Multilayer cancer glycomics

Wang, D.

Chapter 1

Introduction

1.1 Glycosylation

Glycosylation is one of the most observed post-translational modifications (PTMs), and its process is defined as the enzyme-catalyzed covalent attachment of carbohydrates **-** also known as **glycans** - to a polypeptide, lipid, polynucleotide or another organic compound. To form **glycoconjugates** (such as glycoproteins and glycolipids), a series of sequential actions of glycosyltransferases are typically involved using specific sugar nucleotide donor as substrates [1].

Glycans are one of the four major macromolecules of life, the other three classes are deoxyribonucleic acid (DNA), ribonucleic acid (RNA) and proteins. Glycans are composed of monosaccharides which are covalently linked via glycosidic bonds [2, 3]. Monosaccharides can be grouped into two types: aldoses and ketoses, depending on the location of the carbonyl group. Aldoses have a carbonyl group at the terminal of the carbon chain (an aldehyde group), while ketoses carry a carbonyl group at an inner carbon forming a ketone group. Generally, free monosaccharides (sugars) exist in linear or ring forms, the latter often consisting of five carbons and one oxygen in the ring [4]. Ring formation produces a chiral anomeric center at C-1 position of the monosaccharide aldoses to form an α- or β-anomer which can be interconverted [2]. Depending on the position of glycosidic oxygen at the anomeric center, covalent glycosidic linkages between two monosaccharides are generated in two forms: α- and β-linkages, resulting in different monosaccharide stereoisomers [2]. In addition, monosaccharides may occur as D- and L-enantiomers which are assigned based on the configuration of the hydroxyl group at the C-5 position [5]. In humans, except for L-fucose (Fuc), the most common naturally occurring monosaccharides are in the D-configuration (**Figure 1**, D-glucose (Glc), D-galactose (Gal), D-mannose (Man), *N*-acetyl-D-glucosamine (GlcNAc), *N*-acetyl-D-galactosamine (GalNAc) and *N*-acetylneuraminic acid (Neu5Ac)).

In eukaryotes, membrane proteins and secretory proteins go through the endoplasmic reticulum (ER)- Golgi pathway, in which the majority of the glycan synthesis takes place. Proteins can be modified with *N*-glycans, *O*-glycans, glycosylphosphatidylinositol (GPI) anchors and glycosaminoglycan chains. All these modifications as well as the glycosylation of lipids (glyco(sphingo)lipids) occur in the ER and Golgi. Notably, preassembly of *N*-glycan and GPI precursor occurs at the

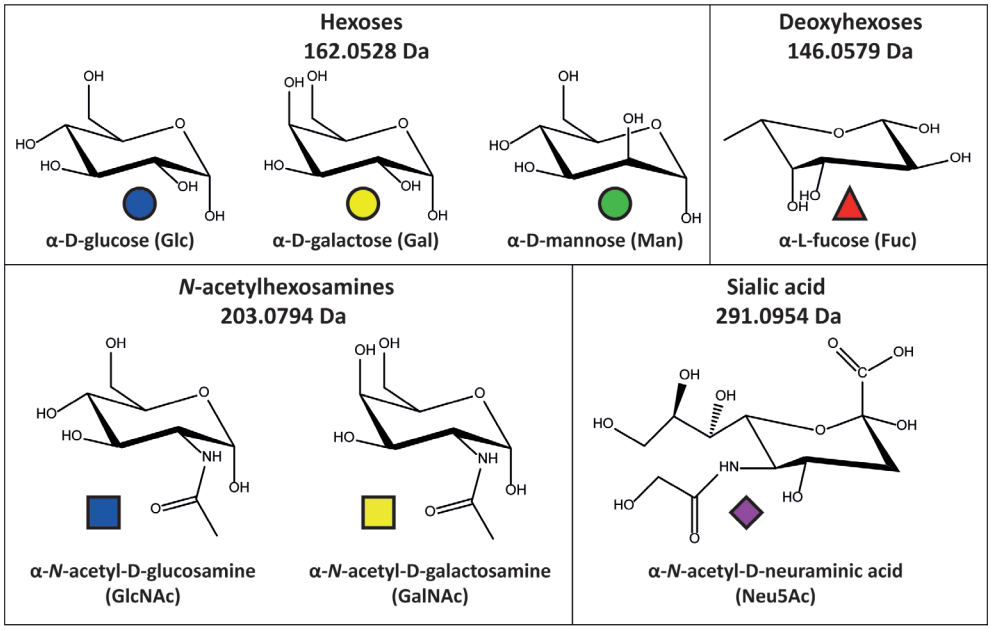

Figure 1. Most common monosaccharides structure in human.

cytoplasmic side of the ER membrane, followed by further modifications at the luminal side of the ER and Golgi [6, 7]. The conformation of proteins can be affected by glycosylation resulting in the changes in protein stability, activity and interactions [8-10]. The glycan portion on secreted glycoproteins influences solubility, hydrophilicity and charge of the protein, and may protect the protein from proteolysis and decrease or mediate intermolecular interactions [1]. Glycosylation on membrane proteins such as receptors, adhesion molecules and channels facilitates the folding and promotes their stability [1].

1.2 *N*-glycan

N-glycosylation is a process in which *N*-glycans are covalently linked to an asparagine (Asn) residue of a protein via an *N*-glycosidic linkage (*N*-glycan linkage). A consensus sequence of Asn-X-serine (Ser)/Threonine (Thr) termed “sequon” is required for the attachment of *N*-glycans to a protein, where “X” can be any amino acid with exception of proline [11]. *N*-glycans are found on many glycoproteins including secreted and membrane-bound glycoproteins at Asn-X-Ser/Thr. Approximately 70% of proteins contain these sequons, out of which ~ 70% are glycosylated [11]. Before the *N*-glycan can be transferred to a protein on the luminal reticulum membrane,

the first step in the biosynthesis of the *N*-glycan precursor occurs on the cytoplasmic face of the ER membrane to generate Man5GlcNAc2-P-P-dolichol (Dol (lipid), with P indicating phosphate). Subsequently, the molecule is turned over to the luminal side of the ER membrane for the next synthesis steps. The resulting *N*-glycan precursor is transferred to the nascent protein in the ER followed by a series of reaction in ER and Golgi by membrane-bound glycosides and glycosyltransferases [12].

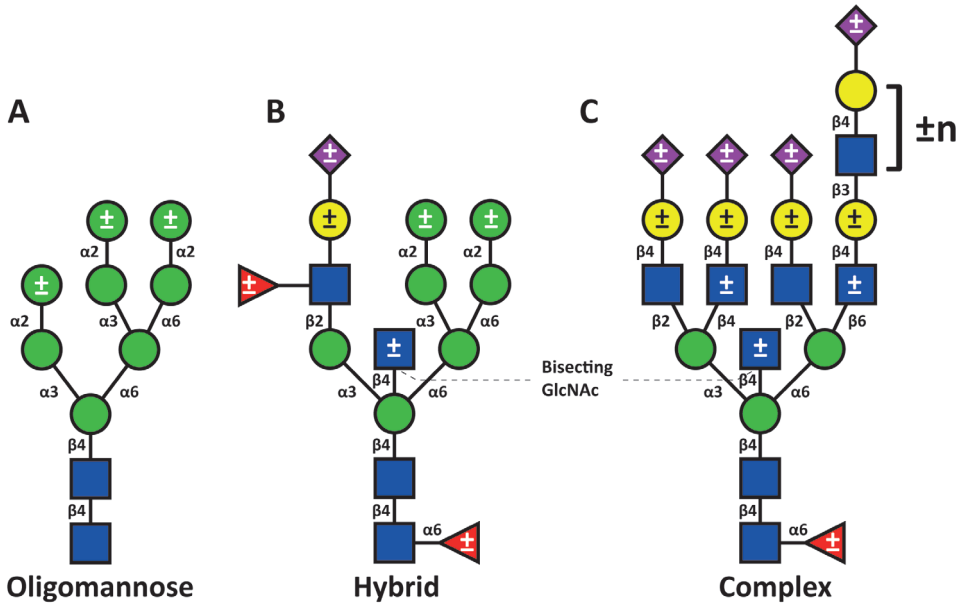

Figure 2. Three common types of *N*-glycans on human proteins [11]. (A) Oligomannosidic *N*-glycans can carry five up to nine mannoses with or without glucose at the terminal of the Manα1-3 arm. **(B)** Hybrid and **(C)** complex type *N*-glycans have antennae of varying length and may carry a core or antennary fucose, bisecting GlcNAc and sialic acid attached to the galactose which is indicated by ± . The addition of one or more LacNAc chains to the antenna, usually to the 6-antenna of complex glycan indicated with ±n.

There are three major types of *N*-glycans and all share a common core motif Manα 1-6 (Manα1-3) Manβ1-4GlcNAcβ1-4GlcNAcβ1-Asn-X-Ser/Thr (**Figure 2**). The most simplistic class is the oligomannosidic type, where only mannose residues are linked to the core sequence and is generated in the ER (**Figure 2A**). After trimming down the mannoses on the Manα1-3 arm the hybrid type class is generated, with one or two antenna extended by attachment of GlcNAc to Manα1-3 arm and on the other arm (Manα1-6 arm) the remaining mannoses are attached to the core (**Figure 2B**). The last class is the complex type (**Figure 2C**) which varies in the number of antennae, and in all cases, the antennae starts with a GlcNAc attached to the core

and, in some cases, with the attachment of a GlcNAc residues to the C-4-position of the branching mannose residue, known as a bisecting GlcNAc (can also occur on the hybrid type). Further modifications include addition of a fucose residue to the reducing-terminal GlcNAc and non-reducing terminal GlcNAc and Gal, which form the so-called core-fucosylation and antenna fucosylation, and addition of one or more *N*-acetyllactosamine (Galβ1-4GlcNAc, LacNAc) to the antenna, in most cases this is to the 6-antenna.

1.3 *O*-glycan

Another common glycan class present on proteins is *O*-GalNAc linked glycosylation, which starts with an α-linked GalNAc covalently bonded to the side chain -OH of a serine (Ser) or threonine (Thr), normally found in mucous secretions and at the cell surface. Next to *O*-GalNAc glycans, there are several kinds of non-mucin *O*-glycans including α-linked *O*-fucose, β-linked *O*-xylose, α-linked *O*-mannose, β-linked *O*-GlcNAc (*N*-acetylglucosamine), α- or β-linked *O*-galactose and α- or β-linked *O*-glucose glycans [13]. In this thesis, we mainly focused on mucin type *O*-glycans (*O*-GalNAc). Unlike *N*-glycans, mucin *O*-glycans are formed by a combination of five common monosaccharides instead of seven, including GalNAc, GlcNAc, Gal, Fuc and Neu5Ac, but lack Man and Glc [14]. In total, there are four common *O*-glycan core structures (**Figure 3**). Core 1 consists of Galβ1-3GalNAcα1-Ser/Thr, core 2 of GlcNAcβ1-6(Galβ1-3)GalNAcα1-Ser/Thr, core 3 of GlcNAcβ1-3GalNAcα1-Ser/Thr and core 4 of GlcNAcβ1-6(GlcNAcβ1-3)GalNAcα1-Ser/Thr. Importantly, many terminal structures on *O*-glycans are antigenic or represent recognition sites for lectins [14]. The sialylated or sulfated Lewis (Le) antigens are considered as ligands for selectins. Moreover, the simplest mucin *O*-glycan is the Tn antigen and is formed by solely a GalNAc residue linked to the Ser or Thr and is also known to be antigenic. When a Gal is added in β1-3 linkage the T antigen is formed [14]. Both, Tn and T antigen can be extended by a sialic acid residue to generate sialylated-Tn (sTn) or -T(sT) antigens, respectively. Core 1 and core 2 *O*-glycans can be found on many glycoproteins and mucins, while core 3 and core 4 *O*-glycans are mostly found on secreted mucins of mucin-secreting tissues [13]. Mucins are known for their "variable number of tandem repeat (VNTR)" regions, which is rich in proline residues promoting *O*-glycosylation and various repeats of Ser and/or Thr residues [14]. Elongation of *O*-glycans can contain an i or I antigen, which is defined as a repetition

of Galβ1-3GlcNAcβ1- or GalNAcβ1-4GlcNAcβ1- (LacdiNAc) motifs, respectively. There are hardly any reports on Neu5Acα2-6Gal linkages on *O*-glycans with α2-6 sialylation being restricted to the innermost GalNAc (Neu5Acα2-6GalNAcα1-) and the α2-3 sialylation is generally found on Gal of *O*-glycans [14].

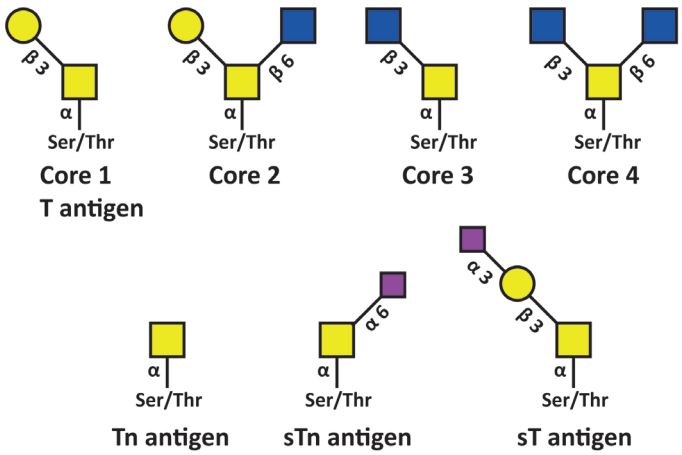

Figure 3. **Four common *O*-glycan cores and related antigens.**

A number of diseases get involved in the alteration of expression of mucin gene as well as aberrant mucin *O*-glycans structure and properties [13]. Tn antigen and sTn antigen are common cancer-associated structures which might result from the block of core 1 synthesis due to the abnormal expression of T synthase and lack function of Cosmc chaperone which is required by T synthase for its full activity [13]. Core 2 *O*-glycans are found to be involved in tumor progression and high abundance of sialylated core 1 and core 2 *O*-glycans were observed in many leukemia and tumor cells. In cancer, the abnormal expression of *O*-glycans can be attributed to the altered expression of expression or changed activities of relevant glycosyltransferases (GTs) [13].

1.4 GSL glycan

Glycosphingolipids (GSLs) are formed by attachment of glycans *via* glycosidic linkage to the C-1 hydroxyl group of a ceramide lipid moiety, which consists of an amino alcohol that is attached to a fatty acid in amide linkage. GSLs are heterogenic compounds, varying in their lipid portions regarding length of the alkyl chain, hydroxylation of the ceramide and saturation status of both sphingosine and fatty

acid moieties. In addition, GSLs show glycan structural diversity in number and type of monosaccharides, the order of monosaccharides, anomeric configuration, branching and linkage positions (**Figure 4**) [15, 16]. Even though ceramide diversity contributes to the heterogeneity of GSLs, classification is mainly based on its glycosylation. On the basis of the first sugar that is attached to the ceramide, a galactosylceramide (GalCer, with a β-linked Gal) or a glucosylceramide GlcCer, with a β-linked Glc are formed. Notably, GalCer is rarely further elongated and is the major GSL class found in the brain where sialylated GalCer (Neu5Acα2-3GalβCer, GM4) is expressed as well (**Figure 4**) [17]. However, most large GSLs glycans are produced by elongation of GlcCer. The major core structures in vertebrates includes lacto-series (Galβ1-3GlcNAcβ1-3Galβ1-4Glcβ1-Cer, Lc4Cer), neolacto-series (Galβ1-4GlcNAcβ1-3Galβ1-4Glcβ1-Cer, nLc4Cer), ganglio-series (Galβ1-3GalNAcβ1-4Galβ1-4Glcβ1-Cer,Gg4Cer) and globo-series (GalNAcβ1-3Galα1-4Galβ1-4Glcβ1-Cer, Gb4Cer). Majority of sialylated GSLs belong to ganglio-series in which most species carry one or more sialic acids (**Figure 4**). In mammals, there are two sialylated forms found on *N*-, *O*- and GSL glycans, namely the Neu5Ac and *N*-glycolylneuraminic acid (Neu5Gc) [17]. However, due to a mutation of the enzyme responsible for hydroxylating cytidine monophospho (CMP)-NeuAc to CMP-NeuGc, only Neu5Ac is found in humans [18]. However, the expression of Neu5Gc was detected in some disease such as acute myeloid leukemia (AML), which might result from the mutation of relevant genes in AML cell lines [19].

Expression of GSLs is related to the enzymes' expression and intracellular distribution for their synthesis and levels of nucleotide sugar donors used for biosynthesis of GSLs by GTs including uridine diphosphate (UDP)-Gal, UDP-Glc, UDP-GlcNAc, UDP-GalNAc and CMP-Neu5Ac. The competition between GTs for same substrates contributes to the heterogeneity and relative expression level of GSLs. GSLs are involved in certain biological processes including regulation of signal transduction and cell-cell recognition [21, 22].

1.5 Altered glycosylation in cancer

Abnormality in glycosylation has become a well-known hallmark of cancer [23]. Altered expression of glycans is involved in the development and progression of cancer with regarding to cell signaling and communication, invasion of tumor cells, cell-matrix interaction, tumor angiogenesis, immune modulation and metastasis

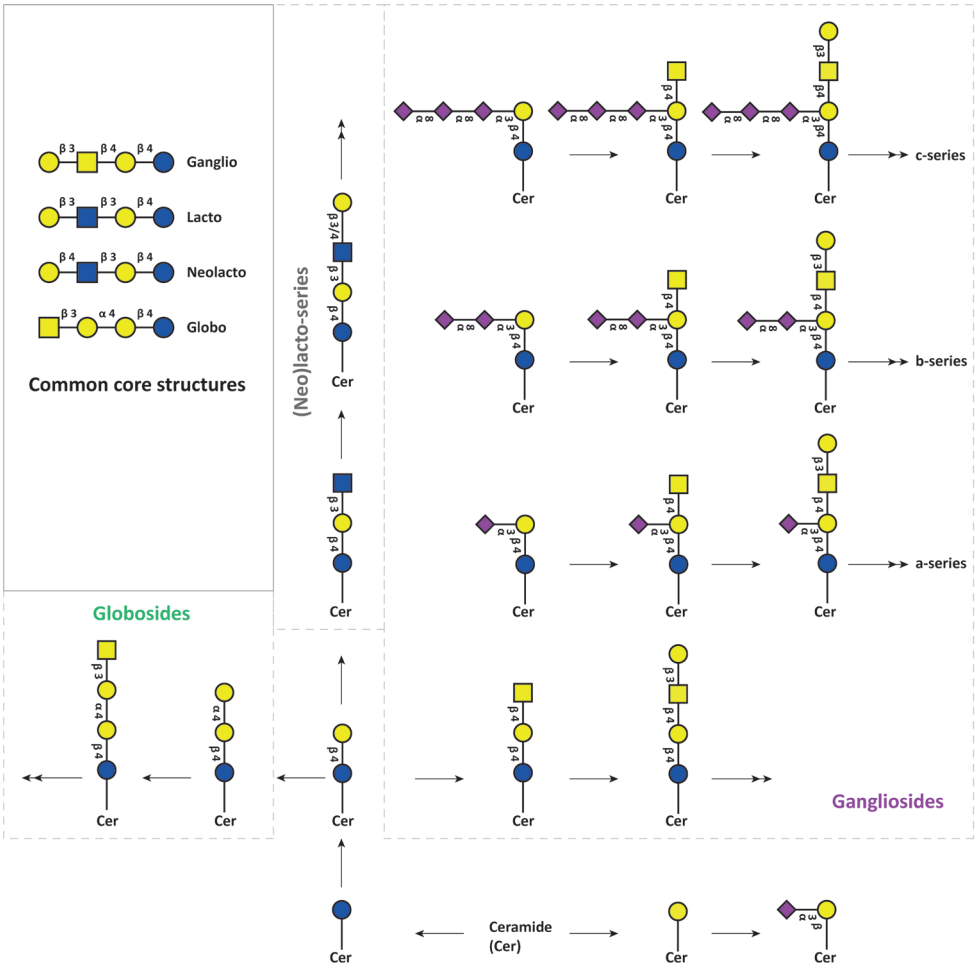

Figure 4. Four common of core structure of GSL glycans and relevant biosynthetic pathway. Adjusted with permission from *Zhang, T. et al* [20].

formation [24]. Tumor-associated changes in glycans may be due to both incomplete synthesis and neo-synthesis processes [25]. Incomplete synthesis commonly occurs in the early stages of cancer leading to the production of truncated glycans, for instance, the expression of sTn in gastrointestinal cancer [26]. Neo-synthesis often takes place in the advanced stages of cancer such as *de novo* expression of sialyl Lewis A (sLe^A) antigen in cancers [27]. Moreover, previous studies suggested that abnormal glycosylation results from the dysregulation of the glycan biosynthetic pathway at different levels consisting of epigenetics, transcription, post-transcription, expression of important chaperones, changes in glycosidase activity, abundance and

availability of the sugar nucleotide donors and altered sugar nucleotide transporter activity [28, 29]. To gain further insights into the potential regulatory mechanism of glycans, and to discover tumor-associated glycans for novel targeted treatment of disease including cancer, further investigations are needed that explore the glycosylation profile of biological samples such as cell lines, tissues, plasma and serum.

1.6 Colorectal cancer

Colorectal cancer (CRC) ranks the third most commonly diagnosed cancer (10% incidence) and the second leading cause of cancer death (9.4% mortality) worldwide [30]. An elevated intake of animal-source foods and a more desk-bound lifestyle results in reduced physical activity and an enhanced prevalence of excess body weight which contributes to CRC risk [31]. Other risk factors include smoking, alcohol consumption, whereas sufficient consumption of whole grains, fiber and dairy products as well as intake of calcium supplements contribute to a decreased risk [32]. Albeit population screening is a popular applied method and traditional therapeutic treatments for CRC -consisting of chemotherapy, radiation therapy and surgery- are used for the patient, the majority of CRC cases are still detected at an advanced stage which often results in an ineffective treatment [30]. As stated earlier, altered glycosylation has been considered a hallmark of cancer [33], which is not surprising as glycans play essential roles in several aspects of cancer cell biology consisting of cellular adhesion, cellular signaling and proliferation, extracellular matrix interactions and proximal and distal communication [23, 34, 35], which underlie critical cancer hallmarks including invasion, angiogenesis, metastasis and development of cancer features such as the modulation of the immune response [23, 34, 35].

Sialyl Lewis$^{A/X}$ (sLe$^{A/X}$) antigens play a functional role in the adhesion, invasion and metastasis of CRC *via* their interaction with E-selectin [36] Moreover, their expression was highly increased in liver metastases of CRC which has potential as a prognostic factor for patients at an advanced stage of CRC [37-39] and was found to be associated with poor survival in patients with CRC [40]. Elevated concentrations of carbohydrate antigen 19-9 (CA19-9) (also known as sLe^{A} tumor antigen) is related with a worse prognosis for patients at stage IV CRC and is just as a predictive tool

for prognosis in combination with the measurement of carcinoembryonic antigen (CEA; another tumor marker related to CRC) [41]. Altered glycosylation of integrinsin CRC cells may contribute to a more invasive or metastatic phenotype in CRC [42]. Significantly increased *N*-glycan-specific α2-6 sialyltransferase was revealed in metastasizing CRC and increased Galβ1-3GalNAc-specific sialyltransferase was found at early stages of CRC. Meanwhile, the expression of α2-6 sialylation of glycoproteins was upregulated during carcinogenesis of CRC [43]. Previous studies reported that β-galactoside α2-6 sialyltransferase (ST6GAL1) is positively associated with CRC invasion and migration [44, 45]. Overexpression of β1-4 galactosyltransferase IV (B4GALT4) was involved in poor prognosis and metastasis of CRC [46]. However, another enzyme β1-4 galactosyltransferase III (B4GALT3), responsible for poly-*N*-acetyllactosamine synthesis, is negatively associated with CRC invasiveness and metastasis [47]. Interestingly, β1-3 *N*-acetylglucosaminyltransferase 8 (B3GNT8) was overexpressed in CRC and promotes CRC invasion, which is responsible for the biosynthesis of poly-*N*-acetyllactosamine on β1-6 branched *N*-glycans upregulated in CRC and associated with progression and metastasis of CRC [47-49]. A decrease of bisecting *N*-glycans was found in CRC [50]. The aforementioned findings demonstrate the potential of developing targeting tumor associated carbohydrate antigens (TACAs) which would start a new era of cancer therapy [51, 52].

1.7 Acute myeloid leukemia

Acute myeloid leukemia (AML) is a malignancy of the stem cell precursors of the myeloid lineage covering platelets, red blood cells and white blood cells, which results from genetic alteration (such as gene mutation and chromosomal rearrangement) leading to neoplastic changes and clonal proliferation [53]. The traditional therapeutic principles include the standard regime of multiple courses of induction chemotherapy, allogenic hematopoietic stem cell transplantation and palliative care [54]. Advanced genomics technologies provide insight into abnormalities at the gene level. The balanced chromosomal rearrangements produce in-frame chimeric fusion genes and recurrently target genes responsible for hematopoietic transcription factors (such as RUNX1), epigenetic regulators and components of the nuclear pore complex [55]. The generation of chimeric fusion genes due to translocation or insertion evens in the hematopoietic stem/progenitor cell and is the initial step in the pathogenesis of a significant proportion

of AML in younger adults [55-60]. The mutation in thereceptor tyrosine kinase fms-like tyrosine kinases 3(FLT3) plays an essential role in normal hematopoiesis and was found in a third of the AML cases that consisted of in-frame duplication (*FLT3-ITD*) and point mutation (*FLT3-TKD*) [61]. Gene CCAAT/enhancer binding protein a (*CEBPA*) encoding a myeloid transcription factor was mutated in approximately 10% of the AML cases [62]. The most common molecular lesion in AML is the mutation of nucleophosmin NPM1 which can promote self-renewal of hematopoietic progenitors associating with expanded myelopoiesis and leading to development of AML [63]. *RUNX1* is a common somatic target in AML which results in familial platelet disorder predisposing to AML [64]. Moreover, other genes that are responsible for encoding signaling pathway components are *RAS*, *Ckit*, *CBL* and *NF1*, and epigenetic regulators such as *DNMT3A*, *TET2*, *CREBBP/KAT3A* and *EZH2*), as well as components of the splicing machinery. *BCOR* and *KMT2D* are found to be recurrently mutated in AML [55, 65, 66]. Owing to the heterogeneity of AML, classification seems important for selection of treatment, assessment of risk and stratification [67]. Based on the morphological appearance and cytochemical characteristics, AML has been classified into eight groups (M0-M7) according to the French-American-British (FAB) classification [68].

The interaction of $sLe^{A/X}$ antigen with E-selectin in bone marrow niche are associated with the homing of leukemia cells and chemoresistance in AML [69, 70]. Another report demonstrated that the aberrant cell surface sialylation patterns is involved in the development of multidrug resistance (MDR) of AML cells by relevant GTs ST3GAL5 and SI8SIA4 regulating the activity of *PI3K/Akt* signaling and the expression of P-gp and MRP1 [71]. Besides, the abnormality of *N*-glycosylation (e.g. CD79a, CD82, CD95 and MPL) can also affect the leukemia cell functions [24]. CD82 modulates adhesion and bone marrow homing of AML cells [72]. Inhibition of CD82 glycosylation enhances the bone marrow homing of AML cells and the molecular packing of *N*-cadherin [72]. Thus, exploration of the glycosylation features of AML cells involved in the homing process to the bone marrow might be relevant for future studies. Several GSLs are considered as markers of hematological malignancies, for instance, Gg3 (GalNAcβ1-4Galβ1-4Glcβ1-ceramide) in Hodgkin lymphoma and Gb3 (Galα1-4Galβ1-4Glcβ1-ceramide) for Burkitt lymphoma [73, 74]. Thus investigation of phenotype of AML cells such as glycosylation could be an approach towards the development of new treatments of AML.

1.8 Glycosylation analysis

To gain a complete characterization of glycosylation, various analytical techniques are required in order to elucidate the diversity of glycans and to further study their role in disease development and progression. Heterogeneity of glycans are attributed to the composition and sequence of monosaccharides, linkage information and anomeric configuration, leading to the presence of isomers. To reduce the complexity of glycans and achieve isomeric separation, capillary electrophoresis (CE) and liquid chromatography (LC) separation techniques can be used, such as reversed phase (RP) chromatography, hydrophilic interaction liquid chromatography (HILIC) and porous graphitized carbon (PGC). As a detection method, fluorescence and UV detection are often applied for the analysis of complex glycans [75]. However, to prompt detection after chromatographic or electrophoretic separation, a derivatization of the glycans is required to introduce a chromatophore or fluorophore on glycans prior to separating the glycans by LC (LC-FLD (fluorescence detection)) and CE (CE-LIF (laser-induced fluorescence)), respectively [75]. In addition, these derivatizationstrategies can also enhance glycan separation on the hydrophobic stationary phases [76, 77], and to enhance ionization for mass-spectrometric detection [75]. Common labels for LC-MS measurements include 2-aminobenzamide (2-AB), 2-aminobenzoic acid (2-AA), procainamide (ProA), aminoxyTMT labeling. HILIC and PGC are able to separate isomeric and isobaric structures without the need of glycan labeling [77]. In regard to HILIC, analytes are retained on the hydrophilic stationary phase via hydrogen bonding, ionic interaction and dipole-dipole interactions [78]. In general, the larger the glycans the later they elute. Charges and steric properties of the glycan will also have an effect on the retention of the analytes, leading to the separation of isomers. To be noted, the specific limitation of HILIC is the precipitation of the polar glycans when an initial high amount of organic modifier is used, this can be avoided by application of PGC column [78]. Namely, separation by PGC is on the basis of hydrophobic and polar interactions and present superior separation capability both in linkage and positional isomers compared to other LC principles [79-82].

To gain more confidence in the structural assignments of the glycans, often mass spectrometry is applied as detection technique which provides higher sensitivity than fluorescence and more information such as fragment ions for characterization of structural glycans [75]. Especially hyphenated with a separation technique, such

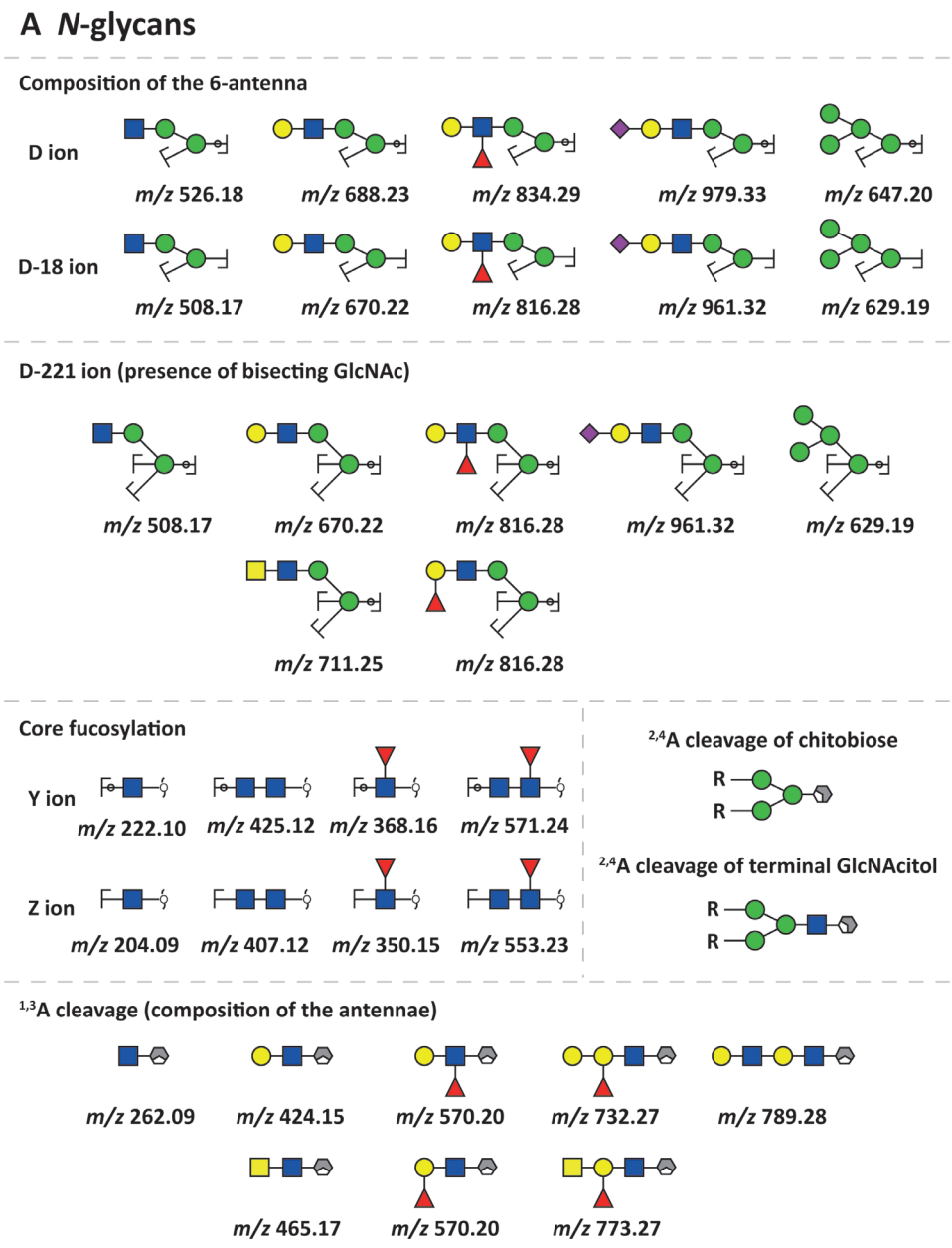

Figure 5. (A) Signature and feature ions in negative mode for glycan structural determination. Adjusted with permission from *Arun V. Everest-Dass et al* [83]. Copyright © 2013 American Chemical Society.

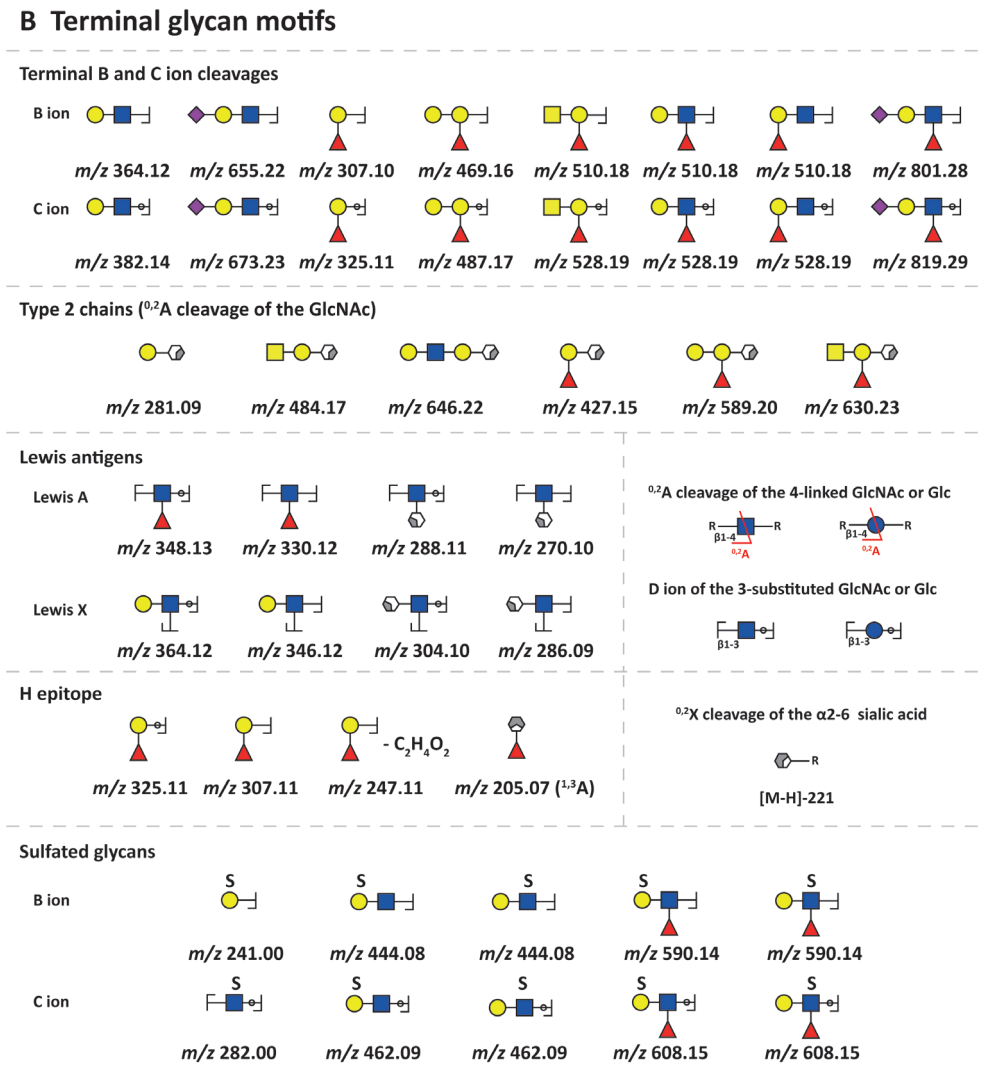

Figure 5. (B) Signature and feature ions in negative mode for glycan structural determination. Adjusted with permission from *Arun V. Everest-Dass et al* [83]. Copyright © 2013 American Chemical Society.

as LC, the composition, sequence and topology of the glycans can be identified [83]. However, to enable analysis by an MS, the analytes in solution need to beconverted into gas-phase ions. The most common ionization methods are matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI). MALDI-MS, generally with time of flight analyzer (TOF), is also known as a “soft” ionization technique as the chemical bonds of analytes do not break during MALDI

ionization [84] and this is a simple technique for rapid glycan profiling with a high mass [83]. However, to prevent the loss of sialic acids during MALDI ionization in positive mode a derivatization of the acidic groups of the sialylated glycans needs to be performed [85]. Stabilization and neutralization of sialic acids can be made by permethylation or by specific carboxyl group derivatization, which, next to stabilization, also improves the sensitivity of mass spectrometry in positive mode and provides more information in MS/MS for structural elucidation [75] but normally requires puresamples, long reaction times and strict conditions. To overcome these obstacles, a robust, rapid and high-throughput sialic acid stabilization was established, linkage-specific sialic acid esterification [86]. In ESI-MS, samples can be infused directly by either an off-line ESI-source or an on-line ESI-source coupled to LC or CE. In order to gain structural identification, fragmentation information is needed. To the best of our knowledge, more fragment ions are produced via MS/MS collision induced dissociation (CID) of negatively charged ions for the structural identification of glycans. Under the negative mode, the obtained series of C-type fragment ions provide sequence information, but also partial linkage information is obtained that aid in distinguishing isomeric structures, for instance, the double D-type cleavages were dominated in MS/MS for 3-linked GlcNAc or Glc and cross-ring fragmentation of ^{0,2}A-type for 4-linked GlcNAc or Glc give partial linkage information, which are essential for differentiation of type 1 and 2 chains and to confirm blood group H, Le^A, Le^X, Le^B and Le^Y determinants [87], which have been summarized in **Figure 5** [83]. Next to retention time on PGC column and diagnostic ions produced in MS/MS, exoglycosidase digestion experiments as well as what is known about the biosynthetic pathway of glycans contribute to the identification of glycans obtained from a biological source. In the thesis, we mainly employed PGC-nanoLC-ESI-MS/MS in negative mode to investigate the glycomic layer of cancer cell lines and tissues.

1.9 Scope of the thesis

While glycosylation is known to be relevant for malignant processes in both AML and CRC, it is poorly explored, both at the tissue and cancer cell line level. Especially, at an in-depth level where glycans are structural identified and glycan isomers can be differentiated from each other. Also, the association of glycosylation with transcriptomics is rarely explored. This thesis addresses these shortcomings by

exploring the cancer and cancer cell line glycome using PGC-nanoLC-ESI-MS and linking it to transcriptomics as well as clinical data, providing new insights into the glycobiology of AML and CRC.

In **Chapter 1** the general concept of glycosylation is introduced. Especially, altered glycosylation in different cancers is described with a main focus on AML and CRC. Additionally, an overview is provided of various analytical methods that can be utilized for the characterization of glycans.

To gain a deeper understanding of the role of glycosylation in CRC, various glycan classes in CRC cell lines and tissues were explored (**Chapters 2-5**). **Chapter 2** particularly focused on the in-depth analysis of the total *N*-glycome of 25 CRC cell lines. A total of 139 *N*-glycans could be identified on the basis of retention time on PGC, diagnostic ions in MS/MS spectra, the biosynthetic pathway and digestion experiments. After exploring the *N*-glycome of CRC cell lines, we aimed to investigate the GSL glycome (**Chapter 3**). Here, we explored the expression and regulation of GSLs glycans, revealing a high diversity of GSL glycans which reflected the CRC cellular differentiation phenotype. To gain insights in the regulation of GSL glycans, the correlation of expression of GSLs glycans with corresponding glycosyltransferases (GTs) and transcription factors were examined. As the terminal glycan structure generally plays an essential role in development and progression of cancer, and are common features within the different glycan classes, we integrated the obtained *N*-, *O*- (previous study) and GSLs glycomic data of the CRC cell lines (**Chapter 4**). This provided a clear overview of how common structures are distributed and provided insights in their potential regulation in CRC cell lines, by correlating the glycome with relevant GTs and TFs as well as the correlation of GTs and TFs in CRC tissues. In **Chapter 5**, we further extended our study by exploring in the *N*-glycome of CRC tissues and paired colon mucosa samples (controls). A striking diversity was found of the *N*-glycome between CRC, stroma and normal colon mucosa by using laser capture microdissection and analysis by PGC-nanoLC-ESI-MS.

Chapters 6 and **7** explore the glycome of various glycan classes in relation to AML. In order to gain better insights into the expression and regulation of GSL glycans in AML, GSLs glycosylation was investigated in 19 AML cell lines by PGC-nanoLC-ESI-MS/MS (**Chapter 6**). Interestingly, differences in expression of GSL glycans were found that could be related to the differentiation of AML classes. In **Chapter 7**,

integration of *N*-, *O*- and GSLs glycomic data of AML cell lines was performed. This provided insights in the association between glycomic and transcriptomic data and the possible regulation of glycans in AML.

A general discussion of the work described in the previous chapters is provided in **Chapter 8**, providing insights in remaining challenges in regard to glycan analysis, how to investigate the regulation of glycans in development of CRC and AML as well as perspectives on the application of tumor-associated glycans as new targets for novel therapy.

1.10 References

1. Colley KJ, V.A., Haltiwanger RS, et al. Cellular Organization of Glycosylation. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 4th edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022. Chapter 4. Available from: https://www.ncbi.nlm.nih.gov/books/NBK579926/ doi: 10.1101/glycobiology.4e.4.
2. Varki A, K.S.H.B.a.O.I.V.A., Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 4th edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022. Chapter 1. Available from: https://www.ncbi.nlm.nih.gov/books/NBK579927/ doi: 10.1101/glycobiology.4e.1.
3. Seeberger PH. Monosaccharide Diversity. In: Varki A, C.R., Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 4th edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022. Chapter 2. Available from: https://www.ncbi.nlm.nih.gov/books/NBK579981/ doi: 10.1101/glycobiology.4e.2.
4. McNaught, A.D., *International Union of Pure and Applied Chemistry and International Union of Biochemistry and Molecular Biology. Joint Commission on Biochemical Nomenclature. Nomenclature of carbohydrates.* Carbohydr Res, 1997. **297**(1): p. 1-92.
5. Seeberger PH. Monosaccharide Diversity. In: Varki A, C.R., Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 4th edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022. Chapter 2. Available from: https://www.ncbi.nlm.nih.gov/books/NBK579981/ doi: 10.1101/glycobiology.4e.2.
6. Stanley P, S.H., Taniguchi N. N-Glycans. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology. 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009. Chapter 8. Available from:https://www.ncbi.nlm.nih.gov/books/NBK1917/.
7. Komath SS, F.M., Hart GW, et al. Glycosylphosphatidylinositol Anchors. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 4th edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022. Chapter 12. Available from: https://www.ncbi.nlm.nih.gov/books/NBK579963/

doi: 10.1101/glycobiology.4e.12.
8. Varki, A., *Biological roles of oligosaccharides: all of the theories are correct.* Glycobiology, 1993. **3**(2): p. 97-130.
9. Wormald, M.R. and R.A. Dwek, *Glycoproteins: glycan presentation and protein-fold stability.* Structure, 1999. **7**(7): p. R155-60.
10. Van Breedam, W., et al., *Bitter-sweet symphony: glycan-lectin interactions in virus biology.* FEMS Microbiol Rev, 2014. **38**(4): p. 598-632.
11. Stanley P, M.K., Lewis NE, et al. N-Glycans. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 4th edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022. Chapter 9. Available from: https://www.ncbi.nlm.nih.gov/books/NBK579964/ doi: 10.1101/glycobiology.4e.9.
12. Kornfeld, R. and S. Kornfeld, *Assembly of asparagine-linked oligosaccharides.* Annu Rev Biochem, 1985. **54**: p. 631-64.
13. Brockhausen I, S.H., Stanley P. O-GalNAc Glycans. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology. 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009. Chapter 9. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1896/.
14. Brockhausen I, W.H., Hagen KGT, et al. O-GalNAc Glycans. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 4th edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022. Chapter 10. Available from: https://www.ncbi.nlm.nih.gov/books/NBK579921/ doi: 10.1101/glycobiology.4e.10
15. Karlsson, H., A. Halim, and S. Teneberg, *Differentiation of glycosphingolipid-derived glycan structural isomers by liquid chromatography/mass spectrometry.* Glycobiology, 2010. **20**(9): p. 1103-16.
16. Schnaar RL, S.A., Stanley P. Glycosphingolipids. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology. 2nd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009. Chapter 10. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1909/.
17. Schnaar RL, S.R., Tiemeyer M, et al. Glycosphingolipids. In: Varki A, Cummings RD, Esko JD, et al., editors. Essentials of Glycobiology [Internet]. 4th edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2022. Chapter 11. Available from: https://www.ncbi.nlm.nih.gov/books/NBK579905/ doi: 10.1101/glycobiology.4e.11.
18. Irie, A., et al., *The molecular basis for the absence of N-glycolylneuraminic acid in humans.* J Biol Chem, 1998. **273**(25): p. 15866-71.
19. Wang, D., et al., *Glycosphingolipid-glycan signatures of acute myeloid leukemia cell lines reflect hematopoietic differentiation.* J Proteome Res, 2022. **21**(4): p. 1029-1040.
20. Zhang, T., et al., *The Role of Glycosphingolipids in Immune Cell Functions.* Front Immunol, 2019. **10**: p. 90.
21. D'Angelo, G., et al., *Glycosphingolipids: synthesis and functions.* FEBS J, 2013.

280(24): p. 6338-53.
22. Hakomori, S., *Structure, organization, and function of glycosphingolipids in membrane.* Curr. Opin. Hematol. , 2003. **10**: p. 16–24.
23. Pinho, S.S. and C.A. Reis, *Glycosylation in cancer: mechanisms and clinical implications.* Nat Rev Cancer, 2015. **15**(9): p. 540-55.
24. Pang, X., et al., *Multiple Roles of Glycans in Hematological Malignancies.* Front Oncol, 2018. **8**: p. 364.
25. Hakomori S, K.R., *Glycosphingolipids as tumor-associated and differentiation markers* J Natl Cancer Inst., 1983. **71**: p. 231–51.
26. Marcos, N.T., et al., *ST6GalNAc-I controls expression of sialyl-Tn antigen in gastrointestinal tissues.* Front Biosci (Elite Ed), 2011. **3**(4): p. 1443-55.
27. Kannagi, R., et al., *Current relevance of incomplete synthesis and neo-synthesis for cancer-associated alteration of carbohydrate determinants--Hakomori's concepts revisited.* Biochim Biophys Acta, 2008. **1780**(3): p. 525-31.
28. Stowell, S.R., T. Ju, and R.D. Cummings, *Protein glycosylation in cancer.* Annu Rev Pathol, 2015. **10**: p. 473-510.
29. Dall'Olio, F., et al., *Mechanisms of cancer-associated glycosylation changes.* FBL, 2012. **17**(2): p. 670-699.
30. 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, 2021. **71**(3): p. 209-249.
31. Siegel, R.L., et al., *Colorectal cancer statistics, 2020.* CA Cancer J Clin, 2020. **70**(3): p. 145-164.
32. World Cancer Research Fund/American Institute for Cancer Research. The Continuous Update Project Expert Report 2018. Diet, N., Physical Activity and Cancer: Colorectal Cancer. Accessed October 23, 2020. wcrf.org/sites/ defau lt/ files/ Color ectal -cance r-report.pdf, .
33. Mereiter, S., et al., *Glycosylation in the Era of Cancer-Targeted Therapy: Where Are We Heading?* Cancer Cell, 2019. **36**(1): p. 6-16.
34. Lau, K.S., et al., *Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation.* Cell, 2007. **129**(1): p. 123-34.
35. RodrÍguez, E., S.T.T. Schetters, and Y. van Kooyk, *The tumour glyco-code as a novel immune checkpoint for immunotherapy.* Nature Reviews Immunology, 2018. **18**(3): p. 204-211.
36. Ugorski, M. and A. Laskowska, *Sialyl Lewis(a): a tumor-associated carbohydrate antigen involved in adhesion and metastatic potential of cancer cells.* Acta Biochim Pol, 2002. **49**(2): p. 303-11.
37. Nakayama T, W.M., Katsumata T, Teramoto T, Kitajima M. , *Expression of sialylLewis(a) as a new prognostic factor for patients with advanced colorectal carcinoma.* Cancer, 1995 Apr 15. **75(8):2051-6**
38. Yamada, N., et al., *Increased expression of sialyl Lewis A and sialyl Lewis X in liver metastases of human colorectal carcinoma.* Invasion Metastasis, 1995. **15**(3-4): p. 95-102.

39. Schiffmann, L., et al., *A novel sialyl Le(X) expression score as a potential prognostic tool in colorectal cancer.* World J Surg Oncol, 2012. **10**: p. 95.
40. 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, 1993. **53**(15): p. 3632-7.
41. Kawamura, H., et al., *Prognostic Role of Carcinoembryonic Antigen and Carbohydrate Antigen 19-9 in Stage IV Colorectal Cancer.* Anticancer Res, 2022. **42**(8): p. 3921-3928.
42. von Lampe, B., A. Stallmach, and E.O. Riecken, *Altered glycosylation of integrin adhesion molecules in colorectal cancer cells and decreased adhesion to the extracellular matrix.* Gut, 1993. **34**(6): p. 829-36.
43. Kemmner, W., D. Kruck, and P. Schlag, *Different sialyltransferase activities in human colorectal carcinoma cells from surgical specimens detected by specific glycoprotein and glycolipid acceptors.* Clin Exp Metastasis, 1994. **12**(3): p. 245-54.
44. Park, J.J., et al., *Sialylation of epidermal growth factor receptor regulates receptor activity and chemosensitivity to gefitinib in colon cancer cells.* Biochem Pharmacol, 2012. **83**(7): p. 849-57.
45. Park, J.J. and M. Lee, *Increasing the alpha 2, 6 sialylation of glycoproteins may contribute to metastatic spread and therapeutic resistance in colorectal cancer.* Gut Liver, 2013. **7**(6): p. 629-41.
46. Chen, W.S., et al., *Tumor beta-1,4-galactosyltransferase IV overexpression is closely associated with colorectal cancer metastasis and poor prognosis.* Clin Cancer Res, 2005. **11**(24 Pt 1): p. 8615-22.
47. Chen, C.H., et al., *beta-1,4-Galactosyltransferase III suppresses beta1 integrin-mediated invasive phenotypes and negatively correlates with metastasis in colorectal cancer.* Carcinogenesis, 2014. **35**(6): p. 1258-66.
48. Ishida, H., et al., *A novel beta1,3-N-acetylglucosaminyltransferase (beta3Gn-T8), which synthesizes poly-N-acetyllactosamine, is dramatically upregulated in colon cancer.* FEBS Lett, 2005. **579**(1): p. 71-8.
49. Jiang, Z., et al., *beta3GnT8 Promotes Colorectal Cancer Cells Invasion via CD147/MMP2/Galectin3 Axis.* Front Physiol, 2018. **9**: p. 588.
50. Zhang, D., et al., *Mass spectrometry analysis reveals aberrant N-glycans in colorectal cancer tissues.* Glycobiology, 2019. **29**(5): p. 372-384.
51. Costa, A.F., et al., *Targeting Glycosylation: A New Road for Cancer Drug Discovery.* Trends Cancer, 2020. **6**(9): p. 757-766.
52. Nativi, C., F. Papi, and S. Roelens, *Tn antigen analogues: the synthetic way to "upgrade" an attracting tumour associated carbohydrate antigen (TACA).* Chem Commun (Camb), 2019. **55**(54): p. 7729-7736.
53. Pelcovits A, N.R., *Acute Myeloid Leukemia: A Review.* R I Med J, 2020 Apr 1. **103(3)**: p. 38-40.
54. Roboz, G.J., *Novel approaches to the treatment of acute myeloid leukemia.* Hematology Am Soc Hematol Educ Program, 2011. **2011**: p. 43-50.

55. Grimwade, D., A. Ivey, and B.J. Huntly, *Molecular landscape of acute myeloid leukemia in younger adults and its clinical relevance.* Blood, 2016. **127**(1): p. 29-41.
56. Rucker, F.G., et al., *TP53 alterations in acute myeloid leukemia with complex karyotype correlate with specific copy number alterations, monosomal karyotype, and dismal outcome.* Blood, 2012. **119**(9): p. 2114-21.
57. Lindsley, R.C., et al., *Acute myeloid leukemia ontogeny is defined by distinct somatic mutations.* Blood, 2015. **125**(9): p. 1367-76.
58. Micol, J.B., et al., *Frequent ASXL2 mutations in acute myeloid leukemia patients with t(8;21)/RUNX1-RUNX1T1 chromosomal translocations.* Blood, 2014. **124**(9): p. 1445-9.
59. Groschel, S., et al., *Mutational spectrum of myeloid malignancies with inv(3)/t(3;3) reveals a predominant involvement of RAS/RTK signaling pathways.* Blood, 2015. **125**(1): p. 133-9.
60. Green, C.L., et al., *Prognostic significance of CEBPA mutations in a large cohort of younger adult patients with acute myeloid leukemia: impact of double CEBPA mutations and the interaction with FLT3 and NPM1 mutations.* J Clin Oncol, 2010. **28**(16): p. 2739-47.
61. Levis, M., *FLT3 mutations in acute myeloid leukemia: what is the best approach in 2013?* Hematology, 2013. **2013**(1): p. 220-226.
62. Pabst, T., et al., *Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein-α (C/EBPα), in acute myeloid leukemia.* Nature Genetics, 2001. **27**(3): p. 263-270.
63. Vassiliou, G.S., et al., *Mutant nucleophosmin and cooperating pathways drive leukemia initiation and progression in mice.* Nat Genet, 2011. **43**(5): p. 470-5.
64. Lam, K. and D.E. Zhang, *RUNX1 and RUNX1-ETO: roles in hematopoiesis and leukemogenesis.* Front Biosci (Landmark Ed), 2012. **17**(3): p. 1120-39.
65. Wouters, B.J. and R. Delwel, *Epigenetics and approaches to targeted epigenetic therapy in acute myeloid leukemia.* Blood, 2016. **127**(1): p. 42-52.
66. Shirai, C.L., et al., *Mutant U2AF1 Expression Alters Hematopoiesis and Pre- mRNA Splicing In Vivo.* Cancer Cell, 2015. **27**(5): p. 631-43.
67. Dohner, H., D.J. Weisdorf, and C.D. Bloomfield, *Acute Myeloid Leukemia.* N Engl J Med, 2015. **373**(12): p. 1136-52.
68. Bennett, J.M., et al., *Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group.* Br J Haematol, 1976. **33**(4): p. 451-8.
69. Krause, D.S., et al., *Selectins and their ligands are required for homing and engraftment of BCR-ABL1+ leukemic stem cells in the bone marrow niche.* Blood, 2014. **123**(9): p. 1361-71.
70. Barbier, V., et al., *Endothelial E-selectin inhibition improves acute myeloid leukaemia therapy by disrupting vascular niche-mediated chemoresistance.* Nat Commun, 2020. **11**(1): p. 2042.
71. Ma, H., et al., *Modification of sialylation is associated with multidrug resistance*

in human acute myeloid leukemia. Oncogene, 2015. **34**(6): p. 726-740.
72. Marjon, K.D., et al., *Tetraspanin CD82 regulates bone marrow homing of acute myeloid leukemia by modulating the molecular organization of N-cadherin.* Oncogene, 2016. **35**(31): p. 4132-40.
73. Kniep, B., et al., *Gangliotriaosylceramide (asialo GM2), a glycosphingolipid marker for cell lines derived from patients with Hodgkin's disease.* The Journal of Immunology, 1983. **131**(3): p. 1591-1594.
74. Mangeney, M., et al., *Apoptosis Induced in Burkitt's Lymphoma Cells via Gb3/CD77, a Glycolipid Antigen1.* Cancer Research, 1993. **53**(21): p. 5314-5319.
75. Ruhaak, L.R., et al., *Glycan labeling strategies and their use in identification and quantification.* Anal Bioanal Chem, 2010. **397**(8): p. 3457-81.
76. Zhou, S., et al., *Direct comparison of derivatization strategies for LC-MS/MS analysis of N-glycans.* Analyst, 2017. **142**(23): p. 4446-4455.
77. Vreeker, G.C. and M. Wuhrer, *Reversed-phase separation methods for glycan analysis.* Anal Bioanal Chem, 2017. **409**(2): p. 359-378.
78. Wuhrer, M., A.R. de Boer, and A.M. Deelder, *Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry.* Mass Spectrom Rev, 2009. **28**(2): p. 192-206.
79. Ruhaak, L.R., A.M. Deelder, and M. Wuhrer, *Oligosaccharide analysis by graphitized carbon liquid chromatography-mass spectrometry.* Anal Bioanal Chem, 2009. **394**(1): p. 163-74.
80. Fan, J.Q., et al., *High-Performance Liquid Chromatography of Glycopeptides and Oligosaccharides on Graphitized Carbon Columns.* Analytical Biochemistry, 1994. **219**(2): p. 224-229.
81. Pabst, M., et al., *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, 2007. **79**(13): p. 5051-7.
82. Jensen, P.H., et al., *Structural analysis of N- and O-glycans released from glycoproteins.* Nat Protoc, 2012. **7**(7): p. 1299-310.
83. Everest-Dass, A.V., et al., *Structural feature ions for distinguishing N- and O-linked glycan isomers by LC-ESI-IT MS/MS.* J Am Soc Mass Spectrom, 2013. **24**(6): p. 895-906.
84. Ng, E.W., M.Y. Wong, and T.C. Poon, *Advances in MALDI mass spectrometry in clinical diagnostic applications.* Top Curr Chem, 2014. **336**: p. 139-75.
85. Bielik, A.M. and J. Zaia, *Historical overview of glycoanalysis.* Methods Mol Biol, 2010. **600**: p. 9-30.
86. Reiding, K.R., et al., *High-throughput profiling of protein N-glycosylation by MALDI-TOF-MS employing linkage-specific sialic acid esterification.* Anal Chem, 2014. **86**(12): p. 5784-93.
87. Chai, W., V. Piskarev, and A.M. Lawson, *Negative-ion electrospray mass spectrometry of neutral underivatized oligosaccharides.* Anal Chem, 2001. **73**(3): p. 651-7.