Terminal Galactose as Cancer Recognition Marker: Computing Analysis With Implications of Vicinal Sugars, Linkage and Anomery

AFSHAN KALEEM, AHMED KHURSHID, ISHTIAQ AHMAD, EVELYNE WALKER-NASIR, SADIA KHAN, M. IQBAL CHOUDHARY, A. R. SHAKOORI AND NASIR-UD-DIN*

Institute of Molecular Sciences and Bioinformatics, Lahore, Pakistan (AK, AK, IA, EWN, ND), HEJ Research Institute of Chemistry, University of Karachi, Karachi, Pakistan (SK, MIC, ND) and School of Biological Sciences, University of the Punjab, New Campus, Lahore, Pakistan (ARS)

Abstract.- Carbohydrate structures on proteins play vital role as recognition markers in several diseases including cancer. The carbohydrate chains are dramatically altered in cancer cells compared to normal cells, both in structure and quantity. The mucin O-glycans show several cancer associated structures like T- (Gal-GalNAc-Ser/Thr) and Tn- antigen (GalNAc-Ser/Thr). Terminal galactose (Gal) in mucin type O-glycan oligosaccharide structures is known to act as a recognition marker for several cancer-associated lectins like galectin and mistletoe lectins. This study describes the role of terminal Gal, inclusive of its anomery, linkage and that of sugar residues vicinal to terminal Gal in oligosaccharide structures of glycoproteins, as an epitope or as a recognition marker in cancer.

Key words: O-linked terminal galactose, cancer recognition marker.

INTRODUCTION

Glycoproteins play key role in several biological processes such as embryogenesis (Schachter, 2002), immunity (Rudd et al., 2001) and are known to mediate several pathological conditions (Birkle et al., 2003). In nature more than half of eukaryotic proteins are glycosylated, and around 90% of these glycoproteins are N-linked (Apweiler et al., 1999). Different structures of sugar chains are found on glycoproteins. The contribution of these structures in health and disease are largely undefined. The glycosylation process produces a substantial multiplicity of chemical structures, and makes large complex structures, in contrast to other cellular macromolecules such as proteins, DNA and RNA, which form linear chains.

The O-linked oligosaccharides on many glycoproteins are predominantly found on secreted and membrane bound mucins. Mucins are large extracellular proteins that are heavily glycosylated with complex oligosaccharides. Alterations in mucin expression or in the glycosylation accompany the development of cancer and influence cellular growth, differentiation, transformation, adhesion, invasion and immune surveillance. Mucins are used as diagnostic markers in cancer, and are continuously investigated as therapeutic targets for cancer. Oncogenic transformation is often associated with changes in glycosylation of glycoproteins and glycolipids in cell membranes. This leads to the incomplete glycosylation of the core carbohydrate structures resulting in the formation of pancarcinoma carbohydrate antigens: T/TF antigen (Galβ1-3GalNAc), Tn (GalNACαThr/Thr) and sialyl Tn antigens (NeuAcα2-6GalNAc) (Ref?). These structures show limited distribution in normal adult tissues (Springer, 1984).

Carbohydrate structures have been found to be major antigenic determinants (Feizi and Childs, 1987). Different sugar moieties have different functions in the terminal (or non-reducing) position of glycoproteins. Sialic acid and its derivates at the terminal positions of the oligosaccharides of glycoproteins determine the half-lives of many circulating glycoproteins and play a role in cellular behavior (Wang, 2005). GlcNAc has been shown to acts as a phagocytic marker during apoptosis (Orlando and Pittman, 2006). Fucosylated...
haptoglobin is a novel marker for pancreatic cancer (Okuyama et al., 2005).

Terminal sugar moieties act as recognition for lectin (major carbohydrate binding proteins). The European mistletoe lectins are Gal and GalNAc specific lectins, and has been widely used in cancer by inducing apoptosis of cancerous as well as normal cells (Park et al., 2001). Another β-galactoside binding lectin is the family of Galactians (Hernandez and Baum, 2002). These lectins play a key role in cell death, eukaryotic development, immune system homeostasis, and in control of tumorigenesis (Hernandez, and Baum, 2002). The liver is a dominant site for metastasis from primary colorectal tumors, and has been documented to be fatal (Rohlf et al., 1999). Hepatocytes express the liver-specific terminal Gal and GalNAc binding receptor ASGP-R that plays a major role in the clearance of desialylated proteins from the general circulation. This receptor is also found on human hepatoma and human colonic adenocarcinoma cells. When Gal is used as a blocking agent it prevents the settling of metastatic cells in the liver (Rohlf et al., 1999), suggesting that the sugar Gal also can play a protective role against cancer metastasis.

Terminal Gal in oligosaccharide chains have shown to have several functions. Removal of non-reducing Gal of the O-linked oligosaccharides of the glycoprotein zona pellucida 3 in mouse are essential for sperm–egg binding (Mengerink and Vacquier, 2001). Furthermore terminal Gal is also blood group B determinant in humans (Viitala et al., 1981). The protozoan parasite Trypanosoma brucei that causes African sleeping sickness in humans contains terminal Gal(α1-3)Gal moieties on its variant surface glycoprotein, which is highly immunogenic in humans (Pingel et al., 1999).

Terminal Gal is utilized as a marker in several diseases such as in bone resorption in serum of patients with Paget disease, where β-1-Gal is linked to hydroxylysines in collagen (Al-Dehaimi et al., 1999) and Gal-α(1,3)gal is a cell-associated epitope responsible for hyper acute rejection of porcine whole-organ xenografts in primates (McPherson et al., 2000).

Cancerous cells express glycoproteins with Gal on non-reducing end of oligosaccharide chain. Thus it is becoming more important to define these oligosaccharide structures to fight the exhausting battle against cancer. Sulfated terminal Gal moiety on colonic mucins expressed in cancer cells (Tsuiji et al., 1998), α-linked terminal Gal glycoproteins on Ehrlich ascites tumor cells (Takagaki et al., 1993), β-linked terminal Gal in breast cancer (Desai et al., 1995) and the very common disaccharide Gal-GalNAc, a precancerous marker (Shamsuddin et al., 1995; Baldus et al., 2000), are among the many non-reducing Gal containing oligosaccharides expressed in various types of cancer. Characterization of tumor markers are know becoming crucial for diagnosis of diseases like cancer, and vaccination in form of glycopeptides against specific glycoepitope like Tn/STn-MUC1, which has been shown to be able to override tolerance in human MUC1 transgenic mice and induce humoral immunity with high specificity for MUC1 cancer specific glycoforms (Tarp et al., 2007).

The current study was based on isolating data and searching function of specific oligosaccharide structures containing terminal Gal that are specific determinants as cancer marker. The terminal Gal recognition marker, epitopes, are usually short chain oligosaccharides in glycoproteins in cancer. The effects of anomery, linkages, and the vicinal sugar residues are significantly important for epitopes. Computational approaches are emerging in defining the role of sugar residues with multiple hydroxyl groups and complexities of different linkages and anomeric diversities.

MATERIALS AND METHODS

Data collection

Several carbohydrate databases with structural and functional information are available on the internet. The Consortium for Functional Glycomics (CFG) is a large research database that is developed to elucidate the roles of carbohydrate-protein interactions in cell communication at the cell surface. The scientific cores produce a variety of resources and services for investigators performing experiments that address the goals of the CFG. Information regarding glycan-binding proteins, glycan structures, and glycosyltransferases are also
available at CFG website: http://www.functionalglycomics.org/.

Terminal galactose data

We have collected all the structures containing terminal Gal from the CFG website. Total 7088 structures containing Gal glycans (%Gal%) were retrieved. Out of these 2834 oligosaccharide structures containing terminal Gal were extracted manually.

For the analysis of terminal Gal as a cancer and/or tumor markers 268 O-linked glycoproteins with terminal Gal oligosaccharide structures were found. For the purpose of analyzing the terminal Gal residue in glycoproteins for their association with cancer we gathered the information about vicinal sugars, their linkages and anomery to determine the requirement of terminal Gal as a cancer marker (Table I).

To elucidate the function of the terminal Gal containing proteins in cancer we searched the Glycoepitope database (http://www.glyco.is.ritsumei.ac.jp/epitope/). Glycoepitope is an integrated database containing information about carbohydrate antigens like glyco-epitopes and antibodies, and has been assembled as a compact encyclopedia. We found 9 cancer-associated structures containing terminal Gal in various form of cancer (Table II).

RESULTS AND DISCUSSION

The carbohydrate structures of cancer and tumor glycoproteins often become modified. Especially cancer-associated structures like the T/TF-antigen (Shamsuddin et al., 1995; Springer, 1984), Tn-antigen (Springer, 1984) and Sialyl-Tn antigen (Julien et al., 2001) etc. are over expressed in different types of cancers and tumors. Tumor markers are often found in the blood, urine or body tissues and are produced either by the tumor itself or by the body in response to the presence of cancer or certain benign (non cancerous) conditions. Several tumor marker are widely utilized for the diagnosis of different forms of cancer and tumors, like CA19-9 (Hanisch et al., 1984) that is used as prognostic indicator in advanced colorectal carcinoma (Kouri et al., 1992, Kouri et al., 1993) and CA125 (Wong et al., 2003), a protein that is over expressed up to 80% in ovarian cancer (Högberg and Kågedal, 1992). These tumor marker sugar residues are residing on glycoproteins and/or glycolipids with a high degree of glycosylation that may play a crucial role in promoting modulation of the immune response in patients with cancer.

In an effort to define the role of terminal Gal residues containing glycoproteins as specific cancer marker, we searched the CFG and Glycoepitope database and observed that several terminal Gal containing oligosaccharide structures are implicated as specific marker for different cancers and tumors (Table II).

From the database different oligosaccharide structures were retrieved having only non-reducing Gal (structure 1-3, Table II) and oligosaccharide structures with non-reducing Gal and Fuc (structure 4-7, 9, Table II) and non-reducing Gal and Neu5Ac (structure 8, Table I). Sulfated terminal Gal structures were found as well (structure 6-7, Table II). Sulfated terminal Gal structures (3'-Sulfo Lewis a and 3'-Sulfo Lewis x) have been detected in cancer cells as well as in surrounding nonmalignant epithelia in human colon cancer tissues (Yamori et al., 1989; Loveless et al., 1998; Izawa et al., 2000) and the 3'-sulfo-Lewis x epitope has been found to be a major carbohydrate motif in metastatic human colon carcinoma (Capon et al., 1997). These findings indicate that 3'-sulfated Lewis epitopes expression modulates tumor progression, in the case of human colon cancer. Another interesting oligosaccharide (structure 3) containing terminal α1-3 linked Gal was found to be a recognition marker on Ehrlich tumor cell membrane in BALB/c mice (Takagaki et al., 1993). The addition of this structure Gal α1-3 Gal is catalyzed by α1,3-galactosyl transferase that is only partially transcribed in human even though a complete gene is present in human (Lantéri et al., 2002), enabling a natural production of antibodies to α-Gal in humans (and old world monkeys as well as apes). The Lewis x blood group and the SSEA-1 determinant is normally found on glycolipids. Lewis x determinants and SSEA-1 expressed on morula stage embryos are known to act as a recognition marker and play a significant role during the course of embryogenesis (Eggens et al., 1989). Furthermore SSEA-1 together with SSEA-3 and
SSEA-4 is used as a marker for characterization of chicken primordial germ cells (Jung et al., 2005).

Table I.- The classification of 268 terminal Gal oligosaccharide structures related to its vicinal sugars, family and sub family, anomery and linkages.

<table>
<thead>
<tr>
<th>Terminal</th>
<th>Vicinal</th>
<th>Family</th>
<th>Sub Family</th>
<th>Anomery - Linkage</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core 1</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Core 2</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>Core 3</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Not Classified</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Core 1</td>
<td>68</td>
<td>68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core 2</td>
<td>48</td>
<td>1</td>
<td>4</td>
<td>9</td>
<td>49</td>
</tr>
<tr>
<td>Core 3</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Core 8</td>
<td>7</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core 9</td>
<td>1</td>
<td>1</td>
<td>32</td>
<td>2</td>
<td>38</td>
</tr>
<tr>
<td>Not Classified</td>
<td>3</td>
<td>1</td>
<td>32</td>
<td>2</td>
<td>38</td>
</tr>
<tr>
<td>Core 1</td>
<td>1</td>
<td>14</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core 2</td>
<td>4</td>
<td>18</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core 3</td>
<td>3</td>
<td>16</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core 4</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Core 6</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Not Classified</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>14</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>3</td>
<td>15</td>
<td>179</td>
<td>2</td>
</tr>
</tbody>
</table>

These structures have also been found on cancer associated cells on glycoproteins (Eggen et al., 1989; Ozawa et al., 1985). Both structures contain terminal Gal as well as terminal Fuc, so it is not possible to elucidate from the literature whether terminal Gal or Fuc or both terminal sugars act as an epitope.

The T/TF tumor marker Galβ1-3GalNAc is a well known marker in different cancers like that of colon, colorectal, gastric, liver, prostate and breast (Fig. 1). TF expression is highly specific for tumor cells and is virtually absent in healthy tissue, perhaps making this epitope an ideal tumor marker. The antigen behaves as an onco-fetal carcinoma-associated antigen and is expressed in neonatal colon and meconium (Picard and Feizi, 1983; Hounsell et al., 1985), but in normal adult epithelium it is thought to be masked by additional glycosylation (fucosylation and/or sialylation) (Lance and Lev, 1991; Campbell et al., 2001). Adenocarcinomas of most tissues (lung, breast, ovary, uterine, pancreas, stomach and liver (metastic adenocarcinoma) express Gal-GalNAc (Shamsuddin et al., 1995), suggesting that the Gal-GalNAc epitope is an ubiquitous marker of cancer. In Table I, there are 68 oligosaccharide structures with terminal Gal in core 1, 48 structures in core 2, 1 structure in core 9 and 32 structures that are not defined to any core structure category (a total of 149 oligosaccharide structures) with same anomery (β), linkage (1-3) and vicinal sugar (GalNAc) as the described T/TF-tumor marker. These defined 149 oligosaccharide structures are either directly α-linked to the protein or α/β-linked to a carbohydrate.
structure. Furthermore the Gal-(β1-3)-GalNAc cells, the O-glycosylation is prematurely terminated, structure found on tumor cells are not effected by vicinal oligosaccharide structure, because in tumor cells, the O-glycosylation is prematurely terminated, Table II.- Determination of O-linked terminal Gal containing oligosaccharide structures as cancer and/or tumor marker.

<table>
<thead>
<tr>
<th>Name of Antigen</th>
<th>Structure</th>
<th>Diseases &amp; Cell Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 T/TF Antigen</td>
<td>Galβ1-3GalNAcα1-Ser/Thr</td>
<td>Breast (nipple aspirate fluid) cancer, colon, colorectal cancer, esophageal cancer, liver cancer, ovary cancer, prostate cancer, stomach cancer</td>
</tr>
<tr>
<td>2 i-Antigen</td>
<td>Galβ1-4GlcNAcβ1-3Galβ1-4GlcNAcβ1-1-R</td>
<td>human leukaemic cells, teratocarcinomas of mouse</td>
</tr>
<tr>
<td>3 Alpha Galactosyl Epitope</td>
<td>Galα1-3Galβ1-4GlcNAc-R</td>
<td>Ehrlich tumor cell membrane</td>
</tr>
<tr>
<td>4 Lewis x</td>
<td>Galβ1 4 GlcNAcβ1-1-R Fuco1 3</td>
<td>acute leukemia lymphoma Renal cell carcinoma</td>
</tr>
<tr>
<td>5 SSEA-1</td>
<td>Galβ1 4 GlcNAcβ1-3Galβ1-4GlcNAcβ1-3Galβ1-1-R Fuco1 3</td>
<td>embryonal carcinoma cells</td>
</tr>
<tr>
<td>6 3'-Sulfo Lewis a</td>
<td>Fuco1 4 GlcNAcβ1-1-R HSO3-3Galβ1 3</td>
<td>colon cancer, colonic epithelia</td>
</tr>
<tr>
<td>7 3'-Sulfo Lewis x</td>
<td>HSO3-3Galβ1 4 GlcNAcβ1-1-R Fuco1 3</td>
<td>colon cancer, colonic epithelia, ovarian cystadenoma</td>
</tr>
<tr>
<td>8 GD1b</td>
<td>Galβ1-3GalNAcβ1 4 Galβ1-4Glcβ1-1-Cer Neu5Acα2-BNeu5Acα2</td>
<td>leukemia cells (Rat)</td>
</tr>
<tr>
<td>9 Blood group B type 2</td>
<td>Galα1 3 Galβ1-4GlcNAcβ1-1-R Fuco1 2</td>
<td>Human colon cancer</td>
</tr>
</tbody>
</table>

and short carbohydrate precursors such as Tn, T/TF and sialyl T antigen are synthesized that marks the
tumor cell (Vlad et al., 2002). These results suggest that the T/TF antigen is widely found on glycoproteins, and when this disaccharide is directly linked to the protein it may act as a cancer marker. Another important structure is Gal-(β1-4)-GlcNAc structure found in human leukaemias (Table II).

This disaccharide structure was found on terminal position in 69 different oligosaccharide structures (Table I). In Table II several different cancer related structures were found containing this terminal disaccharide structure (structure 2-5 and structure 6-7 (sulfated)). This structure is associated with colon, ovarian, leukemia, and embryonal carcinoma, suggesting that this structure is also an important cancer marker.

Several lectins like peanut agglutinin lectin bind to the T/TF-antigen and are mitogenic for colorectal cancer cell lines (Singh et al., 2006). Another important lectin is mistletoe lectin I (ML-I) that induces apoptosis of tumor cells by triggering a series of signaling events resulting in modification of anti- and proapoptotic proteins (Khwaja et al., 2007). Mistletoe lectin is used as adjuvant cancer therapy (Franz et al., 1981) and has shown promising results as a candidate for treatment of cancer. Especially in Germany and Austria up to two third cancer patients receive alternative therapies, primarily mistletoe extracts (Kienle et al.,
Mistletoe lectins are sugar-binding proteins and have been found to have many important properties and functions for cancer treatment. ML-I is a Gal specific lectin that binds to terminal Gal moieties on glycoproteins (like asialofetuin) and thereby induce its immunomodulatory and anti-tumor effect. Furthermore ML-I has shown to have a much higher affinity towards the disaccharide sequences Galβ1-2Gal and Ga1β1-3Gal rather than Gal (Lee et al., 1992). Besides being a promising anti-cancer agent, mistletoe lectin has also shown to have other traditional usage like prevention of hypertension, epilepsy, degenerative inflammation of the joints etc. (Loeper, 1999).

The aberrant glycosylation pattern in colon cancer and precancer are unclear. The synthesis of TF-antigen is based on the composition of the peptide (Ser/Thr), where the addition of GalNAc to the predetermined peptide occurs in the cis-golgi, and Gal is added in the trans-golgi network by glycosyltransferases (Campbell et al., 2001). In a study by Yang and colleges (1994), the changes in expression of the relevant glycosyl-, sialyl-, and sulfo-transferases in colon cancer was investigated, but it was found to correlate relatively poorly with the changes in carbohydrate expression. It was speculated perhaps an altered arrangement of transferases within the Golgi, might be responsible. Recently Brockhausen (2006) described the arrangement of different glycosyltransferases in the golgi. The enzymes that synthesize mucin cores 1 and 2 are mainly in the cis-golgi, whereas terminally acting enzymes are mainly in the trans-golgi. In colon cancer an alteration of these glycosyltransferase contributes to abnormal glycosylation of proteins and lipids associated with neoplastic transformation (Egea et al., 1993).

Terminal Gal moieties in the oligosaccharide chain of glycoproteins have several functions and may induce an immunomodulatory response or may act as a recognition marker. In general the ability of a sugar to act as an epitope or a recognition marker depends on the neighboring or vicinal sugar, the anomery and the linkage (Fig. 2). Depending on these factors the terminal Gal containing oligosaccharide structure can act as a recognition marker or an epitope for cancer or other diseases. In conclusion defining these factors including vicinal sugars, their linkage type and anomery is important to determine the role of terminal Gal in different cancers and other diseases.

ACKNOWLEDGEMENT


REFERENCES


EGGENS, I., FENDERSON, B., TOYOKUNI, T., DEAN, B.,...


(Received 21 December 2007, revised 9 January 2008)