INSTITUTO DE CIÊNCIAS BIOMÉDICAS ABEL SALAZAR
UNIVERSIDADE DO PORTO

BIOMARKERS FOR EARLY DETECTION OF GASTRIC CANCER BASED ON AUTOANTIBODIES SIGNATURES

DIANA CAMPOS

Dissertação de Mestrado em Oncologia

2010

DIANA CAMPOS

BIOMARKERS FOR EARLY DETECTION OF GASTRIC CANCER BASED ON AUTOANTIBODIES SIGNATURES

Dissertação de Candidatura ao grau de Mestre Oncologia submetida ao Instituto de Ciências Biomédicas de Abel Salazar da Universidade do Porto.

Orientador – Doutor Celso Reis

Categoria – Investigador; Professor Auxiliar Convidado

Afiliação – IPATIMUP - Instituto de Patologia e Imunologia Molecular da Universidade do Porto; Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto.

Agradecimentos/Acknowledgments

Gostaria de agradecer ao meu orientador, Doutor Celso Reis, pela oportunidade oferecida, por acreditar no meu potencial e pelo apoio no âmbito desta dissertação.

I would like to thank Professor Henrik Clausen for receiving me as his student and for his effort to help me in this study.

Às "Mucinas" pelo convívio, pela alegria e amizade e sobretudo por me fazerem recusar a acreditar que existam por aí por esse mundo laboratórios que possibilitem que o trabalho flua com tão boa disposição.

Um agradecimento especial à mestre Vânia pela amizade e pelas muitas boleias. À Rita Barros pelas tantas conversas matinais. À Ritinha pelo carinho e por me acompanhar em mais uma etapa "estudantil". À Sofia por tornar a minha adaptação tão fácil. À Ana Maria e à Joana Gomes pelo empenho em me educarem. Ao Bruno, à Catarina e à Dina por todas as *happy hours.*

I also want to thank everyone from the Copenhagen lab, especially Aaron, Catharina and Johannes for all the help and support.

Ainda, um agradecimento especial à Lara pela amizade, pelo grande apoio em Copenhaga e pelo louvor à língua portuguesa.

Por fim, agradeço aos meus pais por todo o apoio, mimo e amizade.

Index

Abbreviations 3

Abstract 5

Resumo 7

Introduction 9

1. Mucins 10
   1.1 Mucin structure 10
   1.2 Mucin and cancer 11

2. *O*- Glycosylation 12
   2.1 Polypeptide α-GalNAc transferases 13
   2.2 Sialyl-Tn antigen 14
   2.3 Core 1 15
   2.4 Core 2 and core 4 16
   2.5 Core 3 16
   2.6 Elongation and Synthesis of terminal structures 17

3. Alterations in mucin glycosylation as biomarker for cancer detection 18

4. Rationale of this work project 19

Objectives 21

Methods 23

1. Synthesis of recombinant mucin glycoprotein fragments in *E.coli* 25

2. Matrix-assisted Laser Desorption/Ionization Mass Spectrometry (MALDI-TOF) 26

3. Synthesis of *O*-glycopeptides 26

4. Mucin *O*-glycopeptide array print and anlysis 27

5. Monoclonal antibodies 28

6. Human sera 28

Results..29

1. *O*-glycosylated mucin proteins and synthetic peptides libraries 30

2. Mucin *O*-glycopeptide array analysis 32

Discussion 39

Conclusion and future perspectives 47

References 49

Abbreviations

Gal: Galactose;

GalNAc: *N*-acetylgalactosamine;

GlcNAc: *N*-acetylglucosamine;

mAb: monoclonal antibody

NeuAc: *N*-acetylneuraminic acid;

ppGalNAcT: polypeptide-N-acetyl-galactosaminyltransferase

PTMs: posttranslational modifications

STn: NeuAcα2,6GalNAcα-O-Ser/Thr;

T: Galβ1,3GalNAcα-O-Ser/Thr;

Tn: GalNAcα-O -Ser/Thr.

TR: tandem repeat

VNTR: variable number of tandem repeat

Abstract

A major group of glycoproteins produced by carcinomas are the membrane and secreted mucins. Changes in the types and levels of mucins expressed as well as their aberrant glycosylation are examples of tumor-associated antigens. These aberrant structures are widely produced by carcinomas but not by normal cells, thus constituting excellent biomarkers.

Since these structures are not exposed to the immune system in normal cells, antibody responses may develop early in carcinogenesis. Therefore circulating autoantibodies directed to specific glycopeptide epitopes elicited are emerging as biomarkers for the early detection of cancer.

Our study addressed the need for the identification of cancer-specific biomarkers to improve early gastric carcinoma detection based on autoantibodies signatures. Our strategy is based on the development of a glycopeptide array constituted by peptides and fusion proteins from human mucins with different cancer-associated *O*-linked glycans. We focused on the identification of cancer-specific glycoforms of mucins expressed in gastrointestinal cells: MUC1, MUC2, MUC5AC and MUC6 by autoantibodies in the serum. This approach could create a cancer-specific glycoform signature expected to have organ specificity and hence good biomarker value to gastric cancer diagnosis.

We have demonstrated the existence of such cancer-associated autoantibodies to the aberrantly *O*-glycosylated mucins in gastric cancer patients, as well as in individuals with precursor lesions of gastric carcinoma, namely intestinal metaplasia. The existence of such autoantibodies was either moderate or not observed in normal control individuals. Glycopeptide structures such as core 3, Lewis c and Tn revealed particular interesting results. Moreover, Tn glycoforms autoantibodies also presented epitope specificity depending on the site of *O*-GalNAc occupancy within the tandem repeat.

The study provided preliminary evidence for the utility of glycopeptide arrays for detection of cancer by autoantibody signatures and the expansion of the targets should improve the specificity and sensitivity of the strategy.

Resumo

As mucinas trans-membranares e secretadas são um dos maiores grupos de glicoproteínas produzidos por carcinomas. Alterações nos tipos e níveis de expressão de mucinas, assim como a sua glicosilação aberrante, são exemplos de antigénios associados a tumores. Estas estruturas aberrantes são abundantemente produzidas por carcinomas mas não por células normais, revelando-se desta forma excelentes biomarcadores.

Já que nas células normais estas estruturas não são expostas ao sistema imune, respostas imunológicas podem-se desenvolver numa fase inicial da carcinogénese. Desta forma, auto-anticorpos desenvolvidos para epitopes aberrantes de glicopéptidos surgem como possíveis biomarcadores para a detecção precoce de cancro.

O nosso estudo abordou a necessidade de identificação de biomarcadores específicos de cancro para a detecção precoce de carcinoma gástrico baseada numa assinatura de auto-anticorpos. A nossa estratégia baseou-se no desenvolvimento de um *array* de glicopéptidos constituído por péptidos e proteínas de fusão de mucinas humanas com diferentes *O*-glicanos associados. O trabalho focou-se na identificação de glicoformas especificas de cancro de mucinas expressas nas células gastro-intestnais (MUC1, MUC2, MUC5Ac e MUC6) por auto-anticorpos séricos. Esta abordagem criaria uma assinatura de glicoformas que fossem órgão-especificas e, assim, apresentando potencial como biomarcador na detecção de cancro gástrico.

Neste estudo, demonstramos a existência de tais auto-anticorpos para mucinas com *O*-glicosilação aberrante em pacientes com cancro gástrico, assim como em indivíduos com lesões precursoras, nomeadamente metaplasia intestinal. A existência destes auto-anticorpos em indivíduos saudáveis era nula ou reduzida. Estruturas de glicopéptidos como core 3, Lewis c e Tn apresentaram resultados particularmente interessantes. Adicionalmente, auto-anticorpos contra a glicoforma Tn apresentaram especificidade dependente do local ocupado no tandem repeat pelo resíduo de O-GalNAc.

Este estudo apresentou evidências preliminares da utilidade de *arrays* de glicopéptidos na detecção de cancro por auto-anticorpos e o uso de outros alvos poderá melhorar a especificidade e sensibilidade desta estratégia.

Introduction

1. Mucins

Mucins are high-molecular-weight *O*-glycosylated proteins that are major components of mucus covering the luminal surface epithelium of the respiratory, gastrointestinal, and reproductive tracts (*1*).

Mucins have a central role in maintaining homeostasis in these harsh conditions. Their general structure and biochemical composition provides protection for the cell surface, and specific molecular structures regulate the local molecular microenvironment near the cell surface. In fact, mucins might serve as cell-surface receptors and sensors, and conduct signals in response to external stimuli that lead to coordinated cellular responses that include proliferation, differentiation, apoptosis and secretion of specialized cellular products (*2-4*).

1.1 Mucin Structure

Early studies to characterize the molecular nature of mucins were difficult to perform due to their biophysical properties: a relatively large mass, a complex biochemical composition (50–80% *O*-linked oligosaccharides) and a tendency to form higher-order structures through polymerization (*5*).

The hallmark of mucins is the presence of repeated peptide stretches called "variable number of tandem repeat" (VNTR). This structural feature is common to all mucins and comprises tandem repeats of identical or highly similar sequences that are rich in serine, threonine and proline residues (*1, 6*). Regions that are rich in serine or threonine are putative *O*-glycan acceptor sites and have an abundance of clustered mucin *O*-glycans. The proline residues appear to facilitate *O*-GalNAc glycosylation. The specific sequence and number of tandem repeats, however, is highly variable among different mucins.

The two main classes of mucins (secreted and cell-surface associated) include different members that have both unique and shared structural features. Cell-surface-associated mucins are bound to cells by an integral transmembrane domain and have relatively short cytoplasmic tails that associate with cytoskeletal elements, cytosolic adaptor proteins and/or participate in signal transduction. They include MUC1, MUC3, MUC4, MUC12, MUC13, MUC16 and MUC17. Secreted mucins show patterns of expression that are restricted to secretory organs and cell types, and present cysteine-rich regions and cystine knots that are responsible for their polymerization and the formation of extremely

large molecules of several million daltons. They include MUC2, MUC5AC, MUC5B, MUC6 and MUC7. There are two classes of secreted mucins-gel-forming (MUC2, MUC5AC, MUC5B, MUC6) and non-gel-forming (MUC7) (*7-22*).

1.2 Mucins and Cancer

Mucins have long been implicated in the pathogenesis of cancer. As sites of tumour growth are often hypoxic, acidic and with proteases and other biologically active factors, it is possible that tumours use mucins to configure the local microenvironment during tumorigenesis. In addition, mucin glycans are involved in processes such as invasion and metastasis (*23*).

Increasing concentrations of mucin-type glycoproteins in serum are correlated with increasing tumour burden and poor prognosis. Altered mucin expression is seen in cancer and precancerous lesions, both at the protein (apomucin) and at the glycosylation levels (*23-25*). Realizing the importance of *O*-glycans in many cell functions, we can then understand the major influence of structures of glycoproteins present on cancer cells surfaces or in their secretions, on the biology, physiology and immunological properties of a cancer cell. In cancer, *O*-glycans play important roles in the attachment and invasion of cancer cells and their survival in the blood stream. Unusual or novel carbohydrate and peptide structures may lead to immune responses in cancer patients. Anti-mucin antibodies have been proposed to have therapeutic potential in cancer (*26*).To exploit the immune response therapeutically, mucin- like antigens with cancer-specific epitopes have been developed as a vaccine (*27, 28*). MUC1 peptides injected into mice delayed tumor growth of mammary epithelial cancer cells (*29*). Additionally, and as will be described in this study, altered antigenicities of mucin peptides and carbohydrates may also be useful for diagnosis.

O-Glycan chains of glycoproteins, as well as tissue and blood group antigens, including Lewis antigens, may be qualitatively and quantitatively altered, and are often truncated and highly sialylated in cancer cells (*1, 30, 31*). These truncated forms result from a marked disorganization of the secretory pathway organelles and alteration in the enzymatic machinery expression and topology. For instance, rearrangement of Golgi enzymes may lead to a failure to synthesize essential intermediates or may block further conversions by premature synthesis of terminal structures, and may account for altered glycosylation in cancer (*32-34*).

2. O-Glycosylation

The *O*-PTM form of protein glycosylation is one of the most diverse forms of posttranslational modification because it involves 50-100 distinct genes, including up to 20 polypeptide GalNAc transferases controlling where the *O*-glycans are attached (*35-37*). The range and chemical composition of structures that are created by branched *O*-linked oligosaccharides are immense, especially when altered by additional post-translational modifications. The functions of *O*-glycans are as diverse as their structures. *O*-Glycans protect underlying proteins as well as epithelial cell surfaces, maintain protein conformations, control active epitopes and antigenicity; they participate in cell adhesion, binding to microbes and in the immune system. They may determine the cell surface expression and function of cell surface receptors and may be involved in growth regulation. *O*-Glycans have also been shown to be involved in blood clotting, embryogenesis, development and cell death (*6, 34*).

O-glycosylation with complex oligosaccharides is crucial to mucin structure and function. Mucin-type oligosaccharides are involved in specific ligand–receptor interactions, confer hydroscopic properties and might bind various small molecules and proteins (*5*).

There is a remarkable structural variation in mucin glycosylation, due to the fact that at least eight different core structures of *O*-glycans exist in mammalian mucins. Core structures 1-6 have been described in human cancer cells (*34*).

The process always begins with GalNAc *O*-glycosidically linked to Ser or Thr residues (Tn antigen), being the only structure common to all *O*-glycans. The Tn antigen may be sialylated to form the sialyl-Tn antigen, sialylα2- 6GalNAc-O-Ser/Thr. Alternatively, core structures may be synthesized (Figure1). Depending on the availability and specificity of glycosyltransferases and sulfotransferases involved, *O*-glycan cores may be elongated and terminated in many ways; they may be short and terminated by sialic acid, or may be elongated. When it comes to *O*-glycosilation regulation factors, gene expression of glycosyltransferases, and of their protein substrates, are the most understandable ones.

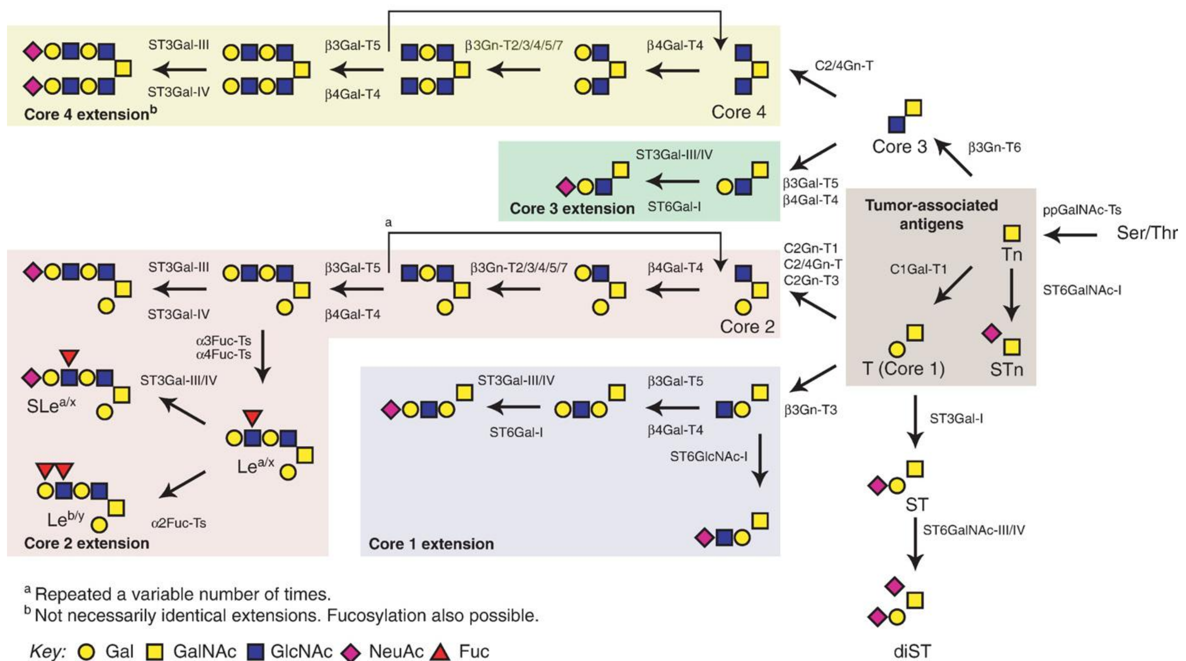

Figure 1 - Biosynthetic pathways of most common mucin-type *O*-glycans. Note that most likely candidate glycosyltransferases are indicated based on *in vitro* activities. From (*38*).

Factors that control the activities of transferases *in vitro* include metal ion concentrations and the presence of certain proteins and lipids, and post-translational modifications of enzyme proteins (*35, 39*).

The peptide moieties of substrates and sugar residues near *O*-glycosylation sites have been shown to regulate primary *O*-glycosylation as well as subsequent glycosylation steps. The sequence of sugar additions in specific Golgi compartments is controlled by the relative activities of glycosyltransferases acting on common acceptor substrates. If these enzymes are present in the same Golgi compartment, competition may take place, and the relative activities determine the quantities of possible products (*38*).

2.1 Polypeptide α-GalNAc-transferases

The first step of mucin *O*-glycosylation is the transfer of *N*-acetylgalactosamine from UDP-GalNAc to serine or threonine residues, which is catalyzed by a polypeptide-N-acetyl-galactosaminyltransferase (ppGalNAcT) (*35*).

The localization of individual GalNAc-transferases may vary depending on the cell type and possibly also the differentiation status. Their expression levels vary considerably between cell types and tissues (*35, 36*). All ppGalNAcTs bind UDP-GalNAc (the donor of

N-acetylgalactosamine), but they often differ in the protein substrates to which they transfer *N*-acetylgalactosamine. Such differences allow ppGalNAcTs to be distinguished (*37*). Many ppGalNAcTs appear to have a hierarchical relationship with one another, such that one enzyme cannot attach an *N*-acetylgalactosamine until an adjacent serine or threonine is glycosylated by a different ppGalNAcT. Thus, coexpression in the same cell of ppGalNAcTs with complementary, partly overlapping acceptor substrate specificities probably ensures efficient *O*-GalNAc glycosylation (*35, 36, 40*).

In cancer…

The Tn antigen (GalNAc-*O*-Ser/Thr), which is uncommon in normal mucins, is often found in mucins derived from tumors (*23, 41*), which suggests that the extension of *O*- glycans beyond the first carbohydrate is blocked in some cancer cells. This and other truncated glycoforms result from a disorganization of the secretory pathway organelles and alterations in the *O*-glycosylation enzymatic "machinery" controling *O*-glycosilation. In fact, it is plausible that in cancer cells, polypeptide α-GalNAc-transferases localization may differ from the normal. This way, if GalNAc is added in a later compartment, substrates are not available for early acting enzymes of the assembly line, and structures will be incomplete. Incomplete chain synthesis occurs in many cancer cells and contributes to the prevalent expression of the Tn antigen (*23, 42, 43*).

2.2 Sialyl-Tn antigen

Another common cancer-associated structure found in mucins is the sialyl-Tn antigen. The sialyl-Tn antigen (Neu5Acα2-6GalNAc-*O*-Ser/Thr), also known as STn and CD175s, is a simple mucin-type carbohydrate antigen which has attracted much attention because it is highly expressed in most gastric (*44*), colorectal (*45*), ovarian (*46*), breast (*47*) and pancreatic (*48*) carcinomas, whereas expression of STn on normal tissues is highly restricted (*41, 49*). In addition, sialyl-Tn expression is associated with carcinoma aggressiveness and poor prognosis (*47, 50*). The mechanisms underlying the appearance of this *O*-glycan are not totally clear and may vary with the tissue. One obvious mechanism is an increase in the level of activity of an α2,6-sialyltransferase that catalyzes the transfer of sialic acid from CMP-Neu5Ac to a GalNAc *O*-linked to serine or threonine residues of a polypeptide (*51*). Previous studies have shown that the expression of the sialyl-Tn antigen in cancer cells is due to ST6GalNAc-I activity (*52*). The sialyl-Tn product structure cannot be further processed and therefore blocks the biosynthesis of any core structure and the posterior elongation (*53*).

Another biosynthetic basis for expression of Tn and STn glycans may relate to somatic mutations in Cosmc, a chaperone for the enzyme core 1 beta1,3-galactosyltransferase enzyme that controls synthesis of the common core 1 *O*-glycosylation pathway (*54-56*). In this scenario, sialyl-Tn accumulation would then result as a side effect of the loss of ability to make the core-1 O-glycan and its extensions along with expression of the specific ST6GalNAc sialyltransferase. Given the frequency of sialyl-Tn accumulation by cancer cells, it is likely that its expression confers some as yet not fully understood advantage to the tumor cells.However, previous studies did show that sialyl-Tn antigen may play a role in cancer cell recognition by the immune system, protecting metastatic cells from degradation in the blood stream (*57*).

The tremendous structural variation of *O*-glycan structures in mucins is due to the existance of at least eight different core structures of *O*-glycans. We will focus on the core structures that are most relevant.

2.3 Core 1

Core 1 occurs on most glycoproteins, and is often substituted with sialic acid residues or with various other sugars. The synthesis of core 1 is, in fact, a prerequisite for the synthesis of sialylated T antigens, elongated core 1 structures, and core 2 structures found in glycoproteins.

Most *O*-glycan structures are based on the Core 1 structure formed by addition of the monosaccharide galactose (Gal) in a β1–3 linkage to GalNAc. This reaction is carried out by a single Core 1 Galtransferase in mammals (*56*). The enzyme synthesizing core 1, is a ubiquitous enzyme that occurs in most mammalian cells and in many species and its activity is present in most cell types (*56*). However, and as mentioned above, T synthase requires a specific molecular chaperone called Cosmc to be exported from the endoplasmic reticulum. Cosmc is an ER protein that appears to bind specifically to T synthase and ensures its full activity in the Golgi. Lack of core 1 synthesis can be due to either defective T synthase or the absence of functional Cosmc chaperone (*54-56*).

In cancer…

Core 1 as a terminal structure is prevalent in cancer cells and their secretions (*58, 59*). In these cases, it is likely that there is an abnormality in either the sialylation of core 1 or on further extension and branching of core 1. This prevents other modifications of core 1. In fact C2GnT1 enzyme converts core 1 to core 2, by the addition of GlcNAc to GalNAc and

is the dominant enzyme expressed in normal breast tissue. Expression of C2GnT1 is low or absent in around 50% of breast cancers, whereas expression of ST3Gal-I which transfers sialic acid to galactose in the core 1 substrate is consistently increased (*60*).

2.4 Core 2 and 4

Core 2 mucin *O*-glycans are branched core 1 structures, where N-acetylglucosamine (GlcNAc) is added in a β1–6 linkage to GalNAc of the core 1 structures by one of three β1-6 *N*-acetylglucosaminyltransferases or C2GnTs (leukocyte type, C2GnT-1 and -3 and mucin type, C2GnT-2) (*61*) (*62*). They are produced in many tissues, including the intestinal mucosa. The synthesis of core 2 *O*-GalNAc glycans is regulated during activation of lymphocytes, cytokine stimulation, and embryonic development. Two enzymes synthesizing core 2 have been well characterized and can be distinguished by their substrate specificities and tissue distributions (*63*). One of the enzymes (C2GnT-2) synthesizing the core 2 structure is also capable of utilizing core 3 as a substrate, this way producing the core 4 structure, where GlcNAc is added in a β1–6 linkage to GalNAc similar to synthesis of the core 2 structure (*61*).

In cancer…

Cancer cells and other diseased tissues have abnormal amounts of core 2 *O*-GalNAc glycans. Up-regulation of the sialyltransferase ST3Gal-I was previous reported in human breast cancer cell lines preventing Core 2 extension (*64*). On the other hand, increased activity of Core 2 C2GnT-1 is positively correlated with progression of prostate cancer (*65*).

2.5 Core 3

Core 3 is restricted in its occurrence to mucins from specialized tissues and is generally believed to be limited to the mucosa of the digestive tract (*66, 67*). In fact, it was first identified in *O*-glycans derived from stomach and colon. Core 3 synthase β3GlcNAc-T6 is an important enzyme in the synthesis of mucn-type *O*-glycans in digestive organs that transfers GlcNAc to GalNAc in a β1–3 linkage. On immunohistochemical analysis, β3GlcNAc-T6 was detected in the Golgi region of both normal gastric and colorectal epithelial cells (*68*).

Core3 *O*-glycans controlled by the Core 3 β3GlcNAc-T6 are typically capped by galactosylation and sialylation or other histo-blood group related structures.

In cancer…

It was demonstrated that the expression of β3GlcNAc-T6 dramatically decreased in cancerous tissues. In fact, in contrast to normal tissues, β3GlcNAc-T6 disappeared in both gastric and colorectal cancer (*68, 69*). The changes to the core structures give rise to phenotypic changes of cancer cells, such as mobility and metastic activity.

Truncated core 3 *O*-glycans have been found in colon cancer in accordance with the expression pattern for the controlling enzyme β3GnT6 (*68*). Whereas this enzyme is downregulated in colon cancer, little is known about its expression in other types of cancer.

Core 1 and core 2 structures can appear due to the disappearance of one of its competitors, core 3 synthase, in cancerous tissues. This shift favors the synthesis of core 1 and may explain the prevalence of the T antigen in cancer tissue (*43*). The loss of core 3 synthase may also play a role for the appearance of STn-glycoform, since the core3 synthase competes for the Tn substrate utilized by ST6GalNAc-I to produce STn glycoforms (*52*).

2.6 Elongation and Synthesis of terminal structures

There are many ways by which O-glycan core structures may be further processed and the ABO and other glycan-based blood groups as well as sialic acids, fucose, and sulfate are common terminal structures in mucins (*24, 34*). Terminal sugars and sulfate groups can have important biological roles. For example, the degree of sialylation is important for the overall properties of O-glycans and glycoproteins; in addition, sialylated structures often block subsequent glycosylation steps. Specific structures can serve as antigenic determinants or ligands for cell adhesion molecules.

Lewis-type blood group antigens, such as sialyl Lewis A (SLe^{a}) and sialyl Lewis X (SLe^{x}), are expressed in cancer cells, mimicking their normal expression on blood cells (monocytes and neutrophils) and also mimicking their potential for migration through binding to endothelial cell selectins (*24*).

The relevance of Lewis sialylated structures in cancer was first revealed in the 1980s, when monoclonal antibodies raised against cancer cells were shown to recognize $SLe^{a/x}$.

overexpression of SLe^x and SLe^a is common in carcinomas of several origins (eg, lung, colon, gastric and pancreas) and is associated with increased metastatic ability (*70*) and poor survival of the patients (*71*).

3. Alterations in mucin glycosylation as biomarkers for cancer detection

In a short balance of the information above, we can conclude by now that glycoconjugate modifications are a universal hallmark of cancer. These alterations are not only visible on the cell surface of cancer cells, and therefore easily accessible to antibodies or lectins as tissue biomarkers, as they are also often shed into the blood stream, and therefore identifiable as serum biomarkers (*24*).

As mentioned above, mucins are the major group of glycoproteins produced by carcinomas and which are densely *O*-glysylated, consequently being major carriers of cancer-associated carbohydrates (*72*). Changes in the types and levels of mucins expressed as well as their aberrant glycosylation are examples of tumor-associated antigens, creating a diverse set of unusual molecular structures found on cancer cells, secretions and blood (*23, 73*). These aberrant structures are widely produced by carcinomas but not by normal cells, thus constituting excellent biomarkers. In fact, many of the current biomarkers used in clinics, in both tissue and serum assays, are based on these carbohydrate modifications (*74*), (*75, 76*).

The rich content in proteins of serum and plasma that reflect diverse physiological or pathological states, and the ease with which this compartment can be sampled, make it a favorite choice for biomarker applications.

Traditional cancer biomarker serum assays rely on the detection of circulating molecules produced by tumour cells and shed into the blood stream. Most cancer biomarker assays in use today including CEA (colorectal cancer) (*74*), CA125 (ovarian cancer) (*20*) and CA15-3 (breast cancer) (*77*) are detecting specific glycoproteins, often mucins, or in the case of CA19-9 and CA72-4 carbohydrates attached to circulating glycoproteins (*75, 76*). However, all current assays suffer from relative low sensitivity and specificity in early detection of cancer, but are useful for monitoring treatment and follow-up for recurrence. The relative low specificity of current assays is due to enhanced shedding of the same glycoproteins in some benign diseases and inflammatory conditions.

In normal cells, truncated immature *O*-glycans may exist as brief biosynthetic intermediates in the early Golgi, and these precursor structures are not exposed to the

immune system. Aberrant surface exposure of these structures by cancer cells can elicit production of autoantibodies directed to specific glycopeptide epitopes (*78*). Therefore, as an alternative to the antigen detection, circulating autoantibodies to glycopeptides epitopes are emerging as biomarkers for the early detection of cancer (*79*). In fact, there are a series of good reasons for autoantibodies to represent novel diagnostic biomarkers.

First, the immune response to tumor-associated antigens occurs at an early stage during tumorigenesis, and therefore autoantibodies can be detected in the asymptomatic stage of cancer (*80*).

Second, autoantibodies against tumor-associated antigens are found in the sera of cancer patients where they are easily accessible to screening.

Third, they can be produced at relatively high concentrations (autoantibodies are amplified by the immune system in response to a single autoantigen) with a long circulation time. Autoantibodies are inherently stable and persist in the serum for a relatively long period of time because they are generally not subjected to the types of proteolysis observed in other proteins, including other biomarkers, which are transiently secreted and may be rapidly degraded or cleared (*81, 82*). Consequently they can be detected with sensitive and specific methods, as illustrated by the detection of high titers of autoantibodies in patients with early stage cancer (*83*).

Recently, it was confirmed the existence of such cancer-associated autoantibodies to the aberrantly *O*-glycosylated MUC1 mucin in patients with breast, ovarian, and prostate cancer at time of diagnosis (*78*). In addition to this, the immune response to these antigens has also been shown to correlate with the progression of malignant transformation (*84, 85*).

Thus, the production of autoantibodies can be detected before any other biomarkers or phenotypic aberrations are observed, rendering such autoantibodies indispensable as biomarkers for early cancer detection.

4. Rationale of this work project

Accordingly to the previous information, as serum biomarkers that manifest prior to the onset of cancer, autoantibodies are highly sought after, and their detection on gastric carcinoma patients sera is the main goal of this work project.

Gastric carcinoma is the second most common cause of cancer deaths worldwide and remains among the top-5 leading cancers in incidence and mortality both in Portugal and in several other European countries (*86*). The poor prognosis and high mortality of gastric carcinoma reflects the late diagnosis due to appearance of clinical symptoms only when the cancer has progressed. To overcome this challenge, the current medical focus has been centered on early detection accomplished through an efficient screening.

Focusing on a single and well known environment such as the gastric environment is of great benefit for the study. Gastric mucus covering the mucosal surface is mainly composed of mucins. MUC1, MUC5AC and MUC6 are expressed in normal gastric mucosa, whereas MUC2 were found to be de novo expressed in gastric carcinomas (*25, 48, 87, 88*). For instance, altered histochemical pattern of MUC5AC (*25*), higher expression level of MUC1 (*89*), increased heterogeneity of mucins, and the exposure of truncated mucin-type carbohydrates have been detected in gastric cancers (*23*).

The focus on this specific environment and on the identification of cancer-specific glycoforms of these particular mucins by autoantibodies would create a cancer-specific glycoform signature providing a powerful tool for the early diagnosis of gastric carcinoma. This actually would allow overcoming one of the most peremptory limitations of this kind of biomarkers, as the use of autoantibodies as stand-alone cancer biomarkers is limited as some mucins have a broad expression by various types of cancer. This is exemplified on the recent cancer-associated autoantibodies to the aberrantly *O*-glycosylated MUC1 mucin study as MUC1 is broadly expressed by most adenocarcinomas (*90*). The reason for this low specificity lies in the heterogenic nature of cancer, whereby different proteins are aberrantly processed or regulated in patients with the same type of cancer. Although a single autoantigen would lack adequate sensitivity and specificity, a panel of antigens may overcome this problem by enabling the development of a cancer-specific glycoform signature. In fact, several studies have reported the high sensitivity and specificity that a panel of carefully selected tumor associated antigens can achieve in cancer diagnosis (*91, 92*).

In conclusion, autoantibodies to cancer-specific glycoforms of gastrointestinal mucins (MUC1, MUC5AC, MUC6 and MUC2) (*25, 87*) is expected to improve organ specificity and hence better biomarker value.

Objectives

The main objective of this study was to develop a new strategy for early detection of gastric cancer based on autoantibodies signatures. To accomplish this we wanted to identify autoantibodies directed to cancer-associated mucin antigens in serum from gastric cancer patients, in individuals with precursor lesions patients and compare with healthy controls.

We proposed that the identification of cancer-specific glycoforms of MUC1, MUC2, MUC5Ac and MUC6 by autoantibodies would create a cancer-specific glycoform signature providing a powerful tool for the early diagnosis of gastric carcinoma.

The following specific aims were proposed:

- Establish a versatile chemoenzymatic approach to produce libraries of cancer associated *O*-glycopeptides and *O*-glycoproteins, producing different glycoforms of MUC1, MUC2, MUC5Ac and MUC6. The glycoforms included are: Tn (GalNAc-Ser/Thr), STn (Neu5Acα2-6GalNAc-*O*-Ser/Thr), T (Galβ1-3 GalNAc-Ser/Thr), core 3 (GlcNAcβ1-3GalNAcα-O-Ser/Thr) and Lewis c (Galβ1-3GlcNAcβ1-3GalNAcα-O-Ser/Thr).

- Develop a microarray platform using *O*-glycopeptide and *O*-glycoproteins libraries for the detection of autoantibodies in human serum. The collection of sera samples from individuals with gastric carcinoma, gastritis, intestinal metaplasia and control normal samples were tested in this novel microarray platform.

Materials and Methods

In order to develop a sensitive detection of gastric cancer by detecting autoantibodies signatures, we used a series of complementary research lines based on *O*-glycosylated mucin proteins and synthetic peptides for building compound libraries of aberrant glycopeptide epitopes (Figure 2). In order to achieve this, we developed a microarray platform displaying *O*-glycopeptide libraries covering MUC1, MUC2, MUC5AC, and MUC6 different glycoforms. In this study we used a glycan array format that uses standard robotic printing technology with demonstrated applicability to antibody detection in crude human serum (*78, 93*).

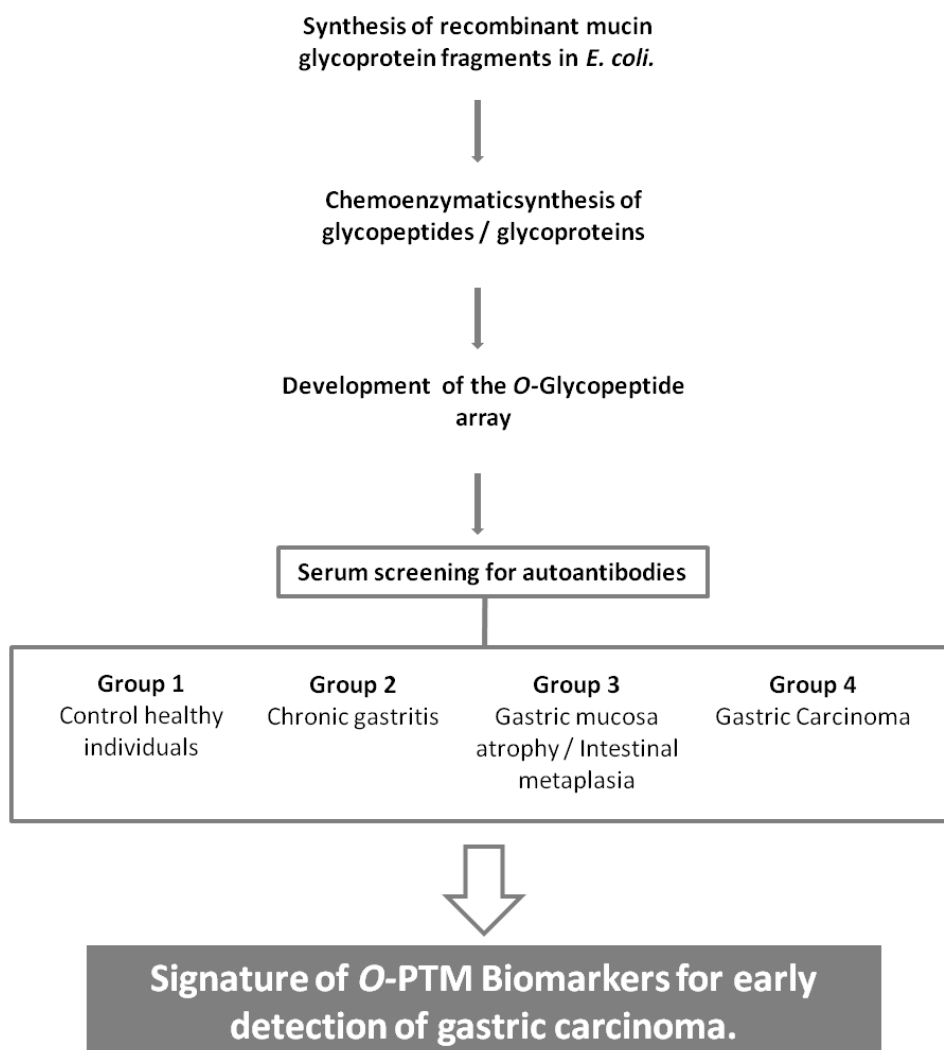

Figure 2 - Schematic representation of this study research plan. From O-glycopeptide libraries to detection of cancer-specific mucin glycoforms by autoantibodies in serum, this way creating a cancer-specific glycoform signature and providing a powerful tool for the early diagnosis of gastric carcinoma.

1. Synthesis of recombinant mucin glycoprotein fragments in *E. coli:*

N- or C- terminally 6×His and T7 tagged recombinant fragments of MUC2, MUC5AC and MUC6 (Table 1) were produced in *E. coli*. The MUC2 construct includes 118 amino acid residues corresponding to the tandem repeat domain with 39 potential glycosylation sites. The MUC5AC construct includes 222 amino acid residues corresponding to the tandem repeat domain with 118 potential glycosylation sites. The MUC6 construct includes 114 amino acid residues corresponding to the tandem repeat domain with 31 potential glycosylation sites.

Mucin	Sequence	Mw
MUC5Ac	MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRIPTSSTTSTPQTSTTSASTTSITSGPGTTPS PVPTTSTTSAPTTSTTSAATTSTISAPTTSTTSAPTTSTTSASTASKTSGLGTTPSPIPTTSTTSPPT TSTTSASTASKTSGPGTTPSPVPTTSTIFAPRTSTTSASTTSTTPGPGTTPSPVPTTSTASVSKTS TSHVSISKTTHSQAAALEHHHHHH	21,6 kDa
MUC2	MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRDPNSSSVDKLDIEFLQPGGSVQCCECVTQ PTTMTTTTTENPTPTPITTTTTVTPTPTPTSTQSTTPTPITTTNTVTPTPTPTGTQT	12 kDa
MUC6	MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRISTSLVTPSTHTVITPTHAQMATSASNHSAP TGTIPPPTTLKATGSTHTAPPITPTTSGTSQAHSSFSTNKAAALEHHHHHH	11,8 kDa

Table 1 - Sequences and molecular weight (Mw) of the mucins used in the present work.

The all process to get purified recombinant fragments is schematically illustrated in Figure 3. Gene sequences were inserted into the bacterial expression vectors and Rosetta2 (Novagen) was the expression host. Overnight cultures were diluted 1:100, and induced 4 h at 37°C by the addition of IPTG to a final concentration of 0.1 mM. Cell lysates were nickel purified using NiNTA agarose (Qiagen) as described by the manufacturer, and the eluted fractions were analyzed by SDS-PAGE on NuPAGE Bis-Tris 8-12% acryl amide gels (Novex) stained with Coomassie. Eluted fractions of recombinant mucin fragments were HPLC purified before and after in vitro O-glycosylation (see below). Samples were diluted in 0.1 % TFA (triflouroacetic acid), loaded onto a Zorbax 300SB-C18 column mounted on an Agilent 1100 HPLC system, and eluted in a 40-minute linear gradient from 0-80% acetonitrile. Eluted fractions were lyophilized and resuspended in water and mass confirmed by MALDI-TOF mass spectrometry.

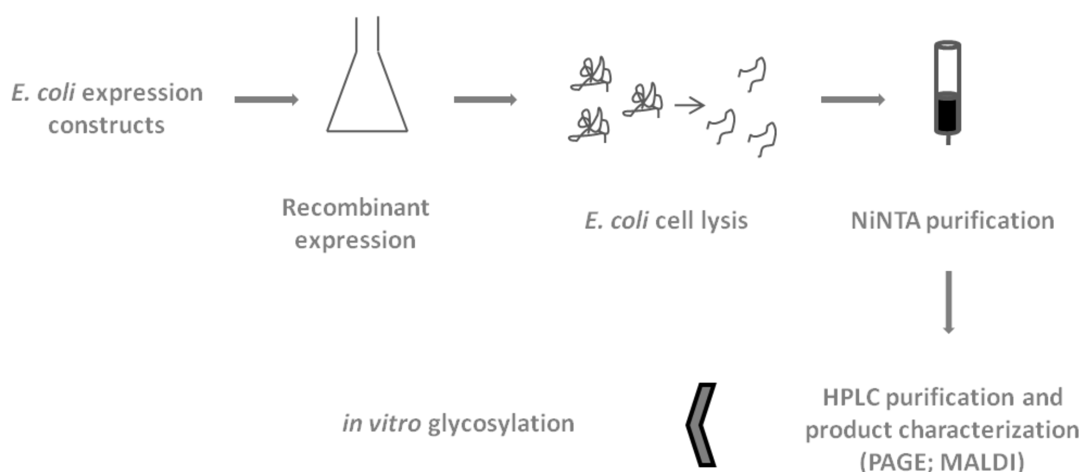

Figure 3 - Synthesis of recombinant mucin glycoprotein fragments workflow. From DNA construct to glycoprotein.

2. Matrix-assisted Laser Desorption/Ionization Mass Spectrometry (MALDI-TOF)

MALDI-TOF mass spectrometry was performed on a Voyager-DE™ PRO workstation (Applied Biosystems). The MALDI matrix was 2,5-dihydroxybenzoic acid (Sigma) dissolved in 2:1 mixture of 0.1% TFA in 30% aqueous acetonitrile. Samples dissolved in 0.1% TFA to a concentration of ~1 pmol/µL were prepared for analysis by placing 1 µL of sample solution on a probe tip followed by 1 µL of matrix. All mass spectra were obtained in the linear mode.

3. Synthesis of *O*-glycopeptides:

Peptides and recombinant fragments were *O*-glycosylated *in vitro* using recombinant glycosyltransferases as previously described (*94, 95*). Synthetic peptides included 19-mer MUC5AC and 19-mer MUC2. A panel of MUC1 peptides was also added, as MUC1 *O*-glycopeptides were used in a previous study, with great success in autoantibodies detection approach for other types of carcinomas, (*78*). Peptides were glycosylated *in vitro* using purified recombinant glycosyltransferases polypeptide GalNAc-T1, GalNAc-T2 and GalNAc-T3, used individually and to direct GalNAc *O*-glycan occupancies on peptides and recombinant proteins. Further peptide occupancie with GalNAc was achieved with GalNAc-T12 using GalNAc-T3 glycopeptide product as substrate. GalNAc glycosylation of the peptides was performed in a reaction mixture (1 mg peptide/mL) containing 25 mM cacodylate buffer (pH 7.4), 10 mM MnCl2, 0.25% Triton X-100, and 2 mM UDP-GalNAc.

The chemically synthesized glycopeptides with single Tn glycans were similarly elongated with human Drosophila Core-1 β3GalT (*96*), human Core3 β3GlcNAc-T6 (*66*) and β3Gal-T5 (*96*) to produce T, core3 and Lewis c structures. Sialylation of MUC2 recombinant protein was performed using ST6 GalNAc-I enzyme (*97*) and MUC2 GalNAc-T2 product as substrate, in a reaction mixture (1 mg peptide/mL) containing 20 mM Bis–Tris buffer (pH 6.5), 20 mM ethylenediaminetetraacetic acid (EDTA), 1 mM dithiothreitol, and 2 mM cytidine-5′-monophospho- N-acetylneuraminic acid sodium salt (Sigma, Brondby, Denmark). Glycosylation was monitored using MALDI-TOF mass spectrometry and all glycopeptides were HPLC purified and characterized by MALDI-TOF, as described above.

4. Mucin *O*-glycopeptide array print and analysis

(Glyco)peptides and control structures were printed on Schott Nexterion® Slide H MPX 48 (Schott AG, Mainz, Germany), as Blixt and colleagues have previously described (*93*). Triplicates of all compounds were printed in 150 mM sodium phosphate pH 8.5 with 0.005% CHAPS and printed on a BioRobotics MicroGrid II spotter (Genomics Solution) using Stealth 3B Micro Spotting Pins (Telechem International ArrayIt Division). As Blixt and colleagues previously described (*98*), slides were incubated for 1 h in a humidified hybridization chamber with 75% relative humidity and stored until use at -20ºC. Prior to use the microarrays were blocked for 1 h with 25 mM ethanolamine in 100 mM sodium borate pH 8.5. Human sera diluted from 1:5–1:50 or mAbs (1 µg/ml or hybridoma supernatants undiluted) were incubated in a closed container with gentle agitation for 1 h, washed three times in PBS with 0.05% Tween-20 (PBS-T) and followed by 1 h incubation with appropriate secondary antibodies. Human IgG antibodies were detected with Cy3-conjugated goat anti-human IgG (Fc specific) diluted 1:1000 in PBS-T. Murine mAbs were detected with Cy3-conjugated goat anti-mouse IgG (H+L) (Jackson ImmunoResearch Laboratories, Inc.) diluted 1:1000 in PBS-T. After incubation with secondary antibodies the slides were washed 3 times in PBS-T, and after the final wash, slides were rinsed shortly in H_2O, dried by centrifugation (200×*g*) and scanned in a ProScanArray HT Microarray Scanner (PerkinElmer) followed by image analysis with ProScanArray Express 4.0 software (PerkinElmer). The overview of the microarray is represented in Figure 4. The mean value of relative fluorescence intensity (RFU) was used. For comparison, slides were scanned with identical scanning parameters. Data were analyzed and plotted using Microsoft Excel or GraphPad Prism software.

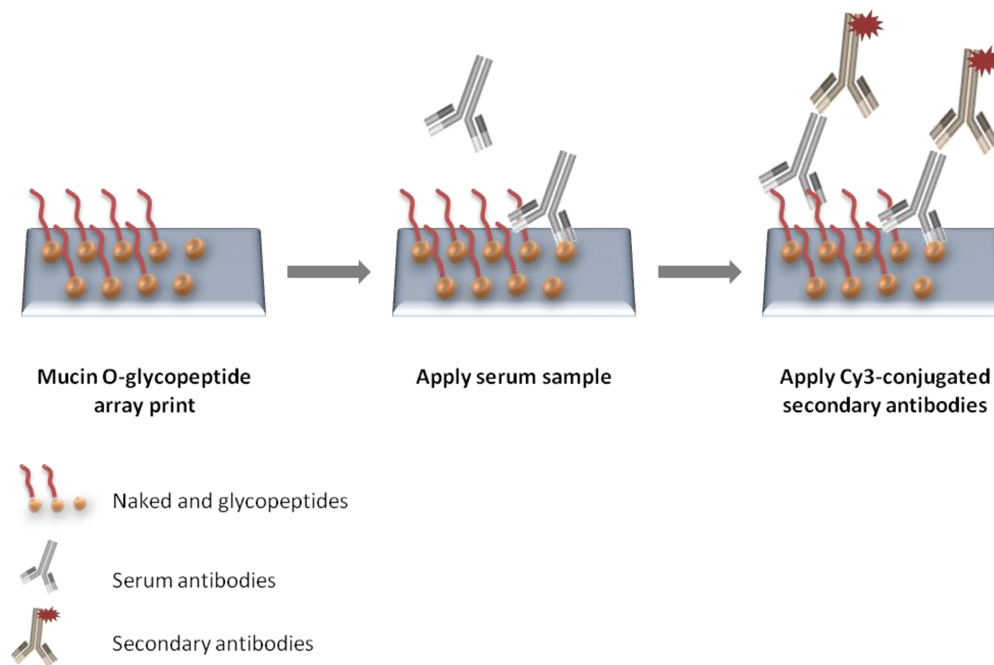

Figure 4 - Overview of mucin microarray assays performed to detect autoantibodies in human serum.

5. Monoclonal antibodies:

Printing of the glycopeptides was confirmed using mAbs against MUC2 Tn (PMH1) (*99*), MUC5AC (CLH2) (*25*), MUC6 (CLH5) (*88*), as well as carbohydrate-specific antibody targeting STn (TKH2) (*49*). Lectin VVA was also used to assess the Tn epitope (*100*).

6. Human sera:

Sera samples to be evaluated for reactivity in the array include serum from individuals with normal gastric mucosa without *Helicobacter pylori* infection confirmed by biopsie (n=8); individuals with chronic gastritis/ *Helicobacter pylori* infection (n=8); Individuals with gastric mucosa atrophy / Intestinal metaplasia (n=9) and individuals with gastric carcinoma (early or advanced) (n=10).

Serum samples were collected from Hospital de São João and Hospital de Santo António in Porto, Portugal. The use of these samples has been approved by the local Ethical committee.

All sera were collected and stored following the same standard operating procedure that involved clotting for 1 hour at room temperature followed by 1 hour incubation on ice. Samples were then freezed and kept at -80^{0}C.

Results

1. *O*-glycosylated mucin proteins and synthetic peptides libraries

A versatile chemoenzymatic approach to produce libraries of cancer associated *O*-glycoproteins was estabilihed, producing different glycoforms of MUC1, MUC2, MUC5Ac and MUC6.

One of the strategies used on the production of compound libraries was chemo-enzymatic synthesis of short synthetic glycopeptides. Other strategy comprised the production of recombinant *E.coli* of large mucin polypeptide fusion proteins derived from MUC2, MUC5Ac and MUC6 (Figure 5). The recombinant proteins and synthetic peptides were glycosylated using purified human recombinant GalNAc-transferase T1, T2 and T3 to produce the cancer-associated Tn (GalNAc-Ser/Thr) glycoforms with different densities. For MUC6 recombinant fragment a mixture of the three enzymes was used, but the different GalNAc-transferases were also used individually and not as an enzyme mixture. This strategy allows the production of different *O*-glycan attachment patterns (*37*), this way creating different epitopes, specific for each GalNAc-T activity.

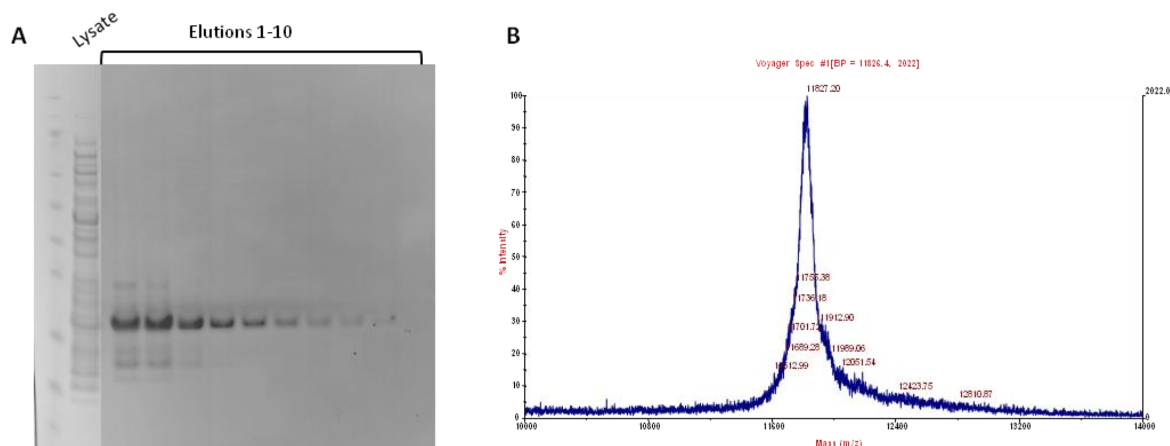

Figure 5 - Recombinant fragment of MUC6 is shown as example of the recombinant mucin proteins production. Similar results were obtained with the remaining mucin fragments. Expression cultures were NiNTA and analyzed by Coomassie PAGE (A), HPLC purified followed by MALDI-TOF mass spectrometry analysis (B).

GalNAc attachments in recombinant protein sequences were strictly controlled as indicated by mass spectrometry. The analysis of MALDI-TOF profiles showed that GalNAc-T2 added thirty five GalNAc residues to recombinant MUC5AC and nineteen GalNAc residues to recombinant MUC2. GalNAc-T3 was used to add sixty eight GalNAc residues to recombinant MUC5AC, twenty seven GalNAc residues to recombinant MUC2 and twenty GalNAc residues to recombinant MUC6. To extend the occupancie of GalNAc

residues, a mixture of GalNAc-T1, GalNAc-T2 and GalNAc-T3 resultued on the addition of twenty four GalNAc residues to recombinant MUC6. Also, sequential use of GalNAc-T3 and GalNAc-T12 added eighty six GalNAc residues to recombinant MUC5Ac. GalNAc-T3 glycosylation of 19-mer- MUC2 and MUC5Ac peptides resulted in addition of seven GalNAc residues on both peptides.

Further glycosylation of GalNAc residues with sialic acid to form STn was achieved with recombinant murine ST6GalNAc-I. Evaluation of number of sialic acid residues attached was made by MALDI-TOF analysis and may be underestimated because of the labile nature of this sugar linkage. However, some residual Tn reactivity was detected in sialylated glycosylation products, which was expected due to the very high density of Tn structures on the recombinant proteins. Purification of each glycosylation product followed by re-glycosylation with ST6GalNAc-I did not prevent such Tn exposure suggesting that the incomplete Tn-elongation was due to accumulation of by-products and impurities during *in vitro* glycosylation.

Figure 6- Biosynthetic pathways of the initial steps in the most common cancer-associated mucin-type O-glycans with indication of glycosyltransferases responsible for each synthetic step (A). Adapted from (*38*). MALDI-TOF profiles represent the chemoenzymatic synthesis of some Tn and STn mucin recombinant proteins. MUC5AC recombinant fragment was glycosylated using GalNAc-T3 with addition of sixty eight GalNAc residues (B). Recombinant MUC2 was glycosylated with GalNAc-T2 with addition of nineteen GalNAc residues. The Tn MUC2 glycoform was further elongated to make cancer associated glycoform STn (C).

The Tn glycoforms were further elongated by other glycosyltransferases to make cancer-associated glycoforms core 3 (GlcNAcβ1-3GalNAcα-O-Ser/Thr), T (Galβ1-3 GalNAc-Ser/Thr) and Lewis c (Galβ1-3GlcNAcβ1-3GalNAcα-O-Ser/Thr), using β3GlcNAc-T6, core 1 β3GalT and β3Gal-T5, respectively.

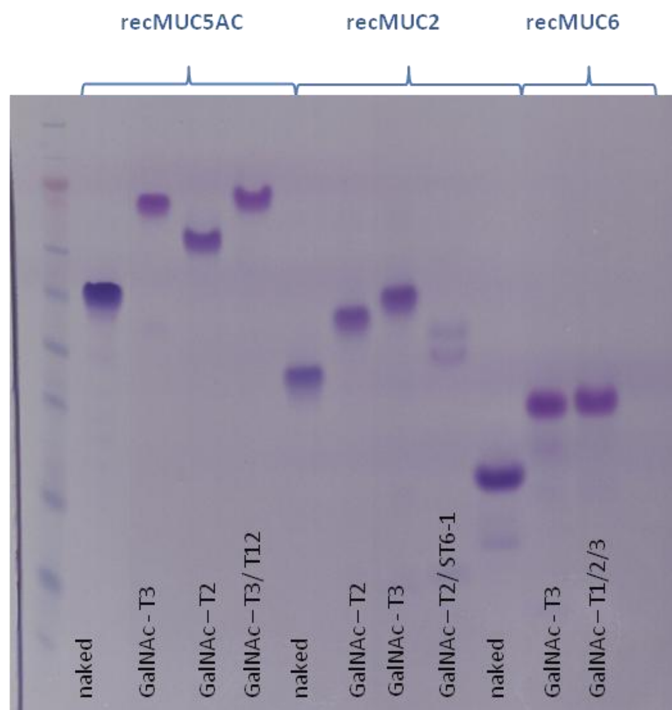

Figure 7 - Coomassie SDS - PAGE analysis of recombinant (rec) mucin proteins produced in *E. coli* and *in vitro* glycosylated using purified recombinant GalNAc-transferase T1, T2, T3 and T12 and sialyltransferase ST6 GalNAc-I, producing a panel of different mucin glycoforms.

2. Mucin *O*-glycopeptide array analysis

The comprehensive glycopeptide mucin array was used to probe for serum autoantibodies directed to glycopeptides epitopes in mucins known to be expressed in gastric cancer.

First, quality and consistency of glycopeptides libraries printed on slides were evaluated by lectins and monoclonal antibodies to defined structures (Figure 8). These tests were also made in order to define the correct peptides concentration printed on the array.

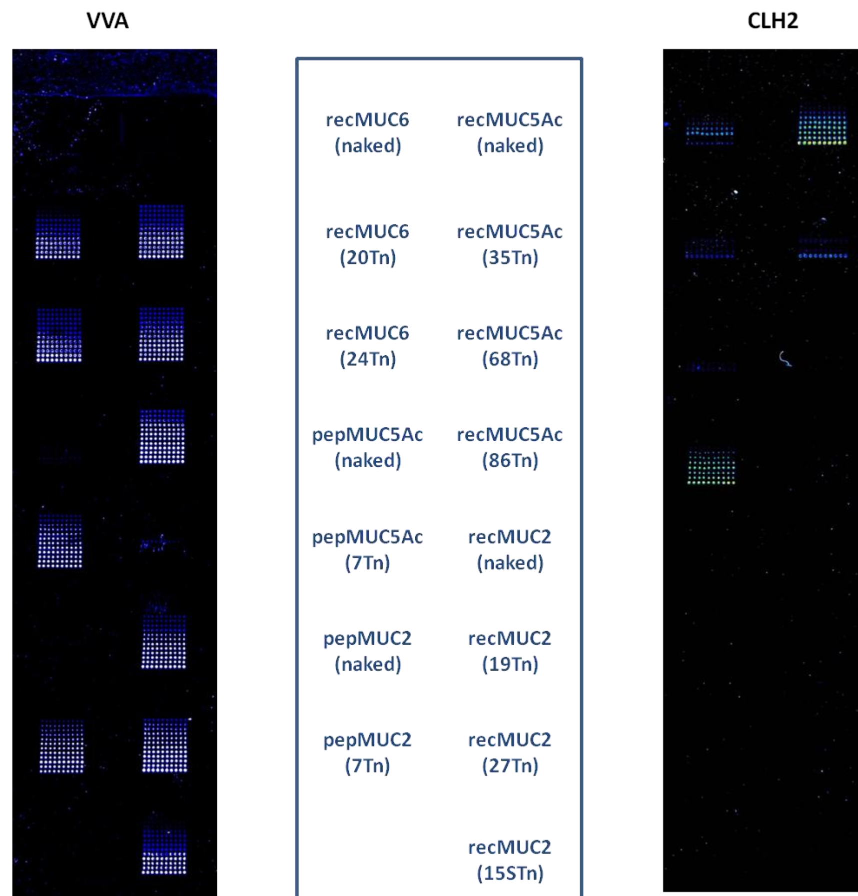

Figure 8 - Examples of fluorescent images of glycopeptide array test with bound monoclonal antibody CLH2 specific for the peptide tandem repeat of MUC5Ac and lectin VVA specific for carbohydrate structure Tn (GalNAc). In the center is the schematic localization of the peptides. Each peptide has a serial of double dilutions in order to define a proper print concentration.

The mucin *O*-glycopeptide array was then probed for serum IgG autoantibodies. The array was analyzed with sera from individuals with normal gastric mucosa without *Helicobacter pylori* infection confirmed by biopsie (n=8); disease controls from patients with chronic gastritis/ *Helicobacter pylori* infection (n=8) and patients with gastric mucosa atrophy / Intestinal metaplasia (n=9), and sera from gastric cancer patients (n=10). Specific IgG autoantibodies against different peptides and glycoforms epitopes were identified in the different sets of sera (Figure 9). Variable titers of antibodies from 1:5 to 1:50 (not shown) were obtained for the small peptides which reactivity values were significantly decreased with a 1:50 serum dilution. This may be due to lower titers of serum autoantibodies for these small peptides, which this way can only be found in lower serum diltions.

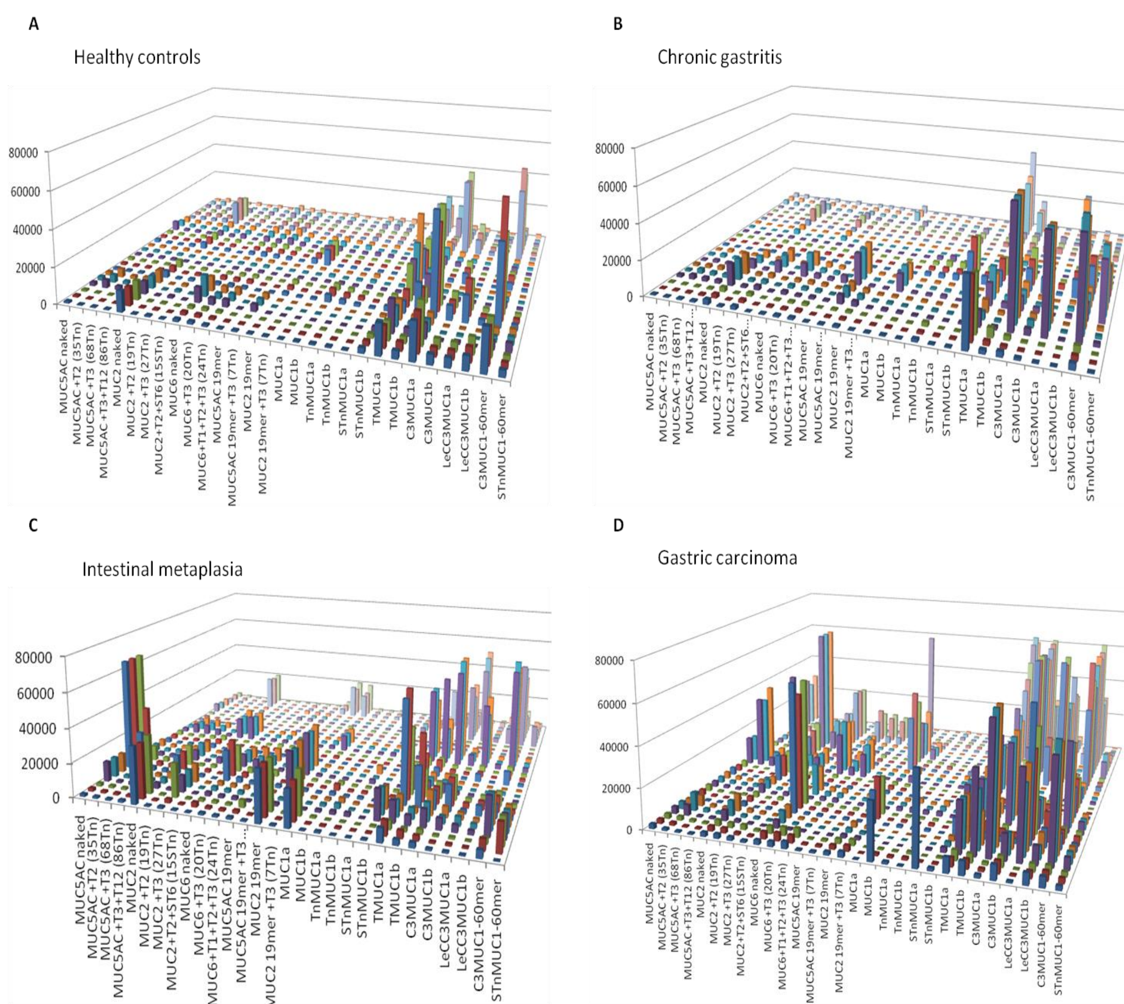

Figure 9- IgG autoantibodies to different peptides in different groups of analyzed serum. Quantification of the antibody responses for glycopeptide targets in each individual is shown in a 3-dimensional bar graph for (A) healthy controls (n=8), (B) chronic gastritis patients (n=8), (C) Intestinal metaplasia patients (n=9), and (D) gastric cancer patients (n=10). Bars indicate the IgG reactivity in serum from each patient in triplicates (y-axis) on one compound (x-axis) expressed as relative fluorescence units (z-axis). Serum was dltuted 1:5.

When analyzing the different glycoforms specific IgG autoantibodies, we assumed positivity for patients with a relative fluorescent value higher than 3 standard deviation over the mean of the control group. In general, we can find a larger reactivity of gastric carcinoma patients serum autoantibodies against the different mucin forms than the ones from control individuals.

Results for MUC5Ac mucin (Figure 10) revealed that all control individuals tested negative or had very low levels of reactivity. The MUC5Ac naked form detected 2/10 of the carcinoma patients and 5/9 of the intestinal metaplasia patiensts. The recombinant MUC5Ac + GalNAc-T2 form detected 1/8 of gastritis patients, 4/9 of intestinal metaplasia and 5/10 of carcinoma patients. Recombinant MUC5Ac + GalNAc-T3 glycoform detected 3/9 of intestinal metaplasia and 3/10 of carcinoma patients. Recombinant MUC5Ac + GalNAc-T3 + GalNAc-T12 form detected 1/9 of intestinal metaplasia and 1/10 of

carcinoma patients. As for the 19mer MUC5Ac peptide, the plot distribution is different from recombinant mucin. This may suggest that some autoantibodies detected for the recombinant mucin react with an epitope that is not present on the smaller pepide. There is also occurrence of some samples reactivity for the 19mer peptide and not in the larger protein, wich may indicate a reactivity with either N- or C-terminal of the peptide. The naked peptide detected 2/9 intestinal metaplasia patients and GalNAc-T3 glycosylated peptide detected 2/8 gastritis patients and 2/9 intestinal metaplasia patients. However, some serum samples were broadly reactive to different MUC5Ac glycoforms. For instance, a serum sample (number 6 on the DOT-PLOT) from one cancer patient was reactive with naked mucin and with all glycoforms, with decreasing activity with higher sugar content, suggesting the epitope is, in fact, on the mucin backbone. This also occurs on some intestinal metaplasia serum samples. However, there are four carcinoma patients serum that have autoantibodies reactive only with the glycosylated forms with one of them being specific for the recombinant MUC5Ac + GalNAc-T2 and not reactive with any other mucins like MUC2 and MUC6 naked- or glycoforms.

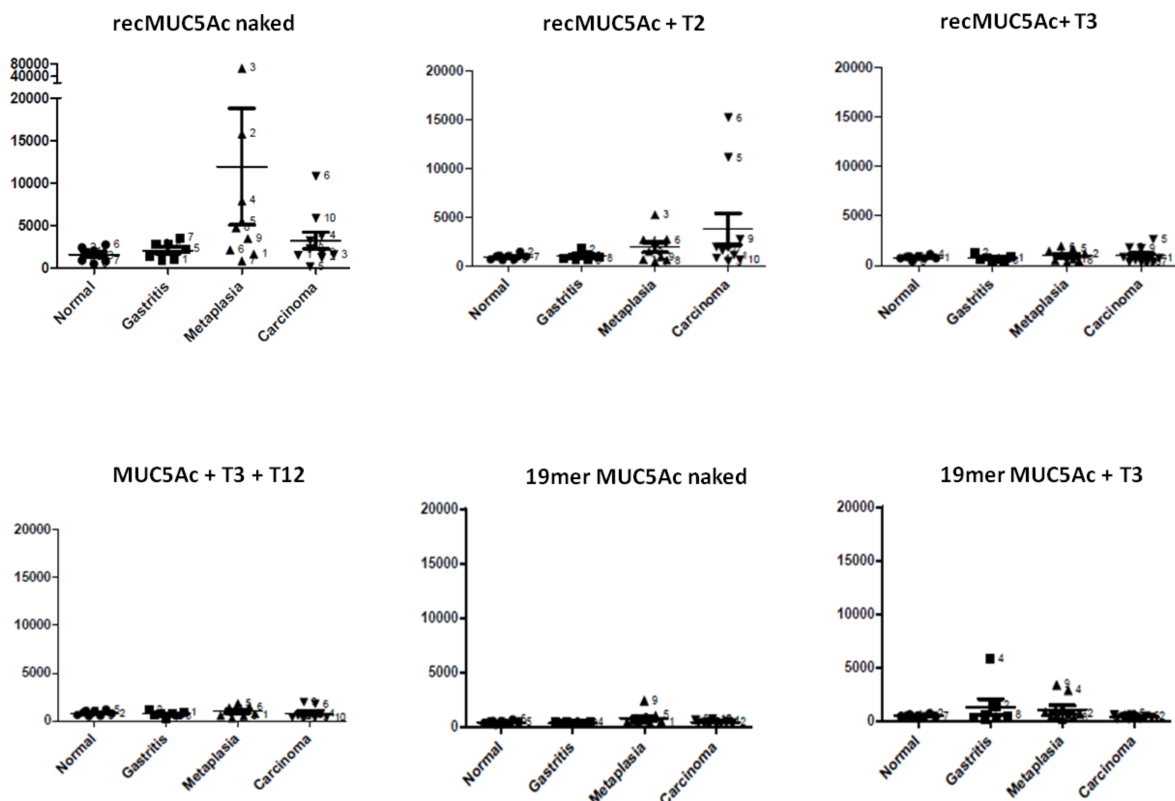

Figure 10 - MUC5Ac target glycoforms of 19mer peptide (19mer MUC5Ac) and recombinant protein (recMUC5Ac) are shown as DOT-PLOT diagram. The different glycoforms were obtained using enzymes GalNAc-T2 (+T2), GalNAc-T3 (+T3) and GalNAc-T12 (+T12). The non-gycosylated (naked) forms are also analyzed. Each dot represents the relative fluorescent unit for each individual in the four groups of patients examined: healthy controls (n=8), chronic gastritis patients (n=8), intestinal metaplasia patients (n=9), and gastric cancer patients (n=10). Patients with relative fluorescent value higher than 3 standard deviation over the mean of the control group were considered as positive. Serum was diluted 1:50.

When analyzing the results concerning the MUC2 mucin (Figure 11) we also see a pattern that indicates a higher reactivity for gastric carcinoma serum samples than for healthy controls. In fact, 6/10 different individuals presented autoantibodies for any of MUC2 naked and glycol forms. Nevertheless, taking to account all mucins glycoforms we don't see any glycoform specificity.

However, there are some interesting MUC2 mucin results regarding intestinal metaplasia condition, with 3/9 positive results in which two samples autoantibodies are specific for the GalNAc-T3 glycosylated form.

No significant reactivity was found for the STn glycoform.

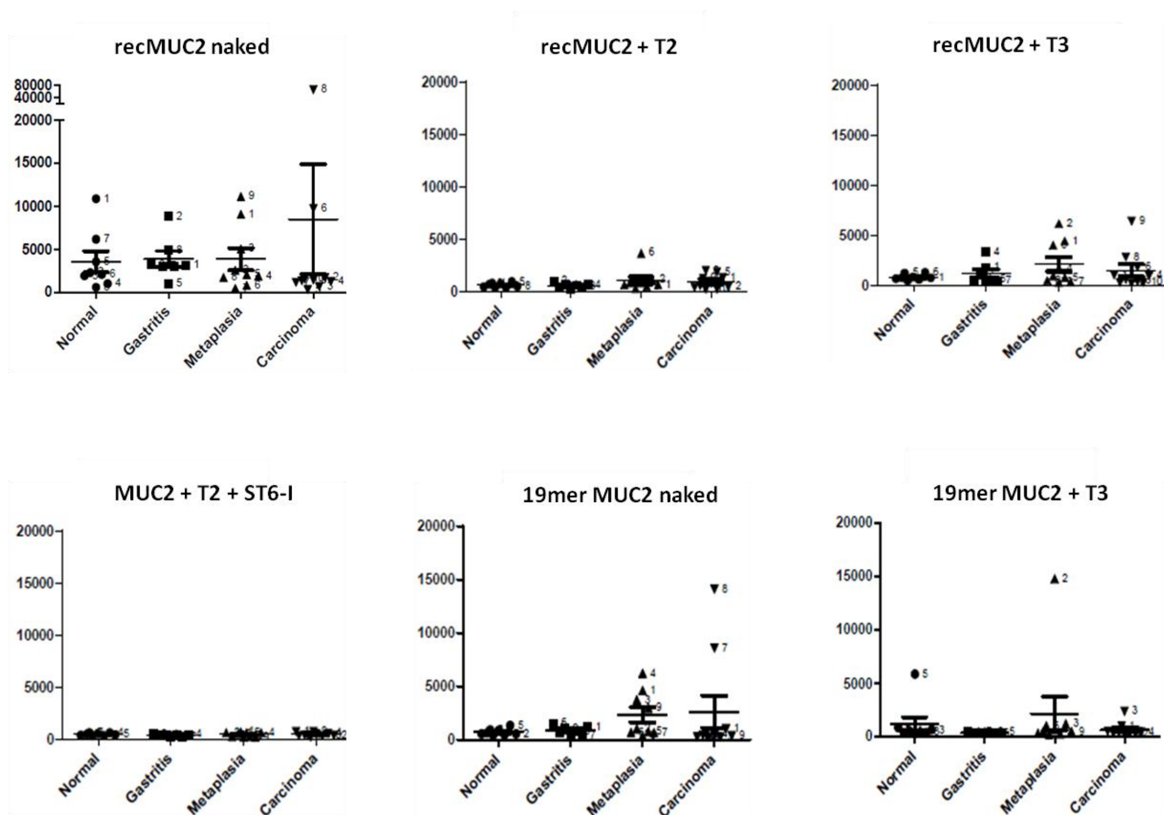

Figure 11- MUC2 target glycoforms of 19mer peptide (19mer MUC2) and recombinant protein (recMUC2) are shown as DOT-PLOT diagram. The different glycoforms were obtained using enzymes GalNAc-T2 (+T2), GalNAc-T3 (+T3) and ST6GalNAc-I (ST6-I). The non-gycosylated (naked) forms are also analyzed. Each dot represents the relative fluorescent unit for each individual in the four groups of patients examined: healthy controls (n=8), chronic gastritis patients (n=8), intestinal metaplasia patients (n=9), and gastric cancer patients (n=10). Patients with relative fluorescent value higher than 3 standard deviation over the mean of the control group were considered as positive. Serum was diluted 1:50.

As for the MUC6 recombinant protein (Figure 12), the MUC6 naked form detected 2/8 of the gastritis patients, 1/9 of the intestinal metaplasia patiensts, and 4/10 of carcinoma

patients. Of this four carcinoma patients serum, two of them are also reactive with the glycosylated forms, which indicates a backbone protein epitope.

The recombinant MUC6 glycoforms detected 4/10 carcinoma patients, with no specific reactivity for any of the two different glycoforms, and that were also reactive for other mucin glycoform, namely a MUC2 glycoform.

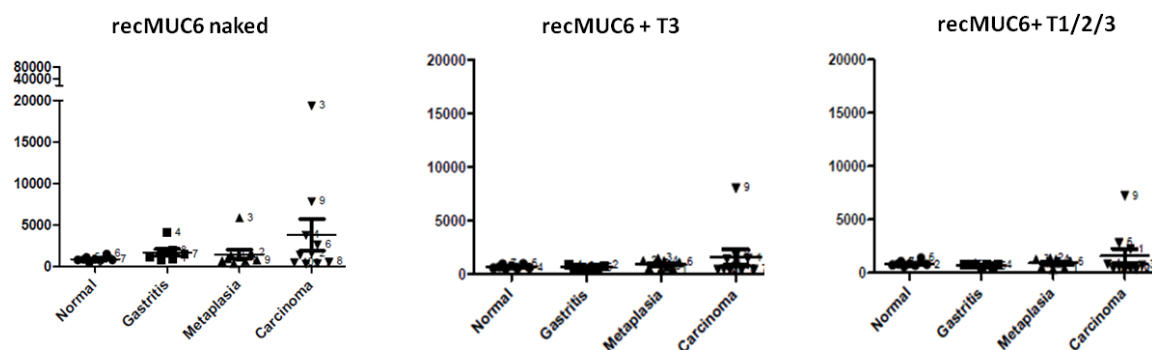

Figure 12 - MUC6 target glycoforms recombinant protein (recMUC6) are shown as DOT-PLOT diagram. The different glycoforms were obtained using enzymes GalNAc-T3 (+T3) and a mixture of GalNAc-T1, GalNAc-T2 and GalNAc-T3 (+T1/2/3). The non-gycosylated (naked) form is also analyzed. Each dot represents the relative fluorescent unit for each individual in the four groups of patients examined: healthy controls (n=8), chronic gastritis patients (n=8), intestinal metaplasia patients (n=9), and gastric cancer patients (n=10). Patients with relative fluorescent value higher than 3 standard deviation over the mean of the control group were considered as positive. Serum was diluted 1:50.

Lewis c-MUC1 was the most selective of the MUC1 targets detecting 5/10 of the cancer patients, while STn-MUC1 detected 1/10 and T-MUC1 detected 3/10 (Figure12). Core 3-MUC1 peptides haven't shown any significant carcinoma samples positivity. In fact, there is broader reactivity with this structure between all the groups analyzed. An inherent problem with many cancer biomarkers identified to date is precisely the low discrimination between cancer and inflammatory conditions. Nevertheless, there is a higher number of carcinoma patients serum with high reactivity than the number of healthy controls or gastritis and intestinal metaplasia conditions reactive samples. No significant reactivity for Tn-MUC1 was found for any of the conditions.

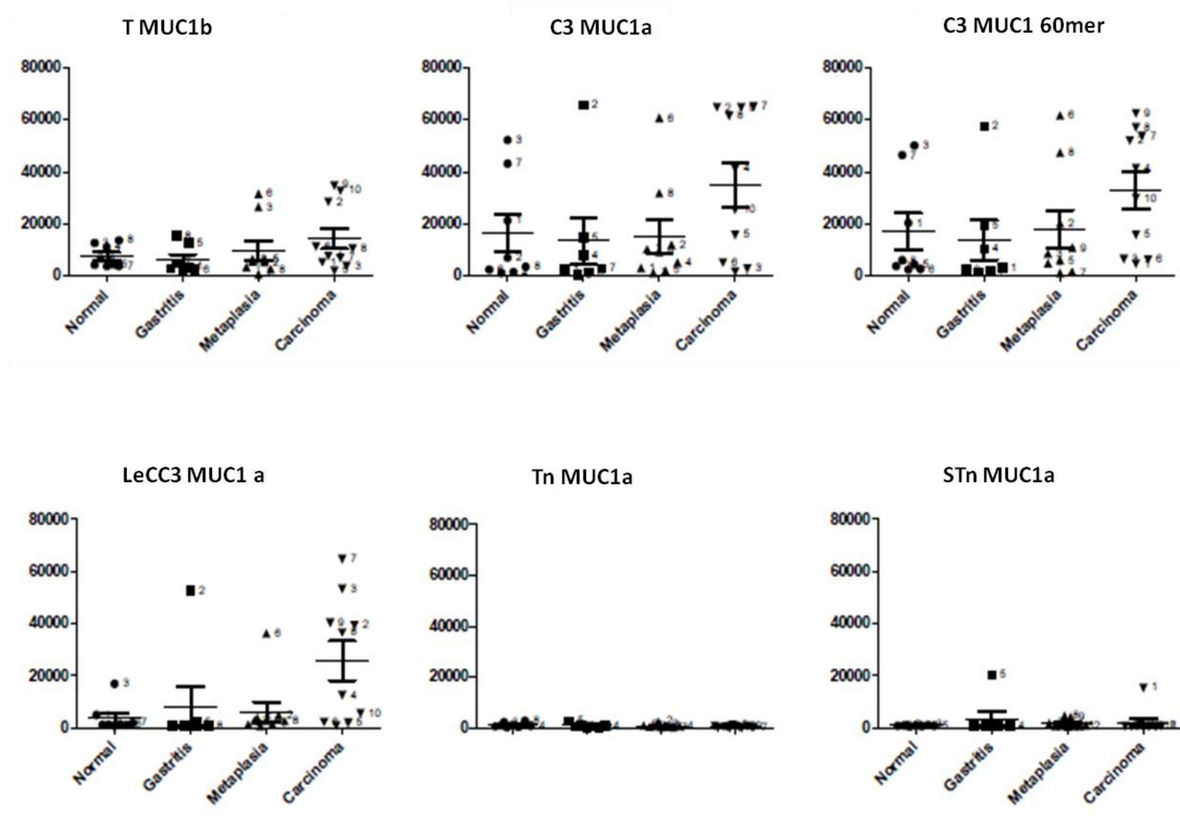

Figure 13 - DOT-PLOT diagrams presenting autoantibody reactivity against core 3 (C3), T, Tn, LewisC (LeCC3) and STn MUC1 peptides. Each dot represents the relative fluorescent unit for each individual in the four groups of patients examined: healthy controls (n=8), chronic gastritis patients (n=8), intestinal metaplasia patients (n=9), and gastric cancer patients (n=10). Patients with relative fluorescent value higher than 3 standard deviation over the mean of the control group were considered as positive. Serum was diluted 1:5.

Discussion

Early cancer detection remains one of the most important goals in cancer diagnosis. Circulating autoantibodies appear to be ideal serological biomarkers because they generally represent amplified signals to low antigen exposure, they have long circulation time, and methods for detection are simple and sensitive provided the target antigens are available. Autoantibodies are induced by exposure to cancer neoantigens and are emerging as promising biomarkers for the early detection of cancer (*79, 101-103*). The mechanism of induction of cancer immunity is not fully understood. However, it is generally accepted that auto-reactive immune responses in cancer patients are induced by overexpression, mutations, missfolding, or aberrant degradation of self-proteins (*104-106*). Aberrant PTMs of self-proteins have also been demonstrated to induce autoantibody production in cancer and a number of autoimmune diseases (*107*). Previous studies suggest that cancer-associated aberrant *O*-glycosylation of proteins such as truncated Tn, STn and truncated core 3 structures leads to the induction of autoantibodies. The development of such autoantibodies would be explained by the *de novo* expression of the aberrant immunogenic glycopeptide epitopes in cancer cells (*94, 108*).

Autoantibody assays have not yet found wide use for diagnosis of cancer and this may be due to lack of well-defined antigens to probe for disease-associated autoantibodies (*101,109*). In order to evaluate and employ such a strategy, the prevalence that the various structures appear and the specific proteins carrying such structures need to be better characterized.

Previous glycoprofiling results using monoclonal antibodies with well-defined specificities showed that circulating MUC1 and MUC16 does not include nonsialylated glycoforms and opens for the possibility that circulating autoantibodies directed to such glycoforms participate in their removal in addition to lectin receptors (*110*). This is in agreement with Varki and coworkers' (*111*) "tip of the iceberg" theory, proposing that circulating tumor glycoprotein antigens, including cancer mucins without sialic acids, are removed by scavenger receptors and hence inherently represent insensitive biomarkers. This selective clearance provides a plausible mechanism by which aberrant glycoforms of mucins and *O*-glycoproteins are presented to and stimulate the immune system (*112*).

The primary aim of this study was to develop a discovery platform for selective detection of cancer-associated autoantibodies directed to aberrant *O*-glycopeptide epitopes. This should happen without interference by anticarbohydrate hapten antibodies. Detection of autoantibodies to *O*-glycopeptide combination epitopes, which include both the peptide backbone as well as cancer-associated posttranslational modification hapten structures, is complicated by the presence of rather ubiquitous natural antibodies to the carbohydrate

haptens. It is well established that non-self carbohydrate structures are immunogenic in man, and natural antibodies to the cancer-associated truncated carbohydrate haptens T, Tn, and STn are among the most widely studied antibodies (*58*). Although titers of antibodies to Tn and STn appear to be increased in cancer patients (*23, 41*), essentially all individuals have IgM antibodies with some specificity to these structures. Analysis of IgM antibodies reveals broad reactivity with all glycoforms of the glycopeptides, which makes it impossible to discern potential *O*-glycopeptide–specific antibodies. It has been shown that some anti-carbohydrate antibodies result from exposure to related structures found in gut microflora; however, expression of aberrant glycosylated glycoproteins in cancer are believed to contribute as well (*58*). The clinical significance of these antibodies in relation to cancer is largely unknown, and they have not emerged as truly useful biomarkers of cancers. In contrast, antibodies directed to combined *O*-glycopeptide epitopes, which would have specificity for distinct proteins carrying aberrant cancer *O*-glycans, are expected to have high affinity, be class switched to IgG, and be absent in healthy individuals (*94*). Thus, the IgM carbohydrate hapten antibodies don't allow to discern antibodies with distinct *O*-glycopeptide specificity. This fact stresses the importance of using assay methods that allow selective detection of combined glycopeptide epitopes. Previous results confirmed that IgM antibodies yielded broad carbohydrate reactivity without discernable *O*-glycopeptide specificities in both healthy controls and cancer patients. In striking contrast, IgG antibodies from healthy control sera were essentially unreactive with the glycopeptides, whereas specific IgG antibodies were identified in sera from newly diagnosed breast, ovarian and prostate cancer patients (*78*).

For this reason we limited the detection to IgG subclass, trying to exclude carbohydrate hapten antibodies interference.

Two strategies applicable to production of compound libraries were used: i) Chemo-enzymatic synthesis of short 15-60-mer glycopeptides, and ii) Recombinant *E.coli* production of larger mucin polypeptide fusion proteins derived from MUC2, MUC5Ac, and MUC6 combined with enzymatic *O*-linked glycosylation to produce the cancer-associated *O*-glycans Tn, STn, truncated core 3 (GlcNAcβ1-3GalNAcα-O-Ser/Thr) and Lewis c (Galβ1-3GlcNAcβ1-3GalNAcα-O-Ser/Thr) with different densities.

The use of recombinant proteins allow for a larger probability of exposing the desired epitope and also allow for some specific conformation only found in larger peptides. Both strategies used a panel of recombinant UDP-GalNAc: polypeptide α-N-acetylgalactosaminyltransferases (GalNAc-Ts) with partially different substrate specificities to produce different *O*-glycan attachment patterns. Furthermore, we used a

toolbox of recombinant glycosyltransferase enzymes that allow enzymatic synthesis of diverse *O*-glycan structures with some degree of control of sites of attachment of *O*-glycans at serine and threonine residues in peptides (*40, 113, 114*). The Tn-glycoforms were elongated with sialyltransferase ST6GalNAc-I, β3GlcNAc-T6 and β3Gal-T5 resulting in production of STn, core 3 and Lewis c glycoforms, respectively (*108*). The different glycoforms were verified by MALDI-TOF mass spectrometry and SDS-PAGE.

To accommodate large glycopeptide libraries, we chose to use a slide microarray format on NHS-activated hydrogel slides, for minimum consumption of compounds, and as Blixt and colleagues have shown (*93*), these provide a remarkably low background for detection of human serum antibodies.

Quality and consistency of glycopeptides libraries printed on slides were evaluated by lectins and monoclonal antibodies to defined structures. The comprehensive glycopeptide mucin array was then used to probe for serum autoantibodies directed to glycopeptides epitopes in mucins.

The detection of antibodies reactive with recombinant MUC5Ac and MUC6 were present in sera from gastric carcinoma patients, individuals with intestinal metaplasia and with gastritis. On the other hand, autoantibodies from healthy controls serum presented much lower or null reactivity. Mucins MUC5Ac and MUC6 are characteristically expressed by normal gastric mucosa. Their level of expression and their glycosylation have been shown to be altered during intestinal metaplasia. It is conceivable that in some patients these mucins may be exposed to the immune system raising a consequent immune response. In addition the expression of MUC5Ac has been shown to be a characteristic of every early gastric carcinoma, regardless of the histologic type of the tumors, in contrast with the absence of MUC5Ac expression in almost half of the advanced carcinomas, and suggests that all gastric carcinomas retain at least some cells with a "gastric" phenotype during tumour progression (*25*). The exposure by tumor cells of these underglycosylated mucins leads to generation of autoantibodies as detected in our study.

The evaluation of recombinant naked MUC2, however, presents a broader reactivity among the different serum groups. Curiously, this broader reactivity was not represented with MUC2 smaller peptide, suggesting that some autoantibodies detected for the recombinant mucin reacted with an epitope that is not present on the smaller peptide. There was also occurrence of some samples reactivity for the 19mer peptide and not in the larger protein, which may indicate reactivity with either N- or C-terminal of the peptide. We can also see four positive sera (4/9_44%) for MUC2 naked peptide in intestinal metaplasia cases.

The human intestinal mucin MUC2 is a major product of goblet cells of the colon and small intestine. Taking into account that in intestinal metaplasia of the stomach MUC2 expression is closely associated with the goblet cell phenotype (*87, 88*), these results are then consistent with the detection of intestinal metaplasia. The presence of antibodies directed to this intestinal mucin MUC2 in sera from gastric carcinoma patients is also observed in the present study, detected in nearly 60% of gastric carcinomas samples.

Previous studies have shown that MUC2 also serves as the major carrier of the sialyl-Tn antigen in IM and gastric carcinomas (*115*). In the present study we did not detect any autoantibody reactivity for the STn MUC2 glycoform in this assay. The lack of such reactivity may stem from the high level of occupancy of the tandem repeat. We cannot exclude that an alternative chemoenzymatic strategy leading to a moderate level of occupancy of the MUC2 tandem repeat with STn antigen may generate the required epitope for the autoantibodies.

The sialyl-Tn antigen (Neu5Acα2-6GalNAc-*O*-Ser/Thr), also known as STn and CD175s, is a simple mucin-type carbohydrate antigen which has attracted much attention because it is highly expressed in most gastric (*44*), colorectal (*45*), ovarian (*46*), breast (*47*) and pancreatic (*48*) carcinomas, whereas no expression is observed in the respective normal tissues. In fact, a common feature in carcinoma cells is the activation of an alternative pathway corresponding to the early α2,6-sialylation of GalNAcα-*O*-Ser/Thr leading to the formation of the sialyl-Tn antigen. The genetic and biosynthetic basis for expression of STn glycans may relate to somatic mutations in Cosmc, a chaperone for the enzyme C1β3Gal-T enzyme that controls synthesis of the common core 1 O-glycosylation pathway ; however, overexpression of the ST6GalNAc-I sialyltransferase responsible for STn synthesis can also play a role (*51, 52*).

It will be interesting to evaluate if other mucins, such as MUC5Ac and MUC6, can also carry the sialyl-Tn antigen in gastric carcinoma cases. It will also be relevant to assess if there is a different biological significance of sialyl-Tn carried by different apomucins or if the observed role played by sialyl-Tn antigen in cancer-cell invasion is independent from the mucin carrier.

In gastric tissues, glycosylation changes, also include exposure of simple mucin-type carbohydrate antigens such as T and Tn, and have been shown in premalignant IM lesions and gastric carcinomas (*23*). When we analyze the MUC2 Tn glycoform we find two serum samples with specific reactivity for the mucin glycosylated with GalNAc-T3 in intestinal metaplasia condition. This is observed both in the recombinant and peptide structures. It means that these autoantibodies are specific for an epitope present only with

GalNAc-T3 sites occupancy. This is also observed in a serum sample from a carcinoma patient this time for the recombinant MUC5Ac+ GalNAc-T2 glycoform, and proves the interest in applying this single enzyme glycosylation approach against an enzyme mixture strategy. It also confirms MUC5Ac as a plausible carrier of carbohydrate antigens.

We have also detected IgG antibodies to peptide epitopes. The antibodies found to react with both naked and glycosylated mucin are directed to the core protein as their reactivity is more intense for the unglycosylated structure, with decreased reactivity whit increasing sugar content.

The present study evaluate sera in two diferent dilutions (1:5 and 1:50), demonstrating quite variable titers of antibodies. This variability of titers was particularly observed for the MUC1 smaller peptides. In most sera the autoantibody reactivity was reduced with the 1:50 dilution. This observation may result from a smaller titer of autoantibodies for these structures.

Nevertheless, when we analyze 1:5 dilution results we see some cancer-specific reactivities for Lewis c MUC1 glycoforms with a approximated 50% positivity. As for the core 3 structure we observe a broader reactivity between normal individuals and benign conditions.

Three distinct MUC1 *O*-glycopeptide epitopes with different *O*-glycan structures (Tn, STn, and core 3) reactive with IgG were identified before in prostate, breast, and ovarian cancer patients (*78*) . We did not detect significant reactivity for the Tn and STn glycoforms in this assay. More interesting is the common finding of antibodies to truncated core 3 MUC1.

The core 3 glycosylation pathway is generally believed to be limited to the mucosa of the digestive tract, and truncated core 3 *O*-glycans have been found in colon cancer (*116*) in accordance with the expression pattern for the controlling enzyme β3GnT6 (*68*).

Core 3 is restricted in its occurrence to mucins from specialized tissues and it was first identified in *O*-glycans derived from stomach and colon. Core 3 synthase is an important enzyme in the synthesis of mucn-type *O*-glycans in digestive organs. On immunohistochemical analysis, β3GlcNAc-T6 was detected in the Golgi region of both normal gastric and colorectal epithelial cells. It is of interest that, in the stomach, it was restricted to the foveolar epithelia. This finding indicated that mucins produced by foveolar epithelia cells contain core 3 structures, but mucins derived from subglandular epithelia do not. It was demonstrated that the expression of β3GlcNAc-T6 dramatically decreased in cancerous tissues, and the changes to the core structures give rise to phenotypic changes

of cancer cells, such as mobility and invasion phenotype. In fact, in contrast to normal tissues, β3GlcNAc-T6 is not detected in both gastric and colorectal cancer. Structural analysis of *O*-glycans released from mucins isolated from colorectal cancer are in agreement with this revealed changes in branching and elongation of *O*-glycans, including the presence of truncated core 3 *O*-glycans terminated with βGlcNAc (*116*) similarly to the structure targeted by glycopeptide specific autoantibodies identified in the present study.

The many enzymes involved in synthesizing the core structure of *O*-glycan may compete with each other for the GalNAc-residue on mucins as substrate.

Core 1 and core 2 structures can appear due to the disappearance of one of its competitors, the core 3 synthase, in cancerous tissues. The loss of core 3 synthase may also play a role for the appearance of STn-MUC1 in colorectal cancer, since the core 3 synthase competes for the Tn substrate utilized by ST6GalNAc-I to produce STn glycoforms. However, we didn't detect reactivity for this structure.

The results with the Lewis c (Galβ1-3GlcNAcβ1-3GalNAcα-O-Ser/Thr) also seem promising, as they reveal agreater carcinoma specificity. Type 1 chain oligosaccharides found in *O*-glycans, as well as in glycolipids, contain the distinctive Galβ1-3GlcNAc disaccharide as their core structure. It is synthesized by β1,3- galactosyltransferases (β1,3Gal-Ts), a family of enzymes whose genes have been cloned (1), with β3Gal-T5 exhibiting a marked preference for the *O*-linked core3 (*117*). The functional role of type 1 chains is not known, but several studies have indicated that some of them are differentially expressed in cancer (*118*). However, it would be important to further confirm the results we obtained in this study and it should also be interesting to apply this glyco-structure to other mucins in a future work.

In the present study, informative glycopeptide epitopes were found in few structures, however the library only includes a small part of the protein core of these mucins and only certain variations of *O*-glycan patterns and structures.

It is very important to understand the timing and dynamics of such autoantibodies. It was found in previous studies that antibodies may disappear as the disease progresses and/or after treatment. With tumor progression, immunosupression may occur during the course of disease. These patients may be unable to respond to the antigens on the tumor, thus antibody levels may not correlate with clinicopathologic parameters.

This would be in agreement with studies of p53 autoantibodies that may reappear at relapse (*119*). Other studies of autoantibodies to p53 have shown that these may occur

before diagnosis of cancer (*120*), which provides support for using autoantibodies as sensitive and early biomarkers.

Nevertheless, the variation in autoantibody levels between patients is noteworthy and indicates important biological variations in antibody responses among individuals. First, there could be variations in amount of expressed antigen between individuals. Secondly, a subject's ability to recognize and present the different glycoforms may vary. Another factor potentially playing a role is variance in local stromal factors such as the secretion of TGF-β among other factors known to down-regulate the immune response to cancer targets causing immunological escape. It is also to take into account adsorption of circulating antibodies by large tumor and metastasis mass and possible immune complex formation which would render the antibodies undetectable.

The autoantibody microarray assay described here for the simple and non-invasive identification of patients with gastric cancer could be a significant advantage over current screening techniques by its detection of immunologically amplified signals present at early stages of the disease. In addition the autoantibody microarray assay may also contribute for the detection of other lesions of the gastric mucosa, such as intestinal metaplasia.

It is important to point out that these were preliminary results and we are confident that combining this strategy with a larger number of epitope targets from other *O*-glycoproteins altered in cancer, we will eventually reach acceptable levels of sensitivity.

The present study will also allow us to overcome one major limitation in the use of these biomarkers for cancer screening: an obvious need to increase sensitivity and organ specificity of biomarker assays.

For instance, it was already confirmed the existence of such cancer-associated autoantibodies to the aberrantly *O*-glycosylated MUC1 mucin in patients with breast, ovarian, and prostate cancer at time of diagnosis (*78*). In addition to MUC1, gastric carcinomas express MUC2, MUC5Ac and MUC6 with more restricted organ- and cancer-type expression patterns (*25, 87*). By focusing in these gastric-related molecules, the identification of cancer-specific glycoforms of these mucins by autoantibodies would create a cancer-specific glycoform signature providing a powerful tool for the early diagnosis of gastric carcinoma and other gastric diseases.

The concept of autoantibody signatures can be extended to many other posttranslational protein modifications that may be altered in disease and may serve as biomarkers.

Conclusions

In this study we have detected IgG autoantibodies against different mucin proteins and glycoforms in serum from gastric carcinoma patients, whereas control individuals tested negative or had very low levels of reactivity.

We have detected autoantibodies for both core protein and glycoprotein epitopes and some of them revealed specific reactivity for epitopes generated with a single enzyme activity. This was the case of MUC2 glycosylated with GalNAc-T3 which generated a glycoform detected by autoantibodies in the sera from intestinal metaplasia patients. Similarly, GalNAc-T2 glycosylated MUC5Ac generated a glycoform detected by autoantibodies in gastric carcinoma patients. This supports our strategy of single enzyme sites occupancy.

Our approach of adding recombinant proteins in the *O*-glycopeptide array also revealed to be successful as many epitopes were only present in these fragments and not on the small peptides.

In the future, the expansion of the targets should improve the specificity and sensitivity of the strategy. For instance, glycoforms structures like Lewis c found in MUC1, for example, could be applied for other mucins.

References

1. Kim, Y. S., Gum, J., Jr., and Brockhausen, I. (1996) Mucin glycoproteins in neoplasia, *Glycoconj J 13*, 693-707.

2. Forstner, J. F. (1978) Intestinal mucins in health and disease, *Digestion 17*, 234-263.

3. Van Klinken, B. J., Dekker, J., Buller, H. A., and Einerhand, A. W. (1995) Mucin gene structure and expression: protection vs. adhesion, *Am J Physiol 269*, G613-627.

4. Hattrup, C. L., and Gendler, S. J. (2008) Structure and function of the cell surface (tethered) mucins, *Annu Rev Physiol 70*, 431-457.

5. Carlstedt, I., Sheehan, J. K., Corfield, A. P., and Gallagher, J. T. (1985) Mucous glycoproteins: a gel of a problem, *Essays Biochem 20*, 40-76.

6. Varki, A. (1993) Biological roles of oligosaccharides: all of the theories are correct, *Glycobiology 3*, 97-130.

7. Gum, J. R., Byrd, J. C., Hicks, J. W., Toribara, N. W., Lamport, D. T., and Kim, Y. S. (1989) Molecular cloning of human intestinal mucin cDNAs. Sequence analysis and evidence for genetic polymorphism, *J Biol Chem 264*, 6480-6487.

8. Gendler, S. J., Lancaster, C. A., Taylor-Papadimitriou, J., Duhig, T., Peat, N., Burchell, J., Pemberton, L., Lalani, E. N., and Wilson, D. (1990) Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin, *J Biol Chem 265*, 15286-15293.

9. Lan, M. S., Batra, S. K., Qi, W. N., Metzgar, R. S., and Hollingsworth, M. A. (1990) Cloning and sequencing of a human pancreatic tumor mucin cDNA, *J Biol Chem 265*, 15294-15299.

10. Ligtenberg, M. J., Vos, H. L., Gennissen, A. M., and Hilkens, J. (1990) Episialin, a carcinoma-associated mucin, is generated by a polymorphic gene encoding splice variants with alternative amino termini, *J Biol Chem 265*, 5573-5578.

11. Wreschner, D. H., Hareuveni, M., Tsarfaty, I., Smorodinsky, N., Horev, J., Zaretsky, J., Kotkes, P., Weiss, M., Lathe, R., Dion, A., and et al. (1990) Human epithelial tumor antigen cDNA sequences. Differential splicing may generate multiple protein forms, *Eur J Biochem 189*, 463-473.

12. Porchet, N., Dufosse, J., Audie, J. P., Duperat, V. G., Perini, J. M., Nguyen, V. C., Degand, P., and Aubert, J. P. (1991) Structural features of the core proteins of human airway mucins ascertained by cDNA cloning, *Am Rev Respir Dis 144*, S15-18.

13. Bobek, L. A., Tsai, H., Biesbrock, A. R., and Levine, M. J. (1993) Molecular cloning, sequence, and specificity of expression of the gene encoding the low molecular weight human salivary mucin (MUC7), *J Biol Chem 268*, 20563-20569.

14. Dufosse, J., Porchet, N., Audie, J. P., Guyonnet Duperat, V., Laine, A., Van-Seuningen, I., Marrakchi, S., Degand, P., and Aubert, J. P. (1993) Degenerate 87-

base-pair tandem repeats create hydrophilic/hydrophobic alternating domains in human mucin peptides mapped to 11p15, *Biochem J 293 (Pt 2)*, 329-337.

15. Toribara, N. W., Roberton, A. M., Ho, S. B., Kuo, W. L., Gum, E., Hicks, J. W., Gum, J. R., Jr., Byrd, J. C., Siddiki, B., and Kim, Y. S. (1993) Human gastric mucin. Identification of a unique species by expression cloning, *J Biol Chem 268*, 5879-5885.

16. Meezaman, D., Charles, P., Daskal, E., Polymeropoulos, M. H., Martin, B. M., and Rose, M. C. (1994) Cloning and analysis of cDNA encoding a major airway glycoprotein, human tracheobronchial mucin (MUC5), *J Biol Chem 269*, 12932-12939.

17. Lapensee, L., Paquette, Y., and Bleau, G. (1997) Allelic polymorphism and chromosomal localization of the human oviductin gene (MUC9), *Fertil Steril 68*, 702-708.

18. Shankar, V., Pichan, P., Eddy, R. L., Jr., Tonk, V., Nowak, N., Sait, S. N., Shows, T. B., Schultz, R. E., Gotway, G., Elkins, R. C., Gilmore, M. S., and Sachdev, G. P. (1997) Chromosomal localization of a human mucin gene (MUC8) and cloning of the cDNA corresponding to the carboxy terminus, *Am J Respir Cell Mol Biol 16*, 232-241.

19. Morgan, S. M., Samulowitz, U., Darley, L., Simmons, D. L., and Vestweber, D. (1999) Biochemical characterization and molecular cloning of a novel endothelial-specific sialomucin, *Blood 93*, 165-175.

20. Yin, B. W., and Lloyd, K. O. (2001) Molecular cloning of the CA125 ovarian cancer antigen: identification as a new mucin, MUC16, *J Biol Chem 276*, 27371-27375.

21. Gum, J. R., Jr., Hicks, J. W., Crawley, S. C., Dahl, C. M., Yang, S. C., Roberton, A. M., and Kim, Y. S. (2003) Initiation of transcription of the MUC3A human intestinal mucin from a TATA-less promoter and comparison with the MUC3B amino terminus, *J Biol Chem 278*, 49600-49609.

22. Itoh, Y., Kamata-Sakurai, M., Denda-Nagai, K., Nagai, S., Tsuiji, M., Ishii-Schrade, K., Okada, K., Goto, A., Fukayama, M., and Irimura, T. (2008) Identification and expression of human epiglycanin/MUC21: a novel transmembrane mucin, *Glycobiology 18*, 74-83.

23. David, L., Nesland, J. M., Clausen, H., Carneiro, F., and Sobrinho-Simoes, M. (1992) Simple mucin-type carbohydrate antigens (Tn, sialosyl-Tn and T) in gastric mucosa, carcinomas and metastases, *APMIS Suppl 27*, 162-172.

24. Reis, C. A., Osorio, H., Silva, L., Gomes, C., and David, L. (2010) Alterations in glycosylation as biomarkers for cancer detection, *J Clin Pathol 63*, 322-329.

25. Reis, C. A., David, L., Nielsen, P. A., Clausen, H., Mirgorodskaya, K., Roepstorff, P., and Sobrinho-Simoes, M. (1997) Immunohistochemical study of MUC5AC expression in human gastric carcinomas using a novel monoclonal antibody, *Int J Cancer 74*, 112-121.

26. Kotera, Y., Fontenot, J. D., Pecher, G., Metzgar, R. S., and Finn, O. J. (1994) Humoral immunity against a tandem repeat epitope of human mucin MUC-1 in

sera from breast, pancreatic, and colon cancer patients, *Cancer Res 54*, 2856-2860.

27. Denton, G., Sekowski, M., and Price, M. R. (1993) Induction of antibody responses to breast carcinoma associated mucins using synthetic peptide constructs as immunogens, *Cancer Lett 70*, 143-150.

28. Graham, R. A., Burchell, J. M., and Taylor-Papadimitriou, J. (1996) The polymorphic epithelial mucin: potential as an immunogen for a cancer vaccine, *Cancer Immunol Immunother 42*, 71-80.

29. Ding, L., Lalani, E. N., Reddish, M., Koganty, R., Wong, T., Samuel, J., Yacyshyn, M. B., Meikle, A., Fung, P. Y., Taylor-Papadimitriou, J., and et al. (1993) Immunogenicity of synthetic peptides related to the core peptide sequence encoded by the human MUC1 mucin gene: effect of immunization on the growth of murine mammary adenocarcinoma cells transfected with the human MUC1 gene, *Cancer Immunol Immunother 36*, 9-17.

30. Kim, Y. J., and Varki, A. (1997) Perspectives on the significance of altered glycosylation of glycoproteins in cancer, *Glycoconj J 14*, 569-576.

31. Brockhausen, I., Schutzbach, J., and Kuhns, W. (1998) Glycoproteins and their relationship to human disease, *Acta Anat (Basel) 161*, 36-78.

32. Hakomori, S. (2001) Tumor-associated carbohydrate antigens defining tumor malignancy: basis for development of anti-cancer vaccines, *Adv Exp Med Biol 491*, 369-402.

33. Hakomori, S. (2002) Glycosylation defining cancer malignancy: new wine in an old bottle, *Proc Natl Acad Sci U S A 99*, 10231-10233.

34. Brockhausen, I. (1999) Pathways of O-glycan biosynthesis in cancer cells, *Biochim Biophys Acta 1473*, 67-95.

35. Clausen, H., and Bennett, E. P. (1996) A family of UDP-GalNAc: polypeptide N-acetylgalactosaminyl-transferases control the initiation of mucin-type O-linked glycosylation, *Glycobiology 6*, 635-646.

36. Ten Hagen, K. G., Fritz, T. A., and Tabak, L. A. (2003) All in the family: the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases, *Glycobiology 13*, 1R-16R.

37. Hassan, H., Bennet, E. P., Mandel, U., Hollingsworth, M. A., Clausen, H., Ernst, Hart, and Sinay. (2000) O-Glycan Occupancy is Directed by Substrate Specificities of Polypeptide GalNAc-Transferases, In *Wiley-VCH chapter "Carbohydrates in Chemistry and Biology - a Comprehension Handbook"*, pp 273-292.

38. Tarp, M. A., and Clausen, H. (2008) Mucin-type O-glycosylation and its potential use in drug and vaccine development, *Biochim Biophys Acta 1780*, 546-563.

39. Zhang, Y., Wang, P. G., and Brew, K. (2001) Specificity and mechanism of metal ion activation in UDP-galactose:beta -galactoside-alpha -1,3-galactosyltransferase, *J Biol Chem 276*, 11567-11574.

40. Hassan, H., Reis, C. A., Bennett, E. P., Mirgorodskaya, E., Roepstorff, P., Hollingsworth, M. A., Burchell, J., Taylor-Papadimitriou, J., and Clausen, H. (2000) The lectin domain of UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-T4 directs its glycopeptide specificities, *J Biol Chem 275*, 38197-38205.

41. Cao, Y., Stosiek, P., Springer, G. F., and Karsten, U. (1996) Thomsen-Friedenreich-related carbohydrate antigens in normal adult human tissues: a systematic and comparative study, *Histochem Cell Biol 106*, 197-207.

42. Cao, Y., Schlag, P. M., and Karsten, U. (1997) Immunodetection of epithelial mucin (MUC1, MUC3) and mucin-associated glycotopes (TF, Tn, and sialosyl-Tn) in benign and malignant lesions of colonic epithelium: apolar localization corresponds to malignant transformation, *Virchows Arch 431*, 159-166.

43. Springer, G. F., Desai, P. R., Wise, W., Carlstedt, S. C., Tegtmeyer, H., Stein, R., and Scanlon, E. F. (1990) Pancarcinoma T and Tn epitopes: autoimmunogens and diagnostic markers that reveal incipient carcinomas and help establish prognosis, *Immunol Ser 53*, 587-612.

44. Baldus, S. E., and Hanisch, F. G. (2000) Biochemistry and pathological importance of mucin-associated antigens in gastrointestinal neoplasia, *Adv Cancer Res 79*, 201-248.

45. Itzkowitz, S. H., Bloom, E. J., Kokal, W. A., Modin, G., Hakomori, S., and Kim, Y. S. (1990) Sialosyl-Tn. A novel mucin antigen associated with prognosis in colorectal cancer patients, *Cancer 66*, 1960-1966.

46. Kobayashi, H., Terao, T., and Kawashima, Y. (1992) Serum sialyl Tn as an independent predictor of poor prognosis in patients with epithelial ovarian cancer, *J Clin Oncol 10*, 95-101.

47. Leivonen, M., Nordling, S., Lundin, J., von Boguslawski, K., and Haglund, C. (2001) STn and prognosis in breast cancer, *Oncology 61*, 299-305.

48. Kim, G. E., Bae, H. I., Park, H. U., Kuan, S. F., Crawley, S. C., Ho, J. J., and Kim, Y. S. (2002) Aberrant expression of MUC5AC and MUC6 gastric mucins and sialyl Tn antigen in intraepithelial neoplasms of the pancreas, *Gastroenterology 123*, 1052-1060.

49. Kjeldsen, T., Clausen, H., Hirohashi, S., Ogawa, T., Iijima, H., and Hakomori, S. (1988) Preparation and characterization of monoclonal antibodies directed to the tumor-associated O-linked sialosyl-2----6 alpha-N-acetylgalactosaminyl (sialosyl-Tn) epitope, *Cancer Res 48*, 2214-2220.

50. Miles, D. W., Happerfield, L. C., Smith, P., Gillibrand, R., Bobrow, L. G., Gregory, W. M., and Rubens, R. D. (1994) Expression of sialyl-Tn predicts the effect of adjuvant chemotherapy in node-positive breast cancer, *Br J Cancer 70*, 1272-1275.

51. Sewell, R., Backstrom, M., Dalziel, M., Gschmeissner, S., Karlsson, H., Noll, T., Gatgens, J., Clausen, H., Hansson, G. C., Burchell, J., and Taylor-Papadimitriou, J. (2006) The ST6GalNAc-I sialyltransferase localizes throughout the Golgi and is responsible for the synthesis of the tumor-associated sialyl-Tn O-glycan in human breast cancer, *J Biol Chem 281*, 3586-3594.

52. Marcos, N. T., Pinho, S., Grandela, C., Cruz, A., Samyn-Petit, B., Harduin-Lepers, A., Almeida, R., Silva, F., Morais, V., Costa, J., Kihlberg, J., Clausen, H., and Reis, C. A. (2004) Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen, *Cancer Res 64*, 7050-7057.

53. Brockhausen, I., Yang, J., Dickinson, N., Ogata, S., and Itzkowitz, S. H. (1998) Enzymatic basis for sialyl-Tn expression in human colon cancer cells, *Glycoconj J 15*, 595-603.

54. Schietinger, A., Philip, M., Yoshida, B. A., Azadi, P., Liu, H., Meredith, S. C., and Schreiber, H. (2006) A mutant chaperone converts a wild-type protein into a tumor-specific antigen, *Science 314*, 304-308.

55. Ju, T., and Cummings, R. D. (2005) Protein glycosylation: chaperone mutation in Tn syndrome, *Nature 437*, 1252.

56. Ju, T., Brewer, K., D'Souza, A., Cummings, R. D., and Canfield, W. M. (2002) Cloning and expression of human core 1 beta1,3-galactosyltransferase, *J Biol Chem 277*, 178-186.

57. Ogata, S., Maimonis, P. J., and Itzkowitz, S. H. (1992) Mucins bearing the cancer-associated sialosyl-Tn antigen mediate inhibition of natural killer cell cytotoxicity, *Cancer Res 52*, 4741-4746.

58. Springer, G. F. (1984) T and Tn, general carcinoma autoantigens, *Science 224*, 1198-1206.

59. Avichezer, D., Springer, G. F., Schechter, B., and Arnon, R. (1997) Immunoreactivities of polyclonal and monoclonal anti-T and anti-Tn antibodies with human carcinoma cells, grown in vitro and in a xenograft model, *Int J Cancer 72*, 119-127.

60. Dalziel, M., Whitehouse, C., McFarlane, I., Brockhausen, I., Gschmeissner, S., Schwientek, T., Clausen, H., Burchell, J. M., and Taylor-Papadimitriou, J. (2001) 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.

61. Schwientek, T., Nomoto, M., Levery, S. B., Merkx, G., van Kessel, A. G., Bennett, E. P., Hollingsworth, M. A., and Clausen, H. (1999) Control of O-glycan branch formation. Molecular cloning of human cDNA encoding a novel beta1,6-N-acetylglucosaminyltransferase forming core 2 and core 4, *Journal of Biological Chemistry 274*, 4504-4512.

62. Schwientek, T., Yeh, J. C., Levery, S. B., Keck, B., Merkx, G., van Kessel, A. G., Fukuda, M., and Clausen, H. (2000) Control of O-glycan branch formation. Molecular cloning and characterization of a novel thymus-associated core 2 beta1, 6-n-acetylglucosaminyltransferase, *J Biol Chem 275*, 11106-11113.

63. Kuhns, W., Rutz, V., Paulsen, H., Matta, K. L., Baker, M. A., Barner, M., Granovsky, M., and Brockhausen, I. (1993) Processing O-glycan core 1, Gal beta 1-3GalNAc alpha-R. Specificities of core 2, UDP-GlcNAc: Gal beta 1-3 GalNAc-R(GlcNAc to GalNAc) beta 6-N-acetylglucosaminyltransferase and CMP-sialic acid: Gal beta 1-3GalNAc-R alpha 3-sialyltransferase, *Glycoconj J 10*, 381-394.

64. Whitehouse, C., Burchell, J., Gschmeissner, S., Brockhausen, I., Lloyd, K. O., and Taylor-Papadimitriou, J. (1997) A transfected sialyltransferase that is elevated in breast cancer and localizes to the medial/trans-Golgi apparatus inhibits the development of core-2-based O-glycans, *J Cell Biol 137*, 1229-1241.

65. Hagisawa, S., Ohyama, C., Takahashi, T., Endoh, M., Moriya, T., Nakayama, J., Arai, Y., and Fukuda, M. (2005) Expression of core 2 beta1,6-N-acetylglucosaminyltransferase facilitates prostate cancer progression, *Glycobiology 15*, 1016-1024.

66. Iwai, T., Inaba, N., Naundorf, A., Zhang, Y., Gotoh, M., Iwasaki, H., Kudo, T., Togayachi, A., Ishizuka, Y., Nakanishi, H., and Narimatsu, H. (2002) Molecular cloning and characterization of a novel UDP-GlcNAc:GalNAc-peptide beta1,3-N-acetylglucosaminyltransferase (beta 3Gn-T6), an enzyme synthesizing the core 3 structure of O-glycans, *Journal of Biological Chemistry 277*, 12802-12809.

67. Cheng, L., Tachibana, K., Zhang, Y., Guo, J., Kahori, T. K., Kameyama, A., Wang, H., Hiruma, T., Iwasaki, H., Togayachi, A., Kudo, T., and Narimatsu, H. (2002) Characterization of a novel human UDP-GalNAc transferase, pp-GalNAc-T10, *FEBS Lett. 531*, 115-121.

68. Iwai, T., Kudo, T., Kawamoto, R., Kubota, T., Togayachi, A., Hiruma, T., Okada, T., Kawamoto, T., Morozumi, K., and Narimatsu, H. (2005) 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.

69. Yang, J. M., Byrd, J. C., Siddiki, B. B., Chung, Y. S., Okuno, M., Sowa, M., Kim, Y. S., Matta, K. L., and Brockhausen, I. (1994) Alterations of O-glycan biosynthesis in human colon cancer tissues, *Glycobiology 4*, 873-884.

70. Borsig, L., Wong, R., Hynes, R. O., Varki, N. M., and Varki, A. (2002) Synergistic effects of L- and P-selectin in facilitating tumor metastasis can involve non-mucin ligands and implicate leukocytes as enhancers of metastasis, *Proc Natl Acad Sci U S A 99*, 2193-2198.

71. Baldus, S. E., Zirbes, T. K., Monig, S. P., Engel, S., Monaca, E., Rafiqpoor, K., Hanisch, F. G., Hanski, C., Thiele, J., Pichlmaier, H., and Dienes, H. P. (1998) Histopathological subtypes and prognosis of gastric cancer are correlated with the expression of mucin-associated sialylated antigens: Sialosyl-Lewis(a), Sialosyl-Lewis(x) and sialosyl-Tn, *Tumour Biol 19*, 445-453.

72. Hollingsworth, M. A., and Swanson, B. J. (2004) Mucins in cancer: protection and control of the cell surface, *Nat.Rev.Cancer 4*, 45-60.

73. Pinho, S., Marcos, N. T., Ferreira, B., Carvalho, A. S., Oliveira, M. J., Santos-Silva, F., Harduin-Lepers, A., and Reis, C. A. (2007) Biological significance of cancer-associated sialyl-Tn antigen: modulation of malignant phenotype in gastric carcinoma cells, *Cancer Lett 249*, 157-170.

74. Goldstein, M. J., and Mitchell, E. P. (2005) Carcinoembryonic antigen in the staging and follow-up of patients with colorectal cancer, *Cancer Invest 23*, 338-351.

75. Safi, F., Schlosser, W., Kolb, G., and Beger, H. G. (1997) Diagnostic value of CA 19-9 in patients with pancreatic cancer and nonspecific gastrointestinal symptoms, *J Gastrointest Surg 1*, 106-112.

76. Ychou, M., Duffour, J., Kramar, A., Gourgou, S., and Grenier, J. (2000) Clinical significance and prognostic value of CA72-4 compared with CEA and CA19-9 in patients with gastric cancer, *Dis Markers 16*, 105-110.

77. Ebeling, F. G., Stieber, P., Untch, M., Nagel, D., Konecny, G. E., Schmitt, U. M., Fateh-Moghadam, A., and Seidel, D. (2002) Serum CEA and CA 15-3 as prognostic factors in primary breast cancer, *Br J Cancer 86*, 1217-1222.

78. Wandall, H. H., Blixt, O., Tarp, M. A., Pedersen, J. W., Bennett, E. P., Mandel, U., Ragupathi, G., Livingston, P. O., Hollingsworth, M. A., Taylor-Papadimitriou, J., Burchell, J., and Clausen, H. (2010) Cancer biomarkers defined by autoantibody signatures to aberrant O-glycopeptide epitopes, *Cancer Res 70*, 1306-1313.

79. Lu, H., Goodell, V., and Disis, M. L. (2008) Humoral immunity directed against tumor-associated antigens as potential biomarkers for the early diagnosis of cancer, *J Proteome Res 7*, 1388-1394.

80. Fernandez Madrid, F. (2005) Autoantibodies in breast cancer sera: candidate biomarkers and reporters of tumorigenesis, *Cancer Lett 230*, 187-198.

81. Caron, M., Choquet-Kastylevsky, G., and Joubert-Caron, R. (2007) Cancer immunomics using autoantibody signatures for biomarker discovery, *Mol Cell Proteomics 6*, 1115-1122.

82. Nam, M. J., Madoz-Gurpide, J., Wang, H., Lescure, P., Schmalbach, C. E., Zhao, R., Misek, D. E., Kuick, R., Brenner, D. E., and Hanash, S. M. (2003) Molecular profiling of the immune response in colon cancer using protein microarrays: occurrence of autoantibodies to ubiquitin C-terminal hydrolase L3, *Proteomics 3*, 2108-2115.

83. Disis, M. L., Pupa, S. M., Gralow, J. R., Dittadi, R., Menard, S., and Cheever, M. A. (1997) High-titer HER-2/neu protein-specific antibody can be detected in patients with early-stage breast cancer, *J Clin Oncol 15*, 3363-3367.

84. Stoler, D. L., Chen, N., Basik, M., Kahlenberg, M. S., Rodriguez-Bigas, M. A., Petrelli, N. J., and Anderson, G. R. (1999) The onset and extent of genomic instability in sporadic colorectal tumor progression, *Proc Natl Acad Sci U S A 96*, 15121-15126.

85. Spiotto, M. T., Reth, M. A., and Schreiber, H. (2003) Genetic changes occurring in established tumors rapidly stimulate new antibody responses, *Proc Natl Acad Sci U S A 100*, 5425-5430.

86. Terry, M. B., Gaudet, M. M., and Gammon, M. D. (2002) The epidemiology of gastric cancer, *Semin Radiat Oncol 12*, 111-127.

87. Reis, C. A., David, L., Correa, P., Carneiro, F., de Bolos, C., Garcia, E., Mandel, U., Clausen, H., and Sobrinho-Simoes, M. (1999) Intestinal metaplasia of human stomach displays distinct patterns of mucin (MUC1, MUC2, MUC5AC, and MUC6) expression, *Cancer Res 59*, 1003-1007.

88. Reis, C. A., David, L., Carvalho, F., Mandel, U., de Bolos, C., Mirgorodskaya, E., Clausen, H., and Sobrinho-Simoes, M. (2000) Immunohistochemical study of the expression of MUC6 mucin and co-expression of other secreted mucins (MUC5AC and MUC2) in human gastric carcinomas, *J Histochem Cytochem 48*, 377-388.

89. Reis, C. A., David, L., Seixas, M., Burchell, J., and Sobrinho-Simoes, M. (1998) Expression of fully and under-glycosylated forms of MUC1 mucin in gastric carcinoma, *Int J Cancer 79*, 402-410.

90. Lau, S. K., Weiss, L. M., and Chu, P. G. (2004) Differential expression of MUC1, MUC2, and MUC5AC in carcinomas of various sites: an immunohistochemical study, *Am J Clin Pathol 122*, 61-69.

91. Zhang, J. Y. (2004) Tumor-associated antigen arrays to enhance antibody detection for cancer diagnosis, *Cancer Detect Prev 28*, 114-118.

92. Koziol, J. A., Zhang, J. Y., Casiano, C. A., Peng, X. X., Shi, F. D., Feng, A. C., Chan, E. K., and Tan, E. M. (2003) Recursive partitioning as an approach to selection of immune markers for tumor diagnosis, *Clin Cancer Res 9*, 5120-5126.

93. Blixt, O., Head, S., Mondala, T., Scanlan, C., Huflejt, M. E., Alvarez, R., Bryan, M. C., Fazio, F., Calarese, D., Stevens, J., Razi, N., Stevens, D. J., Skehel, J. J., van Die, I., Burton, D. R., Wilson, I. A., Cummings, R., Bovin, N., Wong, C. H., and Paulson, J. C. (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins, *Proc Natl Acad Sci U S A 101*, 17033-17038.

94. Tarp, M. A., Sorensen, A. L., Mandel, U., Paulsen, H., Burchell, J., Taylor-Papadimitriou, J., and Clausen, H. (2007) Identification of a novel cancer-specific immunodominant glycopeptide epitope in the MUC1 tandem repeat, *Glycobiology 17*, 197-209.

95. Wandall, H. H., Irazoqui, F., Tarp, M. A., Bennett, E. P., Mandel, U., Takeuchi, H., Kato, K., Irimura, T., Suryanarayanan, G., Hollingsworth, M. A., and Clausen, H. (2007) The lectin domains of polypeptide GalNAc-transferases exhibit carbohydrate-binding specificity for GalNAc: lectin binding to GalNAc-glycopeptide substrates is required for high density GalNAc-O-glycosylation, *Glycobiology 17*, 374-387.

96. Muller, R., Hulsmeier, A. J., Altmann, F., Ten Hagen, K., Tiemeyer, M., and Hennet, T. (2005) Characterization of mucin-type core-1 beta1-3 galactosyltransferase homologous enzymes in Drosophila melanogaster, *FEBS J 272*, 4295-4305.

97. Kurosawa, N., Hamamoto, T., Lee, Y. C., Nakaoka, T., Kojima, N., and Tsuji, S. (1994) Molecular cloning and expression of GalNAc alpha 2,6-sialyltransferase, *J Biol Chem 269*, 1402-1409.

98. Blixt, O., Clo, E., Nudelman, A. S., Sorensen, K. K., Clausen, T., Wandall, H. H., Livingston, P., Clausen, H., and Jensen, K. J. (2010) A High-throughput O-Glycopeptide Discovery Platform for Seromic Profiling, *J Proteome Res*.

99. Reis, C. A., Sorensen, T., Mandel, U., David, L., Mirgorodskaya, E., Roepstorff, P., Kihlberg, J., Hansen, J. E., and Clausen, H. (1998) Development and characterization of an antibody directed to an alpha-N-acetyl-D-galactosamine glycosylated MUC2 peptide, *Glycoconj J 15*, 51-62.

100. Nishiyama, T., Matsumoto, Y., Watanabe, H., Fujiwara, M., and Sato, S. (1987) Detection of Tn antigen with Vicia villosa agglutinin in urinary bladder cancer: its relevance to the patient's clinical course, *J Natl Cancer Inst 78*, 1113-1118.

101. Chatterjee, M., Wojciechowski, J., and Tainsky, M. A. (2009) Discovery of antibody biomarkers using protein microarrays of tumor antigens cloned in high throughput, *Methods Mol Biol 520*, 21-38.

102. Anderson, K. S., and LaBaer, J. (2005) The sentinel within: exploiting the immune system for cancer biomarkers, *J Proteome Res 4*, 1123-1133.

103. Wandall, H., Blixt, O., Tarp, M., Pedersen, J., Bennett, E., U, M., Ragupathi, G., Livingston, P., Hollingsworth, M., Taylor-Papadimitriou, J., Burchell, J., and Clausen, H. (2009) Auto-antibody signatures to aberrant O-glycopeptide epitopes serve as undiscovered biomarkers of cancer *Cancer Research, Accepted manus*.

104. Mintz, P. J., Kim, J., Do, K. A., Wang, X., Zinner, R. G., Cristofanilli, M., Arap, M. A., Hong, W. K., Troncoso, P., Logothetis, C. J., Pasqualini, R., and Arap, W. (2003) Fingerprinting the circulating repertoire of antibodies from cancer patients, *Nat Biotechnol 21*, 57-63.

105. Winter, S. F., Minna, J. D., Johnson, B. E., Takahashi, T., Gazdar, A. F., and Carbone, D. P. (1992) Development of antibodies against p53 in lung cancer patients appears to be dependent on the type of p53 mutation, *Cancer Res 52*, 4168-4174.

106. Tan, H. T., Low, J., Lim, S. G., and Chung, M. C. (2009) Serum autoantibodies as biomarkers for early cancer detection, *FEBS J*.

107. Lolli, F., Mulinacci, B., Carotenuto, A., Bonetti, B., Sabatino, G., Mazzanti, B., D'Ursi, A. M., Novellino, E., Pazzagli, M., Lovato, L., Alcaro, M. C., Peroni, E., Pozo-Carrero, M. C., Nuti, F., Battistini, L., Borsellino, G., Chelli, M., Rovero, P., and Papini, A. M. (2005) An N-glucosylated peptide detecting disease-specific autoantibodies, biomarkers of multiple sclerosis, *Proc Natl Acad Sci U S A 102*, 10273-10278.

108. Sorensen, A. L., Reis, C. A., Tarp, M. A., Mandel, U., Ramachandran, K., Sankaranarayanan, V., Schwientek, T., Graham, R., Taylor-Papadimitriou, J., Hollingsworth, M. A., Burchell, J., and Clausen, H. (2006) Chemoenzymatically synthesized multimeric Tn/STn MUC1 glycopeptides elicit cancer-specific anti-MUC1 antibody responses and override tolerance, *Glycobiology 16*, 96-107.

109. Chapman, C., Murray, A., Chakrabarti, J., Thorpe, A., Woolston, C., Sahin, U., Barnes, A., and Robertson, J. (2007) Autoantibodies in breast cancer: their use as an aid to early diagnosis, *Ann Oncol 18*, 868-873.

110. Chen, S., LaRoche, T., Hamelinck, D., Bergsma, D., Brenner, D., Simeone, D., Brand, R. E., and Haab, B. B. (2007) Multiplexed analysis of glycan variation on native proteins captured by antibody microarrays, *Nat Methods 4*, 437-444.

111. Wahrenbrock, M. G., and Varki, A. (2006) Multiple hepatic receptors cooperate to eliminate secretory mucins aberrantly entering the bloodstream: are circulating cancer mucins the "tip of the iceberg"?, *Cancer Res 66*, 2433-2441.

112. Napoletano, C., Rughetti, A., gervig Tarp, M. P., Coleman, J., Bennett, E. P., Picco, G., Sale, P., da-Nagai, K., Irimura, T., Mandel, U., Clausen, H., Frati, L., Taylor-Papadimitriou, J., Burchell, J., and Nuti, M. (2007) Tumor-associated Tn-MUC1 glycoform is internalized through the macrophage galactose-type C-type lectin and delivered to the HLA class I and II compartments in dendritic cells, *Cancer Res. 67*, 8358-8367.

113. Wandall, H. H., Hassan, H., Mirgorodskaya, E., Kristensen, A. K., Roepstorff, P., Bennett, E. P., Nielsen, P. A., Hollingsworth, M. A., Burchell, J., Taylor-Papadimitriou, J., and Clausen, H. (1997) Substrate specificities of three members of the human UDP-N-acetyl-alpha-D-galactosamine:Polypeptide N-acetylgalactosaminyltransferase family, GalNAc-T1, -T2, and -T3, *J Biol Chem 272*, 23503-23514.

114. Schwientek, T., Bennett, E. P., Flores, C., Thacker, J., Hollmann, M., Reis, C. A., Behrens, J., Mandel, U., Keck, B., Schafer, M. A., Haselmann, K., Zubarev, R., Roepstorff, P., Burchell, J. M., Taylor-Papadimitriou, J., Hollingsworth, M. A., and Clausen, H. (2002) Functional conservation of subfamilies of putative UDP-N-acetylgalactosamine:polypeptide N-acetylgalactosaminyltransferases in Drosophila, Caenorhabditis elegans, and mammals. One subfamily composed of l(2)35Aa is essential in Drosophila, *J Biol Chem 277*, 22623-22638.

115. Conze, T., Carvalho, A. S., Landegren, U., Almeida, R., Reis, C. A., David, L., and Soderberg, O. (2010) MUC2 mucin is a major carrier of the cancer-associated sialyl-Tn antigen in intestinal metaplasia and gastric carcinomas, *Glycobiology 20*, 199-206.

116. Robbe-Masselot, C., Herrmann, A., Maes, E., Carlstedt, I., Michalski, J. C., and Capon, C. (2009) Expression of a core 3 disialyl-Le(x) hexasaccharide in human colorectal cancers: a potential marker of malignant transformation in colon, *J Proteome Res 8*, 702-711.

117. Zhou, D., Berger, E. G., and Hennet, T. (1999) Molecular cloning of a human UDP-galactose:GlcNAcbeta1,3GalNAc beta1, 3 galactosyltransferase gene encoding an O-linked core3-elongation enzyme, *Eur J Biochem 263*, 571-576.

118. Amado, M., Almeida, R., Carneiro, F., Levery, S. B., Holmes, E. H., Nomoto, M., Hollingsworth, M. A., Hassan, H., Schwientek, T., Nielsen, P. A., Bennett, E. P., and Clausen, H. (1998) A family of human beta3-galactosyltransferases. Characterization of four members of a UDP-galactose:beta-N-acetyl-glucosamine/beta-nacetyl-galactosamine beta-1,3-galactosyltransferase family, *J Biol Chem 273*, 12770-12778.

119. Regele, S., Vogl, F. D., Kohler, T., Kreienberg, R., and Runnebaum, I. B. (2003) p53 autoantibodies can be indicative of the development of breast cancer relapse, *Anticancer Res 23*, 761-764.

120. Trivers, G. E., De Benedetti, V. M., Cawley, H. L., Caron, G., Harrington, A. M., Bennett, W. P., Jett, J. R., Colby, T. V., Tazelaar, H., Pairolero, P., Miller, R. D., and Harris, C. C. (1996) Anti-p53 antibodies in sera from patients with chronic obstructive pulmonary disease can predate a diagnosis of cancer, *Clin Cancer Res 2*, 1767-1775.