Review

Mucin gene and antigen expression in biliopancreatic carcinogenesis

Y. S. Kim,1,2 J. R. Gum, Jr.,1,3 S. C. Crawley,1 G. Deng1 & J. J. L. Ho1
Departments of 1Medicine, 2Pathology and 3Anatomy, University of California, San Francisco, Gastrointestinal Research Laboratory, Department of Veterans Affairs Medical Center, San Francisco, CA, USA

Summary

Mucins are high molecular weight glycoproteins which are heavily glycosylated with many carbohydrate side chains. In epithelial cancers such as biliopancreatic cancer, both quantitative and qualitative alterations in carbohydrate and polypeptide moieties of mucin glycoproteins occur. These changes in mucin glycoproteins are one of the most common phenotypic markers of biliopancreatic carcinogenesis and may play an important pathobiological role. The expression of some of the sialylated carbohydrate antigens appears to correlate with a poor prognosis and increased metastatic potential in biliopancreatic cancer. The increased exposure of peptide epitopes of mucin glycoproteins in biliopancreatic cancer appears to be due to either abnormal glycosylation and/or altered levels of mucin gene transcription. In addition, dysregulation of tissue specific mucin gene expression occurs in biliopancreatic cancer. This information is currently being exploited for further elucidation of the molecular mechanisms involved in carcinogenesis, tumor progression and metastasis, and the development of novel methods of diagnosis and therapy of biliopancreatic cancer.

Key words: carcinogenesis, mucin antigen, mucin gene

Introduction

Mucins are heavily glycosylated, generally high molecular weight proteins that are synthesized by the epithelial cells of the gastrointestinal, respiratory and genito-urinary tract. Epithelial mucins consist of protein backbone structures with many carbohydrate side chains linked O-glycosidically. The carbohydrates may make up 40 to 80% of the total weight of mucin glycoproteins. Carbohydrate side chains are varied in their composition, sequence, and anomeric linkage and length and therefore are thought to play an important role in the diverse biological properties of cells such as protective and adhesive functions. Recently, using molecular cloning techniques, several mucin genes have been identified and characterized, and the deduced structure of mucin protein backbone (apomucins) are beginning to be elucidated [1,2]. The elucidation of the apomucin structure is beginning to yield important information on cell and tissue-specific expression of mucin genes, and potential functional role of various mucin genes and gene products in normal and diseased organs.

Considerable alterations in both carbohydrate and protein moieties of mucin glycoproteins have been reported in epithelial cancers such as biliopancreatic cancer [1-5]. These alterations in mucin glycoproteins have been suggested to play important roles in cell-cell, cell substratum interaction, growth regulation and differentiation. Further elucidation of the immunochemical, biochemical, and molecular biological mechanisms involved in the altered expression of glycosyltransferases responsible for the addition of carbohydrate residues in carbohydrate side chains, and of the level and the pattern of mucin gene expression in biliopancreatic cancers will be extremely important in furthering our understanding of the phenotypic, molecular and cell biological changes that occur in biliopancreatic carcinogenesis. This information also will be helpful in developing strategies for early detection, diagnosis, monitoring of recurrence, prognostic assessment and therapy of biliopancreatic cancer.

The mucin gene family

Contrary to the commonly held view that mucin represents a fairly homogenous molecule, at least nine human mucin genes have been identified to date [1,2]. Recent data indicates that mucins may either be secreted or membrane associated. The dominant and common structural feature of all types of epithelial mucins is the central tandem repeat region consisting of a repetitive peptide unit which is rich in proline and threonine and/or serine to which carbohydrate side chains are linked O-glycosidically. This heavily glycosylated central tandem repeat region is flanked on either side by less glycosylated, non-repeat region. The tandem repeat regions of mucin genes differ considerably among different mucins and different species, while less glycosylated regions tend to be relatively well conserved. Polymorphism of mucin genes frequently occurs due to the occurrence of variable numbers of tandem repeats. Different mucin genes have distinct chromosomal localization, although some cluster in the same chromosomal loci.

Domain structures of apomucin and their possible function

Mucin glycoproteins may be broadly divided into four main types. The first type is characterized by MUC1, a small membrane-associated mucin with a cytoplasmic tail which interacts with cytoskeletal structures [5]. It is a type 1 membrane glycoprotein with N-terminal region containing variable hydrophobic signal sequences as a consequence of alternate splicing. The C-terminal region contains a transmembrane domain and a cytoplasmic tail that contains tyrosine phosphorylation sites. The cytoplasmic tail also contains SXXXXXSSL motif that binds β-catenin in...
association with adhesion of epithelial cells. Ectodomain of MUC1 is rigid, due in part to polypeptide β-turn helix protruding more than 200 nm above the cell surface and is thought to interfere with cellular adhesion by steric hindrance.

The second type is represented by a very large secreted mucin without transmembrane domain such as MUC2, which has a small cysteine-rich domain in the carboxy terminus involved in dimerization and four cysteine-rich domains with similarity to D-domains of von Willebrand factor [6]. Specific cysteine residues in the D-domains are thought to be involved in linking the dimers to form multimers, which are in part responsible for the viscous nature of the mucin. Since the viscosity of the mucin contributes significantly to protective functions for the epithelial cells, mutation in carboxy terminal cysteine-rich domain or D-domains could lead to synthesis of mucin unable to form multimers, resulting in the formation of defective protective barriers.

The third type is represented by MUC3 and MUC4, which have both the membrane spanning and two EGF-like domains [7,8]. Both MUC3 and MUC4 are much higher molecular weight mucins than MUC1, having large ectodomains protruding from the cell surface. This type of mucin may constitute glyocalyx, which has been an ill-defined entity for many years. The presence of EGF-like domain indicates the possible role of these mucins in ligand or a growth factor-like function for MUC3 and MUC4. In fact, MUC4 has been shown to act as a ligand for ErB2/Neu receptor tyrosine kinase. The formation of the complex between MUC4 and ErB2 has been reported to potentiate the effect of neurugulin on the tyrosine phosphorylation of both ErB2 and ErB3, leading to activation of growth signaling pathways [8]. The fourth type is represented by the small salivary mucin, MUC7, which lacks both the membrane spanning and dimer forming domain and therefore is unable to form multimers. Although hydrophilic carbohydrate side chains may provide some degree of viscosity, this type of mucin should be much less viscous than MUC2 type of mucin but capable of carrying out lubricant function. Apparently soluble splice variants lacking membrane spanning domains have been reported for MUC1 and MUC3. Interestingly, MUC5 AC, MUC5B and MUC6 share similar domain structure with MUC2, and these four mucin genes are clustered on chromosome 11P15.5 in a region of some 400 Kb, whereas the five other mucin genes are dispersed randomly throughout the human genome [9]. It is not known at the present time how these mucin genes have evolved or how they are regulated.

Cell and tissue specific expression of mucin genes

Recent immunohistochemical and in situ hybridization studies indicate that mucin genes are expressed in a highly cell and tissue specific manner. For example, in the human intestine, MUC2 is expressed in goblet cells while MUC3 is predominantly expressed in absorptive cells [10]. These mucins are not expressed in the human stomach or normal respiratory tract. In the stomach, MUC5 AC is expressed in surface mucous cells, whereas MUC6 is expressed in mucous neck and pyloric gland cells. MUC5 AC and MUC6 are not expressed in the intestine, although MUC5 AC is also expressed in the respiratory tract [11]. In the pancreas, duct and ductules express both MUC1 and MUC6 albeit MUC3 is heterogenously and weakly expressed in interlobular ducts. MUC2, MUC4 and MUC5 AC are not expressed in normal pancreas [12]. In the gallbladder, MUC3 is expressed in the surface epithelium while MUC6 is expressed in the gland. Bile duct expressed both MUC3 and MUC6. However, the normal gallbladder and bile duct do not express MUC1, MUC2, MUC4 and MUC5 AC. Thus, mucin genes appear to be expressed in a relatively tissue and cell-specific manner in normal tissue, although the same cells may express more than one mucin gene.

Alteration of mucin glycoproteins in cancer

The changes that occur in mucin glycoprotein in cancer may be broadly divided into two general types: aberrant glycosylation and dysregulation of mucin genes [1,2]. Aberrant glycosylation involves incomplete synthesis resulting in deletion of normally expressed antigens with or without exposure of core sugar and peptide structures, inappropriate expression of antigenic structures not normally present, increased level of expression over that found in normal tissue and modification of existing structures in the oligosaccharide side chains of mucin glycoproteins. Dysregulation of mucin genes involves altered levels of expression of tissue and cell-specific mucin genes and/or altered pattern of expression of mucin genes.

Aberrant glycosylation resulting in the expression of cancer-associated antigens

Aberrant glycosylation results in the expression of a variety of a tumor associated carbohydrate antigen and apomucin (protein backbone structure) antigens. In mucin glycoproteins, many oligosaccharide side chains are linked O-glycosidically through N-acetylgalactosamine (GalNAc) to threonine or serine in the polypeptide backbone. The carbohydrate side chains may be divided into three regions, core, backbone and peripheral regions. All three regions provide a specific recognition site for carbohydrate-specific antibodies and can serve as tumor markers recognized by monoclonal antibodies and lectins [1,2]. Accumulation and increased expression of core region carbohydrates with shorter carbohydrate side chains occur as a result of either increased synthesis along a specific pathway or a defect in the elongation pathways of carbohydrate side chains. In pancreatic cancer, an increased expression of core region carbohydrates such as Tn, T, sialyl Tn and sialyl T antigens occur. T and Tn antigens are weakly expressed in pancreatic ducts and ductules, whereas acinar cells express high levels of these antigens. By contrast, sialyl Tn and sialyl T antigens are not expressed in normal pancreas and therefore they may serve as pancreatic tumor markers [2,13]. Modification of existing structures in backbone or peripheral regions of the carbohydrate side chains may lead to the expression of tumor associated antigens such as Le' related antigens (e.g., extended Le' and sialyl Le' antigens). These antigens are highly expressed in pancreatic cancer but not in normal pancreas [14]. An example of overexpression of carbohydrate antigens in
pancreatic cancer tissue that are expressed in normal ductal and ductular cells of the pancreas is Le^a and Le^-related antigens such as sialyl Le^a antigens. Several monoclonal antibodies that recognize the sialyl Le^-related epitope have proven useful as serum markers in the follow-up of patients with pancreatic cancer. These are CA19-9, CA50, SPan-1, DuPan 1 antigens which are highly elevated in the sera of patients with pancreatic cancer compared to control subjects [15]. This may be due, in part, to the loss of polarity of the cancer cells, increased shedding into the tissue stroma caused either by increased protease activities and/or cytokines, and to increased angiogenesis. However, these antigens also become elevated in chronic pancreatitis and patients with cancer of other organs and lack specificity and sensitivity. Therefore, these tests are not very useful for early detection or for routine screening of patients for pancreatic cancer in the general population. However, these antigens appear to be useful as adjunct in the diagnosis and monitoring patients with pancreatic cancer for recurrence of the disease after surgical resection or chemotherapy.

There are also other monoclonal antibodies that recognize as yet unknown epitopes in mucin glycoproteins that show potential for in the radioimaging and radio-immuno or chemo-immunotherapy. For example, a monoclonal antibody generated against a human pancreatic cancer xenograft mucin, Nd2, reacts with over 80% of pancreatic cancer and to a lesser degree with gastric and colorectal cancers, but does not react with normal pancreas and other tissues [16]. Moreover, the epitope recognized by Nd2 is not detectible in the sera of patients with pancreatic cancer, and radiolabeled Nd2 localize well to xenografts of a human pancreatic cancer cell line and in tumors of patients with pancreatic cancer [17,18]. However, the epitope structure of Nd2 has not yet been determined pancreatic tumors in athymic nude mice although available data indicate that both carbohydrate and peptide structures are involved. These promising studies indicate the potential utility of Nd2 in the diagnosis and treatment of pancreatic cancer.

It is also of interest that the cluster configuration of antigens such as Tn and sialyl Tn has been found to be more immunogenic than non-cluster configurations. However, little is known about the clustered carbohydrate antigen epitopes on normal and cancerous cells.

**Functional significance of aberrant glycosylation**

A family of molecules known as selectins are adhesion molecules that mediate calcium dependent cell-cell interactions among leukocytes, platelets and endothelial cells. The ligands for three selectins, L, E and P are mostly mucin type glycoproteins. Sialylated and/or sulfated Le^- or Le^a related structures are the most common glycan recognition sites for the selectins [4]. In vitro experiments suggest that these antigens are potential mediators of vascular adhesion and extravasation of metastatic cells. It has also been demonstrated that tumor metastasis could also be redirected by E-selectin expression in transgenic mice or inhibited by soluble E-selectins. It was also shown that growth and metastasis formation is reduced in P-selectin deficient mice. Recently, carcinoma mucins have been reported to have distinct binding sites for L, E, and P selectins which are involved in pathological interactions of cancer cells with leukocytes, platelets and endothelial cells [19,20]. The expression of sialyl Tn and T antigens and sialyl Le^- and related antigens in the epithelial cancer cells have been associated with poor prognosis of the patient with these tumors [1,2]. Increased sialylation has long been observed in cancer cells more prone to metastasis. Sialidase-reliable membrane-associated sialic acid and sialyl transferase activities were 2- to 3-fold higher in the metastatic cell lines compared to parental cell lines. A decrease in sialic acid content of glycoproteins caused by treatment with sialidase or O-glycosylation inhibitor such as benzyl α-GalNAc results in a decrease in metastatic potential of cancer cells [20].

**Altered mucin gene expression in cancer**

As mentioned previously, mucin genes are expressed in a relatively cell and tissue-specific fashion. In normal pancreas, MUC1 and MUC6 are the predominant mucin genes that are expressed in duct and ductular cells. Acinar cells do not express mucin genes MUC1 through MUC6 [12]. In pancreatic cancer, overexpression of MUC1 and ectopic expression of MUC3, MUC4 and MUC5 AC are observed. In normal gall bladder, MUC3 is expressed by surface epithelial cells while glands express MUC6. In gall bladder adenocarcinoma, over expression of MUC1 and ectopic expression of MUC4 and MUC5 A0C occur.

Recent immunohistochemical and in situ hybridization studies also indicate that MUC1 mucin was highly expressed in invasive ductal carcinomas of the pancreas (IDC) and invasive cholangiocarcinomas (ICC) having a poor prognosis, whereas it was rarely expressed in intraductal papillary mucinous tumors (IPMT) of the pancreas and bile duct cyst adenocarcinomas (BCDD) of the liver having a favorable prognosis. Conversely, MUC2 mucin was rarely expressed in IDC and ICC, while it was highly expressed in IPMT and BDCC. These results suggest that the expression pattern of MUC1 and MUC2 mucins may serve as a prognostic indicator in the histologically different neoplasms of the pancreas and the bile duct [3,21]. The differential expression of MUC1 and MUC2 mucins genes by the neoplasms of pancreas and bile duct may represent the different cell lineage-associated carcinogenesis in the tumors of different histological types, since MUC1 is normally expressed by ductal and ductular cells of the pancreas and MUC2 may be expressed by goblet cells of the larger ducts. The molecular mechanisms involved in altered regulation of mucin genes in cancer cells are not well understood, but several possibilities can be considered. For example, cytokines produced by tumor cells or adjacent inflammatory cells may induce or express mucin genes via signal transduction pathway leading to activation of specific transcriptional factors that bind to the regulatory elements in the promoter region of various mucin genes. Recent studies also indicate that the degree of methylation of CpG islands in the 5'-flanking region of mucin genes may play an important role in the epigenetic regulation of mucin gene expression [22]. For example, hypermethylation of CpG islands in the 5'-flanking region of MUC2 was observed in a colon cancer cell line expressing a very low level of MUC2 mRNA, while hypomethylation of MUC2 5'-flanking region was seen in a mucinous colon cancer cell line expressing high levels of MUC2 mRNA.
The possible function of altered mucin gene expression

There is considerable evidence that supports the notion that MUC1 over expression results in an increase in cell surface, rigid mucin glycoproteins that inhibit cell-cell and cell substratum adhesive interactions resulting in metastasis. MUC1 over expression by cancer cells also causes inhibition of cancer cell killing by cytotoxic T-lymphocytes, and plays a role in metastasis by inhibiting tumor cell adhesion and in escaping from immune surveillance. The anti-adhesive effects caused by over expression of MUC1 are likely to be due to reduction in integrin-mediated interaction with the extracellular matrix. The high density of the extended rigid structure of MUC1 physically hinders interaction of ligands with their receptors [23].

Although only limited data is available at the present time, the ectopic expression of membrane-associated large mucin glycoproteins on the cell surface such as MUC3 and MUC4 may have even more significant effect on cell-cell and cell substratum interaction, since these are considerably larger molecules than MUC1. In fact, muc4, a rodent homologue of MUC4 has been shown to play a role in the progression of rat mammary adenocarcinoma by affecting cancer cell adhesion and metastasis [24]. In addition, both MUC3 and MUC4 have two EGF-like domains that may serve as ligands for ErbB and ErbB3. The complex formed between MUC3 and/or MUC4 with ErbB2 and/or ErbB3 may potentiate the effect of neuregulin on the activation of ErbB2 and/or ErbB3 via tyrosine phosphorylation [8]. The activation of ErbB2 and/or 3 leads to activation of signaling pathways involved in growth and proliferation of cancer cells.

In normal mucin glycoprotein, the tandem repeats in the central tandem repeat region are heavily glycosylated, there are many carbohydrate side chains per molecule and each chain is long. With malignant transformation, the tandem repeats are more sparsely glycosylated, carbohydrate chains may be much shorter and/or modified in the outer region. Thus, the resultant modified sugar structures, or exposed inner sugar core structures, or protein backbone moiety may serve as tumor markers and also may be involved in various biological properties of cancer. In fact, the epitopes, present in the tandem repeat peptide region of MUC1 as well as of other apomucins are exposed in cancer cells due to sparse glycosylation and shortening of carbohydrate side chains. These peptide epitopes may be shed into circulation and serve as tumor associated antigens or may serve as a target on the cancer cell surface for cytotoxic T-cells. In the case of MUC1, both the over expression and the increased exposure of the peptide epitopes that are normally covered by long carbohydrate side chains lead to increased expression of peptide epitopes that can serve as a target for immunotherapy [25,26].

Clinical applications

As mentioned before, no mucin glycoprotein antigenic markers have been shown to be sensitive or specific enough to be clinically useful as diagnostic serological markers for epithelial cancers. Although high levels of mucin glycoprotein antigens such as Le^ related antigens are relatively specific and sensitive for pancreatic cancer (about 80%), the serum assay for these antigens is used mainly to follow the course of the patients after surgical resection and/or chemotherapy [15]. CA15-3 is also elevated in the blood of some patients with breast, ovary and pancreatic cancer. This assay employs two monoclonal antibodies directed to epitopes on the MUC1 apomucin. However, it lacks the sensitivity and specificity to be clinically useful. Monoclonal antibodies against mucin-related antigens may also be used for targeting radio isotopes (I^131, I^111In, T^125) for imaging and therapy or for targeting cytotoxic drugs. Much of the work in this field has been carried out with MAb B7.3 directed to sia1y Tn and Tn epitope [27]. Recently, another MAb directed to epitopes on mucin glycoproteins, Nd2 developed in our laboratory, has also been used particularly for pancreatic cancer. I^111In-labeled Nd2 MAb was found to detect relatively small pancreatic cancer in 67% of the cases [18]. Furthermore, I^131-labeled Nd2 MAb suppressed the growth of pancreatic cancer xenografts in nude mice [28]. Nd2 MAb conjugate to adriamycin injected either I.V. or intra-tumorally into a athymic nude mice suppressed growth of pancreatic cancer xenograft tumors, while adriamycin did not [29]. Nd2 MAb drug conjugates bind to cell surface antigenic epitopes, endocytosed and then cleaved by the lysophagosomal proteases releasing the cytotoxic drugs. Nd2 MAb also was shown to exhibit antibody dependent cellular cytotoxicity (ADCC) effect on pancreatic cancer cells.

Cytotoxic T lymphocytes obtained from the draining lymph nodes of patients with breast and pancreatic cancers recognize the cell surface MUC1 tandem repeat peptide sequence and kill the target tumor cells expressed in the epitope in a non major histocompatibility complex restricted fashion. These MAb's and the synthetic mucin fragments (carbohydrate and mucin polypeptides) are currently being evaluated for their potential as tumor vaccines [25,26]. While gene therapy may provide a new therapeutic approach, in order to achieve clinical efficacy, proper gene delivery systems that possess both high target cell specificity and gene transduction efficiency are required. In in vivo experiments using breast cancer cell lines that over express MUC1, adenoviral vector system containing the MUC1 promoter β-galactosidase has been shown to direct efficient and selective expression of heterologous genes in MUC1 positive breast carcinomas. Moreover, in an intraperitoneal breast cancer metastasis model, i.p. injection of adenovirus containing DF3 herpes simplex virus thymidine kinase followed by ganciclovir treatment resulted in inhibition of tumor growth [30]. Recently, transduction of recombinant retrovirus containing the herpes simplex virus thymidine kinase gene coding region cloned downstream of chimeric MUC1/ERBB2 promoters resulted in an increase in ganciclovir sensitivity of target cells. These results suggest that a tumor selective therapy may be achieved by using the transcriptional regulatory regions of pancreatic cancer cell associated genes to drive the expression of suicide genes [31].

Future directions

Alterations in both the carbohydrate and protein moieties of mucin glycoproteins in pancreatic cancer and biliary cancer have been amply documented. However, limited data are available on the biochemical and molecular mechanisms
involved in regulation of the mucin gene family, and of the multi-glycosyltransferase system in pancreatic biliary carcinogenesis. In addition, further studies are needed to clarify the structure-function relationship of carbohydrate and peptide moieties of mucin glycoproteins of pancreatic cells in various stages of carcinogenesis. Mucin associated carbohydrates and peptides in epithelial cancer cells are being exploited for immunohistochemical diagnosis and prognostic assessment, serological detection, radio immuno- and radio-immunotherapy. The synthesis of cancer-associated oligosaccharides and polypeptide backbone structures or their mimetics as glycan interaction inhibitors of selectin mediated adhesions, or as tumor vaccine may play an important role in the treatment of pancreatic cancer. Promising progress is being made in gene therapy for pancreatic cancer with respect to proper target cell specificity and gene transduction deficiency.

Thus, further elucidation of the structure, biology and molecular mechanisms involved in the regulation of the expression of cancer-associated mucin glycoproteins and the development of transgenic, hybrid oncogene and/or knock-out (mucin genes and/or glycosyltransferases) mice will yield very important and useful data leading not only to furthering our knowledge of pancreatic or biliary carcinogenesis, but also the development of effective diagnosis and therapeutic methods of biliary pancreatic cancer.

Acknowledgements

This work was supported by the Department of Veterans Affairs Medical Research Service, VA/DOD Merit Award and by grant CA24321 from the U.S. Public Health Service, National Cancer Institute.

References


Correspondence to:
Young S. Kim, M.D.
Gastrointestinal Research Laboratory (151M2)
Department of Veterans Affairs Medical Center
4150 Clement St.
San Francisco, CA 94121, U.S.A.
e-mail: youngk@itsa.ucsf.edu