Unraveling mucin type o-glycosylation signatures of colorectal cancer

Madunić. K.

INTRODUCTION

Chapter 1

This introduction will cover colorectal cancer (CRC) pathogenesis and current treatment strategies for targeted pharmacotherapy of CRC. Moreover, it will introduce glycosylation and give a detailed overview of the biosynthesis of mucin type *O*-glycosylation, as well as changes in the expression observed in cancer. Finally, an overview will be provided regarding the methodologies being used for the analysis of mucin type *O*-glycans, with a specific focus on separation techniques as well as mass spectrometry as a detection technique.

COLORECTAL CANCER

PATHOGENESIS OF COLORECTAL CANCER

CRC is one of the most common types of cancers with nearly two million new cases in 2020 worldwide[1]. In addition, CRC is the leading cause of cancer related deaths[1], as patients are often asymptomatic until the late stages of the disease. CRC develops from precancerous adenoma, by the accumulation of somatic and germline mutations over a longer period of time. Among others, mechanisms such as chromosomal instability (CIN), microsatellite instability (MSI), and epigenetic instability also known as CpG island methylator phenotype (CIMP) have been linked to CRC pathogenesis[2]. According to The Cancer Genome Atlas (TCGA), which contains genomic data from a large sample set of CRC patients, a differentiation between hypermutated and non-hypermutated CRC can be made[3]. The hypermutated CRC subtype (15% of CRC patients) is characterized as MSI, with immune cell infiltration and better prognosis in the early stages. The prevalent non-hypermutated subtype is characterized by CIN, worse prognosis, and resistance to immunotherapy. Despite insights into molecular pathogenesis, this stratification provides limited clinical benefit for outcome prediction and therapeutic intervention. To provide a more accurate patient prognosis, stratification based upon the four consensus molecular subtypes (CMS) revealed promising results[4]. For this purpose, multiple molecular markers were combined which were derived from transcriptomic profiling of a large sample set of CRC patients (**Figure 1**). CMS 1 tumors were predominantly classified as MSI and showed high immune infiltration in the tumor microenvironment. This was associated with a better patient prognosis, however worse survival after relapse. Both CMS 2 and CMS 3 tumors showed strong epithelial differentiation signatures. CMS 2 is characterized by

CMS1 MSI immune	CMS2 Canonical	CMS3 Metabolic	CMS4 Mesenchymal
14%	37%	13%	23%
MSI, CIMP high, hypermutation	SCNA high	Mixed MSI status, SCNA low, CIMP low	SCNA high
BRAF mutations		*KRAS* mutations	
Immune infiltration and activation	WNT and MYC activation	Metabolic deregulation	Stromal infiltration, TGF-β activation, angiogenesis
Worse survival after relapse			Worse relapse-free and overall survival

Figure 1. Proposed classification of CRC reflecting the gene expression-based molecular subtypes. *Figure adapted from Guinney et al.*[4]

WNT and MYC activation, whereas the CMS 3 group revealed characteristic metabolic pathway dysregulation[4]. In contrast, the mesenchymal CMS 4 tumors were found to be characterized by stromal infiltration, TGF-β activation and upregulation of epithelial to mesenchymal transition (EMT) resulting in worse overall patient prognosis[4].

TARGETED THERAPY FOR CRC

Conventional therapies for the treatment of CRC include surgery, radiotherapy and chemotherapy and are associated with various side effects[5]. To increase the efficiency and to decrease the toxicity of CRC treatment, targeted therapies provide new perspectives as it intervenes with the function of specific molecules present on cancer cells or cells in the tumor microenvironment [6]. For example, promising results were achieved using antibodies for immuno-targeted therapies in cancer treatment[7]. Several targets have been approved for the antibody treatment of CRC, namely, the anti-VEGF (bevacizumab), anti-EGFR (cetuximab and panitumumab), anti-PD-1 immune checkpoint (pembrolizumab, nivolumab), and BRAF V600E inhibitors (several combinations of vemurafenib, irinotecan, and cetuximab or panitumumab)[6]. Many more are still undergoing clinical trials[6]. Due to the tumor heterogeneity and relative paucity of the mentioned protein targets

present on cancer cells, only a small part of patients clearly benefit from these treatments[6]. Therefore, better patient stratification and identification of new targets is a crucial aspect for the development of more effective therapies. Since aberrant glycosylation is one of the hallmarks of cancer[8], and cancer associated glycans are present in high abundance on the surface of cancer cells which are linked to many different proteins, it has an immense potential to be used as therapy target[9].

GLYCOSYLATION

The process where sugar molecules are covalently linked to a protein or lipid carrier by specific enzymes is called glycosylation and plays an important role in modulating their function. In fact, the surface of eukaryotic cells is composed of a very dense layer of oligosaccharides, linked to their carriers, which play an important role in cell-cell and cell-extracellular matrix interactions, leading to changes in immune response, differentiation and cell growth[10]. This layer, also known as the glycocalyx, will be extensively studied in this thesis in the context of CRC. Oligosaccharides can be present on various carriers and based upon the glycoconjugate a distinction is made between the different classes such as glycoproteins, glycosaminoglycans (GAGs), glycosylphosphatidylinositol (GPI) anchors, proteoglycans and glycosphingolipids (GSLs)[10]. In regard to glycoproteins, glycans can be bound to the protein via a nitrogen atom to the amino acid asparagine (Asn), with a specific sequence motif of three amino acids (Asn-X-Ser/Thr in which “X” is any amino acid except proline (Pro)) and are known as *N*-linked glycans whereas, in the case where the glycan is attached via an oxygen atom to a serine (Ser) or threonine (Thr) it is referred to as *O*-linked glycosylation. In eukaryotic organisms, the glycans are composed of a limited set of monosaccharide building blocks, which for glycoproteins consist of hexoses (galactose (Gal), mannoses (Man) and glucoses (Glc)), *N*-acetylhexosamines (*N*-acetylgalactosamine (GalNAc), *N*-acetylglucosamine (GlcNAc)), deoxyhexose (fucose (Fuc)), pentose (xylose (Xyl)) and sialic acid (*N*-acetylneuraminic acid (Neu5Ac), and *N*-glycolylneuraminic acid (Neu5Gc)). Next to the various building blocks, the glycans become even more diverse as the monosaccharides can be differently linked to each other, namely, linkages can be formed on different hydroxyl groups in the sugar ring as well as in two anomeric configurations (α and

β). The glycan repertoire is further complicated by modifications such as the addition of acetyl groups, phosphate or sulfate.

O-LINKED GLYCOSYLATION

As stated before, *O*-linked glycosylation occurs on specific amino acids such as Ser and Thr, and in rare cases on tyrosine (Tyr)[11]. However, compared to *N*-glycosylation, no consensus sequence motif has been defined. The biosynthesis of *O*-GalNAc linked glycans is a process that takes place in the Golgi apparatus. The enzymes involved are type II transmembrane proteins, with catalytic domain in the lumen of the Golgi. The activated monosaccharide donors are transported from the cytosol. The oligosaccharide chains can be initiated by a GalNAc, GlcNAc, Fuc, Xyl or Man residues. Although it can also be present on other types of glycoproteins, the main glycosylation type of mucins is the *O*-GalNAc glycosylation, where GalNAc is the first monosaccharide attached to the Ser or Thr[11]. This glycosylation type occurs predominantly in central tandem repeat domains which are rich in Pro, Thr and Ser (PTS). Aberrant mucin expression and mucin type glycosylation have been associated with disease pathologies, such as cancer[11], therefore mucin type *O*-linked glycosylation will be the main focus of this thesis, and it will be described in further detail.

BIOSYNTHETIC PATHWAY OF MUCIN TYPE *O*-GLYCANS

BIOSYNTHESIS OF MAIN MUCIN TYPE GLYCAN CORES

The initial step of adding a GalNAc in α-linkage to Ser or Thr is catalyzed by polypeptide-*N*-acetylgalactosaminyltransferases (GALNT1-20), a family of 20 different enzymes[12] (**Figure 2**). This step results in the biosynthesis of the Tn antigen. When Neu5Ac is added in the α2-6 position of the core GalNAc, sialyl-Tn antigen is formed. Addition of β1-3-linked galactose by the action of *C1GALT1* forms the T antigen, a core 1 glycan, that can be further extended. The activity of T synthase is dependent on its molecular chaperone COSMC, and high expression of Tn and sialyl-Tn antigen in cancer are often associated with defects in the activity of T synthase or COSMC[13–15]. Next to the addition of a Neu5Ac, the Tn-antigen can also be extended differently, forming four well-known types of *O*-GalNAc glycan cores (1-4) **(Figure 2)**, as well as four less common ones (cores 5-8).

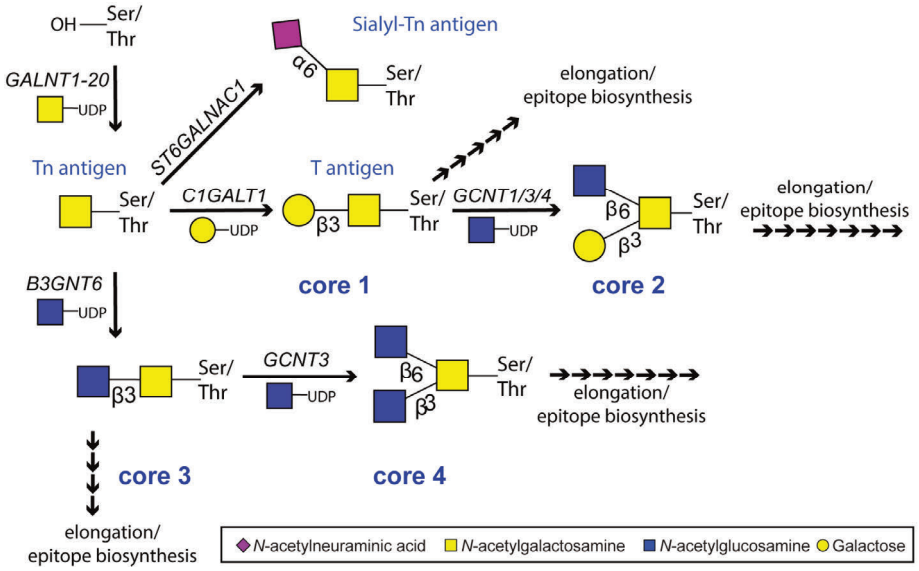

Figure 2. Biosynthesis of the most common mucin type glycan cores.

When the GlcNAc is added in β1-6-linkage to the GalNAc residue of the T antigen, a core 2 glycan is formed. This step is catalyzed by β1-6-*N*-acetylglucosaminyltransferases 1, 3 and 4 (GCNT1, GCNT3 and GCNT4)[14]. GCNT1 is the leukocyte type enzyme (L-type) which synthesizes only core 2 glycans, whereas the mucin(M)-type enzyme (GCNT3) can also synthesize core 4 glycans by adding the β1-6-GlcNAc to a core 3 glycan precursor[16]. The M-type enzyme is found predominately in high mucus secreting organ systems such as the gastrointestinal and respiratory tract. Core 2 and core 4 glycans can be further elongated on both branches by the activity of different glycosyltransferases (GSTs). Addition of β1-3-linked GlcNAc to the Tn structure by the action of β1-3-*N*-acetylglucosaminyltransferase (*B3GNT6*) forms core 3 glycans, that can be further extended into a linear glycan, or transformed into a branched core 4 glycan by the action of M-type β1-6 *N*-acetylglucosaminyltransferase (*GCNT3*)[10]. The expression of core 3 and core 4 glycans is limited to epithelial mucosa of the gastrointestinal and respiratory tract as well as salivary glands, and their expression is downregulated in cancer[17–19] due to downregulation of core 3 synthase (*B3GNT6)*[20]. Apart from the initiation of GalNAc type glycosylation, most GSTs involved in the core formation are pathway specific and do not have close paralogues, which allows to predict cell glycosylation based on the expression of

those enzymes[11]. Core structures are further extended by elongation with Type 1 and 2 chains, and/or capped with terminal motifs such as Lewis type and blood type antigens or by sialylation.

The four less common core types (cores 5-8) are all linear glycan types. Of which cores 5 and 7 contain a GalNAc linked to the core GalNAc in an α1-3- and α1-6-linkage, respectively. Core 6 glycans have a β1-6-linked GlcNAc to the core GalNAc whereas core 8 glycans contain a Gal which is α1-3-linked to the GalNAc.

CORE EXTENSION, ELONGATION OR BRANCHING

Extension of the pre-existing core GlcNAc in branched glycans can form two different types of chains; Type 1 is defined by the addition of the Gal by β1-3 GSTs (*B3GAL1/2/5*) and type 2 by β1-4 galactosyltransferases (*B4GALT1-4*) **(Figure 3)**. Additionally, a GlcNAc can be added to a terminal Gal to form a linear elongation by β1-3-*N*-acetylglucosaminyltransferases (*B3GNT2,3,4,7,8 and 9*) or otherwise a branching point by β1-6-*N*-acetylglucosaminyltransferase (*GCNT2, GCNT7* and *GCNT3*) forming the I-antigen[20,21]. Type 2 motifs form repeating units (polyLacNAc) or may be capped by terminal epitopes. Interestingly, the *GCNT2* is the main enzyme responsible for biosynthesis of the I-antigen on GSLs, whereas the M-type (*GCNT3)* plays a central role in the I-antigen next to core 2 and 4 biosynthesis of mucin type glycans in colon, small intestine, trachea and stomach[16]. A less common elongation is the formation of a LacdiNAc motif (GalNAcβ1-4GlcNAcβ1) by action of β1-4-GalNAc transferase (*B4GALNT3/4*) which can be further modified to a fucosylated LacDiNAc or terminated by sialylation[21,22]. GSTs involved in glycan elongation are considered non-pathway specific and, in combination with overlapping specificities, predictions of glycan expression become far more challenging. However, it is known that *B3GALT5* preferably acts on core 3 *O*-glycans, and *B3GNT3* prefers core 1 and 2 *O*-glycans as substrates[10].

BIOSYNTHESIS OF TERMINAL EPITOPES

GSTs responsible for the biosynthesis of terminal epitopes act on distinct types of glycans. The α1-2-fucosyltransferases *FUT1* and *FUT2* synthesize the blood group H antigens on *N*-glycans, *O*- glycans and GSLs by adding fucose to the terminal Type 1, 2, 3 and 4 linked galactose residues **(Figure 3).** Galactose is β1-3-linked to a GalNAc residue in Type 3 H antigen, either to reducing end GalNAc in mucin

type *O*-glycans, or to inner GalNAc residues in glycolipids. If those are globo-series glycolipids, it is referred to as Type 4[10]. The *H* allele encodes for the Type 2 and Type 4 specific *FUT1* transferase, forming blood group H in erythrocytes, whereas the *Se* allele encodes for *FUT2* transferase, forming the Type 1 and Type 3 H-antigen in epithelia of the gastrointestinal, respiratory and reproductive tract. Individuals with inactive *FUT2* gene, termed non-secretors, do not express soluble forms of blood group ABH epitopes[19]. When α1-3-linked GalNAc (by *A3GALNT*) or α1-3-linked Gal (by *A3GALT1*) is added to the galactose, A or B blood group antigens are formed, respectively. In secretor-positive individuals, formation of Lewis $(Le)^B$ antigen is possible by action of α1-4 fucosyltransferase *FUT3* on an already formed Type 1 H antigen. Le^A antigen is formed by the action of the same GSTs on a Type 1 LacNAc substrates. Similarly, Le^X and Le^Y antigens are formed

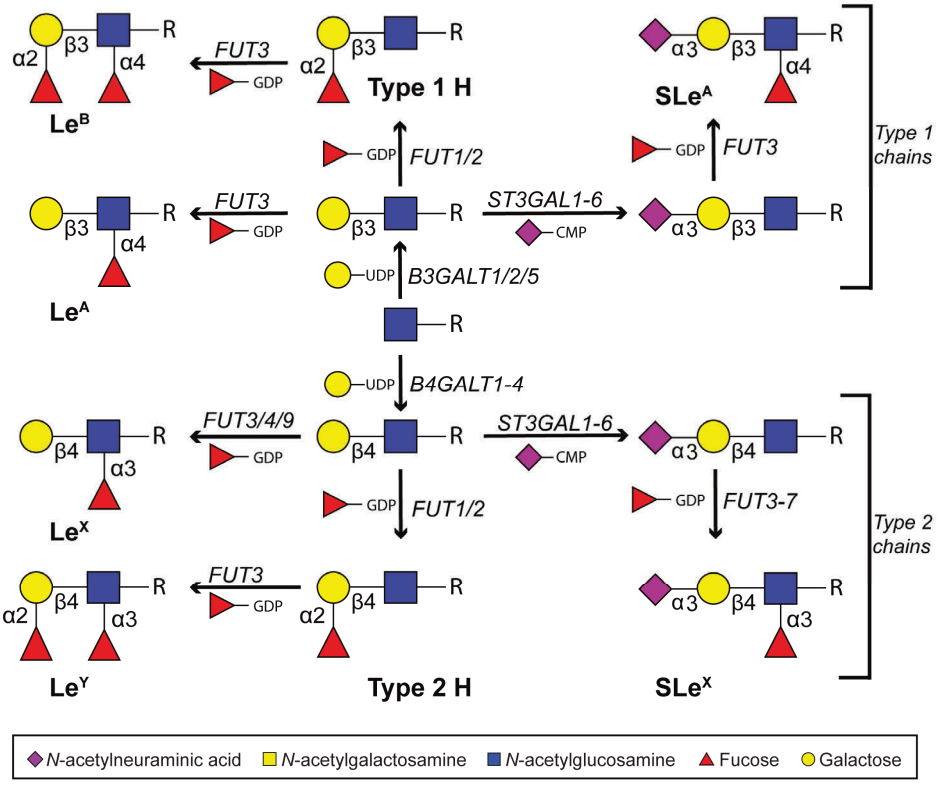

Figure 3. Common core extensions and terminal epitopes.

by action of α1-3-fucosyltransferases encoded by *FUT3, FUT4 or FUT9* on Type 2 chains and *FUT3* on Type 2 blood group H antigens, respectively[20].

A family of α2-3-sialyltransferases (*ST3GAL1-6)* is responsible for terminal sialylation of Gal, which is in competition with terminal fucosylation as described above (**Figure 3**). Fucosylation of the GlcNAc residue on sialylated substrates is still possible by the activity of *FUT3* or *FUT3/4/5/6/7* forming the sLe^A or sLe^X antigens, respectively. *ST3GAL1* is involved in the sialylation of core 1 substrates and, therefore, important for the biosynthesis of sialyl T and disialyl T antigens. On the other hand, α2-6 sialylation on *O*-glycans was described mostly in the context of the core GalNAc modification by the activity of a family of α2-6-sialyltransferases (*ST6GALNAC1-6*). *ST6GALNAC 1* and *2* are important for the biosynthesis of sialyl-6T, disialyl T and cancer associated sialyl-Tn antigen. Sialylation blocks further elongation of Gal by other monosaccharides other than the Neu5Ac, which can be added to form polysialic acids by the activity of a family of α2-8-sialyltransferases (*ST8SIA1-6*). Additionally, a specific epitope found in the colon named Sda antigen is formed by addition of a GalNAc to a Gal which is already substituted with an α2-3-NeuAc by β1-4-GalNAc transferase encoded by *B4GALNT2*[24]. This epitope is often found on core 3 glycans in the large intestine[24]. A similar antigen, named Cad, shares the same epitope sequence, however, it is carried by a core 1 *O*-glycan with a α2-6-linked Neu5Ac on the core GalNAc. Colonic glycans often show sulfation on the C3 position of Gal or at the C6 of the GlcNAc. The sulfation of GlcNAc is mediated by *CHST2,4-7* sulfotransferases, where *CHST2,4,5* and *7* are responsible for sulfation of the innermost GlcNAc in core 3 glycans. In addition, *CHST2* and *4* are involved in the sulfation of Lewis antigens[24]. The C3 sulfation of Gal is mediated by *GAL3ST2,3* and *4*, whereas the C6 sulfation of Gal is mediated by *CHST1* or *3*[25]. Sialic acid acetylation is also a common modification, which occurs on the level of the activated sugar donor CMP-Neu5Ac and is incorporated into the glycans by action of different sialyltransferases in the Golgi.

MUCIN TYPE GLYCOSYLATION IN CANCER

Mucins are a family of large secreted or membrane-bound glycoproteins, that play an important role in the epithelial cell homeostasis, acting as a protective shield[25].

The secreted mucins (MUC2, MUC5AC, MUC5B, MUC6, MUC7, and MUC19) play an important role in mucus rich tissues such as the gastrointestinal and respiratory tract. Of which, MUC1, MUC3, MUC4, MUC12, MUC16 and MUC17 are transmembrane mucins, which are also involved in cell signaling pathways. In CRC, low expression of MUC2 was reported in classical adenocarcinoma, while a high expression was reported in mucinous adenocarcinoma[26]. Additionally, higher MUC5AC expression was found in adenocarcinoma, while limited or no expression in normal colon mucosa[26]. The changes in the glycosylation that occur with cancer have shown to have an impact on the function of mucins, by changing their interaction with glycan binding proteins such as galectins[27], selectins[28] and siglecs[29]. These interactions can affect various cell processes including cell proliferation, adhesion, migration and immune surveillance which play an essential role in cancer pathogenesis[30–33]. Moreover, the very dense distribution of glycans in clustered PTS regions of mucins, as well as the possibility to target aberrant glycans present on different proteins on the cell surface, makes them attractive candidates for glycan-based immunotherapy as well as image guided surgery (IGS)[34].

The characteristic glycan alterations in cancer include specific aberrant expression of incomplete carbohydrate structures or de novo expression of carbohydrate antigens also known as tumor-associated carbohydrate antigens (TACAs). The expression of truncated structures such as Tn, sialyl-Tn, T and sialyl-T antigens was extensively described for various cancers including colon cancer[13,14]. Tn and Sialyl-Tn expression have been linked to somatic mutations in the chaperone essential for the function of core 1 synthase, COSMC (*C1GALTC1*) enabling the elongation of Tn antigen into core 1 structure, T antigen. However, this mechanism is applicable for a part of tumors expressing these antigens, and the fact that some tumors express both truncated and elongated glycans supports the hypothesis that there are other mechanisms involved such as silencing by promoter hypermethylation of COSMC[15,35]. Nevertheless, T synthase knock outs resulted in increased Tn antigen expression which enhanced cell proliferation, migration and invasiveness in colon cell cultures[36]. Additionally, the forced expression of truncated glycans directly induced oncogenic features in an organotypic epithelial tissue model[15].

The downregulation of core 3 and core 4 structures in cancer has been associated with both downregulation of core 3 synthase and upregulation of core 1 synthase. However, no study so far demonstrated the expression of those glycan types in patient derived cancer cells *vs* normal colon epithelia. The downregulation of core 3 synthase was described in colon and pancreatic adenocarcinoma and associated with poor patient prognosis and metastatic potential[17,18,37,38]. Moreover, the silencing of core 1 synthase induced increased expression of Tn, Sialyl-Tn and core 3 glycans in colon cancer cells[39]. On the other hand, the expression of the core 4 structures is dictated by the availability of core 3 precursors, and the activity of the C2GnT-M enzyme, encoded by the *GCNT3* gene, expressed in mucus secreting tissues[40]. *GCNT3* expression is downregulated in colon, ovarian and pancreatic cancer[41–43], and its transfection reduced cell proliferation, invasion and adhesion in CRC cells[40].

High expression of C2GnT-L type β1-6 GlcNAc transferase, which is responsible for biosynthesis of core 2 glycans, was described extensively in relation to cancer[44–47]. The formation of the 6-branch creates another scaffold for the activity of β1-3/4-galactosyltransferases creating Type 1 or Type 2 extensions, which have been shown to switch from primarily Type 1 in normal colon mucosa to Type 2 in CRC due to upregulation of B4GALT4[48]. Type 2 chains are precursors of sLe^X, and it has been shown that sLe^X expression is higher in CRC, whereas Le^A (Type 1) chains characteristic for normal mucosa are downregulated. The expression of β1-3-galactosyltransferase A Type 1 chain synthase, encoded by *B3GALT5,* has shown to be downregulated in colon cancer compared to normal colon mucosa[49]. It has also been previously described that the upregulation of $sLe^{X/A}$ antigens in CRC is due to the downregulation of *B4GALNT2* which encodes the β1-4-GalNAc transferase which synthesizes the Sda-antigen characteristic for the normal colon mucosa[24]. Additionally, the expression of sulfated (s)Le antigens was found to be high in normal colon epithelia, contrary to elevated sLe^X expression in CRC. Moreover, sLe^X antigens have been associated with promoting metastasis and showed correlation with poor patient prognosis[31,52]. Further research suggested that it is related to the interaction with selectins[53].

O-GLYCAN ANALYSIS

Dysregulation of mucin type *O*-glycosylation affects signaling pathways involved in cancer progression, therefore targeting mucin type glycosylation has a great clinical potential; however, it has remained an analytical challenge to enable an in-depth exploration as has been achieved for to *N*-glycosylation[53]. This is mainly due to the complexity and high heterogeneity of this type of glycosylation. Moreover, as it is not template driven, the cell glycosylation is a result of a complex interplay between different enzyme expression (GSTs and glycosidases), their competition, location within the endoplasmic reticulum and Golgi and the substrate availability. Additionally, compared to *N*-glycan analysis, *O*-glycan analysis brings along its own specific challenges as there is no universal enzyme that releases all types of *O*-glycans from glycoproteins. Therefore, the release is often performed with a chemical approach in alkaline conditions, called β-elimination, generally under reducing conditions to minimize glycan degradation (peeling). The non-reducing β-elimination is performed to enable reducing end glycan labeling, since the reduction results in non-modifiable glycan alditols. Additionally, hydrazinolysis can be used in the same manner, however, it requires special handling conditions due to the flammability of the chemicals used[54]. Additionally, *O*-glycans do not share a common core structure as *N*-glycans and can, therefore, express even more structural isomers for the same monosaccharide composition. This requires sophisticated analytical tools to be used for their analysis.

Glycan characterization from complex biological samples can be performed using a variety of methodologies that can be targeted or untargeted. Targeted approaches such as lectin and antibody binding assays evaluate the expression of known antigens whereas untargeted techniques such as mass spectrometry (MS) allow for discovery of novel glycan and glycoprotein targets.

GLYCOMICS USING MASS SPECTROMETRY

Glycan structure analysis can be performed in different ways, depending on the depth of the information needed. If the focus is on the in-depth structure sequencing of the glycan, techniques such as MS are often used. The most common ionization techniques applied for glycan analysis are matrix assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI), both can be

performed in positive and negative ionization mode. MALDI-MS is often used for analysis of released glycans; however, when measured in positive ionization mode, stabilization of sialic acids by permethylation or carboxyl group derivatization is necessary to avoid the partial loss of sialic acids due to in source and metastable decay. The obtained mass profiles provide insights into the possible monosaccharide compositions but is unable to provide information about structural isomers originating from positional or linkage variations. This can be partially overcome by tandem MS spectra (MS/MS), supplying more information about the glycan sequence. ESI can also generate positively or negatively charged ions while the most informative fragment ions for the elucidation of glycan structures are obtained by MS/MS collision induced dissociation (CID) of negatively charged precursors. In addition, the glycans do not undergo monosaccharide rearrangements, such as described in CID of positively charged precursors[55]. Nevertheless, sulfate modification rearrangements have been described in negative mode[56].

NEGATIVE IONIZATION MODE COLLISION INDUCED DISSOCIATION

CID fragmentation predominantly induces glycosidic C-type ions derived from non-reducing end, as well as A-type cross ring non-reducing end fragments, compared to glycosidic reducing end Y- and non-reducing end B-type ions (**Figure 4**). The pioneering work from Harvey *et al.,* set the groundwork for patterns in fragmentation of deprotonated *N*-glycan alditols (reviewed in [57]) where specific D- and E-type ions reveal the composition of the α1-6- and α1-3-linked glycan antennae, respectively. Specific negative ion fragmentation of mucin type glycans was described extensively by the Karlsson group and is reviewed elsewhere[58–60]. Briefly, linear glycans, such as core 1 and core 3 yield a B-ion from the loss of reducing end GalNAcol ($[M - H]^{1-}$ - 223) which decreases in intensity with the size of the glycan. Additionally, linear structures show a characteristic $[M - H]^{1-}$ - 108 ion, from the partial loss of side groups from the reducing end GalNAcol ($C_3H_8O_4$) which is never present in the branched structures with GlcNAc β1-6-linked to the core (cores 2 and 4). Dominant ions in the MS/MS spectra of the branched core 2 and core 4 are the Z-fragment ions from the glycosidic cleavage in between the core GalNAc and C3 branch. Additionally, a characteristic cross ring $A^{0,4}$ cleavage of the GalNAcol gives information about the composition of the C6 branch of the

glycan. Both core 1 and core 2 glycan fragmentation gives Y- and Z-ions at *m/z* 384^{1-} and 366^{1-}, respectively, whereas the core 2 additionally generate YY-, YZ-

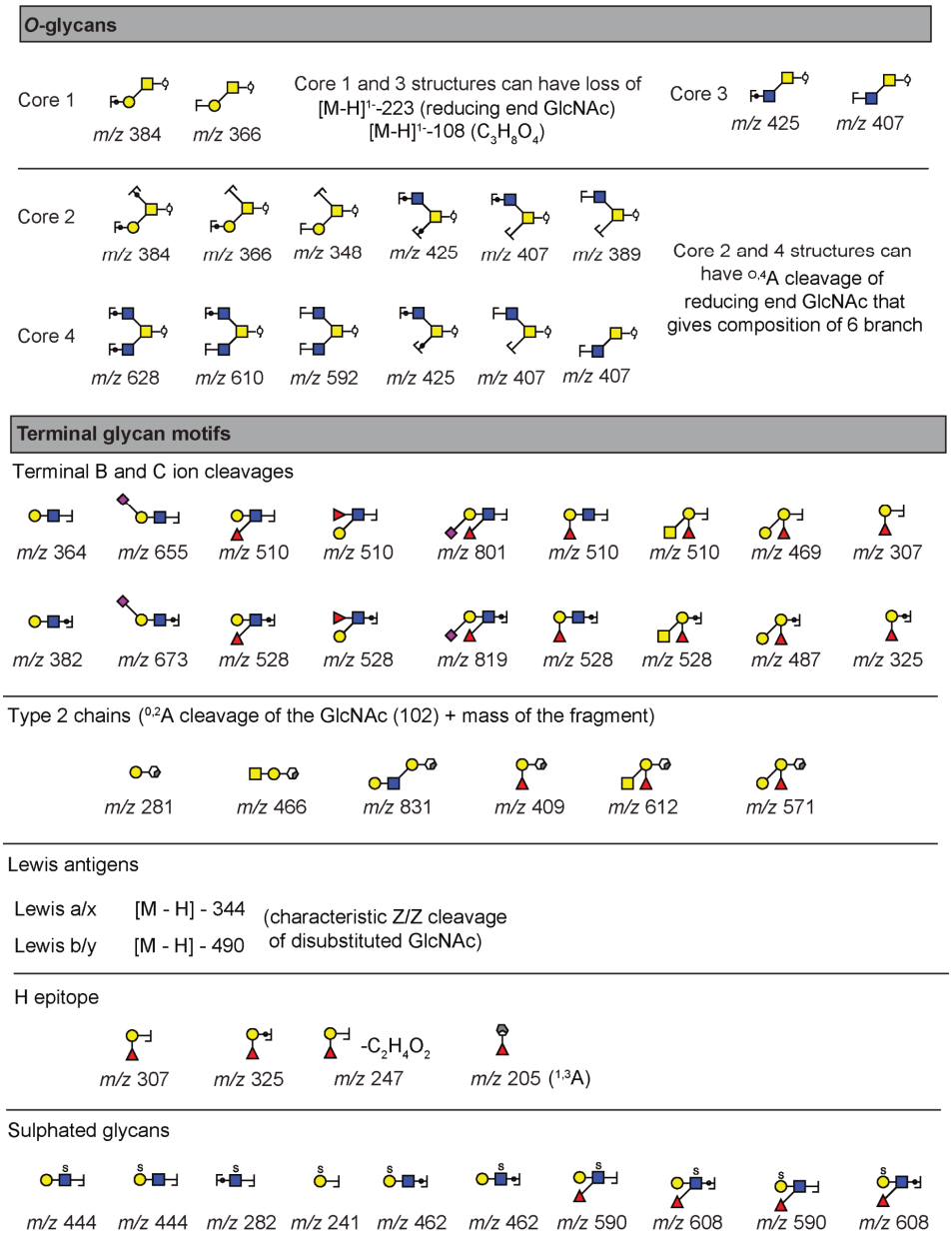

Figure 4. Characteristic fragment ions in negative mode tandem mass spectra analysis of reduced mucin type O-glycans. *Reprinted with permission from Everest-Dass et al.*[60]

and ZZ-ions ions at *m/z* 425^{1-}, $407^{1-,}$ and 389^{1-}. Type 2 LacNAc chains give specific cross ring $A^{0,2}$ cleavage, together with a concurrent ion arising from loss of water $A^{0,2} - H_2O$. While there are no diagnostic ions for Type 1 structures, the absence of the characteristic Type 2 cleavages can be used as an indication. Moreover, negative ion mode fragmentation spectra produce fragment ions enabling determination of the fucose position. Lewis type fucosylation gives rise to specific elimination of both C3 and C4 substituents of GlcNAc residue generating Z/Z and Z/Z- CH_2O ions. These ions are also present in the spectra of Le^B and Le^Y structures, although at lower intensity. These ions can also form an additional loss of acetyl and acetate groups from other acetylated monosaccharides (Z/Z-C_2H_2O and Z/Z-$C_2H_4O_2$). Unfortunately, no diagnostic ions are formed in MS/MS to distinguish between Le^X and Le^A, as well as Le^B and Le^Y. Blood group H determinants induce weak B- and C-type cleavages, at *m/z* 307^{1-}, *m/z* 325^{1-}, respectively. Additional loss of acetate (B2-$C_2H_4O_2$) at *m/z* 247^{1-}, and the cross ring fragment of galactose ($A^{1,3}$) at *m/z* 205^1 indicate an α1-2-linked fucose, however, those fragments can only be observed in MS/MS for small glycan structures. Moreover, $A^{0,2}$ cleavage with a loss of water from the GlcNAc at *m/z* 409^{1-}, is characteristic for Type 2 blood group H antigen. The Z- and Y-ions at *m/z* 495^{1-} and 513^{1-}, respectively, are diagnostic for α2-6-sialylation of the GalNAc and a cross ring $X^{0,2}$ can give additional indication of the presence of the α2-6-linked Neu5Ac. Determining the position of the sulfate modification on a particular monosaccharide is a very challenging task. The ions indicating the position of a sulfate on either Gal or GlcNAc are B-ions at *m/z* 444^{1-} which in case of fucosylation shifts to *m/z* 590^1. The location of the sulfate can be determined by MS/MS for very small structures, or by MS^3, where fragment ions at *m/z* 241^{1-} and 282^{1-} are indicative of its linkage to Gal and GlcNAc, respectively.

Of note, it has been shown that negative ion mode fragmentation performed by ion trap (IT) instruments produces reproducible spectra, even if instruments from different laboratories and manufacturers are used[61]. Therefore, a repository and a database Unicarb DR and DB, respectively, aims to collect a large number of annotated glycan spectra to facilitate spectral matching and automatic assignments of glycan structures[62].

However, ESI alone is unable to distinguish between glycan isomers, therefore coupling ESI with online separation techniques is often a preferred choice, particularly for the analysis of mucin type *O*-glycans.

ISOMERIC SEPARATION TECHNIQUES

Several liquid chromatography (LC) separation techniques are used for isomeric separation of glycans from complex samples including reverse phase (RP), hydrophilic interaction liquid chromatography (HILIC), porous graphitized carbon (PGC), and mesoporous graphitic carbon (MGC). RP-C18-MS was used recently for the analysis of both *N*- and *O*-linked permethylated glycans and demonstrated baseline separation of glycan isomers, together with mobile phase optimization by adding lithium which limited adduct heterogeneity and simplified data analysis[63]. New approaches include the use of micro pillar array (qPAC) columns for the analysis of permethylated glycans, demonstrating isomeric separation of *N*- and *O*-glycans, although the diversity of *O*-glycans tested was very limited[64]. Unlike RP, HILIC-LC can be used for the separation of non-permethylated, native or reducing end labeled glycans, and has recently been demonstrated as a valuable tool also for the isomeric separation of *N*-glycans[65,66]. Despite increasing efficiency of HILIC and RP stationary phases in separating glycan isomers, due to a specific combination of hydrophobic and polar interactions of carbohydrates with the planar surface of the graphite, PGC outstands with its superior results in separating both linkage and positional glycan isomers[67].

The usage of PGC chromatography was pioneered by the laboratory of prof. Packer for glycan desalting and purification prior to MS analysis using the material for solid phase extraction (SPE)[68]. It was thereafter employed for HPLC analysis or oligosaccharides where it showed excellent separation power for structural and linkage isomers[69]. Moreover, the glycans can be analyzed without prior derivatization, in their native form or after reduction thereby avoiding separation of reducing end α- and β-anomers. The advantages include shorter sample preparation, less sample loss due to additional purification steps and incomplete derivatization. Nevertheless, the PGC-LC was also used for the analysis of permethylated glycans, enabling separation at higher temperatures[70,71]. Coupling of PGC-LC with ESI in positive ion mode is used less frequently[72], although the signal intensities of neutral and negatively charged glycan species are more

comparable to UV detection LC of labeled glycans[73]. Despite the bias in ionization of negatively charged species, PGC is more often coupled with negative ion ESI, which results in MS/MS spectra that contain informative cross-ring cleavages useful for determining glycan structures[69]. Specific elution orders are observed for *N*-linked glycans such as bisected *N*-glycans elute prior to their branched isomers, core fucosylated glycans elute later than their antenna fucosylated isomer pairs, and finally α2-6-linked sialylated glycans elute earlier than their α2-3-linked sialylated isomers. However, in the case of specifically mucin type *O*-glycans, elution patterns are difficult to predict, as they do not retain based on the glycan size, as exemplified by the linear trisaccharide Fucα1-2Galβ1-3GalNAcol which elutes around 30 minutes after the branched trisaccharide Galβ1-3[GlcNAcβ1-6]GalNAcol (unpublished observation).

In summary, PGC-ESI-MS/MS in negative ion mode allows the analysis of mucin type glycosylation in sensitive and an in-depth manner, allowing chromatographic separation of different linkage and positional isomers, and their structural characterization via negative ion mode CID. In this way cancer specific glycan signatures can be revealed giving new insights into the cancer specific glycosylation pathways originating from the changes in the cell biosynthetic machinery.

SCOPE OF THE THESIS

The scope of this thesis was to investigate mucin type glycomic signatures of CRC in order to explore new potential therapeutic targets for immunotherapy.

Previous studies from Kolarich *et. al.*,[69] showed the potential of using polyvinylidene difluoride (PVDF) membrane-based protein immobilization strategy to facilitate sequential release of both *N*- and *O*-glycans from the same sample, followed by analysis with PGC-LC-MS/MS allowing isomeric separation and in-depth structural identification by negative ion mode CID. In **Chapter 2** we optimized their approach, allowing a higher throughput for sample preparation using 96-well plates enabling a robust and sensitive analysis of a greater number of samples at a time. In **Chapter 3** we explored the effects of different polar protic solvents for dopant enriched nitrogen gas in order to improve the sensitivity for the analysis of different *N*- and *O*-glycan species and to evaluate its effect on the charge state distribution.

In **Chapter 4** we applied the methodology developed in Chapter 1 and characterized the *O*-glycomes of 26 CRC cell lines. This work was a continuation of the previous work from our group characterizing the same set of cell lines for their *N*-glycomes[74]. Furthermore, the *O*-glycan phenotypes were associated with the GST expression relevant for their biosynthesis. Inspired by the associations of the cell line glycomes and cell differentiation, we evaluated in **Chapter 5** the changes in cell line glycome upon butyrate stimulation and spontaneous differentiation in culture. Moreover, we explored the association between the glycome changes with the changes in the cell proteome. **In Chapter 6** we optimized a previously developed protocol for *N*- and *O*-glycomics from formalin fixed paraffin embedded tissues and applied it for the analysis of CRC tissues and their patient matched colon mucosa controls. This allowed us to identify CRC specific glycan signatures, which showed association with the differential regulation of specific biosynthetic pathways in cancer previously described on the transcriptomic level.

Finally, **Chapter 7** gives a general discussion about the technical challenges encountered related to the analysis of *O*-glycans, evaluation of different models for studying glycomic changes in cancer, the association between transcriptomic signatures and glycomic profiles, as well as future perspectives on how our results could be used in clinical applications, as well as the remaining challenges to be addressed.

REFERENCES

1. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J. Clin. 71, 209–249 (2021).
2. Chung, D. C. The Genetic Basis of Colorectal Cancer: Insights Into Critical Pathways of Tumorigenesis. doi:10.1053/gast.2000.16507.
3. Angelova, M. et al. Characterization of the immunophenotypes and antigenomes of colorectal cancers reveals distinct tumor escape mechanisms and novel targets for immunotherapy. Genome Biol. 16, 64 (2015).
4. Guinney, J. et al. The consensus molecular subtypes of colorectal cancer. Nature Medicine (2015) doi:10.1038/nm.3967.
5. Ghani, S. et al. Recent developments in antibody derivatives against colorectal cancer; A review. Life Sci. 265, 118791 (2021).
6. Piawah, S. & Venook, A. P. Targeted therapy for colorectal cancer metastases: A review of current methods of molecularly targeted therapy and the use of tumor biomarkers in the treatment of metastatic colorectal cancer. *Cancer* **125**, 4139–4147 (2019).
7. Ayyar, B. V., Arora, S. & O'Kennedy, R. Coming-of-Age of Antibodies in Cancer Therapeutics. *Trends Pharmacol. Sci.* **37**, 1009–1028 (2016).
8. Holst, S., Wuhrer, M. & Rombouts, Y. *Glycosylation characteristics of colorectal cancer*. vol. 126 203–256 (Elsevier Inc., 2015).
9. Mereiter, S., Balmaña, M., Campos, D., Gomes, J. & Reis, C. A. Glycosylation in the Era of Cancer-Targeted Therapy: Where Are We Heading? *Cancer Cell* vol. 36 6–16 (2019).
10. Stanley P, Cummings RD. Structures Common to Different Glycans. 2017. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 14.
11. Schjoldager, K. T., Narimatsu, Y., Joshi, H. J. & Clausen, H. Global view of human protein glycosylation pathways and functions. *Nat. Rev. Mol. Cell Biol.* **21**, 729–749 (2020).
12. Schjoldager, K. T.-B. G. & Clausen, H. Site-specific protein O-glycosylation modulates proprotein processing — Deciphering specific functions of the large polypeptide GalNAc-transferase gene family. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1820**, 2079–2094 (2012).
13. Cornelissen, L. A. M. *et al.* Tn Antigen Expression Contributes to an Immune Suppressive Microenvironment and Drives Tumor Growth in Colorectal Cancer. *Front. Oncol.* **10**, 1–15 (2020).
14. Matsumoto, T. *et al.* Tn antigen expression defines an immune cold subset of mismatch-repair deficient colorectal cancer. *Int. J. Mol. Sci.* **21**, 1–13 (2020).
15. Ju, T. *et al.* Human tumor antigens Tn and sialyl Tn arise from mutations in Cosmc. *Cancer Res.* **68**, 1636–1646 (2008).
16. Yeh, J. C., Ong, E. & Fukuda, M. Molecular cloning and expression of a novel β-1,6-N-acetylglucosaminyltransferase that forms core 2, core 4, and I branches. *J. Biol. Chem.* **274**, 3215–3221 (1999).

17. Doi, N. *et al.* Clinicopathological significance of core 3 O-glycan synthetic enzyme, β1,3-N-acetylglucosaminyltransferase 6 in pancreatic ductal adenocarcinoma. *PLoS One* **15**, e0242851 (2020).
18. An, G. *et al.* Increased susceptibility to colitis and colorectal tumors in mice lacking core 3-derived O-glycans. *J. Exp. Med.* **204**, 1417–1429 (2007).
19. Vavasseur, F., Yang, J. M., Dole, K., Paulsen, H. & Brockhausen, I. Synthesis of O-glycan core 3: Characterization of UDP-GlcNAc: GalNAc-R β3-N-acetyl-glucosaminyltransferase activity from colonic mucosal tissues and lack of the activity in human cancer cell lines. *Glycobiology* **5**, 351–357 (1995).
20. Lee, S. H. *et al.* Core3 O-glycan synthase suppresses tumor formation and metastasis of prostate carcinoma PC3 and LNCaP cells through down-regulation of α2β1 integrin complex. *J. Biol. Chem.* **284**, 17157–17169 (2009).
21. Saarinen, J., Welgus, H. G., Flizar, C. A., Kalkkinen, N. & Helin, J. N-Glycan structures of matrix metalloproteinase-1 derived from human fibroblasts and from HT-1080 fibrosarcoma cells. *Eur. J. Biochem.* **259**, 829–840 (1999).
22. Dell, A. *et al.* Structural Analysis of the Oligosaccharides Derived from Glycodelin, a Human Glycoprotein with Potent Immunosuppressive and Contraceptive Activities (∗). *Journal of Biological Chemistry* **270**, 24116–24126 (1995).
23. Dotz, V. & Wuhrer, M. Histo-blood group glycans in the context of personalized medicine. *Biochimica et Biophysica Acta - General Subjects* **1860**, 1596–1607 (2016).
24. Groux-Degroote, S. *et al.* B4GALNT2 gene expression controls the biosynthesis of Sda and sialyl Lewis X antigens in healthy and cancer human gastrointestinal tract. *Int. J. Biochem. Cell Biol.* **53**, 442–449 (2014).
25. Corfield, A. P. Mucins: A biologically relevant glycan barrier in mucosal protection. *Biochimica et Biophysica Acta - General Subjects* vol. 1850 (2015).
26. Krishn, S. R. *et al.* Mucins and associated glycan signatures in colon adenoma-carcinoma sequence: Prospective pathological implication(s) for early diagnosis of colon cancer. *Cancer Lett.* **374**, 304–314 (2016).
27. Belo, A. I., Van Der Sar, A. M., Tefsen, B. & Van Die, I. Galectin-4 Reduces Migration and Metastasis Formation of Pancreatic Cancer Cells. (2013) doi:10.1371/journal.pone.0065957.
28. Trinchera, M. *et al.* Selectin Ligands Sialyl-Lewis a and Sialyl-Lewis x in Gastrointestinal Cancers. *Biology* **6**, 16 (2017).
29. Rodriguez, E. *et al.* Sialic acids in pancreatic cancer cells drive tumour-associated macrophage differentiation via the Siglec receptors Siglec-7 and Siglec-9. *Nat. Commun.* **12**, 1–14 (2021).
30. Rabinovich, G. A. & Croci, D. O. Regulatory Circuits Mediated by Lectin-Glycan Interactions in Autoimmunity and Cancer. *Immunity* vol. 36 322–335 (2012).
31. Witz, I. P. The selectin-selectin ligand axis in tumor progression. *Cancer Metastasis Rev.* **27**, 19–30 (2008).

32. Bärenwaldt, A. & Läubli, H. The sialoglycan-Siglec glyco-immune checkpoint–a target for improving innate and adaptive anti-cancer immunity. *Expert Opinion on Therapeutic Targets* vol. 23 (2019).
33. Mantuano, N. R., Natoli, M., Zippelius, A. & Läubli, H. Tumor-associated carbohydrates and immunomodulatory lectins as targets for cancer immunotherapy. *Journal for ImmunoTherapy of Cancer* **8**, e001222 (2020).
34. Houvast, R. D. *et al.* Targeting glycans and heavily glycosylated proteins for tumor imaging. *Cancers* vol. 12 1–26 (2020).
35. Ju, T., Aryal, R. P., Kudelka, M. R., Wang, Y. & Cummings, R. D. The Cosmc connection to the Tn antigen in cancer. *Cancer Biomark.* **14**, 63–81 (2014).
36. Dong, X. *et al.* T-Synthase Deficiency Enhances Oncogenic Features in Human Colorectal Cancer Cells via Activation of Epithelial-Mesenchymal Transition. *Biomed Res. Int.* **2018**, (2018).
37. Iwai, T. *et al.* Core 3 synthase is down-regulated in colon carcinoma and profoundly suppresses the metastatic potential of carcinoma cells. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 4572–4577 (2005).
38. Henion, T. R. & Schwarting, G. A. N-Linked polylactosamine glycan synthesis is regulated by co-expression of β3GnT2 and GCNT2. *J. Cell. Physiol.* **229**, 471–478 (2014).
39. Barrow, H., Tam, B., Duckworth, C. A., Rhodes, J. M. & Yu, L. G. Suppression of Core 1 Gal-Transferase Is Associated with Reduction of TF and Reciprocal Increase of Tn, sialyl-Tn and Core 3 Glycans in Human Colon Cancer Cells. *PLoS One* **8**, e59792 (2013).
40. Huang, M. C. *et al.* C2GnT-M is downregulated in colorectal cancer and its re-expression causes growth inhibition of colon cancer cells. *Oncogene* **25**, 3267–3276 (2006).
41. Fernández, L. P. *et al.* The role of glycosyltransferase enzyme GCNT3 in colon and ovarian cancer prognosis and chemoresistance. *Sci. Rep.* **8**, 8485 (2018).
42. González-Vallinas, M. *et al.* Clinical relevance of the differential expression of the glycosyltransferase gene GCNT3 in colon cancer. *Eur. J. Cancer* **51**, 1–8 (2015).
43. Gupta, R. *et al.* Global analysis of human glycosyltransferases reveals novel targets for pancreatic cancer pathogenesis. *Br. J. Cancer* **122**, 1661–1672 (2020).
44. Chen, Z., Gulzar, Z. G., St Hill, C. A., Walcheck, B. & Brooks, J. D. Increased expression of GCNT1 is associated with altered O-glycosylation of PSA, PAP, and MUC1 in human prostate cancers. *Prostate* **74**, 1059–1067 (2014).
45. Dalziel, M. *et al.* The Relative Activities of the C2GnT1 and ST3Gal-I Glycosyltransferases Determine O-Glycan Structure and Expression of a Tumor-associated Epitope on MUC1. *J. Biol. Chem.* **276**, 11007–11015 (2001).
46. St Hill, C. A., Baharo-Hassan, D. & Farooqui, M. C2-O-sLeX glycoproteins are E-selectin ligands that regulate invasion of human colon and hepatic carcinoma cells. *PLoS One* **6**, e16281 (2011).
47. St Hill, C. A., Bullard, K. M. & Walcheck, B. Expression of the high-affinity selectin glycan ligand C2-O-sLeX by colon carcinoma cells. *Cancer Lett.* **217**, 105–113 (2005).

48. Chen, W.-S., Chang, H.-Y., Li, C.-P., Liu, J. M. & Huang, T.-S. Tumor beta-1,4-galactosyltransferase IV overexpression is closely associated with colorectal cancer metastasis and poor prognosis. *Clin. Cancer Res.* **11**, 8615–8622 (2005).
49. Salvini, R., Bardoni, A., Valli, M. & Trinchera, M. β1,3-Galactosyltransferase β3Gal-T5 Acts on the GlcNAcβ1→3Galβ1→4GlcNAcβ1→R Sugar Chains of Carcinoembryonic Antigen and Other N-Linked Glycoproteins and Is Down-regulated in Colon Adenocarcinomas *. *J. Biol. Chem.* **276**, 3564–3573 (2001).
50. Nakamori, S. *et al.* Increased expression of sialyl Lewisx antigen correlates with poor survival in patients with colorectal carcinoma: clinicopathological and immunohistochemical study. *Cancer Res.* **53**, 3632–3637 (1993).
51. Yamada, N. *et al.* Increased expression of sialyl Lewis A and sialyl Lewis X in liver metastases of human colorectal carcinoma. *Invasion Metastasis* **15**, 95–102 (1995).
52. Paschos, K. A., Canovas, D. & Bird, N. C. The engagement of selectins and their ligands in colorectal cancer liver metastases. *J. Cell. Mol. Med.* **14**, 165–174 (2010).
53. Lageveen-Kammeijer, G. S. M., Kuster, B., Reusch, D. & Wuhrer, M. High sensitivity glycomics in biomedicine. *Mass Spectrom. Rev.* e21730 (2021).
54. Kozak, R. P., Royle, L., Gardner, R. A., Fernandes, D. L. & Wuhrer, M. Suppression of peeling during the release of O-glycans by hydrazinolysis. *Anal. Biochem.* **423**, 119–128 (2012).
55. Wuhrer, M., Koeleman, C. A. M., Hokke, C. H. & Deelder, A. M. Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antennae gives rise to misleading fragments. *Rapid Commun. Mass Spectrom.* **20**, 1747–1754 (2006).
56. Kenny, D. T., Issa, S. M. A. & Karlsson, N. G. Sulfate migration in oligosaccharides induced by negative ion mode ion trap collision-induced dissociation. *Rapid Commun. Mass Spectrom.* **25**, 2611–2618 (2011).
57. Harvey, D. J. Negative Ion Mass Spectrometry for the Analysis of N -linked Glycans. *Mass Spectrom. Rev.* **39**, 586–679 (2020).
58. Karlsson, N. G., Schulz, B. L. & Packer, N. H. Structural determination of neutral O-linked oligosaccharide alditols by negative ion LC-electrospray-MSn. *J. Am. Soc. Mass Spectrom.* **15**, 659–672 (2004).
59. Robbe, C., Capon, C., Coddeville, B. & Michalski, J. C. Diagnostic ions for the rapid analysis by nano-electrospray ionization quadrupole time-of-flight mass spectrometry of O-glycans from human mucins. *Rapid Commun. Mass Spectrom.* **18**, 412–420 (2004).
60. Everest-Dass, A. V., Abrahams, J. L., Kolarich, D., Packer, N. H. & Campbell, M. P. Structural feature ions for distinguishing N- and O-linked glycan isomers by LC-ESI-IT MS/MS. *J. Am. Soc. Mass Spectrom.* **24**, 895–906 (2013).
61. Campbell, M. P. *et al.* Validation of the curation pipeline of UniCarb-DB: Building a global glycan reference MS/MS repository. *Biochimica et Biophysica Acta - Proteins and Proteomics* **1844**, 108–116 (2014).
62. Rojas-Macias, M. A. *et al.* Towards a standardized bioinformatics infrastructure for N- and O-glycomics. *Nat. Commun.* **10**, 3275 (2019).

63. Kurz, S., Sheikh, M. O., Lu, S., Wells, L. & Tiemeyer, M. Separation and Identification of Permethylated Glycan Isomers by Reversed Phase NanoLC-NSI-MSn. *Mol. Cell. Proteomics* **20**, 100045 (2021).
64. Cho, B. G., Jiang, P., Goli, M., Gautam, S. & Mechref, Y. Using micro pillar array columns (μPAC) for the analysis of permethylated glycans. *Analyst* **146**, 4374–4383 (2021).
65. Messina, A. *et al.* HILIC-UPLC-MS for high throughput and isomeric N-glycan separation and characterization in Congenital Disorders Glycosylation and human diseases. *Glycoconj. J.* **38**, 201–211 (2021).
66. Moravcová, D., Čmelík, R. & Křenková, J. Separation of labeled isomeric oligosaccharides by hydrophilic interaction liquid chromatography - the role of organic solvent in manipulating separation selectivity of the amide stationary phase. *J. Chromatogr. A* **1651**, 462303 (2021).
67. Pabst, M., Bondili, J. S., Stadlmann, J., Mach, L. & Altmann, F. Mass + Retention Time = Structure: A Strategy for the Analysis of N -Glycans by Carbon LC-ESI-MS and Its Application to Fibrin N -Glycans. *Anal. Chem.* **79**, 5051–5057 (2007).
68. Packer, N. H., Lawson, M. A., Jardine, D. R. & Redmond, J. W. A general approach to desalting oligosaccharides released from glycoproteins. *Glycoconj. J.* **15**, 737–747 (1998).
69. Jensen, P. H., Karlsson, N. G., Kolarich, D. & Packer, N. H. Structural analysis of N- and O-glycans released from glycoproteins. *Nat. Protoc.* **7**, 1299–1310 (2012).
70. Zhou, S., Dong, X., Veillon, L., Huang, Y. & Mechref, Y. LC-MS/MS analysis of permethylated N-glycans facilitating isomeric characterization. *Anal. Bioanal. Chem.* **409**, 453–466 (2017).
71. Zhou, S., Hu, Y. & Mechref, Y. High-temperature LC-MS/MS of permethylated glycans derived from glycoproteins. *Electrophoresis* **37**, 1506–1513 (2016).
72. Hua, S. *et al.* Isomer-specific chromatographic profiling yields highly sensitive and specific potential N-glycan biomarkers for epithelial ovarian cancer. *J. Chromatogr. A* **1279**, 58–67 (2013).
73. Pabst, M. & Altmann, F. Influence of electrosorption, solvent, temperature, and ion polarity on the performance of LC-ESI-MS using graphitic carbon for acidic oligosaccharides. *Anal. Chem.* **80**, 7534–7542 (2008).
74. Holst, S. *et al.* N-glycosylation Profiling of Colorectal Cancer Cell Lines Reveals Association of Fucosylation with Differentiation and Caudal Type Homebox 1 (CDX1)/Villin mRNA Expression. *Mol. Cell. Proteomics* **15**, 124–140 (2016).