The synthesis of potential inhibitors included removal of hydroxyls that are essential for sugar transfer, blocking access to critical hydroxyls by introducing large groups, or preventing the folding of enzymes in the active conformation. Other strategies that could cause loss of activity could focus on the donor binding site and the binding of essential divalent metal ions.

Compounds were either free oligosaccharides or had a hydrophobic group at the reducing end of the chain. Hydroxyl groups were either deleted or replaced by methyl groups or by potentially reactive substituents such as azide and epoxide groups that can react with nucleophilic amino acids, as well as diazirino and iodoacetamido groups that are activated by UV. These modified substrates enhanced our understanding of how GTs bind to their substrates, and that small changes in the stereochemistry, conformation, or electronic, bulky, or hydrophobic character of the sugar ring substituents, depending on their positions, could block enzyme activity by affecting interactions with specific amino acids in the substrate binding and catalytic sites of enzymes. A basic requirement for understanding biosynthetic mechanisms is knowledge of the enzyme structure and substrate specificity.

Several N-glycan derivatives proved to be potent inhibitors of GTs (Table 3). However, due to their large size and hydrophilicity, it is difficult to envision their transport across the membranes to reach the GTs in the Golgi.

Since extended glycan chains carry many cancer-associated antigens, inhibition of glycoprotein extension may be useful for biological studies. Gal is added to the GlcNAc of the N-glycan antennae by a family of enzymes (Table 1). We examined the specificity and inhibition of purified bovine B4GALT1 with a large panel of acceptor substrate derivatives. While GlcNAc $\beta$ -Bn was the standard acceptor, GlcNAc $\alpha$ -Bn was only 5% active, and free GlcNAc showed 29% activity, indicating that it is mainly the  $\beta$  anomer that forms the substrate. Glycopeptide acceptors having GlcNAc $\beta$ 1-3/6GalNAc-glycans were relatively less active than GlcNAc $\beta$ -Bn. Derivatives with the large 2-naphthyl as aglycone could not support activity, and most of these compounds inhibited the enzyme. 1-Thio-N-butyrylGlcNAc $\beta$ (2-naphthyl) inhibited B4GALT1 by 100% at 1 mM with a  $K_i$  of 0.01 mM [18]. However, bicyclic quinolinyll compounds (having nitrogen in the ring) were active substrates. This indicates that the overall structure of the acceptor and the conformation and electronic properties of the aglycone are important for enzyme recognition. The catalytic domain of B4GALT1 has two flexible loops that change from an open to a closed conformation upon binding of UDP-Gal and metal ions [61]. This creates an acceptor binding site and seems to be a mechanism common to several GTs. After the reaction, the oligosaccharide product is released, and the loops revert to the initial conformation, releasing UDP. Thus, additional strategies are required to make inhibitory compounds applicable in biological systems. Future studies will be seeking higher potency and specificity of inhibitors, targeting strategies, and examining the biological effects in cell cultures and in animals before development into human therapeutics.

## 15. Mucin Glycoproteins

Mucins are very large glycoproteins that are heavily O-glycosylated, sometimes with 90% carbohydrates. They may also carry several N-glycans. Mucins form a viscoelastic mucus layer over the epithelia and protect the host from environmental stresses, dehydration, and infections. Humans have 20 or more genes that encode mucins and many other proteins that have mucin-like O-glycosylated domains. Human mucin-type O-glycans all start with a GalNAc residue  $\alpha$ linked to Ser or Thr. Four core structures are common, with several less common structures such as core 5, GalNAc $\alpha$ 1-3GalNAc- (Table 4) [23,30]. These core structures can be sialylated or extended with structures and epitopes similar to those found on complex N-glycans. The O-glycan structures often vary between cell types and in conditions such as inflammation or in cancer and are regio-specific in the intestine. Thus, the expression of genes encoding GTs is delicately regulated, although our knowledge in this field is still marginal [23].

**Table 4.** Glycopeptides synthesized by Hans Paulsen. Paulsen synthesized hundreds of glycopeptides, some with 1 to 23 amino acids and 1 to 3 glycan chains attached to the longer peptides. Glycans were various combinations of GalNAc, core 1 to 7. The table shows one example of each series and the systematic changes to make variants of either amino acids, attachment site of glycan, type of core structure, or number of glycans attached. Bold amino acids carry glycans.

| Different Series with Examples of Glycopeptides                                                                                                                                                        | Variants Synthesized                                     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------|
| AcTPPP                                                                                                                                                                                                 |                                                          |
| 1. APT <sup>GalNAc<math>\alpha</math></sup> S <sup>GalNAc<math>\alpha</math></sup> SS                                                                                                                  | Thr/Ser, position of GalNAc                              |
| 2. APT <sup>Gal<math>\beta</math>3GalNAc<math>\alpha</math></sup> SSS                                                                                                                                  | Thr/Ser, position of core 1                              |
| 3. Ac-T <sup>GalNAc<math>\alpha</math></sup> P-t-Bu                                                                                                                                                    | protected N/C-terminal protection                        |
| 4. APT <sup>GalNAc<math>\alpha</math></sup> -SSSTKKT                                                                                                                                                   | peptide length                                           |
| 5. Ac-VT <sup>GalNAc<math>\alpha</math></sup> P-NH <sub>2</sub>                                                                                                                                        | peptide length, N/C-terminal protection                  |
| 6. Ac-PTTT <sup>GalNAc<math>\alpha</math></sup> PIST-NH <sub>2</sub>                                                                                                                                   | amino acids, position of GalNAc                          |
| 7. Ac-PTPTGTQTPTT <sup>GalNAc<math>\alpha</math></sup> TPITTTTTVTPT-NH <sub>2</sub>                                                                                                                    | number of GalNAc                                         |
| 8. AHGVTS <sup>Gal<math>\beta</math>3GalNAc<math>\alpha</math></sup> APDTRPAGSTAP<br>T <sup>Gal<math>\beta</math>3GalNAc<math>\alpha</math></sup> A                                                    | amino acids, position and number of GalNAc, core 1       |
| 9. PTTTPTTTG                                                                                                                                                                                           | amino acid sequence                                      |
| 10. Ac-PS <sup>GalNAc<math>\alpha</math></sup> S <sup>GalNAc<math>\alpha</math></sup> S <sup>GalNAc<math>\alpha</math></sup> PIST-NH <sub>2</sub>                                                      | amino acids, number, Ser/Thr position of glycans         |
| 11. TTT <sup>GlcNAc<math>\beta</math>3GalNAc<math>\alpha</math></sup> VTPT <sup>GlcNAc<math>\beta</math>3GalNAc<math>\alpha</math></sup> PTG                                                           | number and position of GalNAc, core 3                    |
| 12. TETTSHT <sup>GalNAc<math>\alpha</math></sup> PG                                                                                                                                                    | number and position of GalNAc, core 3, length of peptide |
| 13. TTT <sup>Gal<math>\beta</math>3GalNAc<math>\alpha</math></sup> VTPT <sup>Gal<math>\beta</math>3GalNAc<math>\alpha</math></sup> PTG                                                                 | position and number of core 1                            |
| 14. TTTVTPTPT <sup>GlcNAc<math>\beta</math>6(Gal<math>\beta</math>3)GalNAc<math>\alpha</math></sup> G                                                                                                  | position and number of core 1, 2                         |
| 15. TTT <sup>GlcNAc<math>\beta</math>6(Gal<math>\beta</math>3)GalNAc<math>\alpha</math></sup> VTPT <sup>Gal<math>\beta</math>3GalNAc<math>\alpha</math></sup> PTG                                      | position of core 1, 2                                    |
| 16. TTT <sup>GlcNAc<math>\beta</math>6(GlcNAc<math>\beta</math>3)GalNAc<math>\alpha</math></sup> VTP<br>T <sup>GlcNAc<math>\beta</math>6(GlcNAc<math>\beta</math>3)GalNAc<math>\alpha</math></sup> PTG | position and number of core 4                            |
| 17. TTTVTPT <sup>GlcNAc<math>\beta</math>6GalNAc<math>\alpha</math></sup> PTG                                                                                                                          | position and number of core 6                            |
| 18. TTT <sup>GlcNAc<math>\beta</math>6(GlcNAc<math>\beta</math>3)GalNAc<math>\alpha</math></sup> VTPT <sup>GlcNAc<math>\beta</math>6GalNAc<math>\alpha</math></sup> PTG                                | position and number of core 4, 6                         |
| 19. Ac-PT <sup>GalNAc<math>\alpha</math></sup> T <sup>Gal<math>\beta</math>3GalNAc<math>\alpha</math></sup> T <sup>Gal<math>\beta</math>3GalNAc<math>\alpha</math></sup> PIST-NH <sub>2</sub>          | position and number of GalNAc and core 1                 |
| 20. Ac-PTTT <sup>Gal<math>\beta</math>3GalNAc<math>\alpha</math></sup> PIST-NH <sub>2</sub>                                                                                                            | position of core 1 $\alpha$ or $\beta$ , amino acids     |
| 21. TETTSHT <sup>Gal<math>\beta</math>3GalNAc<math>\alpha</math></sup> PG                                                                                                                              | position and number of core 1, 2, 4, 6                   |
| 22. Ac-ELST <sup>GalNAc<math>\alpha</math></sup> T <sup>GalNAc<math>\alpha</math>3GalNAc<math>\alpha</math></sup> GPG- NH <sub>2</sub>                                                                 | amino acids, number and position of GalNAc, core 5, 7    |
| 23. Ac-ELAT <sup>GalNAc<math>\alpha</math></sup> VGPG-NH <sub>2</sub>                                                                                                                                  | amino acids, GalNAc, core 5, 7                           |

Mucins are characterized by variable number of tandem repeat (VNTR) regions that are rich in Thr/Ser/Pro residues and carry most of the O-glycans. There are three types of mucins in humans: cell surface-bound mucins, small, secreted mucins, and large, gel-forming mucins. For each mucin, several variants have been identified. Small, secreted mucins (MUC7) are relatively soluble and found in the saliva. Cell surface-bound mucins (MUC1,3,4,12,13,15,16,17,20,21,22) have a transmembrane domain and a cytoplasmic tail. These often interact with other cell surface proteins and are involved in cell signalling. Large, gel-forming mucins (MUC2,5AC, 5B,6) have a Cys-rich D domain and are polymerized to several millions of molecular weight.

## 16. Role of Mucins

The role of mucins varies from protection in mucus to cell adhesion and regulation of tumour survival and diagnostic markers in cancer. Some mucins carry hundreds of O-glycans with different structures and lengths up to more than 20 sugars, which are hydrophilic and can be negatively charged due to sialic acid and sulfate esters and bind to metal ions. The properties of mucins are dictated by the abundance of O-glycans, which cause an extended conformation of mucins and prevent protease cleavage. Mucins play an essential role in intestinal health, supporting beneficial microbiota and avoiding pathogens. O-glycans form adhesive receptors for bacteria and other microbes that can degrade them, use glycans as a source of nutrition, and compete for adhesion [62,63]. O-glycans also play a role in fertilization and sperm–egg binding. Since O-glycans often contain epitopes and antigens, e.g., Lewis-type structures, that are recognized by antibodies and lectins, and they have a controlling function in the immune system. Simple O-glycans are often associated with cancer cells, and sialylated O-glycans have been shown to increase in cancer [21]. They regulate the survival and attachment of cancer cells and can be useful biomarkers and basis for vaccines. Some growth factors and cytokine receptors are O-glycosylated. For example, death receptors require O-glycosylation for efficient apoptotic signalling in tumours [64].

There are four major mucin type O-glycan core structures (1 to 4) (Table 4), with core 1 and 2 being mostly universal, while core 3 and 4 are found in mucins from specific tissues such as intestinal or lung tissue [23]. These core structures can be extended by a variety of sugars and epitopes. Core 5 is a rare structure that is not extended but can be sialylated, and it has been found in human meconium and adenocarcinoma [30,65]. Other core structures (6 to 8) are extremely rare. The core structures are expected to have different conformations and thus display their extensions and epitopes differently. For example, cores 2 and 4 have a GlcNAc $\beta$ 1-6 branch that can occupy a significant amount of space due to the rotation of the 1-6 bond. This was established by analyzing core 1 to 6 structures using NMR and energy calculations [66].

The overwhelming majority of mucins are bound to the cell membrane and have multiple roles in signalling and recognition events. These Type I membrane glycoproteins can form complexes with growth factors and tyrosine kinases; they can be receptors for mammalian lectins such as Siglecs and can interact with proteins and cells of the immune system, controlling inflammatory responses and immune responses against tumours. The cytoplasmic C-termini can be phosphorylated and interact with proteins of signal transduction and regulation of survival and cell death signals. The extracellular domains can also be cleaved and then secreted as soluble mucins with a number of additional roles. In cancer, MUC1 expression in particular has been found to be elevated in addition to alterations of their O-glycan structures [67]. For example, simple O-glycan structures Tn and T antigens (Table 2) and their sialylated forms are typically present in increased amounts in cancer, together with anti-Tn and anti-T antibodies. MUC1 and MUC4 have consistently been shown to be highly expressed in cancer. Thus, cancer vaccines based on MUC1 have been

developed and tested in clinical trials, with some success in diminishing tumour load. MUC4 is a marker for pancreatic cancer and is highly immunogenic, and vaccines have been developed to target cells expressing MUC4. MUC16 is also known as ovarian cancer antigen CA125.

## 17. Glycosyltransferases That Assemble Mucin O-Glycans

O-glycans are assembled step by step in the Golgi and are directly attached to the side chain oxygen of Ser or Thr residues in the completed polypeptide chains. There is no known single amino acid sequence common for all O-glycans, but databases allow an estimate of probable O-glycosylation based on statistics of known sequences around O-glycosylation sites of glycoproteins. The first sugar residue of O-glycans (GalNAc) is added by GTs localized to Golgi membranes (Table 1). GALNT occurs as 20 isoenzymes that O-glycosylate peptides in humans. GalNAc is transferred to peptide by a family of at least 20 polypeptide GalNAc-transferases (GALNTs) that are expressed in a cell-type-specific fashion and support the essential function of mucin O-glycans. GALNTs are classified as GT27 enzymes, and most have a lectin domain connected to the catalytic domain by a flexible linker sequence. GALNTs have similar broad specificities for the amino acids near the O-glycosylation site, with a preference for nearby Pro residues that may enhance recognition of the hydroxyls of Ser/Thr. GALNT7 requires a pre-existing GalNAc-O-glycan in the peptide [68]. The crystal structures of GALNT1,2,3,4,7,10,12 have been determined and show complex structures with a GT-A fold in the GT domain. A DxH motif is found to coordinate UDP-GalNAc and a  $Mn^{2+}$  cofactor [69]. Several GALNTs are aberrantly expressed in cancer. For example, the expression and activity of GALNT14 in hepatocellular and gastrointestinal cancer may be a useful disease marker [25].

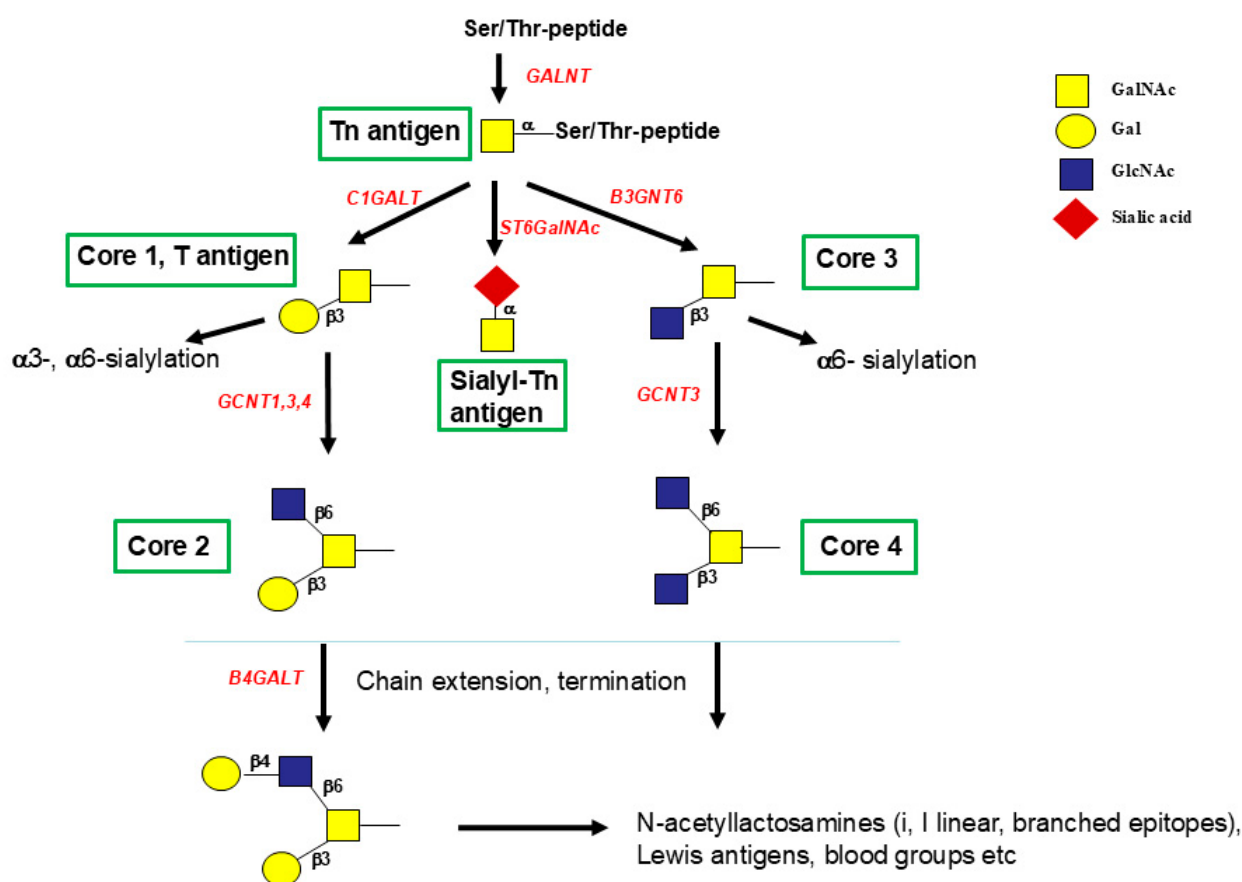
GalNAc $\alpha$ -Ser/Thr (Tn antigen) as well as its sialylated form sialyl $\alpha$ 2-6GalNAc $\alpha$ -Ser/Thr (sialyl-Tn antigen) are epitopes commonly found in cancer. The  $\alpha$ 6-sialylation of GalNAc prevents further processing, with the exception of O-acetylation of Sia, which then masks the sialyl-Tn antigen.

The enzymes that synthesize the four major mucin-type O-glycan core structures (1 to 4) have been characterized and the genes cloned (Figure 4). Crystal structures are available for core 1  $\beta$ 1,3-GalT (C1GALT1) and core 2 ( $\beta$ 1,6-GnT, GCNT1). The expression of C1GALT1 is uniquely dependent on the coexpression of the chaperone COSMC [70], which prevents protein degradation by ensuring proper folding of C1GALT1. In the absence of Cosmc, C1GALT1 is ubiquitinated and targeted to the proteasome for degradation. It has been observed that a number of cancer cells lack core 1 and its further metabolites, and cells express the Tn and sialyl-Tn antigen due to the lack of COSMC. The occurrence of unmodified core 1, the T antigen, is common in cancer and during early development as an oncofoetal antigen.

C1GALT1 acts on a variety of GalNAc $\alpha$ -R acceptor substrates, including glycoproteins and the GalNAc $\alpha$ -glycopeptides, synthesized by the Paulsen group. Control assays for C1GALT1 are usually carried out with GalNAc $\alpha$ -Bn acceptor, which is an acceptor for core 1 and core 3 synthesis and has been used as an O-glycosylation inhibitor in cells since it competes with natural glycoprotein biosynthesis. The problem is that GalNAc $\alpha$ -Bn is converted to a number of benzyl-terminating hydrophobic O-glycan structures, which overwhelm the storage and transport capacity of the cell and induce apoptotic cell death [71]. Since core 1 and 3 are precursors for core 2 and 4 structures (Figure 4) GalNAc $\alpha$ -Bn inhibits the synthesis of core 1 to 4 in glycoproteins.

Derivatives of GalNAc $\alpha$ -Bn were synthesized and tested as potential substrates and inhibitors [72]. Removal or substitution of the 6-hydroxyl of GalNAc slightly reduced C1GALT1 activity. 6-O-(4,4-azo)pentyl- GalNAc $\alpha$ -Bn was photosensitive and inhibited

Gal incorporation into GalNAc $\alpha$ -Bn (Table 3). The results indicate that the enzyme has a relatively broad acceptor specificity and does not require the 6-hydroxyl of GalNAc, but it needs the 3- and axial 4-hydroxyl as essential requirements for binding and activity.



**Figure 4.** O-glycosylation pathways, synthesis of cores 1 to 4. All mucin-type O-glycans are linked through GalNAc $\alpha$  to Ser or Thr. This linkage is formed by one or more of 20 polypeptide GalNAc-transferases (GALNTs) that are expressed in a cell type-specific fashion. GalNAc can be converted to core 1 by  $\beta$ 1,3-Gal-transferase C1GALT1, or to core 3 by  $\beta$ 1,3-GlcNAc-transferase B3GNT6. Unmodified GalNAc is recognized as the cancer-associated Tn antigen. If sialic acid is added to GalNAc by ST6GalNAc1, the O-glycan becomes the Sialyl-Tn antigen, which is not further extended. Core 1 can be converted to core 2 by GCNT1, GCNT3, or GCNT4. Core 3 is converted to core 4 only by GCNT3. Unmodified core 1 is named the T antigen, which is often found in cancer. Sialylation of the T antigen also stops further extension to complex chains.

Yeh et al. [73] cloned a number of  $\beta$ 1,6-GnTs, GCNT1, 2 and 3. GCNT1 introduces a GlcNAc $\beta$ 1-6 branch onto core 1, Gal $\beta$ 1-3GalNAc-R, linked to various hydrophobic groups or peptides, and synthesizes core 2 (Figure 4). This activity increases in many cancer and leukaemia cells [74] and upon cancer cell differentiation. GCNT1 plays an important role in the immune system by providing the scaffold for selectin ligands sialyl-Lewis x, crucial for lymphocyte interactions [75]. In specific breast cancer cells, however, GCNT1 is not expressed [76]. The expression shows large variations among cancer cells derived from prostate cancer [31]. Although the enzyme is an inverting GT with a GT-A fold [33], it does not require  $Mn^{2+}$  as a cofactor. GCNT1 is N-glycosylated and requires several conserved Cys residues for the formation of disulfide bonds. Amino acid reagents that target non-specifically His, Trp, and Cys inhibited purified GCNT1 activity up to 100%.

The groups of Hans Paulsen and Khushi Matta synthesized many derivatives of core 1, which revealed the specificity of GCNT1. Gal $\beta$ 1-3(6-deoxy) GalNAc $\alpha$ -Bn was

found to be a weak inhibitor. Surprisingly, the excellent substrate Gal $\beta$ 1-3GalNAc $\alpha$ -p-nitrophenyl inhibited GCNT1 activity from acute myelogenous leukaemia cells by 87% upon UV irradiation.

Another  $\beta$ 1,6-GnT, GCNT4, has similar activity and 44% sequence identity with GCNT1 but has a selected tissue expression. It is associated with the thymus and downregulated in a number of different cancer cells and tissues [26].

The metal ion-independent  $\beta$ 1,6-GnT, GCNT3, synthesizes both core 2 from core 1 and core 4 from core 3. The enzyme has high activity in colonic tissue, explaining the abundance of core 4 O-glycans in colonic mucins [23]. GCNT3 has a 50.35% sequence identity with GCNT1 and is also classified as GT14, but shows very different inhibition by bis-imidazolium salts.

In many cancer cells, both core 3 and 4 synthesis by B3GNT6 and GCNT3 showed altered activity [77]. GCNT3 is not expressed in normal or cancer-derived prostate cells [31], but is variably expressed in other cancer types [29]. Huang et al. [78] suggested a role for GCNT3 in suppressing cancer.

The crystal structure of murine GCNT1 [79,80] suggests that the DxD motif is replaced by the DE sequence, with Glu320 as the catalytic base. Two positively charged amino acids (Arg378 and Lys 401) take the role of a Mn<sup>2+</sup> ion in stabilizing the UDP leaving group. These amino acids are conserved among these  $\beta$ 1,6-GnTs. The other metal ion-independent  $\beta$ 1,6-GnT, GCNT2, is 37.8% identical to GCNT1 but has a different acceptor specificity and synthesizes the branch of the I antigen [73]. It is not clear which of the Arg residues cooperate with Lys in binding UDP.

GalNAc $\alpha$ -Bn is also an acceptor for  $\beta$ 1,3-GnT, B3GNT6, that synthesize core 3. The enzyme was discovered in rat and human colon tissues that are rich in mucins carrying core 3. Many colon cancer cell lines and colon cancer tissue, however, have very low enzyme activity [77]. While GnTs I, II, and V are variably active in normal and cancer prostate cells, B3GNT6 was not detected [31]. The lack of core 3 synthesis may be one factor promoting core 1 (T antigen) expression.

Rare core structures in human glycoproteins include cores 5 to 8 (Table 4). Core 5, GalNAc $\alpha$ 1,3-GalNAc-, was found in human intestinal fetal mucins and in adenocarcinoma [30,65]. A preliminary report of core 5 synthesis using the GalNAc-mucin acceptor [81] showed that Sia may be added to GalNAc after the formation of core 5. However, the gene of this  $\alpha$ 1,3-GalNAcT has not yet been identified. It is possible that this activity is caused by a variant of a known enzyme. In contrast to O-glycan core structures 1 to 4, core 5 has not been found to be elongated, and its role is unknown. The activities or genes encoding the GTs that synthesize core 6 to 8 structures remain to be identified. It is possible that core 6 in mucins is a degradation product of core 2, due to bacterial hydrolases or due to degradation during O-glycan isolation.

## 18. Elongation of O-Glycans

A wide array of glycosyltransferases can extend core structures in complex pathways and create hundreds of different O-glycan structures with lengths of more than 20 sugar residues in mucins. The Gal and GlcNAc residues of O-glycan core structures can be elongated by enzymes that similarly elongate N-glycans. Very few of these enzymes are specific to O-glycans. This includes  $\beta$ 1,3-GnT, B3GNT3, which adds GlcNAc to Gal of core 1 and 2 [23]. The enzyme can act on mucins or hydrophobic acceptors with core 1 or 2 structures and thus allows the attachment of complex glycans and lymphocyte homing receptors [82]. Core 2 sometimes carries an unusual GlcNAc $\alpha$ 1-4 residue linked to Gal in gastric cells, such as pyloric glands and neck cells [83]. The retaining  $\alpha$ 1,4-GnT A4GNT

may use the large mucin MUC6 as a scaffold. The expression of A4GNT together with MUC6 was found to suppress the proliferation of pancreatic tumour cells [84].

Several SiaTs are specific for mucin-type O-glycans and generally block further extension. All human SiaTs are classified in the inverting GT29 family and have common sialylmotifs that contribute to the binding of substrates [85]. ST6GalNAc1-4 transfer Sia from CMP- $\beta$ Sia in  $\alpha$ 2-6 linkage to GalNAc-mucins with different specificities for core 1 and sialylated core 1. ST6GalNAc1 primarily synthesizes the sialyl-Tn antigen, while ST6GalNAc2 primarily synthesizes the sialyl-T antigen. ST6GalNAc2 can act on GalNAc- and core 1- or 3-peptide acceptors [28]. The crystal structure of ST6GalNAc2 with CMP [86] shows a GT-A variant2 fold [87]. ST6GalNAc3 and 4 utilize core 1 substrates that have already been sialylated by ST3GAL1, but they do not require a peptide in the acceptor. ST6GalNAc3 has a restricted expression mainly in the brain and kidney.

ST3Gal1 specifically adds Sia in  $\alpha$ 2-3 linkage to Gal of core 1 and 2 linked to peptide or a hydrophobic group. The ST3GAL1 product, Sialyl-T antigen, is a common antigen in cancer cells. ST3Gal1 is overexpressed in breast cancer tissues [27] and in prostate cancer cells [31], where it controls proliferation, migration, and apoptosis. Porcine ST3Gal1 has been crystallized and revealed a single Rossmann domain, like other SiaTs, and a  $\beta$ -sheet core that contributes to binding CMP-Sia and His319 as a possible catalytic residue [88].

## 19. Synthesis of Glycopeptides

Because of the importance of mucins and O-glycosylation in human health and diseases such as cancer, the Paulsen group synthesized more than 200 peptides and glycopeptides derived from MUC1 to 4 and other glycoproteins for studies of their biological role and pathways of O-glycan biosynthesis (Table 4). It has been extremely difficult to determine the sites of O-glycan addition in complex mucins and to assign O-glycan structures to specific Thr and Ser residues. Thus, peptides and glycopeptides have been a tremendous tool to discover mechanisms of O-glycosylation, the influence of amino acids near O-glycosylation sites in acceptor substrates for GALNTs, and new cancer-associated and immunogenic epitopes.

Hundreds of glycopeptides have been produced by Paulsen's group, together with the unglycosylated control peptides as standards and controls. The development of solid-phase synthesis of GalNAc $\alpha$ - and Gal $\beta$ 1-3GalNAc $\alpha$ -glycopeptides using protected glycosylamino acids was a breakthrough [89,90]. Novel glycosyl amino acids were synthesized as building blocks for Fmoc-based continuous-flow synthesis using multi-column technology. This achieved synthesis of several glycopeptides simultaneously without significant degradation of the glycan-peptide bond and beta-elimination of the O-glycan. The amino acid sequences had different lengths up to 23 amino acids and were based on those from the VNTR of MUC1, 2, 3, or 4. The glycans attached to specific Thr or Ser residues included one or more GalNAc residues and selected core 1 to 7 structures (Table 4). Glycopeptides were purified, and their sequences and glycosylation were confirmed by HPLC, MS, and NMR. Computational analyses and NMR determined the conformations of glycopeptides that could potentially be recognized by individual enzymes.

To examine the preferred attachment sites for GALNTs, a series of glycopeptides were synthesized that were overall neutral with an acetyl group at the N-terminus and amide at the C-terminus and contained one or more GalNAc $\alpha$ -Thr/Ser or Gal $\beta$ 1-3GalNAc $\alpha$ , or in the unnatural  $\beta$ -linkage to Thr. The glycopeptide reaction products were analyzed by the amount of radioactive sugar transferred and by their HPLC elution times using standard glycopeptides.

Within a series, the positions of glycans and some of the Thr/Ser residues were altered in a systematic fashion to include all possibilities of glycan attachments [91]. In addition,

Pro and Thr were replaced by other amino acids at +1 or -1 positions to determine their role in a GALNT substrate. The smallest peptide used as a positive control was Ac-TPPP, which showed the highest activity as an acceptor for highly purified bovine GALNT1 [92,93]. The small acceptor substrate does not have a C-terminal protection group, but clearly has easy access to the catalytic site. In glycopeptides 8 amino acids long, Pro in the +3 position (towards the C-terminus) supported the activity. None of the Thr residues are exclusive sites for GalNAc addition, but the amino acid sequences and sites of existing glycosylation clearly dictate further O-glycosylation. Substrates with core 1 in  $\beta$ -linkage are less active, likely due to the different conformations compared to the natural  $\alpha$ -linkage.

A series of different lengths and GalNAc attachment to Ser did not show significant activity. It is possible that the slight hydrophobic effect of Thr plays a role in enzyme recognition. However, other members of the GALNT family may show different results and could act on Ser. Generally, site specificity is generally broad, and larger glycopeptides and those with different amino acid sequences remain to be tested as models for mucin synthesis. Glycosylated peptides are more rigid in their conformations compared to peptide alone, and large O-glycans cause crowding. Non-mucin glycoproteins often have glycosylation sites only near the N-terminus, and protein folding may block glycosylation. Mucins have large complex O-glycans, often at adjacent amino acids, and the processing enzymes only have access to their substrates in an unfolded (bottlebrush) conformation of the glycoprotein.

Although there is clearly no consensus sequence for O-glycosylation, TxxP is one of the preferred sites for mucin O-glycosylation by GALNT1. O-glycosylation is extremely complex, with substrate conformation, composition, hydrophobicity, existing glycosylation, as well as overall structure controlling recognition by GALNTs and other GTs.

Subsequently, C1GalT1, GCNT1, and B4GALT1 were examined for their ability to use glycopeptide substrates in a site-specific manner [93–95]. The synthesis of core 1 by partially purified rat liver C1GALT1 was examined with GalNAc-containing glycopeptides containing 1 to 3 GalNAc-Thr. Compared to the GalNAc $\alpha$ -Bn acceptor, all of the GalNAc-glycopeptides were good substrates for the enzyme that could add 1 to 3 Gal residues to each GalNAc. The enzyme was strongly influenced by the overall composition and especially amino acids at the -1 and +1 positions. These studies indicate that in addition to initial glycosylation, the second step in the glycosylation pathways of O-glycans is also controlled by the structure and glycosylation of the peptide core of substrates.

The activity of purified human GCNT1 was influenced by amino acid composition, the position of the core 1 substrate, and other glycans within the glycopeptide. Surprisingly, glycopeptides with core 1 in the unnatural  $\beta$ -linkage were also variably good substrates for GCNT1. The extension of core 2-glycopeptides with purified bovine B4GALT1 also showed significant variability. However, B4GALT1 has a higher activity with GlcNAc $\beta$ -Bn as a substrate than with glycopeptides.

Overall, glycopeptides helped to determine that the control of O-glycosylation is based on a multitude of complex factors, including amino acid position, properties, and charge as well as existing glycosylation. GTs that act close to the peptide (GALNT and C1GALT1) are more specific than those adding the third (GCNT1) or fourth (B4GALT1) sugar residue. Currently, Paulsen glycopeptides are still under investigation for their ability to form GT substrates, to discover novel enzymes, to enzymatically synthesize novel glycopeptides, and as potential inhibitors of bacterial biofilm formation.

## 20. Conclusions

Paulsen developed new synthetic technologies by synthesizing complex large N-glycans, modified glycans, and their fragments. This has been documented in hundreds of publications and presentations (Figure 5). His new synthesis included the discovery

of methods to synthesize substituted Man $\beta$ -R compounds and to initiate solid-phase glycopeptide synthesis. Paulsen's substrates are essential for the discovery of new enzymes, for defining enzyme specificities that control the pathways of N- and O-glycosylation, and for defining the biological roles of glycans. The roles include receptors for lectins and adhesion proteins in eukaryotic systems and in bacteria that use mucins as a niche for survival and nutrition. Disease-specific alterations in GTs could be discovered. The inhibitors that Paulsen synthesized will help in the development of future therapies for cancer and other diseases, where the reduction in complex glycans could be beneficial. Many synthetic routes and analytical tools for complex glycans and glycopeptides have been developed and subsequently adopted worldwide. Future studies are expected to use this information to understand how enzymes recognize their glycan substrates and how this influences pathways of glycoprotein biosynthesis. We are grateful for the hard work of Hans Paulsen and his coworkers and students. He believed that the principles behind the contributions to glycobiology are based on the enjoyment of chemistry!



**Figure 5.** Hans Paulsen at the International Carbohydrate Symposium in Whistler, Canada (2006).

**Funding:** Funding was provided by an NSERC Discovery grant 2022-05389.

**Acknowledgments:** The author thanks the hundreds of students and postdoctoral fellows who worked with Hans Paulsen, along with many collaborators, for their hard work and dedication.

**Conflicts of Interest:** The author declares no conflicts of interest.

## References

1. Heyns, K.; Paulsen, H. Oxidative transformation of carbohydrates. VIII. Catalytic oxidation of meso-inositol to scyllo-meso-inosose. *Chem. Berichte Chem. Eur.* **1953**, *86*, 833–840. [[CrossRef](#)]
2. Heyns, K.; Paulsen, H.; Rudiger, G.; Weyer, J. Configuration and conformation selectivity in catalytic oxidation with oxygen on platinum catalysts. *Fortschritte Der Chem. Forsch.* **1969**, *11*, 285–374.
3. Paulsen, H.; Todt, K. Monosaccharides with a nitrogen-containing ring. XV. Nuclear magnetic resonance spectroscopic studies of hindered rotation of monosaccharides with a nitrogen-containing ring. *Chem. Berichte Chem. Eur.* **1967**, *100*, 3397–3404. [[CrossRef](#)]
4. Paulsen, H.; Hayauchi, Y.; Sinnwell, V. Monosaccharides containing nitrogen in the ring. XXXVII. Synthesis of 1,5-dideoxy-1,5-imino-D-galactitol. *Chem. Berichte Chem. Eur.* **1980**, *113*, 2601–2608. [[CrossRef](#)]
5. Paulsen, H.; Brockhausen, I. From imino sugars to cancer glycoproteins. *Glycoconj. J.* **2001**, *18*, 867–870. [[CrossRef](#)] [[PubMed](#)]
6. Luger, P.; Vangehr, K.; Bock, K.; Paulsen, H. Conformational analysis. Part XXIII. Determination of the conformation of 2-acetamido-1,4,6-tri-O-acetyl-2-deoxy-3-O-(2-acetamido-3,4,6-tri-O-acetyl-2-deoxy- $\alpha$ -D-galactopyranosyl)-2-D-galactopyranose by x-ray structural analysis, NOE measurements and HSEA calculations. *Carbohydr. Res.* **1983**, *117*, 23–38. [[CrossRef](#)]
7. Paulsen, H.; Behre, H.; Herold, C.-P. Acyloxonium ion rearrangement in carbohydrate chemistry. *Fortschritte Der Chem. Forsch.* **1970**, *14*, 472–525.

8. Paulsen, H. Cyclic acyloxonium ions in carbohydrate chemistry. *Adv. Carbohydr. Chem.* **1971**, *26*, 127–195. [\[CrossRef\]](#)
9. Thiem, J.; Rasch, D.; Paulsen, H. Phosphorus-containing carbohydrates, XV. Perkow reaction with  $\alpha$ -acyloxy keto sugars for the synthesis of enol phosphates and their reactions. *Chem. Berichte Chem. Eur.* **1976**, *109*, 3588–3597. [\[CrossRef\]](#)
10. Schuster, O.; Klich, G.; Sinnwell, V.; Kränz, H.; Paulsen, H.; Meyer, B. “Wave-type” structure of a synthetic hexaglycosylated decapeptide: A part of the extracellular domain of human glycophorin A. *J. Biomol. NMR* **1999**, *14*, 33–45. [\[CrossRef\]](#)
11. Depmeier, W.; Jarchov, O.H.; Stadler, P.; Sinnwell, V.; Paulsen, H. Conformation of L-streptose reducing sugars as crystals and in solution. *Carbohydr. Res.* **1974**, *34*, 214–218. [\[CrossRef\]](#)
12. Strumpel, M.; Schmidt, H.-J.; Luger, P.; Paulsen, H. X-ray analysis of 3-O-(6-O-acetyl-2,4-diazido-3-O-benzyl-2,4-dideoxy- $\alpha$ -D-glucopyranosyl)-1,6-anhydro-2,4-diazido-2,4-dideoxy- $\beta$ -D-glucopyranose, a disaccharide having an unusual  $\alpha$ -D-(1 $\rightarrow$ 3) linkage. *Carbohydr. Res.* **1984**, *125*, 185–201. [\[CrossRef\]](#)
13. Paulsen, H.; Reck, F.; Brockhausen, I. Synthese von modifizierten Oligosacchariden der N-Glycoproteine als Substrate für N-Acetylglucosaminyltransferase I [Synthesis of modified oligosaccharides of the N-glycoprotein as substrate for N-acetylglucosaminyltransferase I]. *Carbohydr. Res.* **1992**, *236*, 39–71. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Brockhausen, I.; Möller, G.; Yang, J.M.; Khan, S.H.; Matta, K.L.; Paulsen, H.; Grey, A.A.; Shah, R.N.; Schachter, H. Control of glycoprotein synthesis. Characterization of (1 $\rightarrow$ 4)-N-acetyl-beta-D-glucosaminyltransferases acting on the alpha-D-(1 $\rightarrow$ 3)- and alpha-D-(1 $\rightarrow$ 6)-linked arms of N-linked oligosaccharides. *Carbohydr. Res.* **1992**, *236*, 281–299. [\[CrossRef\]](#)
15. Bai, L.; Li, H. Protein N-glycosylation and O-mannosylation are catalyzed by two evolutionarily related GT-C glycosyltransferases. *Curr. Opin. Struct. Biol.* **2021**, *68*, 66–73. [\[CrossRef\]](#)
16. Breton, C.; Fournel-Gigleux, S.; Palcic, M.M. Recent structures, evolution and mechanisms of glycosyltransferases. *Curr. Opin. Struct. Biol.* **2012**, *22*, 540–549. [\[CrossRef\]](#)
17. Möller, G.; Reck, F.; Paulsen, H.; Kaur, K.J.; Sarkar, M.; Schachter, H.; Brockhausen, I. Control of glycoprotein synthesis: Substrate specificity of rat liver UDP-GlcNAc:Man alpha 3R beta 2-N-acetylglucosaminyltransferase I using synthetic substrate analogues. *Glycoconj. J.* **1992**, *9*, 180–190. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Brockhausen, I.; Benn, M.; Bhat, S.; Marone, S.; Riley, J.G.; Montoya-Peleaz, P.; Vlahakis, J.Z.; Paulsen, H.; Schutzbach, J.S.; Szarek, W.A. UDP-Gal: GlcNAc-R beta1,4-galactosyltransferase—a target enzyme for drug design. Acceptor specificity and inhibition of the enzyme. *Glycoconj. J.* **2006**, *23*, 525–541. [\[CrossRef\]](#)
19. Ramasamy, V.; Ramakrishnan, B.; Boeggeman, E.; Ratner, D.M.; Seeberger, P.H.; Qasba, P.K. Oligosaccharide preferences of beta1,4-galactosyltransferase-I: Crystal structures of Met340His mutant of human beta1,4-galactosyltransferase-I with a pentasaccharide and trisaccharides of the N-glycan moiety. *J. Mol. Biol.* **2005**, *353*, 53–67. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Dohi, T.; Yuyama, Y.; Natori, Y.; Smith, P.L.; Lowe, J.B.; Oshima, M. Detection of N-acetylgalactosaminyltransferase mRNA which determines expression of Sda blood group carbohydrate structure in human gastrointestinal mucosa and cancer. *Int. J. Cancer* **1996**, *67*, 626–631. [\[CrossRef\]](#)
21. Cull, J.; Pink, R.C.; Samuel, P.; Brooks, S.A. Myriad mechanisms: Factors regulating the synthesis of aberrant mucin-type O-glycosylation found on cancer cells. *Glycobiology* **2025**, *35*, cwaf023. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Munkley, J. Aberrant Sialylation in Cancer: Therapeutic Opportunities. *Cancers* **2022**, *14*, 4248. [\[CrossRef\]](#)
23. Brockhausen, I.; Argueso, P. Mucin-type O-glycans: Biosynthesis and functions. In *Comprehensive Glycoscience*, 2nd ed.; Barchi, J., Ed.; Elsevier: Amsterdam, The Netherlands, 2021; pp. 233–252. [\[CrossRef\]](#)
24. Venturi, G.; Gomes Ferreira, I.; Pucci, M.; Ferracin, M.; Malagolini, N.; Chiricolo, M.; Dall’Olio, F. Impact of sialyltransferase ST6GAL1 overexpression on different colon cancer cell types. *Glycobiology* **2019**, *29*, 684–695. [\[CrossRef\]](#)
25. Lin, W.R.; Yeh, C.T. GALNT14: An Emerging Marker Capable of Predicting Therapeutic Outcomes in Multiple Cancers. *Int. J. Mol. Sci.* **2020**, *21*, 1491. [\[CrossRef\]](#)
26. Sun, H.; Chang, J.; Ye, M.; Weng, W.; Zhang, M.; Ni, S.; Tan, C.; Huang, D.; Wang, L.; Du, X.; et al. GCNT4 is Associated with Prognosis and Suppress Cell Proliferation in Gastric Cancer. *Oncol. Targets Ther.* **2020**, *13*, 8601–8613. [\[CrossRef\]](#)
27. Burchell, J.; Poulson, R.; Hanby, A.; Whitehouse, C.; Cooper, L.; Clausen, H.; Miles, D.; Taylor-Papadimitriou, J. An alpha2,3 sialyltransferase (ST3Gal I) is elevated in primary breast carcinomas. *Glycobiology* **1999**, *9*, 1307–1311. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Marcos, N.T.; Pinho, S.; Grandela, C.; Cruz, A.; Samyn-Petit, B.; Harduin-Lepers, A.; Almeida, R.; Silva, F.; Morais, V.; Costa, J.; et al. Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen. *Cancer Res.* **2004**, *6*, 7050–7057. [\[CrossRef\]](#)
29. González-Vallinas, M.; Vargas, T.; Moreno-Rubio, J.; Molina, S.; Herranz, J.; Cejas, P.; Burgos, E.; Aguayo, C.; Custodio, A.; Reglero, G.; et al. Clinical relevance of the differential expression of the glycosyltransferase gene GCNT3 in colon cancer. *Eur. J. Cancer* **2015**, *51*, 1–8. [\[CrossRef\]](#)
30. Kurosaka, A.; Nakajima, H.; Funakoshi, I.; Matsuyama, M.; Nagayo, T.; Yamashina, I. Structures of the major oligosaccharides from a human rectal adenocarcinoma glycoprotein. *J. Biol. Chem.* **1983**, *258*, 11594–11598. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Gao, Y.; Chachadi, V.B.; Cheng, P.W.; Brockhausen, I. Glycosylation potential of human prostate cancer cell lines. *Glycoconj. J.* **2012**, *29*, 525–537. [\[CrossRef\]](#) [\[PubMed\]](#)

32. Schachter, H. Complex N-glycans: The story of the “yellow brick road”. *Glycoconj. J.* **2014**, *31*, 1–5. [\[CrossRef\]](#)
33. Unligil, U.M.; Zhou, S.; Yuwaraj, S.; Sarkar, M.; Schachter, H.; Rini, J.M. X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: Catalytic mechanism and a new protein superfamily. *EMBO J.* **2000**, *19*, 5269–5280. [\[CrossRef\]](#)
34. Sharma, C.B.; Lehle, L.; Tanner, W. N-Glycosylation of yeast proteins. Characterization of the solubilized oligosaccharyl transferase. *Eur. J. Biochem.* **1981**, *116*, 101–108. [\[CrossRef\]](#)
35. Yoo, J.; Mashalidis, E.H.; Kuk, A.C.Y.; Yamamoto, K.; Kaeser, B.; Ichikawa, S.; Lee, S.Y. GlcNAc-1-P-transferase-tunicamycin complex structure reveals basis for inhibition of N-glycosylation. *Nature Struct. Mol. Biol.* **2018**, *25*, 217–224. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Timal, S.; Hoischen, A.; Lehle, L.; Adamowicz, M.; Huijben, K.; Sykut-Cegielska, J.; Paprocka, J.; Jamroz, E.; van Spronsen, F.J.; Körner, C.; et al. Gene identification in the congenital disorders of glycosylation type I by whole-exome sequencing. *Human Mol. Gen.* **2012**, *21*, 4151–4161. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Jensen, J.W.; Schutzbach, J.S. Activation of dolichyl-phospho-mannose synthase by phospholipids. *Eur. J. Biochem.* **1985**, *153*, 41–48. [\[CrossRef\]](#)
38. Reznik, N.; Fass, D. Disulfide bond formation and redox regulation in the Golgi apparatus. *FEBS Lett.* **2022**, *596*, 2859–2872. [\[CrossRef\]](#)
39. Stanley, P.; Narasimhan, S.; Siminovitch, L.; Schachter, H. Chinese hamster ovary cells selected for resistance to the cytotoxicity of phytohemagglutinin are deficient in a UDP-N-acetylglucosamine--glycoprotein N-acetylglucosaminyltransferase activity. *Proc. Nat. Acad. Sci. USA* **1975**, *72*, 3323–3327. [\[CrossRef\]](#)
40. Paulsen, H.; Heume, M.; Nürnberger, H. Synthese der verzweigten Nonasaccharid-Sequenz der “bisected” Struktur von N-Glycoproteinen [Synthesis of a highly branched nonasaccharide sequence of the “bisected” structure of N-glycoproteins]. *Carbohydr. Res.* **1990**, *200*, 127–166. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Paulsen, H.; Peters, T.; Sinnwell, V.; Meyer, B. Konformationsanalyse modifizierter Tetrasaccharidsequenzen vom Typ der N-Glycoproteine--zum Problem der alpha-(1----6)-glycosidischen Bindung [Conformation analysis of modified tetrasaccharide sequences of the N-glycoprotein type--problem of the alpha-(1 to 6)-glycosidic bond]. *Carbohydr. Res.* **1987**, *165*, 251–266. [\[CrossRef\]](#)
42. Reck, F.; Springer, M.; Meinjohanns, E.; Paulsen, H.; Brockhausen, I.; Schachter, H. Synthetic substrate analogues for UDP-GlcNAc: Man alpha 1-3R beta 1-2-N-acetylglucosaminyltransferase I. Substrate specificity and inhibitors for the enzyme. *Glycoconj. J.* **1995**, *12*, 747–754. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Reck, F.; Springer, M.; Paulsen, H.; Brockhausen, I.; Sarkar, M.; Schachter, H. Synthesis of tetrasaccharide analogues of the N-glycan substrate of beta-(1→2)-N-acetylglucosaminyltransferase II using trisaccharide precursors and recombinant beta-(1→2)-N-acetylglucosaminyltransferase I. *Carbohydr. Res.* **1994**, *259*, 93–101. [\[CrossRef\]](#)
44. Reck, F.; Meinjohanns, E.; Springer, M.; Wilkens, R.; Van Dorst, J.A.; Paulsen, H.; Möller, G.; Brockhausen, I.; Schachter, H. Synthetic substrate analogues for UDP-GlcNAc: Man alpha 1-6R beta(1-2)-N-acetylglucosaminyltransferase II. Substrate specificity and inhibitors for the enzyme. *Glycoconj. J.* **1994**, *11*, 210–216. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Paulsen, H.; Springer, M.; Reck, F.; Brockhausen, I.; Schachter, H. Synthese von modifizierten Tetrasacchariden als analoge Akzeptor-Inhibitoren der N-Acetylglucosaminyltransferase II [Synthesis of modified tetrasaccharides as analog acceptor-inhibitors of N-acetylglucosaminyltransferase II]. *Carbohydr. Res.* **1995**, *275*, 403–411. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Kadirvelraj, R.; Yang, J.Y.; Sanders, J.H.; Liu, L.; Ramiah, A.; Prabhakar, P.K.; Boons, G.J.; Wood, Z.A.; Moremen, K.W. Human N-acetylglucosaminyltransferase II substrate recognition uses a modular architecture that includes a convergent exosite. *Proc. Nat. Acad. Sci. USA* **2018**, *115*, 4637–4642. [\[CrossRef\]](#)
47. de-Souza-Ferreira, M.; Ferreira, É.E.; de-Freitas-Junior, J.C.M. Aberrant N-glycosylation in cancer: MGAT5 and β1,6-GlcNAc branched N-glycans as critical regulators of tumor development and progression. *Cell. Oncol.* **2023**, *46*, 481–501. [\[CrossRef\]](#)
48. Brockhausen, I.; Reck, F.; Kuhns, W.; Khan, S.; Matta, K.L.; Meinjohanns, E.; Paulsen, H.; Shah, R.N.; Baker, M.A.; Schachter, H. Substrate specificity and inhibition of UDP-GlcNAc:GlcNAc beta 1-2Man alpha 1-6R beta 1,6-N-acetylglucosaminyltransferase V using synthetic substrate analogues. *Glycoconj. J.* **1995**, *12*, 371–379. [\[CrossRef\]](#)
49. Khan, S.H.; Crawley, S.C.; Kanie, O.; Hindsgaul, O. A trisaccharide acceptor analog for N-acetylglucosaminyltransferase V which binds to the enzyme but sterically precludes the transfer reaction. *J. Biol. Chem.* **1993**, *268*, 2468–2473. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Kanie, O.; Crawley, S.C.; Palcic, M.M.; Hindsgaul, O. Acceptor-substrate recognition by N-acetylglucosaminyltransferase-V: Critical role of the 4"-hydroxyl group in beta-D-GlcpNAc-(1→2)-alpha-D-Manp(1→6)-beta-D-Glcp-OR. *Carbohydr. Res.* **1993**, *243*, 139–164. [\[CrossRef\]](#)
51. Kanie, O.; Crawley, S.C.; Palcic, M.M.; Hindsgaul, O. Key involvement of all three GlcNAc hydroxyl groups in the recognition of beta-D-GlcpNAc-(1→2)-alpha-D-Manp-(1→6)-beta-D-Glcp-OR by N-acetylglucosaminyltransferase-V. *Bioorg. Med. Chem.* **1994**, *2*, 1231–1241. [\[CrossRef\]](#)
52. Nagae, M.; Kizuka, Y.; Mihara, E.; Kitago, Y.; Hanashima, S.; Ito, Y.; Takagi, J.; Taniguchi, N.; Yamaguchi, Y. Structure and mechanism of cancer-associated N-acetylglucosaminyltransferase-V. *Nat. Commun.* **2018**, *9*, 3380. [\[CrossRef\]](#)

53. Osuka, R.F.; Nagae, M.; Ohuchi, A.; Ohno, S.; Yamaguchi, Y.; Kizuka, Y. The cancer-associated glycosyltransferase GnT-V (MGAT5) recognizes the N-glycan core via residues outside its catalytic pocket. *FEBS Lett.* **2023**, *597*, 3102–3113. [[CrossRef](#)]
54. Osuka, R.F.; Yamasaki, T.; Kizuka, Y. Structure and function of N-acetylglucosaminyltransferase V (GnT-V). *Biochim. Biophys. Acta Gen. Subj.* **2024**, *1868*, 130709. [[CrossRef](#)] [[PubMed](#)]
55. Paulsen, H.; Peters, T.; Sinnwell, V.; Heume, M.; Meyer, B. Konformationsanalyse der verzweigten Pentasaccharid-sequenz der “bisected” Struktur von N-Glycoproteinen [Conformational analysis of the double pentasaccharide sequence of the “bisected” structure of N-glycoproteins]. *Carbohydr. Res.* **1986**, *156*, 87–106. [[CrossRef](#)]
56. Brockhausen, I.; Hull, E.; Hindsgaul, O.; Schachter, H.; Shah, R.N.; Michnick, S.W.; Carver, J.P. Control of glycoprotein synthesis. Detection and characterization of a novel branching enzyme from hen oviduct, UDP-N-acetylglucosamine:GlcNAc beta 1-6 (GlcNAc beta 1-2)Man alpha-R (GlcNAc to Man) beta-4-N-acetylglucosaminyltransferase VI. *J. Biol. Chem.* **1989**, *264*, 11211–11221. [[CrossRef](#)] [[PubMed](#)]
57. Sakamoto, Y.; Taguchi, T.; Honke, K.; Korekane, H.; Watanabe, H.; Tano, Y.; Dohmae, N.; Takio, K.; Horii, A.; Taniguchi, N. Molecular cloning and expression of cDNA encoding chicken UDP-N-acetyl-D-glucosamine (GlcNAc): GlcNAc beta 1-6 (GlcNAc beta 1-2)- man alpha 1-R [GlcNAc to man] beta 1,4-N-acetylglucosaminyltransferase VI. *J. Biol. Chem.* **2000**, *275*, 36029–36034. [[CrossRef](#)]
58. Dalle Vedove, E.; Costabile, G.; Merkel, O.M. Mannose and Mannose-6-Phosphate Receptor-Targeted Drug Delivery Systems and Their Application in Cancer Therapy. *Adv. Healthc. Mater.* **2018**, *7*, e1701398. [[CrossRef](#)]
59. Brockhausen, I.; Lehotay, M.; Yang, J.M.; Qin, W.; Young, D.; Lucien, J.; Coles, J.; Paulsen, H. Glycoprotein biosynthesis in porcine aortic endothelial cells and changes in the apoptotic cell population. *Glycobiology* **2002**, *12*, 33–45. [[CrossRef](#)]
60. Kuhn, B.; Benz, J.; Greif, M.; Engel, A.M.; Sobek, H.; Rudolph, M.G. The structure of human  $\alpha$ -2,6-sialyltransferase reveals the binding mode of complex glycans. *Acta Crystal. Sect. D Biol. Crystallogr.* **2013**, *69 Pt 9*, 1826–1838. [[CrossRef](#)] [[PubMed](#)]
61. Qasba, P.K.; Ramakrishnan, B.; Boeggeman, E. Structure and function of beta -1,4-galactosyltransferase. *Curr. Drug Targets* **2008**, *9*, 292–309. [[CrossRef](#)]
62. Inaba, R.; Vujakovic, S.; Bergstrom, K. The gut mucus network: A dynamic liaison between microbes and the immune system. *Sem. Immunol.* **2023**, *69*, 101807. [[CrossRef](#)] [[PubMed](#)]
63. Luis, A.S.; Hansson, G.C. Intestinal mucus and their glycans: A habitat for thriving microbiota. *Cell Host Microbe* **2023**, *31*, 1087–1100. [[CrossRef](#)]
64. Wagner, K.W.; Punnoose, E.A.; Januario, T.; Lawrence, D.A.; Pitti, R.M.; Lancaster, K.; Lee, D.; von Goetz, M.; Yee, S.F.; Totpal, K.; et al. Death-receptor O-glycosylation controls tumor-cell sensitivity to the proapoptotic ligand Apo2L/TRAIL. *Nature Med.* **2007**, *13*, 1070–1077. [[CrossRef](#)]
65. Hounsell, E.F.; Lawson, A.M.; Feeney, J.; Gooi, H.C.; Pickering, N.J.; Stoll, M.S.; Lui, S.C.; Feizi, T. Structural analysis of the O-glycosidically linked core-region oligosaccharides of human meconium glycoproteins which express oncofoetal antigens. *Eur. J. Biochem.* **1985**, *148*, 367–377. [[CrossRef](#)] [[PubMed](#)]
66. Pollex-Krüger, A.; Meyer, B.; Stuike-Prill, R.; Sinnwell, V.; Matta, K.L.; Brockhausen, I. Preferred conformations and dynamics of five core structures of mucin type O-glycans determined by NMR spectroscopy and force field calculations. *Glycoconj. J.* **1993**, *10*, 365–380. [[CrossRef](#)]
67. Müller, S.; Alving, K.; Peter-Katalinic, J.; Zachara, N.; Gooley, A.A.; Hanisch, F.G. High density O-glycosylation on tandem repeat peptide from secretory MUC1 of T47D breast cancer cells. *J. Biol. Chem.* **1999**, *274*, 18165–18172. [[CrossRef](#)] [[PubMed](#)]
68. Bennett, E.P.; Hassan, H.; Hollingsworth, M.A.; Clausen, H. A novel human UDP-N-acetyl-D-galactosamine: polypeptide N-acetylglucosaminyltransferase, GalNAc-T7, with specificity for partial GalNAc-glycosylated acceptor substrates. *FEBS Lett.* **1999**, *460*, 226–230. [[CrossRef](#)]
69. Kato, K.; Hansen, L.; Clausen, H. Polypeptide N-acetylglucosaminyltransferase-Associated Phenotypes in Mammals. *Molecules* **2021**, *26*, 5504. [[CrossRef](#)]
70. Ju, T.; Cummings, R.D. Protein glycosylation: Chaperone mutation in Tn syndrome. *Nature* **2005**, *437*, 1252. [[CrossRef](#)]
71. Patsos, G.; Hebbe-Viton, V.; Robbe-Masselot, C. O-glycan inhibitors generate aryl-glycans, induce apoptosis and lead to growth inhibition in colorectal cancer cell lines. *Glycobiology* **2009**, *19*, 382–398. [[CrossRef](#)] [[PubMed](#)]
72. Brockhausen, I.; Möller, G.; Pollex-Krüger, A.; Rutz, V.; Paulsen, H.; Matta, K.L. Control of O-glycan synthesis: Specificity and inhibition of O-glycan core 1 UDP-galactose:N-acetylglucosamine-alpha-R beta 3-galactosyltransferase from rat liver. *Biochem. Cell Biol.* **1992**, *70*, 99–108. [[CrossRef](#)] [[PubMed](#)]
73. Yeh, J.C.; Ong, E.; Fukuda, M. Molecular cloning and expression of a novel beta-1,6-N-acetylglucosaminyltransferase that forms core 2, core 4, and I branches. *J. Biol. Chem.* **1999**, *274*, 3215–3221. [[CrossRef](#)] [[PubMed](#)]
74. Kuhns, W.; Rutz, V.; Paulsen, H.; Matta, K.L.; Baker, M.A.; Barner, M.; Granovsky, M.; Brockhausen, I. Processing O-glycan core 1, Gal beta 1-3GalNAc alpha-R. Specificities of core 2, UDP-GlcNAc: Gal beta 1-3 GalNAc-R (GlcNAc to GalNAc) beta 6-N-acetylglucosaminyltransferase and CMP-sialic acid: Gal beta 1-3GalNAc-R alpha 3-sialyltransferase. *Glycoconj. J.* **1993**, *10*, 381–394. [[CrossRef](#)]

75. Mardahl, M.; Schröter, M.F.; Engelbert, D.; Pink, M.; Sperandio, M.; Hamann, A.; Syrbe, U. Core 2  $\beta$ 1,6-N-acetylglucosaminyltransferase-I, crucial for P-selectin ligand expression is controlled by a distal enhancer regulated by STAT4 and T-bet in CD4<sup>+</sup> T helper cells 1. *Mol. Immun.* **2016**, *77*, 132–140. [\[CrossRef\]](#)
76. Schwientek, T.; Yeh, J.C.; Levery, S.B.; Keck, B.; Merckx, G.; van Kessel, A.G.; Fukuda, M.; Clausen, H. Control of O-glycan branch formation. Molecular cloning and characterization of a novel thymus-associated core 2  $\beta$ 1, 6-n-acetylglucosaminyltransferase. *J. Biol. Chem.* **2000**, *275*, 11106–11113. [\[CrossRef\]](#)
77. Vavasseur, F.; Yang, J.M.; Dole, K.; Paulsen, H.; Brockhausen, I. Synthesis of O-glycan core 3: Characterization of UDP-GlcNAc:GalNAc-R  $\beta$ 3-N-acetyl-glucosaminyltransferase activity from colonic mucosal tissues and lack of the activity in human cancer cell lines. *Glycobiology* **1995**, *5*, 351–357. [\[CrossRef\]](#)
78. Huang, M.C.; Chen, H.Y.; Huang, H.C.; Huang, J.; Liang, J.-T.; Shen, T.-L.; Lin, N.-Y.; Ho, C.-C.; Hsu, S.-M. C2GnT-M is downregulated in colorectal cancer and its re-expression causes growth inhibition of colon cancer cells. *Oncogene* **2006**, *25*, 3267–3276. [\[CrossRef\]](#)
79. Pak, J.E.; Arnoux, P.; Zhou, S.; Sivarajah, P.; Satkunarajah, M.; Xing, X.; Rini, J.M. X-ray crystal structure of leukocyte type core 2  $\beta$ 1,6-N-acetylglucosaminyltransferase. Evidence for a convergence of metal ion-independent glycosyltransferase mechanism. *J. Biol. Chem.* **2006**, *281*, 26693–26701. [\[CrossRef\]](#)
80. Pak, J.E.; Satkunarajah, M.; Seetharaman, J.; Rini, J.M. Structural and mechanistic characterization of leukocyte-type core 2  $\beta$ 1,6-N-acetylglucosaminyltransferase: A metal-ion-independent GT-A glycosyltransferase. *J. Mol. Biol.* **2011**, *414*, 798–811. [\[CrossRef\]](#)
81. Kurosaka, A.; Funakoshi, I.; Matsuyama, M.; Nagayo, T.; Yamashina, I. UDP-GalNAc:GalNAc-mucin  $\alpha$ -N-acetylgalactosamine transferase activity in human intestinal cancerous tissues. *FEBS Lett.* **1985**, *190*, 259–262. [\[CrossRef\]](#) [\[PubMed\]](#)
82. Yeh, J.C.; Hiraoka, N.; Petryniak, B.; Nakayama, J.; Ellies, L.G.; Rabuka, D.; Hindsgaul, O.; Marth, J.D.; Lowe, J.B.; Fukuda, M. Novel sulfated lymphocyte homing receptors and their control by a Core1 extension  $\beta$ 1,3-N-acetylglucosaminyltransferase. *Cell* **2001**, *105*, 957–969. [\[CrossRef\]](#)
83. Nakayama, J.; Yeh, J.C.; Misra, A.K.; Ito, S.; Katsuyama, T.; Fukuda, M. Expression cloning of a human  $\alpha$ 1, 4-N-acetylglucosaminyltransferase that forms GlcNAc $\alpha$ 1 $\rightarrow$ 4Gal $\beta$  $\rightarrow$ R, a glycan specifically expressed in the gastric gland mucous cell-type mucin. *Proc. Nat. Acad. Sci. USA* **1999**, *96*, 8991–8996. [\[CrossRef\]](#) [\[PubMed\]](#)
84. Yuki, A.; Fujii, C.; Yamanoi, K.; Matoba, H.; Harumiya, S.; Kawakubo, M.; Nakayama, J. Glycosylation of MUC6 by  $\alpha$ 1,4-linked N-acetylglucosamine enhances suppression of pancreatic cancer malignancy. *Cancer Sci.* **2022**, *113*, 576–586. [\[CrossRef\]](#)
85. Datta, A.K. Comparative sequence analysis in the sialyltransferase protein family: Analysis of motifs. *Curr. Drug Targets* **2009**, *10*, 483–498. [\[CrossRef\]](#) [\[PubMed\]](#)
86. Moremen, K.W.; Ramiah, A.; Stuart, M.; Steel, J.; Meng, L.; Forouhar, F.; Moniz, H.A.; Gahlay, G.; Gao, Z.; Chapla, D.; et al. Expression system for structural and functional studies of human glycosylation enzymes. *Nature Chem. Biol.* **2018**, *14*, 156–162. [\[CrossRef\]](#)
87. Grewal, R.K.; Shaikh, A.R.; Gorle, S.; Kaur, M.; Videira, P.A.; Cavallo, L.; Chawla, M. Structural Insights in Mammalian Sialyltransferases and Fucosyltransferases: We Have Come a Long Way, but It Is Still a Long Way Down. *Molecules* **2021**, *26*, 5203. [\[CrossRef\]](#) [\[PubMed\]](#)
88. Rao, F.V.; Rich, J.R.; Rakić, B.; Buddai, S.; Schwartz, M.F.; Johnson, K.; Bowe, C.; Wakarchuk, W.W.; Defrees, S.; Withers, S.G.; et al. Structural insight into mammalian sialyltransferases. *Nature Struct. Mol. Biol.* **2009**, *16*, 1186–1188. [\[CrossRef\]](#)
89. Meldal, M.; Bielfeldt, T.; Peters, S.; Jensen, K.J.; Paulsen, H.; Bock, K. Susceptibility of glycans to beta-elimination in Fmoc-based O-glycopeptide synthesis. *Int. J. Peptide Protein Res.* **1994**, *43*, 529–536. [\[CrossRef\]](#)
90. Paulsen, H.; Peters, S.; Bielfeldt, T.; Meldal, M.; Bock, K. Synthesis of the glycosyl amino acids N  $\alpha$ -Fmoc-Ser[Ac4- $\beta$ -D-Galp-(1 $\rightarrow$ 3)-Ac2- $\alpha$ -D-GalN3p]-OPfp and N  $\alpha$ -Fmoc-Thr[Ac4- $\beta$ -D-Galp-(1 $\rightarrow$ 3)-Ac2- $\alpha$ -D-GalN3p]-OPfp and the application in the solid-phase peptide synthesis of multiply glycosylated mucin peptides with Tn and T antigenic structures. *Carbohydr. Res.* **1995**, *268*, 17–34. [\[CrossRef\]](#)
91. Peters, S.; Bielfeldt, T.; Meldal, M.; Bock, K.; Paulsen, H. Multiple-column solid-phase glycopeptide synthesis. *J. Chem. Soc. Perkin Trans.* **1992**, *1*, 1163–1171. [\[CrossRef\]](#)
92. Brockhausen, I.; Möller, G.; Merz, G.; Adermann, K.; Paulsen, H. Control of mucin synthesis: The peptide portion of synthetic O-glycopeptide substrates influences the activity of O-glycan core 1 UDPgalactose:N-acetyl- $\alpha$ -galactosaminyl-R  $\beta$ 3-galactosyltransferase. *Biochemistry* **1990**, *29*, 10206–10212. [\[CrossRef\]](#) [\[PubMed\]](#)
93. Brockhausen, I.; Toki, D.; Brockhausen, J.; Peters, S.; Bielfeldt, T.; Kleen, A.; Paulsen, H.; Meldal, M.; Hagen, F.; Tabak, L.A. Specificity of O-glycosylation by bovine colostrum UDP-GalNAc: Polypeptide  $\alpha$ -N-acetylgalactosaminyltransferase using synthetic glycopeptide substrates. *Glycoconj. J.* **1996**, *13*, 849–856. [\[CrossRef\]](#) [\[PubMed\]](#)

94. Granovsky, M.; Bielfeldt, T.; Peters, S.; Paulsen, H.; Meldal, M.; Brockhausen, J.; Brockhausen, I. UDPgalactose:glycoprotein-N-acetyl-D-galactosamine 3-beta-D-galactosyltransferase activity synthesizing O-glycan core 1 is controlled by the amino acid sequence and glycosylation of glycopeptide substrates. *Eur. J. Biochem.* **1994**, *221*, 1039–1046. [[CrossRef](#)]
95. Brockhausen, I.; Dowler, T.; Paulsen, H. Site directed processing: Role of amino acid sequences and glycosylation of acceptor glycopeptides in the assembly of extended mucin type O-glycan core 2. *Biochim. Biophys. Acta* **2009**, *1790*, 1244–1257. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.