Figure 4

Starting material (CA11-19 168B)

SDS-PAGE

Stain protein bands with CBB

Excise target bands and cut the bands into smaller pieces

Destain CBB

Reduction

Carboxymethylation

PNGase F digestion

In-gel

Extraction

C18 Sep pak

N-glycans

Permethylation

C18 Sep pak

Mass spectrometry

MALDI-MS

[Continued on next page]
Published.

— with international search report (Art. 21(3))

— with sequence listing part of description (Rule 5.2(a))
METHODS AND COMPOSITIONS FOR SCREENING AND DETECTING CANCER

This application is being filed on 28 April 2015, as a PCT International Patent application and claims priority to U.S. Utility patent application Serial No. 14/264,801, filed April 29, 2014, the entire disclosure of which is incorporated by reference in its entirety.

BACKGROUND

Carcinomas cause millions of deaths annually. Of all cancers, colorectal cancer is the second leading cause of cancer-related deaths in the U.S. Most cases of carcinomas are incurable by chemotherapy and radiation therapy unless detected and treated in the early stages of the disease. The more advanced a cancer is when it is diagnosed; the less likely it is that therapy will be effective. Therefore, despite the advances in cancer research, there remains a need for novel antibodies useful for the early diagnosis and treatment of carcinomas of the colon, and lung.

Generally, antibodies are used as invaluable reagents in diagnostics. In fact, they have played a major role in deciphering the functions of various bio-molecules in biosynthetic pathways. They have also become the reagents of choice for identification and characterization of tumor specific antigens and have become a valuable tool in the classification of cancer. Once tumor-associated antigens have been purified from tissue extracts, such antigens can be used to elicit production of antibodies to the antigen by injection into animals. Monoclonal antibodies can then be produced. Such antibodies are useful both therapeutically and diagnostically.

In vitro diagnostic methods are known in the art and include immunohistological detection of tumor cells (e.g., on human tissue, cells or excised tumor specimens) or serologic detection of tumor-associated antigens (e.g., in blood samples or other biological fluids). Immunohistological techniques involve contacting a biological specimen such as a tumor tissue specimen with the antibody of the invention and then detecting the presence on the specimen of the antibody complexed to its antigen. The formation of such antibody-antigen complexes with the specimen indicates the presence of tumor cells in the tissue. Detection of the antibody on the specimen can be accomplished using techniques known in the art,
such as the immunoperoxidase staining technique, the avidin-biotin (ABC) technique or immunofluorescence techniques (see Ciocca et al, Meth. Enzymol., 121:562-79 (1986); Kimball (ed.), Introduction To Immunology (2nd Ed.), pp. 113-117, Macmillan Publ. Co. (1986)).

Serologic diagnostic techniques involve the detection and quantitation of tumor-associated antigens that have been secreted or "shed" into the serum or other biological fluids of patients thought to be suffering from carcinoma. Such antigens can be detected in the body fluids using techniques known in the art such as radioimmunoassays (RIA) or enzyme-linked immunosorbent assays (ELISA) wherein an antibody reactive with the "shed" antigen is used to detect the presence of the antigen in a fluid sample (see, e.g., Uotila et al., J. Immunol. Methods, 42:1 l (1981); Allum et al., "Monoclonal Antibodies in the Diagnosis and Treatment of Malignant Conditions" Surg. Ann., 18:41-64, 48-51 (1986); Sikora et al. (eds.), Monoclonal Antibodies, pp. 32-52, Blackwell Scientific Publications, (1984)).

However, in some cases, serological techniques have not produced results robust enough to be useful for cancer screening. Serum markers such as CEA, CA19-9 and CA242 have all been shown to have poor sensitivity and specificity for colon cancer. CEA was discovered in 1969 and was believed to be a sensitive and specific marker for colon cancer. However, further studies were unable to produce the original results. It was shown that at a cutoff concentration of 2.5ug/L, CEA screening would yield a sensitivity of 30%-40% and a specificity of 87%. Utilizing these numbers, for every 1 colon cancer patient identified with a CEA-based assay there would be 250 false-positives and 60% of cancers would be missed. Due to these overall poor results these markers are not used for colon cancer screening.

For this reason Fecal Occult Blood (FOB) tests have long been the mainstay of colon cancer screening. Large randomized studies have shown that screening with serial FOB tests reduces mortality from colon cancer. Typical examples of FOB tests come in two types, the guaiac-based tests available under the trade names SKB-Hemoccult II® & Hemocult II SENSA®, and the immune-chemical tests available under the trade names SKB-HemeSelect® and Entrerix-InSure Fit®. Disadvantages of guaiac-based FOB tests include burdensome dietary restrictions, the inconvenient collection process, the limited single application sensitivity, and the costs associated with poor specificity (5-10% rate of false positives). FOB tests have sensitivity of
25% - 40% and specificity of 80%-90%. While the immune-chemical based FOB tests have the advantages of improved compliance due to ease of use and no diet restrictions, sensitivity issues remain.

It is thus apparent that antibodies reactive with an antigen expressed at high levels by a variety of tumors may become useful towards an earlier diagnosis of cancers, the immunological monitoring of cancer patients, as well as for development of improved methods for therapy of cancers. It is also apparent that purified antigens associated with carcinomas derived from specific organs and tissues of the body can be of value for creating such monoclonal antibodies, as well as for creating cancer vaccines. It is also apparent that improved materials and methods for serological assays to identify the presence of cancer in humans would be valuable.

SUMMARY OF THE INVENTION

The disclosure provides immunogenic compositions, antibodies, kits, and methods useful in the early detection of cancers or monitoring the treatment of cancer. In embodiments, the cancers are colon-rectal and lung. Immunogenic compositions include an immunogenic composition comprising an isolated fraction from colorectal and/or liver cancer cell samples having a component of about a 100 kDa glycoprotein that has a membrane bound and a soluble form (e.g. cancer cell membrane bound form), has a UV absorbance peak at about 228nm, an isoelectric point of about 3.5 to 4, a sialic acid content of about 20%, and does not substantially bind to antibodies specific for anti-chymotrypsin (ACT). In embodiments, the 100 kDa protein has an amino acid sequence of CEACAM 5 isoform 2 (SEQ ID NO:1; Figure 14). In embodiments, the 100 kDa glycoprotein has a deletion of an alanine at position 320 and an altered glycosylation pattern as compared to 180 kDa CEA.

In some cases, the immunogenic composition is obtained by a method comprising contacting colorectal and/or liver tumor cells with an acid to form an extract; separating components of the extract by molecular weight to form a first fraction comprising components with a molecular weight of about 60 kDa or greater; isolating components of the first fraction that have a molecular weight of about 75 kDa or greater and less than about 200 kDa, and further isolating a fraction that has a molecular weight of about 100 kDa.
In embodiments, the composition further comprises an adjuvant. In specific embodiments, the adjuvant is selected from the group consisting of alum, muramyl dipeptide, and Freund’s complete adjuvant. Immunogenic compositions are useful for obtaining antibodies. Antigens described herein may also be detectably labeled or attached to a solid substrate for use in a screening assay.

The disclosure includes antibodies that specifically bind a 100 kDa glycoprotein that has a membrane bound and a soluble form, has a UV absorbance peak at about 228nm, an isoelectric point of about 3.5 to 4, a sialic acid content of about 20%, and do not substantially bind to ACT are described herein.

In embodiments, an antibody or antigen binding fragment specifically binds to 100 kDa glycoprotein comprising an amino acid sequence of SEQ ID NO:1 and does not substantially bind to one or more linear peptides consisting of a linear peptide of 15 amino acids from amino acids 1-60 of SEQ ID NO:1 as shown in SEQ ID NO:2, a linear peptide of amino acids 111 to 125 of SEQ ID NO:1 as shown in SEQ ID NO:3 and/or a linear peptide of 15 amino acids from amino acids 150-701 of SEQ ID NO:1 as shown in SEQ ID NO:4. In embodiments, the antibodies do not substantially bind to epitopes in domains of CEACAM5 selected from the group consisting of amino acids 146-237 (A1 Ig-2), amino acids 238-323 (B1 Ig-3), amino acids 324-415 (A2 Ig-4), amino acids 416-498 (B2 Ig-5), amino acids 502-593 (A3 Ig-6) and amino acids 594-677 (B3 Ig-7) and combinations thereof. Amino acid numbering of CEACAM5 corresponds to that of in NP_004354;gI 98986445; Uniprot P06731-1.

In embodiments, an antibody or antigen binding fragment does not substantially cross react with CEA of a molecular weight of 180 kDa, CEACAM 6(NCA), CEACAM5 lacking all glycosylation, and/or 100kDa glycoprotein having a sequence of SEQ ID NO:1 lacking all glycosylation.

In embodiments, an antibody or antigen binding fragment specifically binds to an epitope on a 100 kDa glycoprotein comprising an amino acid sequence of SEQ ID NO:1, the epitope comprising, consisting essentially of, or consisting of amino acids 61-69 HLFGYSWYK (SEQ ID NO:6), 73-77 VDGNR (SEQ ID NO:7) or 69-82 KGERVDGNRQIGY (SEQ ID NO:8), and 96-107 SGREI IYPNASL (SEQ ID NO:9) of SEQ ID NO:1. In embodiments, an antibody specifically binds to an epitope on a 100 kDa glycoprotein comprising an amino acid sequence of SEQ ID
NO:1, the epitope comprising, consisting essentially of, or consisting of amino acids 61-69 (SEQ ID NO:6), 78-98 QIIGYVIGTQQAT PGPAYSGR (SEQ ID NO:10), 96-107 (SEQ ID NO:9), and 127-139 SDLVNEATGQFR (SEQ ID NO:11). Amino acid numbering corresponding to that of SEQ ID NO:1. In embodiments, the epitope is a conformational epitope in the N terminal amino acids (amino acids 1-150; SEQ ID NO:5) of the 100kDa glycoprotein. In embodiments, the antibodies do not substantially bind to epitopes in other domains of CEACAM5.

In embodiments, an antibody or antigen binding fragment thereof comprises heavy chain CDRs comprising: CDRH1: DSYIN (SEQ ID NO:14) or GNTFTDSYIN (SEQ ID NO:15); CDRH2: EIYPNGDVYYNENFK (SEQ ID NO:16); CDRH3: TTVFAY (SEQ ID NO:17) or LCAGSNMITTVFAY (SEQ ID NO:18). In embodiments, an antibody or antigen binding fragment thereof comprises heavy chain CDRs comprising: CDRH1: GYTFTNYWIN (SEQ ID NO:19) or NYWIN (SEQ ID NO:20); CDRH2: NIYPGSTRANYNEKFK (SEQ ID NO:21); CDRH3: YCTRTHSI (SEQ ID NO:22).

In embodiments, an antibody or antigen binding fragment thereof comprises light chain CDRs comprising: CDRL1: RASQDI RNYL (SEQ ID NO:26); CDRL2: YTSRLHS (SEQ ID NO:27) or YTS of SEQ ID NO:27; CDRL3: QQGNTLPW (SEQ ID NO:28). In embodiments, an antibody or antigen binding fragment thereof comprises light chain CDRs comprising: CDRL1: RASQSISSYL (SEQ ID NO:31); CDRL2: AASSLQS (SEQ ID NO:32) of AAS of SEQ ID NO:32; CDRL3: QQTVVAPP (SEQ ID NO:33).

In embodiments, an antibody or antigen binding fragment thereof comprises HCDR1/HCDR2/HCDR3 of SEQ ID NOs: 14/16/17, SEQ ID NOs:15/16/17, SEQ ID NOs:14/16/18, or SEQ ID NOs:15/16/18; and LCDR1/LCDR2/LCDR3 of SEQ ID NOs:26/27/28. In embodiments, an antibody or antigen binding fragment thereof comprises HCDR1/HCDR2/HCDR3 of SEQ ID NOs:20/21/22 or SEQ ID NOs:19/21/22, and LCDR1/LCDR2/LCDR3 of SEQ ID NOs:3 1/32/33.

In embodiments, an antibody or antigen binding fragment thereof comprises a heavy chain variable region having an amino acid sequence of SEQ ID NO: 119 and light chain variable region having a sequence of SEQ ID NO: 120. In other embodiments, an antibody or antigen binding fragment comprises a heavy chain.
variable region having an amino acid sequence of SEQ ID NO: 121 and light chain variable region having an amino acid sequence of SEQ ID NO: 122.

Antibodies can be polyclonal, monoclonal, or recombinant. Antibodies can be an antigen binding fragment such as of Fab, single chain antibodies, scFv, F(ab)_2 and Fv fragments. Antibodies can be human, synthetic, (e.g. phage display), chimeric, or humanized antibodies. In embodiments, antibodies are labeled with a signal generating element. In other embodiments, the label is selected from the group consisting of biotin, fluorescent dyes, chemiluminescent tags, radioactive tags, enzymes, and combinations thereof. In yet other embodiments, the antibody is immobilized on a solid surface.

The disclosure provides immunoassays. In embodiments, the method further comprises testing the subject for a cancer selected from the group consisting of colon-rectal and lung. In embodiments, the testing involves detecting a biomarker of the cancer, imaging of the body of the subject such as a CAT, Pet or MRI scan, and/or conducting a colonoscopy.

In embodiments, the immunoassay is a sandwich assay. An immunoassay comprises a first antibody that specifically binds to a 100 kDa glycoprotein having an amino acid sequence of SEQ ID NO: 1, and does not substantially bind to ACT attached to a solid substrate (e.g. a capture antibody), and a second antibody that specifically binds to a 100 kDa glycoprotein that has an amino acid sequence of SEQ ID NO: 1, that is labeled with a signal generating element. In embodiments, the first antibody is a monoclonal antibody or a polyclonal antibody. In embodiments, the second antibody is a monoclonal antibody or polyclonal antibody.

The disclosure also provides kits. In embodiments, a kit comprises: a) a first antibody that specifically binds to a 100 kDa glycoprotein that has an amino acid sequence of SEQ ID NO: 1 and that does not substantially bind ACT; b) a second antibody that specifically binds to a 100 kDa glycoprotein that has an amino acid sequence of SEQ ID NO: 1 and is labeled with a signal generating element; and c) a calibrator. In embodiments, the first antibody is a monoclonal antibody or a polyclonal antibody. In embodiments, the second antibody is a monoclonal antibody or polyclonal antibody. In embodiments, the first antibody has different epitope specificity than the second antibody. In embodiments, the kit also provides
instructions for identifying a subject as "at risk" for colorectal cancer if the amount of antigen detected is about 6.5 units/ml or greater.

**BRIEF DESCRIPTION OF THE FIGURES**

Fig. 1 shows a summary of all data from the 422 patient clinical study. **Panel A** shows the combined data for all non-cancerous and cancerous patients. The box covers the middle 50% of the data while the line is the median. The whiskers cover the middle 90% of the data (from 5% to 95%). **Panel B** is a receiver operator characteristic (ROC) curve comparing the cancerous vs. non-cancerous samples.

Fig. 2 shows SDS-PAGE and goat anti-CA1 1-19 polyclonal antibodies Conjugate Western of tumor and normal tissue perchloric acid extracts. CODES 1, 2, 7, 8, 15, 23, 25, 32, 38, 39, 41, 42, 45, 50 are Normal Tissue Perchloric acid extracts. **Panels A & B** are the initial investigation of ~50 tumor extracts.

Fig. 3 shows fractionation of a perchloric acid extract of two separate colorectal tumor cell samples on Sepharose B. Absorbance at 280 for each sample is shown by (▲; -i)The fraction of one of the samples were also evaluated in the ELISA assay(4).

Fig. 4 shows methods for N-glycan profiling by MALDI-MS of glycoprotein 100 kDa obtained from an acid extract of a colon cancer cell.

Figure 5A shows localization of target glycoprotein 100 kDa band and an additional band at 45 kDa observed in the CA1 1-19 168B. Figure 5B shows pictures of actual gels used for sample analysis.

Figure 6A shows the MALDI-full mass spectra of the 100 kDa band. Figure 6B shows the MALDI-full mass spectra of the 45 kDa band.

Figure 7 shows the sample preparation methods for O-glycan profiling by MALDI-MS.

Figure 8 shows CA1 1-19 168BNC and CA1 1-19 168BSTK separated by SDS-PAGE for in-gel digestion.

Figure 9 shows NSI-FTMS spectra of O-glycan from the 100 kDa band from the both samples. A) is the profile for antigen sample NC. B) is the profile for antigen sample STK.
Figure 10 A) The MSMS profile of the molecular ion at m/z 936 (NeuAc-HexNAc-HexNAc), the diagnostic b ion is shown at m/z 398, together with the complementary y ions at m/z 561, m/z 659, and m/z 677. B) The MSMS profile of the molecular ion at m/z 953 (Hex-DHex-HexNAc2) (10B) and C) The MSMS profile of the molecular ion at m/z 1157 (Hex2-DHex-HexNAc2) (IOC).

Figure 11 Analysis of binding of monoclonal antibodies specific for the 100 kDa antigen isolated from colon or liver cancer cells designated antibody 5E5-1 and Antibody 5A1-1 to the 100 kDa antigen designated as 168B. A) Analysis of the mix A: 5E5-1/168B, showed that the antigen and the antibody were detected with observed MH+=92.879 kDa (168B-antigen) and MH+=158.498 kDa, (5E5-1 antibody); B) The cross-linking experiment was completed after 180 minutes incubation time with the crosslinking reagent K200. After cross-linking, we detected two additional peaks with MH+=258.945 kDa [5E5-1-B168] and MH+=354.164 kDa [5E5-1-2B168 -binding two antigens]. C) Using Complex Tracker software, we overlaid the control and cross-link spectra. The overlay confirmed the detection of two non covalent complexes [5E5-1·B168] and [5E5-1·2B168].

Figure 12 A) Analysis of the Mix C (both monoclonal antibodies plus antigen): 5E5-1/5A1-1/168B showed that the antigen and the antibody were detected with MH+=92.910 kDa (168B) and MH+=158.988 kDa (5E5, 5A1-1). B) The crosslinking experiment was completed after 180 minutes incubation time with the crosslinking reagent K200. After cross-linking, we detected two additional peaks with MH+=260.155 kDa and MH+=356.012 kDa. C) Using Complex Tracker software, we overlaid the control and cross-link spectra. The overlay confirmed the detection of two non covalent complexes [5E5-1·B168] and [5E5-1·2B168].

Figure 13 A) After competition with trypsin peptides, we did not detect inhibition of the binding of the antibody 5E5-1 on the antigen 168B. B) After competition, we observed two non covalent complexes: [5E5-1·B168] and [5E5-1·2B168] C) Using Complex Tracker software, we overlaid the control and cross-link spectra. The overlay confirmed the detection of two non covalent complexes [5E5-1·B168] and [5E5-1·2B168].

Figure 14 shows the sequence of the 168B antigen. Overlap mapping of the trypsin, chymotrypsin and ASP-N peptides covered, 85.6% of the sequence.
Figure 15 shows the map of the epitope bound by antibody 5E5-1. Amino acids in bold indicate potential points of contact between antibody 5E5-1 and antigen 168B.

Figure 16 shows the map of the epitope bound by antibody 5A1-1. Amino acids in bold indicate potential points of contact between antibody 5A1-1 and antigen 168B.

Figure 17 shows the amino acid sequence of the heavy and light chain of antibody 5E5-1. CDR regions are highlighted in bold. Smaller definitions of CDRs are underlined. N terminal sequence of the heavy and light variable domains are identified by italics.

Figure 18 shows the nucleic acid sequence of the heavy and light chain of antibody 5E5-1.

Figure 19 shows the amino acid sequence of the heavy and light chain of antibody 5A1-1. CDR regions are highlighted in bold. Smaller definitions of CDRs are underlined. N terminal sequence of the heavy and light variable domains are identified by italics.

Figure 20 shows the nucleic acid sequence of the heavy and light chain of antibody 5A1-1.

Figure 21 shows a lateral flow test to detect CA1 1-19 antigen in serum or urine as seen with + line in colorectal (# 2) and lung (# 4 and 6) CA patients and with no + line in normal individuals (1,3,5 and7).

DETAILED DESCRIPTION OF INVENTION

It should be understood that this invention is not limited to the particular methodology, protocols, and reagents, etc., described herein and as such may vary. The terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention, which is defined solely by the claims.

As used herein and in the claims, the singular forms "a," "an," and "the" include the plural reference unless the context clearly indicates otherwise. Thus, for example, the reference to an antibody is a reference to one or more such antibodies, including equivalents thereof known to those skilled in the art. Other than in the operating examples, or where otherwise indicated, all numbers expressing quantities
of ingredients or reaction conditions used herein should be understood as modified in all instances by the term "about." The term "about" when used in connection with numerical values means ± 20% and with percentages means ±1%.

All patents and other publications identified are expressly incorporated herein by reference for the purpose of describing and disclosing, for example, the methodologies described in such publications that might be used in connection with the present invention. These publications are provided solely for their disclosure prior to the filing date of the present application. Nothing in this regard should be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention or for any other reason. All statements as to the date or representation as to the contents of these documents is based on the information available to the applicants and does not constitute any admission as to the correctness of the dates or contents of these documents.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as those commonly understood to one of ordinary skill in the art to which this invention pertains.

For the purposes of this application the following terms shall have the following meanings:

An "epitope" refers to that portion of any molecule capable of being recognized by, and bound by, an antibody (the corresponding antibody binding region may be referred to as a paratope). In general, epitopes comprise chemically active surface groupings of molecules, for example, amino acids or sugar side chains, and have specific three-dimensional structural characteristics as well as specific charge characteristics.

An "antigen" is a molecule or a portion of a molecule capable of being bound by an antibody. An antigen may have one or more than one epitope. An antigen will bind in a highly selective manner with its corresponding antibody and not with the multitude of other antibodies which may be evoked by other antigens.

An "antibody" includes both intact immunoglobulin molecules as well as portions, fragments, peptides and derivatives thereof, such as, for example, Fab, Fab', F(ab')2, Fv, scFv, CDR regions, or any portion or peptide sequence of the antibody that is capable of binding antigen or epitope. An antibody is said to be
"capable of binding" a molecule if it is capable of specifically reacting with the molecule to thereby bind the molecule to the antibody.

Antibody also includes chimeric antibodies, anti-idiotypic (anti-Id) antibodies to antibodies that can be labeled in soluble or bound form, as well as fragments, portions, regions, peptides or derivatives thereof, provided by any known technique, such as, but not limited to, enzymatic cleavage, peptide synthesis, phage display, or recombinant techniques. Antibody fragments or portions may lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. Examples of antibody may be produced from intact antibodies using methods well known in the art, for example by proteolytic cleavage with enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). See e.g., Wahl et al., 24 J. Nucl. Med. 316-25 (1983). Portions of antibodies may be made by any of the above methods, or may be made by expressing a portion of the recombinant molecule. For example, the CDR region(s) of a recombinant antibody may be isolated and subcloned into the appropriate expression vector. See, e.g., U.S. Pat. No. 6,680,053.

A "monoclonal antibody" refers to a homogeneous antibody population involved in the highly specific recognition and binding of a single antigenic determinant, or epitope. This is in contrast to polyclonal antibodies that typically include different antibodies directed against different antigenic determinants. The term "monoclonal antibody" encompasses both intact and full-length monoclonal antibodies as well as antibody fragments (such as Fab, Fab', F(ab')2, Fv), single chain (scFv) mutants, fusion proteins comprising an antibody portion, and any other modified immunoglobulin molecule comprising an antigen recognition site.

Furthermore, "monoclonal antibody" refers to such antibodies made in any number of manners including but not limited to by hybridoma, phage selection, recombinant expression, and transgenic animals.

The term "human antibody" means an antibody produced by a human or an antibody having an amino acid sequence corresponding to an antibody produced by a human made using any technique known in the art. This definition of a human antibody includes intact or full-length antibodies, fragments thereof, and/or antibodies comprising at least one human heavy and/or light chain polypeptide such
as, for example, an antibody comprising murine light chain and human heavy chain polypeptides.

The term "humanized antibody" refers to forms of non-human (e.g. murine) antibodies that are specific immunoglobulin chains, chimeric immunoglobulins, or fragments thereof that contain minimal non-human (e.g., murine) sequences.

Typically, humanized antibodies are human immunoglobulins in which residues from the complementary determining region (CDR) are replaced by residues from the CDR of a non-human species (e.g. mouse, rat, rabbit, hamster) that have the desired specificity, affinity, and capability (Jones et al, 1986, Nature, 321:522-525; Riechmann et al, 1988, Nature, 332:323-327; Verhoeyen et al, 1988, Science, 239:1534-1536). In some instances, the Fv framework region (FR) residues of a human immunoglobulin are replaced with the corresponding residues in an antibody from a non-human species that has the desired specificity, affinity, and capability. The humanized antibody can be further modified by the substitution of additional residues either in the Fv framework region and/or within the replaced non-human residues to refine and optimize antibody specificity, affinity, and/or capability. In general, the humanized antibody will comprise substantially all of at least one, and typically two or three, variable domains containing all or substantially all of the CDR regions that correspond to the non-human immunoglobulin whereas all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody can also comprise at least a portion of an immunoglobulin constant region or domain (Fc), typically that of a human immunoglobulin. Examples of methods used to generate humanized antibodies are described in U.S. Pat. No. 5,225,539.

The term "chimeric antibodies" refers to antibodies wherein the amino acid sequence of the immunoglobulin molecule is derived from two or more species. Typically, the variable region of both light and heavy chains corresponds to the variable region of antibodies derived from one species of mammals (e.g. mouse, rat, rabbit, etc) with the desired specificity, affinity, and capability while the constant regions are homologous to the sequences in antibodies derived from another (usually human) to avoid eliciting an immune response in that species.

As used herein, "alpha-1-antichymotrypsin", or "ACT" refers to a polypeptide having an amino acid sequence the same as that derived from a naturally
occurring noncomplexed polypeptide and that has serine protease inhibitory activity. ACT is also known as SERPINA3, AACT, growth inhibiting protein 24 (GIG24), growth inhibiting protein 25 (GIG25), cell growth inhibiting gene 24/25 protein, and serine proteinase inhibitor clade A, member 3. A representative sequence of ACT is NP_001076/gI 50659080.

As used herein, "carcinoembryonic antigen" or "CEA" refers to a family of polypeptides having an amino acid sequence the same as a naturally occurring polypeptide, and is glycosyl phosphotidyl inositol cell surface anchored glycoprotein. CEA is generally understood to refer to a glycoprotein of a molecular weight of 180 kDa.

As used herein "CEACAM5" refers to a carcinoembryonic antigen that is a member of a family of carcinoembryonic antigens. CEACAM5, without any glycosylation, has an exemplary amino acid sequence found in NP_004354;gI 98986445; Uniprot P06731-1. An isoform of CEACAM5 has an amino acid deletion at amino acid position 320 and has an amino acid sequence of SEQ ID NO: 1. (Exemplary sequence found in Uniprot P06731-2). Unglycosylated CEACAM5 contains a signal peptide of 1-34, a protein of 35-685, and amino acids 686-702 that are removed in the mature form. Within the protein several domains are found at amino acids 35-145 (N terminal Ig-1), amino acids 146-237 (Al Ig-2), amino acids 238-323 (B1 Ig-3), amino acids 324-415(A2 Ig-4), amino acids 416-498 (B2 Ig-5), amino acids 502-593 (A3 Ig-6) and amino acids 594-677 (B3 Ig-7). Amino acid numbers refer to those of found in NP_004354;gI 98986445; Uniprot P06731-1. The amino acid numbers will be shifted one number from that of amino acid reference sequence of SEQ ID NO: 1 because of the deletion of amino acid 320. Antibodies to the domains of CEACAM5 have been characterized as Gold epitope designations: Gold 1: A3B3; Gold 2-A2B2; Gold 3-A3B3; Gold 4-A1B1; and Gold 5-N terminal. CEACAM5 can be produced as a heavily glycosylated form having a molecular weight of 180 kDa and has 28 potential N glycosylation sites.

As used herein, the term "not substantially bind" or "does not bind" can be used interchangeably and mean that the detectable signal from the binding of the antibody to a component in a sample is within one or two standard deviations of the signal generated due to the presence of an unrelated polypeptide control such as bovine serum albumin.
As used herein "specific binding" refers to an antibody that reacts or associates more frequently, more rapidly, with greater duration, with greater affinity, or with some combination of the above to an epitope or protein than with alternative substances, including unrelated proteins. For example, specifically binds means that the detectable signal from the binding of the antibody to a component in a sample is greater than one or two standard deviations of the signal generated due to the presence of an unrelated polypeptide control such as bovine serum albumin. In certain embodiments, "specifically binds" means, for instance, that an antibody binds to a protein with a $K_D$ of about $0.1 \text{ mM}$ or less, but more usually less than about $1 \text{ nM}$. In certain embodiments, "specifically binds" means that an antibody binds to a protein at times with a $K_D$ of at least about $0.1 \text{ µM}$ or less, and at other times at least about $0.01 \text{ µM}$ or less. It is understood that an antibody or binding moiety that specifically binds to a first target may or may not specifically bind to a second target. As such, "specific binding" does not necessarily require (although it can include) exclusive binding, i.e. binding to a single target. Thus, an antibody may, in certain embodiments, specifically binds to more than one target. In certain alternative embodiments, an antibody may be bispecific and comprise at least two antigen-binding sites with differing specificities.

The term "comprising" refers to a composition, compound, formulation, or method that is inclusive and does not exclude additional elements or method steps.

The term "consisting of" refers to a compound, composition, formulation, or method that excludes the presence of any additional component or method steps.

The term "consisting essentially of" refers to a composition, compound, formulation or method that is inclusive of additional elements or method steps that do not materially affect the characteristic(s) of the composition, compound, formulation or method.

The term "isolated" refers to the separation of a material from at least one other material in a mixture or from materials that are naturally associated with the material.

The terms "patient" or "subject" are used interchangeably and refer to any member of Kingdom Animalia. Preferably a subject is a mammal, such as a human, domesticated mammal or a livestock mammal.
The phrase "pharmaceutically acceptable" refers to those compounds, materials, compositions, and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of human beings and animals without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ration.

The phrase "pharmaceutically-acceptable carrier" refers to a pharmaceutically-acceptable material, composition or vehicle, such as a liquid or solid filler, diluent, excipient, solvent or encapsulating material, involved in carrying or transporting the compound or analogue or derivative from one organ, or portion of the body, to another organ, or portion of the body. Each carrier must be "acceptable" in the sense of being compatible with the other ingredients of the formulation and not injurious to the patient. Some examples of materials which may serve as pharmaceutically-acceptable carriers include: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, ethyl cellulose and cellulose acetate; (4) powdered tragacanth; (5) malt; (6) gelatin; (7) talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol; (12) esters, such as ethyl oleate and ethyl laureate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) phosphate buffer solutions; and (21) other non-toxic compatible substances employed in pharmaceutical formulations.

The term "purified" or "to purify" or "substantially purified" refers to the removal of inactive or inhibitory components (e.g., contaminants) from a composition to the extent that 10% or less (e.g., 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, > 1% or less) of the composition is not active compounds or pharmaceutically acceptable carrier.

As used herein, the term "at risk for disease" (e.g., at risk for, cancer, etc.) refers to a subject (e.g., a human) that is predisposed to experiencing a particular disease (e.g., cancer, etc.). This predisposition may be genetic (e.g., a particular genetic tendency to experience the disease, such as heritable disorders), or due to
other factors (e.g., hypertension, age, weight, environmental conditions, exposures to detrimental compounds present in the environment, etc.).

**Immunogenic compositions**

The disclosure provides immunogenic compositions comprising an isolated fraction from colorectal and/or liver cancer cell samples having components with a molecular weight that comprises a 100 kDa glycoprotein that has a membrane bound and a soluble form (e.g. cancer cell membrane bound form), has a UV absorbance peak at about 228nm, an isoelectric point of about 3.5 to 4, a sialic acid content of about 20%, and does not substantially bind to antibodies specific for ACT. In embodiments, the immunogenic composition does not include a substantial amount of ACT, for example, 10% or less of the protein on a per weight basis. In embodiments, immunogenic compositions include a soluble form of the membrane bound form of the 100kDa glycoprotein.

In embodiments, the immunogenic compositions and/or the isolated fraction are obtained by a method comprising contacting colorectal and/or liver cancer cells with an acid to form an extract; separating components of the extract by molecular weight to form a first fraction having components with a molecular weight of about 60 kDa or greater; isolating components of the first fraction that have a molecular weight of about 75 kDa or greater and less than about 200 kDa, and further isolating a 100kDa protein. In embodiments, the acid is selected from the group consisting of trichloroacetic acid, dichloroacetic acid, tribromic acid, trifluoric acid, sodium trichloroacetate, hydrochloric acid, and perchloric acid. In embodiments, the acid concentration ranges from about 1 to about 50%. In embodiments, the components are separated by gel filtration. In embodiments, the gel filtration medium is selected from the group consisting of Sepharose, and Superdex gel filtration medium.

In embodiments, the immunogenic composition comprises a 100 kDa glycoprotein having an amino acid sequence of SEQ ID NO: 1. In embodiments, the 100 kDa has a sequence that does not include the leader sequence of amino acids 1-34 of SEQ ID NO: 1. While not meant to limit the scope of the disclosure it is thought that the 100 kDa is a form of CEACAM5 with a deletion of an amino acid at position 320 and altered glycosylation as compared to the 180 kDa CEA.

Immunogenic compositions are optionally combined with an adjuvant. An immunogenic effective amount of the antigen is that amount that stimulates an
immune response. Immunogenic effective amounts can be determined based on the animal used for immunization. Adjuvants include alum, muramyl peptides, monophosphoryl lipid A, liposomes, incomplete Freunds’ adjuvant, and complete Freund’s adjuvant.

A typical immunization protocol involves injecting an animal such as mice, goats, sheep, or primates with the immunogenic composition. At least one booster is administered about two weeks after the initial immunization.

Alternatively, the immunogenic composition can be attached to a solid substrate and used to bind antibodies produced in animal models or by phage display. Solid surfaces include membranes, multiwell plates, chromatography media, glass slides, latex beads, magnetic beads, microarray chips, capillary tubes and chips. Similarly, the immunogenic composition may be detectably labeled with a signal generating element. Such labels include fluorescent markers, biotin, enzymes, radioactive labels and the like.

**Antibodies**

The present disclosure includes novel antibodies, methods for producing the antibodies, and diagnostic and therapeutic methods employing the antibodies.

In embodiments, an antibody specifically binds to a 100 kDa glycoprotein that has a colorectal cancer membrane bound and a soluble form, has a UV absorbance peak at about 228nm, an isoelectric point of about 3.5 to 4, a sialic acid content of about 20%, and does not substantially bind to 180 kDa glycosylated CEA and/or ACT. In embodiments, antibodies are further screened for binding to human cancer cells, particularly from the lung and/or colon. The antibodies are selected to react with a range of cancers (e.g. solid tumors) while showing essentially no reactivity with normal human tissues or other types of tumors such as lymphomas.

In embodiments, an isolated antibody or antigen-binding fragment thereof specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, but does not substantially bind to glycosylated 180 kDa CEA and/or does not substantially bind to one or more linear peptides consisting of a linear peptide of 15 amino acids from amino acids 1-60 of SEQ ID NO:2, a linear peptide of amino acids 110 to 125 of SEQ ID NO:3, and/or a linear peptide of 15 amino acids from amino acids 150-701 of SEQ ID NO:4. In embodiments, the antibodies are further selected to react with colorectal cancer cells.
In embodiments, the antibody or antigen binding fragment binds to a conformational epitope. In some embodiments, the antibody or antigen binding fragment specifically binds to and/or competes for binding to an epitope comprising, consisting essentially of, consisting of amino acids 61-69 (HLFGYSWYK; SEQ ID NO:6), 73-77 (VDGNR; SEQ ID NO:7) or 69-82 KGERVDGNRQIIGY (SEQ ID NO:8), and 96-107(SGREI IYPNASL; SEQ ID NO:9) of SEQ ID NO:1. In other embodiments, the antibody or antigen binding fragment binds to an epitope comprising, consisting essentially of, consisting of amino acids 61-69(SEQ ID NO:6), 78-98(QII GYVIGTQQAT PGPAYSGR;SEQ ID NO:10), 96-107(SEQ ID NO:9), and 127-139(SDLVNEATGQFR; SEQ ID NO:11) of SEQ ID NO:1. In embodiments, the conformational epitope retains glycosylation.

The antibody or antigen binding fragments thereof does not substantially bind to one or more of the linear peptides of an amino acid sequence selected from the group consisting of SEQ ID NOs:40-101.

In embodiments, the antibody or antigen binding fragment does not substantially bind to epitopes in one or all other domains of CEACAM5 selected from the group consisting of amino acids 146-237 (AI Ig-2), amino acids 238-322 (BI Ig-3), amino acids 324-415(A2 Ig-4), amino acids 416-498 (B2 Ig-5), amino acids 502-593 (A3 Ig-6) and amino acids 594-677 (B3 Ig-7).

In embodiments, an isolated antibody or antigen-binding fragment thereof specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, but does not substantially bind to glycosylated 180kDa CEA, CEACAM6 (NCA) or unglycosylated CEACAM5 and/or does not substantially bind one or more linear peptides consisting of a linear peptide of 15 amino acids from amino acids 1-60 of SEQ ID NO:2, a linear peptide of amino acids 111 to 125 of SEQ ID NO:3, a linear peptide of 15 amino acids from amino acids 150-701 of SEQ ID NO:4, a peptide of amino acids 1-32 including the leader sequence (SEQ ID NO:35), a peptide of amino acids 1-32 without the leader sequence (SEQ ID NO:36), a peptide of amino acids 42-60 (SEQ ID NO:37), a peptide of amino acids 117-127 with the leader(SEQ ID NO:38), a peptide of amino acids 117-127 with the leader (SEQ ID NO:39), and combinations thereof.

In embodiments, an isolated antibody or antigen-binding fragment thereof that specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, but
does not substantially bind to glycosylated 180kDa CEA and comprises: (a) heavy chain CDRs (HCDR1, HCDR2 and HCDR3) from a heavy chain variable region having an amino acid sequence selected from the group consisting of SEQ ID NO:1 19 and SEQ ID NO: 121, and (b) light chain CDRs (LCDR1, LCDR2 and LCDR3) from a light chain variable region having an amino acid sequence selected from the group consisting of SEQ ID NO: 120 and SEQ ID NO: 122.

In other embodiments, an isolated antibody or antigen-binding fragment thereof comprises: a) HCDR1, HCDR2, and HCDR3 selected from the group consisting of SEQ ID NOs: 15/16/17, 14/16/17, 18/20/21 and SEQ ID NO:19/20/21; and b) LCDR1, LCDR2, and LCDR3 selected from the group consisting of SEQ ID NOs:26/27/28 and SEQ ID NOs:31/32/33.

In yet other embodiments, an isolated antibody or antigen binding fragment thereof comprises: a) a heavy chain variable region having an amino acid sequence SEQ ID NO: 119 and a light chain variable region having an amino acid sequence of SEQ ID NO:120; or b) a heavy chain variable region having an amino acid sequence SEQ ID NO:121 and a light chain variable region having an amino acid sequence of SEQ ID NO:122.

Antibodies include monoclonal, polyclonal, human, humanized, chimeric antibodies, anti-idiotypic (anti-Id) antibodies to antibodies that can be labeled in soluble or bound form, as well as fragments, portions, regions, peptides or derivatives thereof, provided by any known technique, such as, but not limited to, enzymatic cleavage, peptide synthesis, phage display, or recombinant techniques.

Antibody fragments or portions may lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody such as, for example, Fab, Fab', F(ab')2, Fv, scFv, CDR regions, or any portion or peptide sequence of the antibody that is capable of binding antigen or epitope. Antibody fragments may be produced from intact antibodies using methods well known in the art, for example by proteolytic cleavage with enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). See e.g., Wahl et al, 24 J. Nucl. Med. 316-25 (1983). Portions of antibodies may be made by any of the above methods, or may be made by expressing a portion of the recombinant molecule. For example, the CDR region(s)
of a recombinant antibody may be isolated and subcloned into the appropriate expression vector. See, e.g. U.S. Pat. No. 6,680,053.

Polyclonal antibodies can be prepared by any known method. Polyclonal antibodies are raised by immunizing an animal (e.g. a rabbit, goat, rat, mouse, donkey, etc) by multiple subcutaneous or intraperitoneal injections of the relevant antigen (a purified peptide fragment, full-length recombinant protein, fusion protein, etc) optionally conjugated to keyhole limpet hemocyanin (KLH), serum albumin, etc. diluted in sterile saline and combined with an adjuvant (e.g. Complete or Incomplete Freund's Adjuvant) to form a stable emulsion. The polyclonal antibody is then recovered from blood, ascites and the like, of an animal so immunized. Collected blood is clotted, and the serum decanted, clarified by centrifugation, and assayed for antibody titer. The polyclonal antibodies can be purified from serum or ascites according to standard methods in the art including affinity chromatography, ion-exchange chromatography, gel electrophoresis, dialysis, etc.

The monoclonal antibodies of the present disclosure can be prepared by hybridoma fusion techniques, (see, Kohler and Milstein, Nature, 256:495-97 (1975); Brown et al, J. Immunol, 127 (2):539-46 (1981); and Yeh et al, Int. J. Cancer, 29:269-75 (1982)). These techniques involve the injection of an immunogen (e.g., purified antigen or cells or cellular extracts carrying the antigen) into an animal (e.g., a mouse) so as to elicit a desired immune response (i.e., production of antibodies) in that animal. The immunogenic composition is injected, for example, into a mouse, and after a sufficient time the mouse is sacrificed and somatic antibody-producing lymphocytes are obtained. Antibody-producing cells may be derived from the lymph nodes, spleens and peripheral blood of primed animals.

Spleen cells are preferred. Mouse lymphocytes give a higher percentage of stable fusions with the mouse myelomas described below. The use of rat, rabbit and frog somatic cells is also possible. The spleen cell chromosomes encoding desired immunoglobulins are immortalized by fusing the spleen cells with myeloma cells, generally in the presence of a fusing agent such as polyethylene glycol (PEG). Any of a number of myeloma cell lines may be used as a fusion partner according to standard techniques; for example, the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/0-Agl4 myeloma lines. These myeloma lines are available from the American Type Culture Collection (ATCC), Rockville, Md.
The resulting cells, which include the desired hybridomas, are then grown in a selective medium, such as HAT medium, in which unfused parental myeloma or lymphocyte cells eventually die. Only the hybridoma cells survive and can be grown under limiting dilution conditions to obtain isolated clones. The supernatants of the hybridomas are screened for the presence of antibody of the desired specificity, e.g., by immunoassay techniques using the antigen that has been used for immunization. Positive clones can then be subcloned under limiting dilution conditions, and the monoclonal antibody produced can be isolated. Various conventional methods exist for isolation and purification of the monoclonal antibodies so as to free them from other proteins and other contaminants. Commonly used methods for purifying monoclonal antibodies include ammonium sulfate precipitation, ion exchange chromatography, and affinity chromatography (see, e.g., Zola et al., in Monoclonal Hybridoma Antibodies: Techniques and Applications, Hurell (ed.) pp. 51-52 (CRC Press 1982)). Hybridomas produced according to these methods can be propagated in vitro or in vivo (in ascites fluid) using techniques known in the art (See, generally, Fink et al, supra, at page 123, FIG. 6-1).

Generally, the individual cell line may be propagated in vitro, for example in laboratory culture vessels, and the culture medium containing high concentrations of a single specific monoclonal antibody can be harvested by decantation, filtration or centrifugation. Alternatively, the yield of monoclonal antibody can be enhanced by injecting a sample of the hybridoma into a histocompatible animal of the type used to provide the somatic and myeloma cells for the original fusion.

Alternatively, monoclonal antibodies can also be made using recombinant DNA methods as described in U.S. Pat. No. 4, 816, 567. The polynucleotides encoding a monoclonal antibody are isolated from mature B-cells or hybridoma cell, such as by RT-PCR using oligonucleotide primers that specifically amplify the genes encoding the heavy and light chains of the antibody, and their sequence is determined using conventional procedures. The isolated polynucleotides encoding the heavy and light chains are then cloned into suitable expression vectors, which when transfected into host cells such as E. coli cells, simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, monoclonal antibodies are generated by the host cells. Also, recombinant monoclonal antibodies or fragments thereof of the desired species

The polynucleotide(s) encoding a monoclonal antibody can further be modified in a number of different manners using recombinant DNA technology to generate alternative antibodies. In some embodiments, the constant domains of the light and heavy chains of, for example, a mouse monoclonal antibody can be substituted 1) for those regions of, for example, a human antibody to generate a chimeric antibody or 2) for a non-immunoglobulin polypeptide to generate a fusion antibody. In some embodiments, the constant regions are truncated or removed to generate the desired antibody fragment of a monoclonal antibody. Site-directed or high-density mutagenesis of the variable region can be used to optimize specificity, affinity, etc. of a monoclonal antibody.

Human or humanized antibodies can be directly prepared using various techniques known in the art. Immortalized human B lymphocytes immunized in vitro or isolated from an immunized individual that produce an antibody directed against a target antigen can be generated (See, e.g., U.S. Pat. No 5,750,373). Human antibodies can also be made in transgenic mice containing human immunoglobulin loci that are capable upon immunization of producing the full repertoire of human antibodies in the absence of endogenous immunoglobulin production. This approach is described in U.S. Pat. Nos. 5,545, 807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; and 5,661,016. Examples of methods used to generate humanized antibodies are described in U.S. Pat. No.5,225, 539.

In embodiments, nucleic acids coding for a heavy chain and light chain variable domains of an antibody of the specificity as described herein are shown in Figures 19 and 20. A nucleic acid coding for a heavy chain variable domain of an antibody as described herein comprises a nucleic acid sequence of SEQ ID NO: 12 or SEQ ID NO:24. A nucleic acid coding for a light chain of an antibody as described herein comprises a nucleic acid sequence of SEQ ID NO:25 or SEQ ID NO:30. In embodiments, the present disclosure includes a host cell comprising a nucleic acid sequence of SEQ ID NO: 12 or SEQ ID NO:24 and/or SEQ ID NO:25 or SEQ ID NO:30.

Screening of antibodies for desired binding specificity is conducted according to standard methods. In embodiments, an antibody specifically binds to a 100 kDa glycoprotein that has a colorectal cancer membrane bound and a soluble form, has a UV absorbance peak at about 228nm, an isoelectric point of about 3.5 to 4, a sialic acid content of about 20%. In embodiments, the antibodies are also screened for binding to the 100 kDa glycoprotein, and for not substantially binding to 180 kDa CEA and/or ACT. In embodiments, antibodies are isolated that bind to a cancer cell extract obtained by a method comprising a) contacting colorectal and/or liver cancer cells with an acid to form an extract; b) separating components of the extract by molecular weight to form a first fraction having components with a molecular weight of about 60 kDa or greater; c) isolating components of the first fraction that have a molecular weight of about 75 kDa or greater and less than about 200 kDa, and d) isolating a 100 kDa fraction.

In some embodiments, the antibodies are further screened for the ability to bind to tissue samples of cancers as compared to normal tissue of the same tissue type. In embodiments, the antibodies are further screened for binding to colon or lung tumor cells and for lack of binding to a non diseased tissue such as colon, or lung tissue. In yet other embodiments, the antibodies are optionally screened for binding to a 100 kDa glycoprotein that has an amino acid sequence of SEQ ID NO:1.

It should be understood that the present invention encompasses the antibodies described herein and any fragments thereof containing the active binding region of the antibody, such as Fab, F(ab)\(_2\) and Fv fragments. Such fragments can be
produced from the antibodies using techniques well established in the art (see, e.g., Rousseaux et al, in Methods EnzymoL, 121:663-69, Academic Press (1986)).

In addition, some embodiments of the present disclosure encompass antibodies that are capable of binding to the same epitope as the antibodies and/or competing with the antibody for binding at that site. These include antibodies having the same epitope specificity as the antibodies that specifically bind to a 100 kDa glycoprotein that has a sequence of SEQ ID NO: 1 as described herein, and does not substantially bind to other linear epitopes of CEACAM5 as described herein. For example, class, isotype and other variants of the antibody of the invention may be constructed using recombinant class-switching and fusion techniques known in the art (see, e.g., Thammana et al, Eur. J. Immunol, 13:614 (1983); Spira et al, J. Immunol. Meth., 74:307-15 (1984); Neuberger et al, Nature, 312:604-08 (1984); and Oi et al., supra). Thus, chimeric antibodies or other recombinant antibodies (e.g., antibody fused to a second protein such as a lymphokine) having the same binding specificity as the antibodies fall within the scope of this invention.

Also included within the scope of the invention are anti-idiotypic antibodies of the antibodies that specifically bind to a 100 kDa glycoprotein that has a colorectal cancer membrane bound and a soluble form, has a UV absorbance peak at about 228nm, an isoelectric point of about 3.5 to 4, a sialic acid content of about 20%, and does not substantially bind to 180 kDa CEA and/or ACT. These anti-idiotypic antibodies can be produced using the antibodies as immunogen and are useful for diagnostic purposes in detecting humoral response to tumors and in therapeutic applications, e.g., in a vaccine, to induce an anti-tumor response in patients (See, e.g., Nepom et al, Cancer And Metastasis Reviews, 6:487-501 (1987); and Lee et al, Proc. Nafl. Acad. Sci. (USA), 82:6286-90 (1985)).

Human monoclonal antibodies may be made by using the 100 kDa antigen of the disclosure, to sensitize human lymphocytes to the antigen in vitro followed by EBV-transformation or hybridization of the antigen-sensitized lymphocytes with mouse or human lymphocytes as described by Borrebaeck et al. (Proc. Nat'l. Acad. Sci. (USA), 85:3995-99 (1988)). Therefore, human monoclonal antibodies or chimeric antibodies that bind the 100kDa antigen are also included within the scope of the present invention.
Methods
The disclosure provides methods of screening, diagnosis, and treatment monitoring of patients having a cancer such as colon cancer or lung cancer.

In embodiments, a method involves combining one or more antibodies that specifically bind to a 100 kDa glycoprotein that has a colorectal cancer membrane bound and a soluble form, has a UV absorbance peak at about 228nm, an isoelectric point of about 3.5 to 4, a sialic acid content of about 20%, and does not substantially bind to 180 kDa CEA and/or ACT in a human tissue or fluid sample; and detecting the binding of the antibody to the 100 kDa glycoprotein antigen in the sample. In embodiments, the presence or amount of the 100 kDa glycoprotein detected is indicative of the presence of or risk of cancer as compared to control sample without cancer. In embodiments, the 100 kDa glycoprotein is increased in the sample as compared to a control sample from a subject not having cancer. In embodiments, subjects having about 6.5 units/ml or greater of the 100 kDa glycoprotein are at increased risk for colorectal cancer or tubular adenomas.

In embodiments, the sample from the subject is a blood or serum sample and the immunoassay is able to detect the risk of or presence of stage 1 colon carcinoma and/or tubular adenomas. In embodiments, the assay detects these conditions with a sensitivity and/or a specificity of at least 70%>, 80%>, 90%>, 95%, and any number in between 70-100%.

In one aspect, an immunoassay comprises one or more of any of the antibodies as described herein. In embodiments, an immunoassay for determining whether a subject is at risk of colorectal cancer comprises: a) combining an antibody or antigen binding fragment thereof with a human tissue or fluid sample from the subject; wherein the antibody or antigen binding fragment specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, but does not substantially bind to glycosylated 180kDa CEA and does not substantially bind to one or more linear peptides consisting of a linear peptide of 15 amino acids from amino acids 1-60 of SEQ ID NO:2, a linear peptide of amino acids 111 to 125 of SEQ ID NO:3, and/or a linear peptide of 15 amino acids from amino acids 150-701 of SEQ ID NO:4; and b) determining the amount of the antigen in the human tissue or fluid by determining the amount of the antibody antigen complex. In embodiments, the antibody is a monoclonal antibody. In other embodiments, the antibody is a capture
antibody attached to a solid substrate. In embodiments, the antibody is detectably labelled. In embodiments, one or more capture antibodies can be utilized.

In other embodiments, an immunoassay further comprises a second antibody that specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, but does not substantially bind to glycosylated 180kDa CEA and does not substantially bind to one or more linear peptides consisting of a linear peptide of 15 amino acids from amino acids 1-60 of SEQ ID NO:2, a linear peptide of amino acids 110 to 125 of SEQ ID NO:3, and/or a linear peptide of 15 amino acids from amino acids 150-701 of SEQ ID NO:4; and b) determining the amount of the antigen in the human tissue or fluid by determining the amount of the antibody antigen complex. In embodiments, the second antibody is a monoclonal antibody. In other embodiments, the second antibody is a capture antibody attached to a solid substrate. In embodiments, the antibody is detectably labelled.

In embodiments, a pair of antibodies in an immunoassay bind to an epitope in the N terminal domain of CEACAM5 (SEQ ID NO:5) and do not bind to epitopes in the other domains of CEACAM5 as described herein, and are used as capture antibodies. In embodiments, an immunoassay further comprises an additional antibody that specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1. In embodiments the additional antibody is detectably labelled. In embodiments, the additional antibody is a polyclonal antibody.

In embodiments, detectable labels include, without limitation, a radionuclide, an enzyme, a fluorescent agent, colloidal gold, and a chromophore.

In embodiments, an assay is a point of care assay. In embodiments, the point of care assay is a lateral flow assay. The lateral flow assay can be a sandwich assay or a competitive assay. In embodiments, a first antibody is deposited on a portion of the surface of the solid surface, and may be referred to as a conjugate pad. The first antibody can be a monoclonal or polyclonal. The first antibody can be detectably labelled with colloidal gold, a fluorescent agent or a chromophore. The first antibody specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1 as described herein. In embodiments, the first antibody is a polyclonal antibody. In embodiments, one or more than one additional antibodies are deposited on the solid substrate as a capture antibody to form a test line or a control line. The additional antibodies include monoclonal and/or polyclonal antibodies that specifically bind to
a 100 kDa glycoprotein comprising SEQ ID NO:1 as described herein to form a test line. In embodiments, one or more capture monoclonal antibodies are utilized. In other embodiments, the antibody is control antibody, for example that binds to IgG to form the control line. In embodiments, the solid surface is a nitrocellulose membrane. In embodiments, the amount of material deposited on a test line can be read by a reader in order to provide a quantitative or semiquantitative amount of the test antigen in the sample.

In embodiments, a method of the disclosure involves a screening method for subjects. A subject may visit a health care facility, for example, as a part of routine checkup, and a sample is obtained from the subject. The sample is then screened for the presence of a 100 kDa glycoprotein in accord with methods as described herein. If an increase of the 100 kDa glycoprotein as compared to control is seen in the sample or if the presence or amount of the antigen exceeds a predetermined cutoff value, the health care worker directs the subject to further cancer screening.

In a specific example, a subject comes in for an annual physical and a blood sample and/or urine sample is taken. The blood sample is tested in accord with the kits and methods as described herein for the presence or amount of the 100 kDa glycoprotein. If an amount of the 100 kDa antigen detected is about 6.5 units /ml or greater as compared to control in the sample, the health care worker directs the subject to further cancer screening, such as a colonoscopy rather than waiting for the recommended time for conducting such routine screening. In addition, the health care worker may direct the subject to more specific diagnostic methods such as a biopsy, CT scan, MRI, PET scan, alone or in conjunction with other biomarker tests.

In embodiments, the immunoassay methods as described herein identifies with 95% sensitivity the Stage I (Earliest stage histologically confirmed) of Colorectal Cancer. In other embodiments, the immunoassay also identifies over 60% of the Tubular Adenoma polyps which if not removed will most likely become cancer. In embodiments, the antibodies in the immunoassay are selected or modified in order to provide an immunoassay for detecting tubular adenoma or stage I colorectal cancer with a sensitivity of at least 70% to 100% including every number in between and a specificity of at least 70%->100%>, including every number in between.
In embodiments, a method comprises monitoring treatment of a patient with cancer for efficacy by a) combining an antibody or antigen binding fragment thereof with a human tissue or fluid sample from the subject; wherein the antibody or antigen binding fragment specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO: 1, but does not substantially bind to glycosylated 180kDa CEA and does not substantially bind to one or more linear peptides consisting of a linear peptide of 15 amino acids from amino acids 1-60 of SEQ ID NO:2, a linear peptide of amino acids 111 to 125 of SEQ ID NO:3, and/or a linear peptide of 15 amino acids from amino acids 150-701 of SEQ ID NO:4; and b) determining the amount of the antigen in the human tissue or fluid by determining the amount of the antibody antigen complex. In embodiments, a decrease in the presence or amount of the antigen detected as compared to a value at the beginning of treatment is indicative of the efficacy of treatment of cancer. In embodiments, the cancer is selected from the group consisting of colon cancer and lung cancer. In embodiments, a treatment is identified as efficacious if it results in a decrease in the amount of the 100 kDa antigen of at least 20% or more. In embodiments, the value at the beginning of treatment is about 6.5 units/ml or greater.

In embodiments, an immunoassay kit is provided for use in the methods described herein. In some embodiments, an immunoassay kit comprises: a) a first and/or second antibody that specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, but does not substantially bind to glycosylated 180kDa CEA and does not substantially bind one or more linear peptides consisting of a linear peptide of 15 amino acids from amino acids 1-60 of SEQ ID NO:2, a linear peptide of amino acids 111 to 125 of SEQ ID NO:3, and/or a linear peptide of 15 amino acids from amino acids 150-701 of SEQ ID NO:4; b) an additional antibody that specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, labeled with a signal generating element; and c) a calibrator.

In embodiments, the antibody or antigen binding fragment binds to a conformational epitope. In some embodiments, the antibody or antigen binding fragment specifically binds to and/or competes for binding to an epitope comprising amino acids 61-69 (SEQ ID NO:6), 73-77 (SEQ ID NO:7), or 69-82 KGERVGDNRQIIGY (SEQ ID NO:8), and 96-107(SEQ ID NO:9) of SEQ ID NO:1. In other embodiments, the antibody or antigen binding fragment specifically
binds to an epitope comprising amino acids 61-69 (SEQ ID NO:6), 78-98 (SEQ ID NO:10), 96-107 (SEQ ID NO:9), and 127-139 (SEQ ID NO:11) of SEQ ID NO:1.

In embodiments, the antibody or antigen binding fragment does not substantially bind to epitopes in one or all other domains of CEACAM5 selected from the group consisting of amino acids 146-237 (Al Ig-2), amino acids 238-322 (BIIg-3), amino acids 324-415 (A2 Ig-4), amino acids 416-498 (B2 Ig-5), amino acids 502-593 (A3 Ig-6) amino acids 594-677 (B3 Ig-7). The amino acid numbers will be shifted one number from SEQ ID NO:1 because of the deletion of amino acid 320. In embodiments, the first and/or second antibody are monoclonal antibodies having an epitope specificity as described herein and the additional antibody is a polyclonal antibody that bind to a number of different epitopes. In embodiments, one or more monoclonal antibodies having different epitope specificity are used as capture antibodies.

In embodiments, the first and second antibody can comprise the CDRs and/or the heavy chain or light chain variable domain sequences as described herein.

In yet other embodiments, a method further comprises displaying the presence or amount of the 100 kDa glycoprotein in the sample on a detectable reader. In yet other embodiments, a method further comprises communicating the presence or amount of the 100 kDa glycoprotein in the sample and/or the presence or absence of cancer to a health care worker. In embodiments, the cancer is selected from the group consisting of as colon cancer and lung cancer. In embodiments, the sample is from a biopsy, a serum, or a blood sample. In embodiments, a detectably labeled antibody is administered in vivo to identify the cancer. In embodiments, the antibody is a human or humanized antibody comprising the heavy and/or light chain CDRS as described herein.

In embodiments, an immunoassay includes at least two capture antibodies and one detectably labelled detection antibody. In embodiments, the at least two capture antibodies are monoclonal antibodies. In embodiments, the monoclonal antibodies bind to a similar or different epitope. In embodiments at least one antibody binds to an epitope comprising, consisting essentially of, or consisting of amino acids 61-69 (SEQ ID NO:6), 73-77 (SEQ ID NO:7), or 69-82 KGERVGDNRQIIGY (SEQ ID NO:8), and 96-107 (SEQ ID NO:9) of SEQ ID NO:1. In other embodiments, the antibody or antigen binding fragment specifically
binds to an epitope comprising consisting essentially of, or consisting of amino acids 61-69 (SEQ ID NO:6), 78-98 (SEQ ID NO: 10), 96-107 (SEQ ID NO:9), and 127-139 (SEQ ID NO:1) of SEQ ID NO:1. In embodiments, the antibodies comprise the heavy chain CDRs (HCDR1, HCDR2 and HCDR3) from a heavy chain variable region having an amino acid sequence selected from the group consisting of SEQ ID NO: 119 and SEQ ID NO: 121, and light chain CDRs (LCDR1, LCDR2 and LCDR3) from a light chain variable region having an amino acid sequence selected from the group consisting of SEQ ID NO: 120 and SEQ ID NO: 122. In embodiments, the capture antibodies are antibody designated as 5A1-1 or 5E5-1. In embodiments, the capture antibodies are deposited on a solid surface in a test strip. In embodiments, the detectably labelled detection antibody is a polyclonal antibody. In embodiments, the detection antibody is deposited on a conjugate pad on a solid surface.

In embodiments, a detectable reader is able to convert the signal from the detectably labeled antibody into a numerical value. The detectable reader employed depends on the label employed in the immunoassay. For example, for fluorescent labels or antibody labeled with an enzyme that converts a substrate labeled with fluorescent label, a detectable reader is a spectrofluorometer. In other embodiments, the detectable reader detects a radioactive label. In embodiments the detectable reader comprises a display and/or a printer. The reader may include software that automatically takes into account background and uses a standard curve based on an appropriate calibrator to provide a numerical value. In embodiments, the numerical value is communicated to a health care worker via a network including a wireless network.

The methods and kits of the disclosure are useful as point of care assays. The methods and assays as described herein can be used in conjunction with other cancer screening methods such as colonoscopy, lung imaging, biopsy, PET scans, magnetic resonance imaging, ct scans, and assay for other biomarkers associated with cancer.

In embodiments, one or more antibodies are selected to provide an immunoassay that can detect the presence or absence of a 100 kDa glycoprotein as described herein in a tissue, urine, or serum sample from a patient having cancer with a specificity and/or sensitivity of at least 70%-100% including every number in
between. In embodiments, the first antibody and/or second antibody is a monoclonal antibody. In embodiments, the additional antibody is a polyclonal antibody.

In embodiments, at least two antibodies are selected in order to form an antibody pair that specifically binds to a 100 kDa glycoprotein in a human tissue or bodily fluid sample in an immunoassay. In embodiments, antibodies that bind to the 100 kDa glycoprotein may bind to the same epitope or different epitopes on the protein.

Immunohistological techniques involve contacting a biological specimen such as a tumor tissue specimen with the antibodies of the invention and then detecting the presence on the specimen of the antibodies complexed to their antigen. The formation of such antibody-antigen complexes with the specimen indicates the presence of tumor cells in the tissue. Detection of the antibodies on the specimen can be accomplished using techniques known in the art, such as the immunoperoxidase staining technique, the avidin-biotin (ABC) technique or immunofluorescence techniques (see, e.g., Ciocca et al, Meth. Enzymol., 121:562-79 (1986); Hellstrom et al, Cancer Research, 46:3917-23 (1986); and Kimball (ed.), Introduction To Immunology (2nd Ed.), pp. 113-117, Macmillan Publ. Co. (1986)).

Serologic diagnostic techniques involve the detection and quantitation of tumor-associated antigens that have been secreted or "shed" into the serum or other biological fluids of patients thought to be suffering from carcinoma. Such antigens can be detected in the body fluids using techniques known in the art such as radioimmunoassays (RIA) or enzyme-linked immunosorbent assays (ELISA) wherein an antibody reactive with the "shed" antigen is used to detect the presence of the antigen in a fluid sample (see, e.g., Uotila et al., J. Immunol. Methods, 42:111 (1981) and Allum et al., "Monoclonal Antibodies in the Diagnosis and Treatment of Malignant Conditions" Surg. Ann., 18:41-64, 48-51 (1986)). These assays, using the antibodies disclosed herein, therefore can be used for the detection in biological fluids of the antigens with which the antibodies react and thus provide for the detection of various carcinomas in human patients. Thus, it is apparent from the foregoing that the antibodies of the invention can be used in most assays involving antigen-antibody reactions. These assays include, but are not limited to, standard RIA techniques, both liquid and solid phase, as well as ELISA assays, lateral flow tests, immunofluorescence techniques, and other immunocytochemical assays (see,
e.g., Sikora et al. (eds.), Monoclonal Antibodies, pp. 32-52, Blackwell Scientific Publications, (1984)).

The antibodies of the disclosure are also useful for in vivo diagnostic applications for the detection of human tumors. One such approach involves the detection of tumors in vivo by tumor imaging techniques using the antibodies labeled with an appropriate imaging reagent that produces detectable signal. Imaging reagents and procedures for labeling antibodies with such reagents are well known (see, e.g., Wensel and Meares, Radio Immunooimaging and Radioimmunootherapy, Esevier, N.Y. (1983); Colcher et al, Meth. EnzymoL, 121:802-16 (1986)). The labeled antibodies may be detected by a technique such as radionuclear scanning (see, e.g., Bradwell et al. in Monoclonal Antibodies for Cancer Detection and Therapy, Baldwin et al. (eds.), pp. 65-85, Academic Press (1985)).

In some embodiments, the antibodies are useful as a therapeutic agent, especially if linked to a toxic molecule. Antibodies that have been screened for binding to cancer cell tissue as described herein are useful in a method for treating a cancer comprising administering an antibody that specifically binds to a 100 kDa glycoprotein that has a colorectal cancer membrane bound and a soluble form, has a UV absorbance peak at about 228nm, an isoelectric point of about 3.5 to 4, a sialic acid content of about 20% , has a sequence of SEQ ID NO: 1, and does not substantially bind to 180 kDa CEA and/or ACT in a subject suspected of having cancer. In embodiments, the antibodies are human or humanized antibodies comprising the heavy chain CDRs and/or light chain CDRS described herein. In embodiments, the cancer is selected from colon cancer and lung cancer. In embodiments, the cancer is colon cancer.

For diagnostic purposes, antibodies can be attached to a solid surface. Solid surfaces, include without limitation, membranes, multiwell plates, chromatography media, glass slides, latex beads, magnetic beads, microarray chips, capillary tubes and chips. Antibodies can also be labeled with a signal generating element. Signal generating elements include dyes, enzymes, radioactive labels, and chemiluminescent reagents.

In some embodiments, a cutoff value for classifying unknown samples as having the disease or not having the disease is identified. In embodiments, a cutoff value is determined by diluting each of plurality of samples from subjects known to
have cancer and detecting the amount of the 100 kDa glycoprotein as compared to a similar set of dilutions from samples subject known not to have cancer. In embodiments, a cutoff is selected that discriminates between bodily fluid samples such as serum samples that are indicative of the presence of cancer as compared to samples known not to have cancer. In embodiments, a cutoff value for the presence of colorectal cancer or tubular adenomas in a serum sample is at least about 6.5 units/ml.

**Kits**

The disclosure also provides kits for diagnosis of cancer, such as colon or lung cancer. The kit comprises at least one antibody that specifically binds to a 100 kDa glycoprotein that has a colorectal cancer membrane bound and a soluble form, has a UV absorbance peak at about 228nm, an isoelectric point of about 3.5 to 4, a sialic acid content of about 20% , and does not substantially bind to 180 kDa CEA and/or ACT. In embodiments, the antibody is labeled with a signal generating element.

In yet other embodiments, the kit comprises a) a first and/or second antibody that specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, but does not substantially bind to glycosylated 180kDa CEACAM5 and does not substantially bind to one or more linear peptides consisting of a linear peptide of amino acids from amino acids 1-60 of SEQ ID NO:2, a linear peptide of amino acids 111 to 125 of SEQ ID NO:3, and/or a linear peptide of 15 amino acids from amino acids 150-701 of SEQ ID NO:4; b) an additional antibody that specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, labeled with a signal generating element; and c) a calibrator. The antibodies may bind to the same epitope or different epitopes. One of the antibodies is optionally attached to a solid surface. Solid surfaces include membranes, beads, magnetic beads, microwell plates, slides, and microrarrays. The other antibody is optionally labeled with a detectable label or signal generating element. Detectable labels or signal generating elements include biotin, fluorescent dyes, chemiluminescent tags, colloidal gold, radioactive tags, and enzymes.

The kit may also optionally comprise a calibrator or standard as a positive or negative control. The standard may comprise a cancer cell extract obtained by a method comprising a) contacting colorectal and/or liver tumor cells with an acid to
form an extract; b) separating components of the extract by molecular weight to
form a first fraction having components of about 60 kDa or greater; c) isolating
components of the first fraction that have a molecular weight of about 75 kDa or
greater and less than about 200 kDa; and d) isolating a protein of 100 kDa molecular
weight. In embodiments, a calibrator is a 100 kDa glycoprotein having an amino
acid sequence of SEQ ID NO: 1. In embodiments, a calibrator has a concentration of
at least 6.5 units per ml or greater. In embodiments, the tumor cell is from colorectal
cancer. In other embodiments, the kit comprises a negative control such as bovine
serum albumin, ACT, and/or 180 kDa CEA.

In embodiments, a kit comprises instructions for conducting an immunoassay
to detect the presence of a 100 kDa glycoprotein. The instructions provide a
standard curve and instructions for diluting the positive and negative control. The
instructions provide a cutoff value of 6.5 units/ml. If the antigen is detected at a level
of 6.5 units/ml or greater the sample is identified as positive for an increased risk of
colorectal or lung cancer.

Also included with the disclosure are point of care kits and assay formats. In
such assays, the biological sample is blood, urine, saliva, and/or tears and testing is
conducted in a doctor's office, emergency room, hospital room, outpatient center,
and the like. In embodiments, an antibody as described herein is attached to a solid
substrate such as a multiwell plate, dipstick, or membrane based test strip. When the
antibody is contacted with the patient sample, if the antigen is detected in the patient
sample, the solid substance changes color in at least one location which can be read
visually. In an alternative embodiment, the antibody may be incorporated in a meter
that draws a blood sample and then if the antigen is detected gives a detectable
signal that is read by the meter.

In order that the invention described herein may be more fully understood,
the following examples are set forth. It should be understood that these examples are
for illustrative purposes only and are not to be construed as limiting the scope of this
invention in any manner.

EXAMPLE 1

A perchloric acid tumor extract prepared from liver and colon cancer tissues
was used to prepare polyclonal and monoclonal antibodies. The HAI test referred to
as the Tennagen assay was developed solely using a polyclonal antibody and was used to diagnose colon cancer. An ELISA test using two specific monoclonal antibodies as capture antibodies and the polyclonal antibody as a signal generating antibody was developed later. A cocktail of the two specific monoclonal antibodies in the ELISA Assay repeatedly have a sensitivity of 95+% to Early Stage Colorectal carcinoma and a specificity of 94+% when serum samples from patients not known to have this condition are analyzed.

Immunohistochemical studies have shown the two monoclonal antibodies attach to the cell membranes of adenocarcinoma of colorectal cancer but do not attach to normal tissue. No attachment to cells in bladder or breast tissue is seen.

**Materials and Methods**

**Immunization**

Lots of several Tennessee Antigen (CA1 1-19), used in the immunization of goats for the production of polyclonal antibodies, were used to immunize the Balb/C mice for the production of the mouse hybridomas. The Tennessee Antigen(CA1 1-19) lots were perchloric acid extracts from adenocarcinoma tissue with purification steps including Sepharose 4B and Sephadex G200 chromatography as used for all antigen approved for immunization of goats and mice after the initial characterization of the antigen.

A large number of monoclonal antibodies were cloned from the hybridomas generated from three of the immunized mice. The hybridomas and monoclonals were initially screened using the approved Tennessee Antigen coated RBCs, a component of the TennaGen™ Assay. The hybridoma and monoclonal supernates were mixed with the Tennessee Antigen coated RBCs and hemagglutination titer was determined. Hybridoma and monoclonal cell culture supernates containing antibodies to Tennessee Antigen (CA1 1-19) resulted in the smooth mat agglutination. The HA assay was used as the screening test for screening large numbers of hybridomas and later monoclonal antibody production in cell culture supernates.

Using the Limiting Dilution technique, 23 monoclonal antibodies were produced from the best hybridomas. With extensive evaluation of the 23 clones, 7 clones were selected for further evaluation in the ELISA assay. Immunohistochemical studies confirmed the monoclonals would bind to
adenocarcinoma of colon and lung. Two monoclonals also had positive staining with squamous cell and adenocarcinoma of the Lung. The negative controls were bladder and breast tissue. The final selection for the capture antibodies was a monoclonal cocktail utilizing 5E5-1 and 5A1-1 following the second blind serum panel from the National Cancer Institute. Summary of NCI evaluation using a Double Blinded Serum Panel supplied by the NCI through Mayo Clinic for the CA1 1-19 Mono-Poly ELISA Sandwich Assay is shown.

**ELISA assay**

The ELISA uses a high binding microtiter plate coated with a cocktail of two mouse monoclonal antibodies, 5E5-1 & 5A1-1, the anti-Tennessee Antigen polyclonal antibody conjugated to alkaline phosphatase for use as the detection antibody and Tennessee Antigen lots approved for use as calibrators are calibrators in the CA1 1-19 ELISA Kit. The ELISA uses neat serum and not perchloric acid extracts.

The protocol for an ELISA is as follows:

**Coating:** each well of a high binding 96 well Easy Wash EIA/RIA plate was coated two hours at room temperature (~21°C) followed by overnight at refrigeration (4-8°C) with 125µL of optimized by titration the solution of monoclonal cocktail (Capture Antibodies) in ELISA Coating Buffer. This equates to approximately 1µg of antibody-per-well.

**Blocking:** after 3 x 300µL washes with PBS T 20 Wash Buffer, the plates were blocked with 300µL-per-well of 1.2% SEA Block in ELISA Coating Buffer for 2 hrs. At room temperature. Microtiter plate is further blocked with 2% Sucrose, and 4% Polyvinylpyrrolidone Buffer for 45 minutes.

**Storage:** Plates were dried by inverting on plastic drying racks for 24 hours at room temperature, then placed in foil pouches along with a dessicant packet and humidity indicator, and then stored at 40°C.

**Sample:** Plates, Calibrators and Sample Buffer were moved to room temperature and allowed to equilibrate. 25µL of calibrators, and controls were added to the 2 microtiter wells of plates along with 25 µl of unknown samples followed by 75µL of Sample Diluting Buffer. Plates were sealed, placed in ziplock bag and incubated at 37 degree C overnight for 2 hrs.
CA1 1-19 plates were washed 6 χ 300 µL with distilled water and then 100 µL of CA1 1-19 Polyclonal alkaline phosphatase conjugate was added to each well. Plates were sealed, placed in ziplock bag and left at 37 degree C for 2 hrs.

CA1 1-19 plates were washed 6x 300 µL with Distilled water and 100 µL CA1 1-19 Substrate was added to each well. Plates were placed at 37 degree C for 30 min.

Measurement: The absorbance of each well at 405 nm was measured immediately using a standard plate reader.

Quantification: The absorbance of the standards was plotted using a point to point curve and the absorbance of the controls and unknowns are read against the curve.

The calibrators for the ELISA test are manufactured following SOP/Procedure using Tennessee Antigen (CA1 1-19) lots approved for use as a calibrator. The unitage is based on the original unitage of a vault reference standard and manufacturing allows for the dilution of antigen lot to obtain the standard required unitage. The Scientific advisory board at the FDA meeting made the statement that Tumor Markers were not pure enough to have a purified uniform standard. The World Health Organization made a similar statement concluding that the Tumor Marker Tests should be reported at "units/ml" and not in a weight such as ng or pc until the protein was completely 100% pure for a reference standard. Many tumor markers such as AFP, B-HCG, CA15-3, CA19-9, CA125, CA27-29, and PAP are reported as U/ml while CEA and PSA are reported in ng/ml.

Immunohistochemistry

The various monoclonal (cloned hybridomas) antibodies and goat antibody created for the TennaGen Assay were further characterized using Immunohistochemistry studies at two different independent facilities, Summa Medical, TX and U of Hawaii, HI. The monoclonal 5E5-1 and 5A1-1 stained adenocarcinoma of the colon and lung as well as squamous cell carcinoma of the lung just as the goat antisera; however, the signal from the 5E5-1 antibody was stronger.

Results

This test configuration was used to measure the CA1 1-19 levels in 422 patients. 222 samples were collected from GI Associates, TN) and 200 samples
were purchased commercially from Equitech Bio, TX. The serum from these patients was assayed in the CA1 1-19 ELISA Assay. Results are summarized below in Figure 1 and Tables 1, 2 and 3. With the use of the monoclonal cocktail of 5E5-1 and 5A1-1 along with improvements in other ELISA reagents, the sensitivity has been greater than 90% with specificity between 84 to 92% depending on the clinical status of the patient.

### Table 1: All cancerous patients histologically confirmed categorized by cancer stage using a 6.5 units/mL cut-off to determine true positive, true negative, false positive, and false negative rates. Serum levels

<table>
<thead>
<tr>
<th>Colon Cancer/Histological Confirmed</th>
<th>Number Samples</th>
<th>True Positive</th>
<th>True Negative</th>
<th>False Positive</th>
<th>False Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>21</td>
<td>21 (100%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stage II</td>
<td>23</td>
<td>23 (100%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stage III</td>
<td>17</td>
<td>17 (100%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Stage IV</td>
<td>11</td>
<td>7 (64%)</td>
<td>0</td>
<td>0</td>
<td>4 (36%)</td>
</tr>
<tr>
<td>Unknown Stage (pathology report not-provided)</td>
<td>9</td>
<td>9 (100%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table 2: Polyps considered to be precancerous - Tubular Adenoma, Adenoma Malignancy and Villous Adenoma. Serum levels

<table>
<thead>
<tr>
<th>DIAGNOSIS POLYPS</th>
<th>Number Samples</th>
<th>True Positive</th>
<th>True Negative</th>
<th>False Positive</th>
<th>False Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tubular Adenoma</td>
<td>53</td>
<td>17 (32%)</td>
<td>0</td>
<td>0</td>
<td>36 (68%)</td>
</tr>
<tr>
<td>Adenoma Malignancy</td>
<td>1</td>
<td>1 (100%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Villous Adenoma</td>
<td>1</td>
<td>1 (100%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hyperplastic</td>
<td>26</td>
<td>0</td>
<td>21 (81%)</td>
<td>5 (19%)</td>
<td>0</td>
</tr>
<tr>
<td>Polyp-not classified</td>
<td>13</td>
<td>0</td>
<td>12 (92%)</td>
<td>1 (8%)</td>
<td>0</td>
</tr>
<tr>
<td>No Pathology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3: All non-cancerous patients categorized by condition using a 6.5 units/mL cut-off to determine true positive, true negative, false positive, and false negative rates. Serum levels.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Number of Samples</th>
<th>True Positive</th>
<th>True Negative</th>
<th>False Positive</th>
<th>False Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>GI Bleeding</td>
<td>36</td>
<td>0</td>
<td>27 (75%)</td>
<td>9 (25%)</td>
<td>0</td>
</tr>
<tr>
<td>Crohn's Disease</td>
<td>10</td>
<td>0</td>
<td>8 (80%)</td>
<td>2 (20%)</td>
<td>0</td>
</tr>
<tr>
<td>Diverticulitis</td>
<td>70</td>
<td>0</td>
<td>62 (89%)</td>
<td>8 (11%)</td>
<td>0</td>
</tr>
<tr>
<td>Ulcerative Colitis</td>
<td>5</td>
<td>0</td>
<td>5 (100%)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hemorrhoids</td>
<td>24</td>
<td>0</td>
<td>22 (92%)</td>
<td>2 (8%)</td>
<td>0</td>
</tr>
<tr>
<td>Change in Bowel Habits</td>
<td>20</td>
<td>0</td>
<td>13 (65%)</td>
<td>7 (35%)</td>
<td>0</td>
</tr>
<tr>
<td>Normal GI</td>
<td>73</td>
<td>0</td>
<td>66 (90%)</td>
<td>7 (10%)</td>
<td>0</td>
</tr>
<tr>
<td>Misc GI Disease</td>
<td>6</td>
<td>0</td>
<td>5 (83%)</td>
<td>1 (17%)</td>
<td>0</td>
</tr>
<tr>
<td>Undiagnosed</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Fig.1 shows a summary of all data from the 422 patient clinical study. Panel A shows the combined data for all non-cancerous and cancerous patients. The box covers the middle 50% of the data while the line is the median. The whiskers cover the middle 90% of the data (from 5% to 95%). Panel B is a receiver operator characteristic (ROC) curve comparing the cancerous vs. non-cancerous samples.

The box & whiskers plot (Fig. 1A) shows that the CA1-19 levels in non-cancerous tissues are lower overall than in cancerous patients. The box and whiskers plot comparing the levels of antigen detected in samples from patients known to have cancer as compared to control samples, establishes a cutoff value for 6.5 units/ml as indicative of the presence of colorectal cancer based on the lower limit of the whisker for the middle 90% of the data. When this data is plotted as a ROC curve (Fig. IB), there is an apparent shift to the upper left quadrant consistent with a positive correlation between CA1-19 and cancer. The sensitivity and
specificity of this test was 95% and 84% respectively when a cut-off of 6.5 units/mL is used. When the non-cancerous patients (Table 3) and cancerous patients (Table 1) are separated out by condition and the 6.5 units/mL cut-off is applied, the conditions giving the highest true positive, false positive, true negative, and false negative rate are seen. The highest false negative rate was seen in tubular adenomas (Table 2). Colon cancer from stages I-III gave 100% detection while stage IV was 64%.

Harvested monoclonal antibody supernates from 5E5-1 and 5A1-1 were purified and used in a cocktail as the capture antibody on high binding microtiter plates for further clinical evaluations. Two different Blind Serum Panels were assayed using two different lots of coated microtiter plates. The new lot used the 2012 cocktail as the capture antibody while the second lot used microtiter plates manufactured in 2009-2010 previously used in the evaluation conducted in 2009-2010. The serum panels were assembled at EDP Biotech and delivered to Dr. B.F. Overholt’s office at Gastrointestinal Associates (GIA) where they were blind coded for assay. The CA1 1-19 (shown as CAI-18) assays were performed and results were taken to GIA where the codes were identified. The results are shown in Table 4 and 5.

Table 4 (Serum Panel A)

<table>
<thead>
<tr>
<th></th>
<th>Disease Present</th>
<th>Disease Absent</th>
<th>Total Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA1-18 Elevated</td>
<td>27</td>
<td>6</td>
<td>33</td>
</tr>
<tr>
<td>CA1-18 Normal</td>
<td>5</td>
<td>33</td>
<td>38</td>
</tr>
<tr>
<td>TOTAL</td>
<td>32</td>
<td>39</td>
<td>71</td>
</tr>
</tbody>
</table>

Table 5 (Serum Panel B)

<table>
<thead>
<tr>
<th></th>
<th>Disease Present</th>
<th>Disease Absent</th>
<th>Total Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA1-18 Elevated</td>
<td>21</td>
<td>11</td>
<td>32</td>
</tr>
<tr>
<td>CA1-18 Normal</td>
<td>1</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>TOTAL</td>
<td>22</td>
<td>40</td>
<td>62</td>
</tr>
</tbody>
</table>

Table 4 above compares Colorectal cancer and Normals confirmed by colonoscopy. Serum Panel A (July 23, 2012) shows a Sensitivity of 84.4% and Specificity of 78.9%. Polyps and Lung Cancer not included in calculation. The CA1 1-19 Lot 2012 showed a sensitivity of 95.5% with specificity of 80.0 percent seen in Serum Panel B using the colorectal and benign GI data. Polyps and Lung
Cancer not included in calculation. When data from Serum Panel A and B are combined, a sensitivity of 92.3% and specificity of 85% is seen in Stage I Colorectal cancer as compared to Benign GI Disease.

**Discussion**

These results show that an ELISA assay with two capture monoclonal antibodies and a detectably labeled polyclonal antibody prepared against the Tennessee antigen repeatedly have a sensitivity of 95+% to Early Stage Colorectal carcinoma and a specificity of 94+ % as measured in serum.

Immunohistochemical studies have shown the two monoclonal antibodies attach to the cell membrane of adenocarcinoma of colorectal cancer but do not attach to normal tissue. No attachment to cells in bladder or breast tissue is seen.

**EXAMPLE 2**

The identity of the antigen that is being detected by the antibody assays described above has been difficult to determine. The perchloric acid carcinoma cell extract CAI 1-19 contains a number of different proteins. Further purification of the perchloric acid extract has identified a fraction of about 60-65 kDa and a fraction of about 100 kDa. However, the 60-65 kDa fraction, when used to immunize animals, has not provided any antibodies that specifically detect early stage colorectal carcinoma.

**Isolation and Characterization of antigens present in the CAI 1-19 preparation.**

Human adenocarcinoma tumor tissue from colon and liver were extracted. The purification procedure involved a perchloric acid extraction with further purification steps including Sepharose 4B® and Sephadex G200® chromatography.

Following SOP guidelines, SDS gel, Western Blot, Immunoelectrophoresis and antigen content/protein determinations were used for all antigen lots. Over 100 different tumor tissues were extracted into individual Tennessee Antigen (CAI 1-19) extract lots.

Amino Acid analysis, UV absorbance at 228 and Molecular Weight determinations have been repeatedly made of different antigen lots and compared to CEA. The results are shown in Table 6.
Similar characterizations were done comparing the 100 kDa antigen to the characteristics of alpha 1 anti-chymotrypsin (ACT). The results are shown in Table 7.

**COMPARISON OF TENNESSEE ANTIGEN, ACT, and CEA**

<table>
<thead>
<tr>
<th>TEST</th>
<th>TENNESSEE ANTIGEN(CA11-19)</th>
<th>180 kDa CEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecular Weight</td>
<td>100,000 Daltons</td>
<td>200,000 Daltons</td>
</tr>
<tr>
<td>UV Absorbance</td>
<td>Peak @ 228 nm</td>
<td>Peak @ 280nm</td>
</tr>
<tr>
<td>AMINO ACID STUDIES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lysine content</td>
<td>GREATER (7.6 moles %)</td>
<td>Less (2.9 moles %)</td>
</tr>
<tr>
<td>Glutamic acid, Glycine, Alanine</td>
<td>More than CEA</td>
<td>Less Than TNAG</td>
</tr>
<tr>
<td>Aspartic acid, Serine, Valine,</td>
<td>Less than CEA</td>
<td>Less than TN AG</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>Less than CEA</td>
<td></td>
</tr>
<tr>
<td>Leucine</td>
<td>More than TN AG</td>
<td></td>
</tr>
<tr>
<td>CARBOHYDRATE ANALYSIS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sialic Acid</td>
<td>More than CEA’s</td>
<td>Less than TN AG</td>
</tr>
<tr>
<td></td>
<td>19.88 moles % (average)</td>
<td>3.6 moles % (CEA WHO)</td>
</tr>
<tr>
<td>SDS Polyacrylamide</td>
<td>Band at 100,000 Daltons</td>
<td>No single band at 100,000 Daltons</td>
</tr>
<tr>
<td>Isoelectric Focusing</td>
<td>3.5 – 4.0 range</td>
<td>2.4,3,0,4.5-6.4 crude CEA prep</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.0 – 3.0 in purified prep</td>
</tr>
<tr>
<td>Immunoelectrophoresis/</td>
<td>No cross reactivity with</td>
<td></td>
</tr>
<tr>
<td>Hemagglutination-inhibition</td>
<td>Gold’s original CEA,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Go’s CEA, WHO-CEA,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Egan CEA, or Roche CEA</td>
<td></td>
</tr>
</tbody>
</table>

A large number of perchloric acid extracts of tumor and normal tissue were removed from -20°C storage and sterile filtered. All the Tennessee Antigen and Normal tissue extracts were perchloric acid extract with no further purification. The
Tennessee Antigen perchloric acid extract was more stable than the Tennessee Antigen Lots that had been further purified for use as immunogen and characterization work. These samples were screened for the presence of CA1 1-19 by SDS-PAGE separation (Fig. 2A) and then goat anti-CA1 1-19 polyclonal Conjugate Western in two different batches (Fig. 2B).

A subset of samples showed clean single band reactivity in the 55-76kDa region while others showed a generic reactivity where the entire protein pool showed a streak in the Western. The only immunizing antigen (#19) available showed good Western reactivity as did another sample (#28) which was deemed reactive enough to be used for amino acid analysis. Selected samples showing a clean single band reactivity were rerun on SDS-PAGE with the goal of excising the Western reactive region for protein identification. 11 gel slices were removed and given to Protein Discovery. They digested the samples, sequenced the released peptides via mass spectrometry, and then probed available databases to determine the protein make-up of the gel slice. Interestingly, ACT was detected 3x more often than any other protein with albumin coming in second.

**Discussion**

The isolated fraction from colorectal and/or liver tumors has components with a molecular weight of 55 to 75 kDa and a component with a 100 kDa. (Fig. 2A) The western blot shows reactivity with a 60 kDa and larger molecular weight proteins. (Fig. 2B). The identity of these bands was studied.

Very early data from two independent institutions showed that the Tennessee Antigen used as the immunogen for antibody production had a molecular weight close to 100 kDa. Two points important to the characterization studies: 1) the immunologically active portion of Tennessee Antigen has 100,000 molecular weight. Only the purified 100,000 MW material used as the immunogen has produced antibodies (monoclonal and polyclonal) that can measure a difference between Early Stage adenocarcinoma of colorectal and lung cancer serum with greater than 90% accuracy. The 60kDa fraction used as the immunogen in rabbit, goat and mouse was not successful in producing a usable antibody. Further use of the 60kDa material as the immunogen for antibody production in chicken and mouse also did not produce a usable antibody for the detection of colorectal and lung cancer when used as the capture antibody in an ELISA. The second significant point
in the Tennessee Antigen Characterization is the documentation that Tennessee
Antigen has a major UV peak at 228 and a small minor peak at 280 which is very
different from Alpha 1 antichymotrypsin (ACT)-UV peak 395, and Alpha 1
antitrypsin-UV peak 296. Previous physical characterizations distinguish this band
of material from 180 kDa CEA as well.

A total of 23 different monoclonal antibodies to Tennessee Antigen were
produced, initially studied to be potentially useful and are stored in cryotanks for
further evaluation for their diagnostic and therapeutic utility. Further
immunohistochemical studies would identify the potential value of the additional 21
monoclonals to Tennessee Antigen. Experiments to further characterize the 100 kDa
antigen involve amino acid sequencing of purified material and epitope mapping of
the monoclonal antibodies.

EXAMPLE 3

Separation of the 100 kDa material and the 60-65 kDa material

Separation of high molecular weight material from low molecular weight material in
adenocarcinoma tumor cell extracts in preparation for amino acid sequencing of
large molecular weight fraction

Materials and methods

Tumor cell material

Two different tumor cell specimens were obtained from adenocarcinoma
from colorectal tissue. One tumor cell specimen is a stage II adenocarcinoma
identified as 167B-80. A second tumor cell specimen is also an adenocarcinoma
stage II identified as 173-81. Control colorectal and tubular adenoma serum
specimens are identified as HF numbers.

Tumor cell extraction

Tumor cell extracts from each tumor cell specimen were prepared by
mincing into very small pieces and extracting the material with 1.4 % perchloric
acid.

Column chromatography

Aqueous Perchloric acid extracts of each tumor cell extract were separated
by column chromatography on Sepharose 4B. Fractions were collected and each
fraction was tested for absorbance at 280 and in the ELISA assay. Fractions
containing peak absorbance at 280 and ELISA activity were pooled and were further separated using Superdex 200 chromatography. Fractions were collected from Superdex 200 column and tested for absorbance at 280 and in the ELISA. Fractions were run on SDS polyacrylamide gel.

5

ELISA

The ELISA uses a high binding microtiter plate coated with a cocktail of two mouse monoclonal antibodies, 5E5-1 and 5A1-1, the anti-Tennessee Antigen polyclonal antibody conjugated to alkaline phosphatase for use as the detection antibody. Large molecular weight and low molecular weight fractions from the column chromatography were tested in the ELISA assay for reactivity with the antibodies that can detect serum samples from patients having cancer.

The protocol for an ELISA is as follows:

Coating: each well of a high binding 96 well Easy Wash EIA/RIA plate was coated two hours at room temperature (~2°C) followed by overnight at refrigeration (4-8°C) with 125µL of optimized by titration the solution of monoclonal cocktail (Capture Antibodies) in ELISA Coating Buffer. This equates to approximately 1µg of antibody-per-well.

Blocking: after 3 x 300µL washes with PBS T 20 Wash Buffer, the plates were blocked with 300pL-per-well of 1.2% SEA Block in ELISA Coating Buffer for 2 hrs. at room temperature. Microtiter plate is further blocked with 2% Sucrose, and 4% Polyvinylpyrrolidone Buffer for 45 minutes.

Storage: Plates were dried by inverting on plastic drying racks for 24 hours at room temperature, then placed in foil pouches along with a dessicant packet and humidity indicator, and then stored at 40°C.

Sample: Plates, Calibrators and Sample Buffer were moved to room temperature and allowed to equilibrate. 25µL of calibrators, and controls were added to the 2 microtiter wells of plates along with 25 µL of unknown samples followed by 75µL of Sample Diluting Buffer. Plates were sealed, placed in ziplock bag and incubated at 37 degree C overnight for 2hrs.

CA1 l-19 plates were washed 6x 300µL with distilled water and then 100uL of CA1 l-19 Polyclonal alkaline phosphatase conjugate was added to each well, Plates were sealed, placed in ziplock bag and left at 37 degree C for 2hrs.
CAI 1-19 plates were washed 6x 300µL with Distilled water and 100µL
CAI 1-19 Substrate was added to each well. Plates were placed at 37 degree C for
30min.

Measurement: The absorbance of each well at 405 nm was measured
immediately using a standard plate reader.

Quantification: The absorbance of the standards was plotted using a point to
point curve and the absorbance of the controls and unknowns are read against the
curve.

The calibrators for the ELISA test were manufactured following
SOP/Procedure using Tennessee Antigen(CAl 1-19) lots approved for use as a
calibrator. The unitage was based on the original unitage of a vault reference
standard and manufacturing allows for the dilution of antigen lot to obtain the
standard required unitage.

The polyclonal antibody does not substantially bind to pools of human serum
180 kDa CEA preparations or CEACAM 6 (NCA).

Results

The results in Figure 3 show that for each tumor cell extract bands of protein
as determined by absorbance at 280 are found in fractions F9 to F13. The peak of the
ELISA activity of the fractions indicate antigenic activity is predominantly found in
fractions F10 to F12. Fractions F9 to F12 were pooled. The pooled fractions from
each of the tumor specimens were then run on Superdex 200 column, (data not
shown) A broad band of protein is found in fraction F7 to F13. Each fraction was
tested in the ELISA and fractions showing peak activity were pooled and run on an
SDS gel. A band was found at around 100kDa (data not shown).

The band at 100 kDa will be analyzed further. The band will be cut out and
analyzed for amino acid sequence.

Discussion

Previous studies have identified an antigen derived from tumor cell extracts
detectable in a CAI 1-19 ELISA assay that is predictive of the presence of colorectal
cancer in blood samples from patients. Characterization of the tumor cell extract
revealed inactive protein fractions at 50-60 kDa and an active fraction at 100kDa. In
order to further characterize the 100 kDa fraction, separation and removal of the
material having a molecular weight of 50-60 kDa was conducted. Separation of the
lower molecular weight material from the higher molecular weight material was accomplished using column chromatography. Fractions that contained predominantly the 100 kDa material and that were active in the ELISA assay are being analyzed for amino acid sequence. This analysis should provide the identity of the antigen.

EXAMPLE 4

N-glycan Profiling of CAll-19 antigen

The sample preparation methods for N-glycan profiling by MALDI-MS are summarized in Figure 4.

SDS-PAGE

The 100 kDa CA1 1-19 material is referred to as 168B antigen and was prepared as described in Example 3. The sample was denatured by adding equal amounts of 2X SDS sample buffer and incubated at 100 °C for 5 min. The mixture then divided into small portions and loaded on each lane. The samples were separated in 7.5% SDS-PAGE gel. After electrophoresis, the resolved proteins were stained with Coomassie Brilliant Blue (CBB). (Figure 5)

In-gel PNGase F digestion and extraction of released N-glycans

In gel PNGase F digestion was performed according to the method of (Kuster et al. Sequencing of N-linked oligosaccharides directly from protein gels: in-gel deglycosylation followed by matrix-assisted laser desorption/ionization mass spectrometry and normal-phase high performance liquid chromatography. Anal. Biochem., (1997) 250, 52–101). Briefly, target glycoprotein bands were excised from CBB-stained gel sheet and the gel bands were further cut into smaller pieces. The gel pieces were then destained alternately with 50 mM ammonium bicarbonate (AmBic) and 100% acetonitrile until the color turned clear. Then proteins in gel were reduced with dithiothreitol (DTT), followed by carboxymidomethylation with iodoacetamide (IAM). DTT and IAM in gel were washed out alternately with 50 mM AmBic and 100% acetonitrile. Dehydrated gel pieces were reswelled with PNGase F solution (PNGase F in 50 mM AmBic) on ice for 45 min initially, and then the digestion was carried out at 37 °C overnight. Released N-glycans were extracted from the gels by stepwise extraction increasing the proportion of acetonitrile in 5% formic acid. The extracts were dried.
N-Glycan Preparation

Released N-glycans were purified from peptides by passage through a Sep-Pak C18 cartridge. The glycan fraction was eluted with 5% acetic acid and dried by lyophilization. The fractions were dried by lyophilization then permethylated based on the method of Anumula and Taylor, (Anumula et al., A comprehensive procedure for preparation of partially methylated alditol acetates from glycoprotein carbohydrates. Anal Biochem, 1992. 203(1): p. 101-108.), prior to purification again with a C18 cartridge. The permethylated N-glycans were then analyzed with matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS).

MALDI-TOF Analysis

MALDI/TOF-MS of the permethylated N-linked glycans was performed in the reflector positive ion mode using a-dihydroxybenzoic acid (DHBA, 20 mg/mL solution in 50% methanol) as a matrix. The spectrum was obtained by using AB Sciex TOF/TOF 5800 MADLI mass spectrometry system.

Results:
Separation of CA1-19 168B by SDS-PAGE

CA1 1-19 168B was separated by SDS-PAGE for in-gel digestion. After electrophoresis, the resolved proteins were stained with CBB and images of gels are shown in Figure 5. Figure 5A shows localization of target glycoprotein 100 KD band and an additional band at 45 KD observed in the CA1 1-19 168B. Figure 5B shows pictures of actual gels used for your sample analysis. Since the 100 KD bands were not clearly visible to observe in Figure 5B, we cut the region that corresponds to the molecular weight of 100 KD and were excised for N-glycan profiling. We also performed the N-glycan analysis on the 45 KD band with the same procedures as a reference.

N-glycan profiling by MALDI/TOF-MS

The permethylated N-glycans from samples provided were profiled by MALDI/TOF MS analysis. The MALDI-full mass spectra of each sample bands are shown in Figure 6A-B.

In-gel digestion and N-glycan analysis from 100 KD bands reveals peaks that may correspond to several glycan structures (m/z 1416.6, m/z 1661.7, and m/z 2039.9). However, the intensity of peaks is almost same as the noise level because
the trace amount of sample concentration (Figure 6A). Although we can clearly
detect glycan associated with the sample and they are biantennary structures with
and without galactose, it is possible that we are not detecting all the glycans present
due to lack of sufficient sensitivity from small amount of sample obtained for 100kD
band.

The primary structures that were detected in 45 kDa bands using
MALDI/TOF-MS were complex type N-glycans such as GlcNAc1Man3GlcNAc2
(m/z 1416.6) and GlcNAc2Man3GlcNAc2 (m/z 1661.7). The complex type N-
glycans with fucose were also observed in 45 kDa band (Figure 6B). It should be
noted that the relative peak intensities for these glycans were much higher than that
of the 100 KD band.

A summary of all results following N-glycan profiling is shown in Table 8.
Table 8 Summary of N-glycans from the CA 11-19 168B detected by MALDI-MS. Gal(o), Galactose; Man(e), Mannose; Fuc(T), Fucose; GlcNAc (■), N-acetylglucosamine;

<table>
<thead>
<tr>
<th>Suggested structure</th>
<th>Observed m/z</th>
<th>45 KDa</th>
<th>100 KDa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Man3GlcNAc2</td>
<td>1171.5</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>GlcNAc1Man3GlcNAc2</td>
<td>1416.6</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>GlcNAc1Man3GlcNAc2Fuc1</td>
<td>1590.7</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>GlcNAc2Man3GlcNAc2</td>
<td>1661.7</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Gal1GlcNAc1Man3GlcNAc2Fuc1</td>
<td>1794.7</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Gal1GlcNAc2Man3GlcNAc2</td>
<td>1865.8</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>GlcNAc2Man3GlcNAc2</td>
<td>1906.8</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Gal1GlcNAc2Man3GlcNAc2Fuc1</td>
<td>2039.8</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Gal2GlcNAc2Man3GlcNAc2</td>
<td>2069.8</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Gal1GlcNAc3Man3GlcNAc2</td>
<td>2110.9</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Gal1GlcNAc3Man3GlcNAc2Fuc1</td>
<td>2284.9</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Gal2GlcNAc3Man3GlcNAc2Fuc1</td>
<td>2489.0</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Discussion

N-glycan analysis from 100 KD bands indicates that peaks correspond to several glycan structures as shown in Table 8. In the SDS-PAGE (Figure 5), we could definitely see the difference between the concentration of 45 KDa bands and 100 KDa. Consequently, we were able to observe much stronger signals from the 45 KD band as compared to the 100 KD band.
EXAMPLE 5

O glycan profiling of CA19-2

The sample preparation methods for O-glycan profiling by MALDI-MS are summarized in Figure 7.

SDS-PAGE

Sample material CA19-2BSTK and CA19-2 168BNC were prepared as described in Example 3. Sample STK differed from NC by being concentrated 5 fold. The samples were denatured by adding equal amounts of 2 X SDS sample buffer and incubated at 100 °C for 5 min. Then the mixture was divided into small portions and loaded on each lane. The samples were separated in 7.5% SDS-PAGE gel. After electrophoresis, the resolved proteins were stained with Coomassie Brilliant Blue (CBB).

In-gel O-glycan Release by Reductive β-Elimination

In gel reductive β-elimination was performed according to the method of Dell, A et al /Dell, A. et al. Mass spectrometry of Carbohydrate-Containing Biopolymers. Methods Enzymol, (1994) 230, 105—132.) Briefly, target glycoprotein bands were excised from CBB-stained gel sheet and the gel bands were further cut into smaller pieces. The gel pieces were then destained alternately with 50 mM ammonium bicarbonate (AmBic) and 100% acetonitrile until the color turned clear. And then, gel pieces were washed with ethyl acetate and were washed with acetonitrile. O-glycans were released from gels by β-elimination. The gels were rehydrated by adding 100 mM sodium hydroxide and 1M sodium borohydrate. The samples were incubated for 18 h at 45 °C. The sample was placed on ice and neutralized with 10% acetic acid. The neutralized samples were loaded onto AG-50WX8 cation exchange column to desalt. Ten percent acetic acid in methanol was added to samples to remove the borate.

O-Glycan Preparation

Released O-glycans were purified from peptides by passage through a Sep-Pak C18 cartridge. The glycan fraction was eluted with 5% acetic acid and dried by lyophilization. The fractions were dried by lyophilization then permethylated based on the method of Anumula and Taylor (cited supra), prior to purification again with a C18 cartridge. The permethylated O-glycan were then analyzed with matrix-
assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS).

**MALDI-TOF Analysis**

MALDI/TOF-MS of the permethylated N-linked glycans was performed in the reflector positive ion mode using a-dihydroxybenzoic acid (DHBA, 20 mg/mL solution in 50% methanol) as a matrix. The spectrum was obtained by using AB Sciex TOF/TOF 5800 MALDI mass spectrometry system.

**Nanospray Ionization FTMS**

NSI-MSMS analysis was determined using a LTQ Orbitrap XL mass spectrometer (ThermoFisher) equipped with a nanospray ion source. Permethylated O-linked glycans were dissolved in ImM NaOH in 50% methanol then infused directly into the instrument at a constant flow rate of 0.5 µL/min. A full FTMS spectrum was collected at 30 000 resolution with 3 microscans. The capillary temperature was set at 210°C and MS analysis was performed in the positive ion mode. For total ion mapping (automated MS/MS analysis), m/z range, 500 to 2000 was scanned with ITMS mode in successive 2 mass unit windows.

**Results:**

**Separation of CA119 168BSTK and CA119 168BNC by SDS-PAGE**

CAI 1-19 168BNC and CAI 1-19 168BSTK were separated by SDS-PAGE for in-gel digestion. See Figure 8. After electrophoresis, the gels were stained with CBB. Figure 8 shows the localization of target glycoprotein 100 KDa of CAI 1-19 168BNC and CAI 1-19 168BSTK. Since the 100 kDa band of CA 11-19 168BSTK was not clearly visible to observe in Figure 8, the gel was cut the region that corresponds to the molecular weight of 100 KDa and gel slices excised for O-glycan profiling.

**O-glycan profiling by MALDI/TOF-MS**

Permethylated O-glycans from the sample were analyzed by MALDI/TOF MS as well as NSI-MS and total ion mapping.

The MALDI-TOF-full mass spectra of 100 KDa sample bands in both samples are not shown in this report because MALDI-TOF analysis did not reveal major glycan signals but instead presented mostly non-glycan signals. The samples were further analyzed with a LTQ Orbitrap XL mass spectrometer. The full FTMS scan of the sample show similar results as MALDI-TOF analysis with some O-glycan
signals. NSI-FTMS spectra of O-glycan from the 100 KDa band from the both samples are shown in Figure 9. The upper spectrum presents O-glycans of CA11-19 168BNC(A) and the lower spectrum shows O-glycans of CA11-19 168BSTK(B). The major O-glycan structures are the core 2 structure (m/z 779) and the core3 with fucose (m/z 953). In the SDS-PAGE (Figure 8), we could definitely see the difference of sample concentration between CA11-19 168BNC and CA11-19 168BSTK. Consequentially, we did not observe the strong O-glycan signals from sample CA11-19 168BNC in sample CA11-19 168BSTK.

To confirm the O-glycan structures, total ion mapping (TIM) was carried out by NSI LTQ Orbitrap XL MS. Using TIM scan method, we were able to confirm several O-glycans which is summarized in Table 9. Both analyses showed a similar trend. CA11-19 168BNC sample shows clearer MSMS spectra than CA11-19 168BSTK sample because of the sample concentration.

The MSMS spectra of the permethylated O-glycan at m/z 936, m/z 953, and m/z 1157 derived from the O-glycan spectrum of CA11-19 168BNC (Figure 9 upper spectrum) are shown in Figure 10 as examples. As explained in Figure 10A, MSMS profile of the molecular ion at m/z 936 (NeuAc-HexNAc-HexNAc), the diagnostic b ion is shown at m/z 398, together with the complementary y ions at m/z 561, m/z 659, and m/z 677. The MSMS profile of the molecular ion at m/z 953 (Hex-DHex-HexNAc2) and m/z 1157 (Hex2-DHex-HexNAc2) are interpreted in Figure 10B and IOC.

The MSMS profile of CA11-19 168BNC and CA11-19 168BSTK showed the same patterns and we present spectra from CA11-19 168BNC in this report.
Table 9. Summary of O-glycans from the CAI 1-19 168B glycoproteins.

<table>
<thead>
<tr>
<th>Structural assignment</th>
<th>Theoretical m/z</th>
<th>Observed m/z</th>
<th>NC</th>
<th>STK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hex-HexNAc</td>
<td>534.29</td>
<td>534.29</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>HexNAc2</td>
<td>575.32</td>
<td>575.32</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>NeuAc-HexNAc</td>
<td>691.36</td>
<td>691.37</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>Hex-HexNAc2</td>
<td>779.41</td>
<td>779.42</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>NeuAc-Hex-HexNAc</td>
<td>895.46</td>
<td>895.47</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>NeuAc-HexNAc2</td>
<td>936.49</td>
<td>936.50</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Hex-DHex-HexNAc2</td>
<td>953.51</td>
<td>953.50</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Hex2-HexNAc2</td>
<td>983.51</td>
<td>983.52</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>NeuAc-HexNAc2</td>
<td>1140.59</td>
<td>1140.60</td>
<td>√</td>
<td>-</td>
</tr>
<tr>
<td>Hex2-DHex-HexNAc2</td>
<td>1157.60</td>
<td>1157.61</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Hex3-HexNAc2</td>
<td>1187.61</td>
<td>1187.63</td>
<td>√</td>
<td>-</td>
</tr>
</tbody>
</table>

Discussion
Our results indicate that the major O-linked glycans from CAI 1-19 168BNC and CAI 1-19 168BSTK are the core 2 structure (m/z 779) and the core3 with fucose (m/z 953). Since the amount of CAI 1-19 168BSTK glycoprotein was smaller than CAI 1-19 168BNC, the full ms of CAI 1-19 168BNC showed more O-glycan signals. It is being noted that the assigned glycan structures in the current report are based on common biosynthetic pathways.

EXAMPLE 6

Amino acid sequence of 100kDa material

The 100 kDa (168B) material was isolated and analyzed for amino acid sequence. Amino acid sequence analysis

One sample (in solution) was received. The proteins were reduced, alkylated, and digested with trypsin using the FASP protocol. The sample was resuspended in 2% acetonitrile/0.1% formic acid and analysis by LC-ESI-MS/MS. A 120 min. gradient was used and MS analysis was performed with an LTQ Orbitrap Velos mass spectrometer. MS spectra were searched using the MASCOT algorithm and the results generated an Excel spreadsheet. Searches using the tryptic peptides were conducted against a Human uniprot database. The protein was identified as having a sequence of CEACAM5 isofrom 2 (SEQ ID NO:1).

Peptide Mass Fingerprint of Antigen 168B

In order to characterize the antigen 168B, we submitted the sample to trypsin, chymotrypsin and ASP-N proteolysis followed by LC-LTQ Orbitrap MS/MS analysis. For the characterization of the antigen, a nano-LC chromatography was processed using a Ultimate 3000 (Dionex) system in line with a LTQ Orbitrap XL mass spectrometer (Thermo).

Sample preparation:

5 μl of the antigen (2.5 μM) was mixed with 20 μl of ammonium bicarbonate (25 mM, pH 8.3). After mixing 2.5 μl of DTT (500 mM) is added to the solution. The mixture was then incubated 1 hour at 55°C. After incubation, 2.5 μl of iodoacetamide (1M) is added before 1 hour of incubation at room temperature in a dark room. After incubation, the solution is diluted 1/5 by adding 120 μl of the buffer used for the proteolysis.
145 µl of the reduced/alkylated antigen was mixed with 2 µl of trypsin (Roche Diagnostic) at ratio 1/20. The proteolytic mixture was incubated overnight at 37°C.

145 µl of the reduced/alkylated antigen was mixed with 2 µl of chymotrypsin (Roche Diagnostic) at ratio 1/20. The proteolytic mixture was incubated overnight at 30°C.

145 µl of the reduced/alkylated antigen was mixed with 2 µl of ASP-N (Roche Diagnostic) at ratio 1/20. The proteolytic mixture was incubated overnight at 30°C.

After proteolysis, 10 µl of the peptide generated by proteolysis was loaded onto a nano-liquid chromatography system (Ultimate 3000, Dionex).

- A 95/05/0.1 H2O/ACN/HCOOH v/v/v
- B 20/80/0.1 H2O/ACN/HCOOH v/v/v
- gradient 5-40 % B in 35 minutes
- injected volume 10 µl
- precolumn 300-µm ID x 5-mm C4 PepMapTM

- precolumn flow rate 30 µl/min
- column 75-µm ID x 5-cm C4 PepMapTM
- column flow rate 200 nl/min

The LTQ orbitrap MS analysis has been performed with the following parameters:

- needle voltage 1.8 V
- capillary voltage 5 V
- µscan MS 1
- µscan MS2 1
- MS range m/z 300-1700

- MS/MS strategy MS+6CID MS/MS
- Min. signal required 500
- Ion isolation window 3 m/z units
- Normalized collision energy 35%
- Default charge state 2

- Activation Q 0.25
- Activation time 30
- Dynamic exclusion ON
- Dynamic exclusion params RC 1, RD 30s, ED 30s
-Charge state screening  ON
-Charge state rejection  ON
-Charge state reject. Params +land unassigned rejected

Validation Filters

5 Peptide Delta CN>0.1
Peptide Xcorrvs Charge state>1.5 (+1), 2.00(+2), 2.50(+3), 3.00(+4)
Peptide Probability<0.001
Peptide Matches=1
Protein Number of Different Peptides>2

10 Protein Exclude this reference: Keratin

Results
The sequence of CEACAM5 isoform 2 was used to analyze the mass spec data of the tryptic peptides. 18 tryptic peptides were identified in the sequence of 168B, covering 29.8%:

15 Identified Tryptic peptides of 168B

<table>
<thead>
<tr>
<th>Peptide</th>
<th>MII+</th>
<th>Position</th>
<th>Significance</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>MESPSAPPHR</td>
<td>1108.52</td>
<td>1-10</td>
<td>94.9</td>
<td>40</td>
</tr>
<tr>
<td>WCIPWQR</td>
<td>988.48</td>
<td>11-18</td>
<td>95.7</td>
<td>41</td>
</tr>
<tr>
<td>LLLTASLLTFWNPPTTAK</td>
<td>1987.12</td>
<td>18-35</td>
<td>96.8</td>
<td>42</td>
</tr>
<tr>
<td>VDGGR</td>
<td>560.27</td>
<td>73-77</td>
<td>95.9</td>
<td>43</td>
</tr>
<tr>
<td>QIIGYVIGTQQATPGPAYSGR</td>
<td>2177.13</td>
<td>78-98</td>
<td>92.4</td>
<td>44</td>
</tr>
<tr>
<td>DVLNVEATGQFR</td>
<td>1465.69</td>
<td>127-139</td>
<td>95.8</td>
<td>45</td>
</tr>
<tr>
<td>VYPELPKSISSNSKPVEDK</td>
<td>2328.2</td>
<td>140-160</td>
<td>94.6</td>
<td>46</td>
</tr>
<tr>
<td>LQLSNGNR</td>
<td>901.48</td>
<td>191-198</td>
<td>99.2</td>
<td>47</td>
</tr>
<tr>
<td>TLTLFNVTR</td>
<td>1064.6</td>
<td>199-207</td>
<td>97.1</td>
<td>48</td>
</tr>
<tr>
<td>NDTASYK</td>
<td>798.36</td>
<td>208-214</td>
<td>98.9</td>
<td>49</td>
</tr>
<tr>
<td>CETQNPVSAR</td>
<td>1104.51</td>
<td>215-224</td>
<td>93.6</td>
<td>50</td>
</tr>
<tr>
<td>LQLSNDNR</td>
<td>959.49</td>
<td>368-375</td>
<td>98.5</td>
<td>51</td>
</tr>
<tr>
<td>TLTLSSVTR</td>
<td>1003.61</td>
<td>376-384</td>
<td>92.2</td>
<td>52</td>
</tr>
<tr>
<td>NSGLYTCQAANSASGHRS</td>
<td>1866.81</td>
<td>470-487</td>
<td>97.2</td>
<td>53</td>
</tr>
<tr>
<td>AYVCGIINVSANSANR</td>
<td>1481.71</td>
<td>567-580</td>
<td>92.5</td>
<td>54</td>
</tr>
<tr>
<td>INGIPQQHTQVLFIAK</td>
<td>1807.02</td>
<td>628-643</td>
<td>95.3</td>
<td>55</td>
</tr>
<tr>
<td>ITPNNNGTYACFVSNLATGR</td>
<td>2113.01</td>
<td>644-663</td>
<td>95.1</td>
<td>56</td>
</tr>
<tr>
<td>NNSrVK</td>
<td>674.38</td>
<td>664-669</td>
<td>96.5</td>
<td>57</td>
</tr>
</tbody>
</table>
The sequence of CEACAM5 isoform 2 was used to analyze the mass spec data of the chymotryptic peptides. 35 chymotryptic peptides were identified in the sequence of B168, covering 65.3%:

**Identified Chymotrypsin peptides of 168B**

<table>
<thead>
<tr>
<th>Peptide</th>
<th>MII+</th>
<th>Position</th>
<th>Significance</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESPSAPPHRW</td>
<td>1163.55</td>
<td>2-11</td>
<td>92.5</td>
<td>58</td>
</tr>
<tr>
<td>ESTPFNVAEGKEVLVLVHNL</td>
<td>2209.5</td>
<td>39-58</td>
<td>94.4</td>
<td>59</td>
</tr>
<tr>
<td>HLFGYSWYK</td>
<td>1199.56</td>
<td>61-69</td>
<td>96.5</td>
<td>60</td>
</tr>
<tr>
<td>KGERVDGNRQIIGY</td>
<td>1604.85</td>
<td>69-82</td>
<td>90.3</td>
<td>61</td>
</tr>
<tr>
<td>VIGTQQATPGPAY</td>
<td>1302.66</td>
<td>83-95</td>
<td>94.5</td>
<td>62</td>
</tr>
<tr>
<td>SGREIYPNASL</td>
<td>1319.69</td>
<td>96-107</td>
<td>96.6</td>
<td>63</td>
</tr>
<tr>
<td>IQNIIQNDTGF</td>
<td>1262.63</td>
<td>109-119</td>
<td>97</td>
<td>64</td>
</tr>
<tr>
<td>VNEEATGQF</td>
<td>994.44</td>
<td>130-138</td>
<td>92.3</td>
<td>65</td>
</tr>
<tr>
<td>TCEPETQDATY</td>
<td>1257.49</td>
<td>166-176</td>
<td>93.5</td>
<td>66</td>
</tr>
<tr>
<td>VNNQLPVSPLRL</td>
<td>1323.73</td>
<td>180-191</td>
<td>94</td>
<td>67</td>
</tr>
<tr>
<td>NVTRNDTASY</td>
<td>1140.52</td>
<td>204-213</td>
<td>96.1</td>
<td>68</td>
</tr>
<tr>
<td>TISPLNTSYSGENLNL</td>
<td>1877.93</td>
<td>241-257</td>
<td>98.1</td>
<td>69</td>
</tr>
<tr>
<td>SCHAASNPPAQY</td>
<td>1245.53</td>
<td>258-269</td>
<td>94.3</td>
<td>70</td>
</tr>
<tr>
<td>IPNITVNSGSY</td>
<td>1278.63</td>
<td>286-297</td>
<td>93.6</td>
<td>71</td>
</tr>
<tr>
<td>TCQAHNSDTGL</td>
<td>1146.48</td>
<td>298-308</td>
<td>90.5</td>
<td>72</td>
</tr>
<tr>
<td>NRTTVTTITVY</td>
<td>1268.68</td>
<td>309-319</td>
<td>92.8</td>
<td>73</td>
</tr>
<tr>
<td>ITSNNSNPVEDEAVAVAL</td>
<td>1787.82</td>
<td>326-342</td>
<td>96.5</td>
<td>74</td>
</tr>
<tr>
<td>TCEPEIQNNTY</td>
<td>1298.55</td>
<td>343-353</td>
<td>97.6</td>
<td>75</td>
</tr>
<tr>
<td>SVTRNDVGPYECAIQNEL</td>
<td>1993.92</td>
<td>381-398</td>
<td>92.6</td>
<td>76</td>
</tr>
<tr>
<td>SVDHDSPVIL</td>
<td>1081.55</td>
<td>399-408</td>
<td>93.3</td>
<td>77</td>
</tr>
<tr>
<td>GPDDPTISPSY</td>
<td>1148.51</td>
<td>413-423</td>
<td>94.5</td>
<td>78</td>
</tr>
<tr>
<td>YRPGVNLSSLSCHAASNPPAQYYS</td>
<td>2331.09</td>
<td>426-447</td>
<td>96.3</td>
<td>79</td>
</tr>
<tr>
<td>IDENIQHTQEL</td>
<td>1467.7</td>
<td>450-461</td>
<td>91.5</td>
<td>80</td>
</tr>
<tr>
<td>ISNITEKNSGL</td>
<td>1175.62</td>
<td>463-473</td>
<td>92.5</td>
<td>81</td>
</tr>
<tr>
<td>VKTITVSAELPKPSISNNSKP</td>
<td>2296.25</td>
<td>490-511</td>
<td>93.8</td>
<td>82</td>
</tr>
<tr>
<td>TCEPEAQNNTY</td>
<td>1256.5</td>
<td>521-531</td>
<td>94.5</td>
<td>83</td>
</tr>
<tr>
<td>VNGQSLPVSPRL</td>
<td>1266.71</td>
<td>535-546</td>
<td>96.2</td>
<td>84</td>
</tr>
<tr>
<td>NVTRNDARAY</td>
<td>1179.58</td>
<td>559-568</td>
<td>98.1</td>
<td>85</td>
</tr>
<tr>
<td>VCGIQNSSVSANRDPVLVL</td>
<td>1859.92</td>
<td>569-586</td>
<td>95.6</td>
<td>86</td>
</tr>
<tr>
<td>GPDPITISPDSSY</td>
<td>1445.67</td>
<td>591-604</td>
<td>95.6</td>
<td>87</td>
</tr>
<tr>
<td>SCHASANPSPQY</td>
<td>1277.52</td>
<td>613-624</td>
<td>90.5</td>
<td>88</td>
</tr>
<tr>
<td>RINGIPQHTQVQL</td>
<td>1503.83</td>
<td>627-639</td>
<td>90.6</td>
<td>89</td>
</tr>
<tr>
<td>IAKITPNNGNTY</td>
<td>1305.67</td>
<td>641-652</td>
<td>92.2</td>
<td>90</td>
</tr>
<tr>
<td>SGTPGSLAGAT</td>
<td>1004.46</td>
<td>676-687</td>
<td>94.8</td>
<td>91</td>
</tr>
</tbody>
</table>
9 peptides were identified by Asp N proteolysis in the sequence of B 168, covering 16.8%:

**Identified ASP-N peptides of 168B**

<table>
<thead>
<tr>
<th>Peptide</th>
<th>MII+</th>
<th>Position</th>
<th>Significance</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTGFYTLHVIKS</td>
<td>1380.7158</td>
<td>116-127</td>
<td>94.5</td>
<td>93</td>
</tr>
<tr>
<td>DAVAFTCEPETQ</td>
<td>13 10.5569</td>
<td>161-172</td>
<td>96.2</td>
<td>94</td>
</tr>
<tr>
<td>DTASYKCETQNPVSARRS</td>
<td>2012.9454</td>
<td>209-226</td>
<td>92.5</td>
<td>95</td>
</tr>
<tr>
<td>DSVILNVLYGP</td>
<td>1189.6463</td>
<td>227-237</td>
<td>94.7</td>
<td>96</td>
</tr>
<tr>
<td>DNRTLTLLSVTRN</td>
<td>1502.8285</td>
<td>373-385</td>
<td>92</td>
<td>97</td>
</tr>
<tr>
<td>DVGPYECGIQNELSV</td>
<td>1622.7366</td>
<td>386-400</td>
<td>98.6</td>
<td>98</td>
</tr>
<tr>
<td>DPVILNVLYGP</td>
<td>1199.667</td>
<td>404-414</td>
<td>93.3</td>
<td>99</td>
</tr>
<tr>
<td>DARAYVCGIQNSVSANRS</td>
<td>1910.9137</td>
<td>564-581</td>
<td>95.7</td>
<td>100</td>
</tr>
<tr>
<td>DTPHISPP</td>
<td>839.4509</td>
<td>593-600</td>
<td>96.1</td>
<td>101</td>
</tr>
</tbody>
</table>

Based on the results obtained, we designed overlap mapping of the trypsin, chymotrypsin and ASP-N peptides (Figure 14). Combining the peptides of Trypsin, Chymotrypsin and ASP-N proteolysis, 85.6% of the sequence is covered. The polypeptide is identified as CEACAM5 isoform 2 as shown in figure 14 (SEQ ID NO:l)

**EXAMPLE 7**

**Antibody characterization: Mass Spectroscopy of Antibodies Cross-linked with Antigen**

**Sample preparation**

Samples provided included Antibody 1 5E5-1: 0.2 mg/ml; 2ml; glycine buffer; Antibody 2 5A1-1: 1..2 mg/ml; 2ml; glycine buffer; and Antigen 168B: 0.8 mg/ml; 100 µl; PBS pH 7.4. Antibody 5E5 and Antibody 5A1-1 have been shown to identify a biomarker for colorectal cancer in serum from patients. The antigen is identified as 168B and is a partially glycosylated form of CEACAM5 as described herein.

Three mAb/Ag interactions samples were prepared with the following concentrations:
A Antigen: 168B Antibody: 5E5-1 Mix 168B/5E5-1

5 microliters of the antigen (4 micromolar) were mixed with 5 microliters of antibody (2 micromolar) to form a mixture of 10 microliter (2 micromolar antigen and 1 micromolar antibody)


5 microliters of the antigen (4 micromolar) were mixed with 5 microliters of antibody (2 micromolar) to form a mixture of 10 microliter (2 micromolar antigen and 1 micromolar antibody)

C Antigen: 168B Antibody: 5A1-1/5E5-1 Mix 168B/5A1-1/5E5-1

5 microliters of the antigen (4 micromolar) were mixed with 2 microliters of each antibody (2 micromolar) to form a mixture of 10 microliter (2 micromolar antigen/ 1 micromolar antibody 5A1-1/ 1 micromolar antibody 5E5-1)

1 μl of the mixture obtained was mixed with 1 μl of a matrix composed of a re-crystallized sinapinic acid matrix (10 mg/ml) in acetonitrile/water (1:1, v/v), TFA 0.1% (K200 MALDI Kit). After mixing, 1 μl of each sample was spotted on the MALDI plate (SCOUT 384). After crystallization at room temperature, the plate was introduced in the MALDI mass spectrometer and analyzed immediately. The analysis has been repeated in triplicate.

Cross-link Experiments

The mixtures prepared above (9 μl left for each mixture) was submitted to cross-linking using CovalX’s K200 MALDI MS analysis kit. 9 μl of the mixture was mixed with 1 μl of K200 Stabilizer reagent (2 mg/ml) and incubated at room temperature. CovalX uses a mixture of amine based crosslinking agents to stabilize the interactions between the antibody and the antigen. After the incubation time (180 minutes), the samples were prepared for MALDI analysis. The samples were analyzed by High-Mass MALDI analysis immediately after crystallization.

High-Mass MALDI MS analysis

The MALDI ToF MS analysis has been performed using CovalX’s HM3 interaction module with a standard nitrogen laser and focusing on different mass ranges from 0 to 2000 kDa.

For the analysis, the following parameters have been applied:

Mass Spectrometer:
Linear and Positive mode
Ion Source 1: 20 kV
Ion Source 2: 17 kV
Lens: 12 kV
Pulse Ion Extraction: 400 ns
HM3:
Gain Voltage: 3.14 kV
Acceleration Voltage: 20 kV

To calibrate the instrument, an external calibration with clusters of Insulin, BSA and IgG has been applied. For each sample, 3 spots were analyzed (300 laser shots per spots). The presented spectrum corresponds to the sum of 300 laser shots. The MS data were analyzed using CovalX's Complex Tracker analysis software version 2.0.

**Results**

Analysis of the mix A: 5E5-1/168B, showed that the antigen and the antibody were detected with observed MH+=92.879 kDa (168B-antigen) and MH+=158.498 kDa. (5E5-1 antibody)(Figure HA). The cross-linking experiment was completed after 180 minutes incubation time with the crosslinking reagent K200. After cross-linking, we detected two additional peaks with MH+=258.945 kDa [5E5-1·B168] and MH+=354.164 kDa [5E5-1·2 B168] (Figure IB). Using Complex Tracker software, we overlaid the control and cross-link spectra. The overlay confirmed the detection of two non covalent complexes [5E5-1·B168] and [5E5-1·2B168](Figure IC).

Analysis of the Mix B: 5A1-1/168B, showed that the antigen and the antibody were detected with MH+=92.743 kDa (168B) and MH+=159.125 kDa.(5A1-l) (data not shown). The cross-linking experiment was completed after 180 minutes incubation time with the crosslinking reagent K200. After cross-linking, we detected two additional peaks with MH+=259.155kDa [5A1-1·B168] and MH+=355.025 kDa [5A1-1·2B168] (data not shown). Using Complex Tracker software, we overlaid the control and cross-link spectra. The overlay confirmed the detection of two non covalent complexes [5A1-1·B168] (antibody binding with a single antigen) and [5A1-1·2B168] (antibody binding to two antigens)(data not shown).

Analysis of the Mix C: 5E5-1/5A1-1/168B showed that the antigen and the antibody were detected with MH+=92.910 kDa (168B) and MH+=158.988 kDa...
The cross-linking experiment was completed after 180 minutes incubation time with the crosslinking reagent K200. After cross-linking, we detected two additional peaks with MH+ = 260.155 kDa and MH+ = 356.012 kDa (figure 12B). Using Complex Tracker software, we overlaid the control and cross-link spectra. The overlay confirmed the detection of two non covalent complexes [5E5-1-B168] and [5E5-1-2B168]. (Figure 12C).

**Discussion**

The sandwich assay indicated that the simultaneous binding of the monoclonal antibodies 5E5-land 5A1-1 on the antigen 168 B was not observed. The analysis indicated that the epitope of these two monoclonal antibodies on the antigen 168B are located in the same close region of the protein.

**EXAMPLE 8**

**Epitope mapping of Antibody 5E5-1**

A competition assay with peptides derived from the antigen was conducted to determine if the epitope was linear.

**Pepsin proteolysis**

In order to determine the nature of the epitope, we performed a proteolysis of 168B antigen with immobilized pepsin. 10 µl of the antigen with a concentration of 2.5 µM were mixed with immobilized pepsin 2.5 µM and incubated at room temperature for 30 minutes. After the incubation time the sample was centrifuged and the supernatant was pipetted. The completion of the proteolysis was controlled by High-Mass MALDI mass spectrometry in linear mode and reflectron mode. The pepsin proteolysis was optimized in order to obtain large amounts of peptides in the 1000-3500 Da range.

**Pepsin peptides/ antibody/ antigen mixing and incubation**

5 µl of the antigen peptides generated by proteolysis was concentrated 10X using speed vac and mixed with 5 µl of 5E5-1 antibody (4 µM) and incubated at 37° for 2 hours.

**Antigen peptides/Antigen/Antibody mixing**

After incubation of the antibodies with the antigen peptides, 5µl of the mixture is mixed with 5µl of the intact antigen (4 µM). After mixing, the mixture contains the following proteins:
168B peptides 5E5-1 Antibody/ 168B antigen peptides Mix 5E5-1 antibody/ 168B peptides/ 168B antigen.

**Interaction analysis**

For the competition assay, the antibody/antigen interaction analysis was performed with the same protocol as described in Example 6.

**Results**

For this experiment, the antigen and the antibody were detected with MH+=92.879 kDa and MH+=158.498 kDa. (Figure 13A). The cross-linking experiment was completed after 180 minutes incubation time with the crosslinking reagent K200. After cross-linking, we detected two additional peaks with MH+=258.945 kDa and MH+=354.164 kDa. Using Complex Tracker software, we overlaid the control and cross-link spectra. (Figure 13B) The overlay confirmed the detection of two non covalent complexes [5E5-1·B168] and [5E5-1-2B168] (Figure 13C).

After competition, we did not detect inhibition of the binding of the antibody 5E5-1 on the antigen 168B. After competition, we observed two non covalent complexes: [5E5-1·B168] and [5E5-1-2B168] (Figure 13).

**Discussion**

The competition assay indicated that peptides of the antigen are not inhibiting the binding Antibody/Antigen. The epitope 5E5-1 on 168B is not linear. The epitope is based on the conformation of the antigen.

**Characterization of the molecular interfaces B168/5E5-1**

In order to determine the epitope of 5E5-1 antibody on 168B antigen with high resolution the antibody/antigen complexes are incubated with deuterated cross-linkers and subjected to multienzymatic cleavage. After enrichment of the cross-linked peptides, the samples are analyzed by high resolution mass spectrometry (nLC-Orbitrap MS) and the data generated are analyzed using XQuest and Stavrox software.

For this analysis, a nLC in combination with Orbitrap mass spectrometry have been used as described above. 5µl of the antigen sample (concentration 4 µM) was mixed with 5µl of the antibody sample (Concentration 2µM) in order to obtain an antibody/antigen mix with final concentration 1µM/2µM. The mixture was incubated at 37°C for 180 minutes. In a first step, 1 mg of dO cross-linker was mixed...
with 1 mg of dl2 cross-linker. The 2 mg prepared were mixed with 1 ml of DMF in order to obtain a 2 mg/ml solution of DSS do/dl2. 10 µl of the antibody/antigen mix prepared previously were mixed with 1 µl of the solution of cross-linker d0/dl2 prepared (2 mg/ml). The solution is incubated 180 minutes at room temperature in order to achieve the cross-linking reaction.

In order to facilitate the proteolysis, it is necessary to reduce the disulfide bond present in this protein. The cross-linked sample was mixed with 20 µl of ammonium bicarbonate (25 mM, pH 8.3). After mixing 2.5 µl of DTT (500 mM) is added to the solution. The mixture is then incubated 1 hour at 55°C. After incubation, 2.5 µl of iodoacetamide (1M) is added before 1 hour of incubation at room temperature in a dark room. After incubation, the solution is diluted 1/5 by adding 120 µl of the buffer used for the proteolysis.

145 µl of the reduced/alkyled cross-linked sample was mixed with 2 µl of trypsin (Roche). The proteolytic mixture was incubated overnight at 37°C.

For this proteolysis, the buffer is Tris-HCL 100 mM, CaCl2 10 mM, pH7.8. 145µl of the reduced/alkyled cross-linked complex were mixed with 2 µl of a-chymotrypsin 200µM and incubated overnight at 30°C.

145 µl of the reduced/alkyled cross-linked sample was mixed with 2 µl of ASP-N (Roche). The proteolytic mixture was incubated overnight at 37°C.

**Results**

After cross-linking, the peptides generated by multi-enzymatic proteolysis are covering 81% of the total antigen sequence. After trypsin digestion of the antibody/antigen cross-linked complex with deuterated d0 dl2, the nLC-orbitrap MS/MS analysis allowed to detect 32 peptides having monolinks. These monolinked peptides have been detected with both Xquest and Stavrox software.

After trypsin digestion of the antibody/antigen cross-linked complex with deuterated d0 dl2, the nLC-orbitrap MS/MS analysis allowed to detect 4 cross-linked peptides between the antigen 168B and the antibody 5E5-1 heavy chain. These cross-linked peptides have been detected with both Xquest and Stavrox software.

**Trypsin Cross-link Identified: 168B/5E5-1**

**Protein 1/Protein 2**

QIIGYVIGTQQATPGPAYSGR-LSGTAGVHSQVQL-al9-b2
Position Protein 1/ Protein2
168B 78-98  5E5-1 HC 11-24 (1-4 of heavy chain variable domain)

Protein 1/Protein 2
QIIGYVGTQQATPGPAYSGR-LSGTAGVHSQVQL-a2  1-b2

Position Protein 1/ Protein2
168B 78-98  5E5-1 HC 11-24 (1-4 of heavy chain variable domain)

Protein 1/Protein 2
SDLVNEEATGQFR-VQLQQSGAD-al3-b6

Position Protein 1/ Protein2
168B 127-139  5E5-1 HC 21-29 (1-9 of heavy chain variable domain)

Protein 1/Protein 2
SDLVNEEATGQFR-VQLQQSGAD-al3-b6

Position Protein 1/ Protein2
168B 127-139  5E5-1 HC 21-29 (1-9 of heavy chain variable domain)

15 After chymotrypsin digestion of the antibody/antigen cross-linked complex with deuterated dO dl2, the nLC-orbitrap MS/MS analysis allowed to detect 4 cross-linked peptides between 168 B antigen and 5E5-1 antibody Heavy Chain and light chain. These cross-linked peptides have been detected with both Xquest and Stavrox software.

Chymotrypsin Cross-link Identified: 168B/5E5-1

Protein 1/Protein 2
HLFGYSWYK-QLLGLLLLCFQGTRC-a6-bl4

Position Protein 1/ Protein2
168B 61-69  5E5-1 LC 6-20

Protein 1/Protein 2
HLFGYSWYK-IQMTQITSS-a9-b8

Position Protein 1/ Protein2
168B 61-69  5E5-1 LC 22-30 (2-9 of light chain variable domain)

Protein 1/Protein 2
SGREIIYPNASL-GVHSQVQLQQS-al-b4
Position Protein 1 / Protein2
168B  96-107  5E5-1  HC  16-26 (1-7 of heavy chain variable domain)

Protein 1/Protein 2
SGREIYPNASL-GVHSVQVLQSY-a3-bl

Position Protein 1 / Protein2
168B  96-107  5E5-1  HC  16-26 (1-7 of heavy chain variable domain)

After ASP-N digestion of the antibody/antigen cross-linked complex we did not
detect any crosslinked peptides between 168B and 5E5-1.

Discussion
Using chemical cross-linking, High-Mass MALDI mass spectrometry and
nLC-Orbitrap mass spectrometry we were able to characterize the interaction
interface between the antigen 168B and the monoclonal antibody 5E5-1. Our
analysis indicates that the epitope of this monoclonal antibody includes the
following amino acids on 168B: 66; 69; 96; 98; 127; 139.
On the antibody the paratope includes the following amino acids: Heavy chain: 6
Light Chain: 9. These results are illustrated in Figure 15.

EXAMPLE 9
A competition assay with peptides derived from the antigen was conducted
to determine if the epitope was linear for antibody 5A1-1.

Pepsin proteolysis
In order to determine the nature of the epitope, we performed a proteolysis of
168B antigen with immobilized pepsin. 10 µl of the antigen with a concentration of
2.5 µM were mixed with immobilized pepsin 2.5 µM and incubated at room
temperature for 30 minutes. After the incubation time the sample was centrifuged
and the supernatant was pipetted. The completion of the proteolysis was controlled
by High-Mass MALDI mass spectrometry in linear mode and reflectron mode. The
pepsin proteolysis was optimized in order to obtain large amounts of peptides in the
1000-3500 Da range.

Pepsin peptides/ antibody/ antigen mixing and incubation
5 µl of the antigen peptides generated by proteolysis was concentrated 10X
using speed vac and mixed with 5 µl of 5A1-1 antibody (4 µM) and incubated at 37°
for 2 hours.
Antigen peptides/Antigen/Antibody mixing

After incubation of the antibodies with the antigen peptides, 5µ1 of the mixture is mixed with 5µ1 of the intact antigen (4 µM). After mixing, the mixture contains the following proteins:


Interaction analysis

For the competition assay, the antibody/antigen interaction analysis was performed with the same protocol as described in Example 6.

Results

For this experiment, the antigen and the antibody were detected with MH+=92.879 kDa and MH+=158.498 kDa. (data not shown). The cross-linking experiment was completed after 180 minutes incubation time with the crosslinking reagent K200. After cross-linking, we detected two additional peaks with MH+=258.945 kDa and MH+=354.164 kDa. Using Complex Tracker software, we overlaid the control and cross-link spectra, (data not shown) The overlay confirmed the detection of two non covalent complexes [5A 1-1·B 168] and [5A 1-1·2B 168] (data not shown).

After competition, we did not detect inhibition of the binding of the antibody 5A1-1 on the antigen 168B. After competition, we observed two non covalent complexes: [5A 1-1·B 168] and [5A 1-1·2B 168] (data not shown). This indicates the epitope is not linear.

Characterization of the molecular interfaces B168/5A1-1

In order to determine the epitope of 5A1-1 antibody on 168B antigen with high resolution the antibody/antigen complexes are incubated with deuterated cross-linkers and subjected to multienzymatic cleavage. After enrichment of the cross-linked peptides, the samples are analyzed by high resolution mass spectrometry (nLC-Orbitrap MS) and the data generated are analyzed using XQuest and Stavrox software.

For this analysis, a nLC in combination with Orbitrap mass spectrometry have been used as described above. 5µ1 of the antigen sample (concentration 4 µM) was mixed with 5µ1 of the antibody sample (Concentration 2µM) in order to obtain an antibody/antigen mix with final concentration 1µM/2µM. The mixture was
incubated at 37°C for 180 minutes. In a first step, 1 mg of dO cross-linker was mixed with 1 mg of dl2 cross-linker. The 2 mg prepared were mixed with 1 ml of DMF in order to obtain a 2 mg/ml solution of DSS do/dl2. 10 µ of the antibody/antigen mix prepared previously were mixed with 1 µ of the solution of cross-linker d0/dl2 prepared (2 mg/ml). The solution is incubated 180 minutes at room temperature in order to achieve the cross-linking reaction.

In order to facilitate the proteolysis, it is necessary to reduce the disulfide bond present in this protein. The cross-linked sample was mixed with 20 µ of ammonium bicarbonate (25 mM, pH 8.3). After mixing 2.5 µ of DTT (500 mM) is added to the solution. The mixture is then incubated 1 hour at 55°C. After incubation, 2.5 µ of iodoacetamide (1M) is added before 1 hour of incubation at room temperature in a dark room. After incubation, the solution is diluted 1/5 by adding 120 µ of the buffer used for the proteolysis.

145 µ of the reduced/alkylated cross-linked sample was mixed with 2 µ of trypsin (Roche). The proteolytic mixture was incubated overnight at 37°C.

For this proteolysis, the buffer is Tris-HCL 100 mM, CaCl2 10 mM, pH7.8. 145µl of the reduced/alkylated cross-linked complex were mixed with 2 µ of α-chymotrypsin 200µM and incubated overnight at 30°C.

145 µ of the reduced/alkylated cross-linked sample was mixed with 2 µ of ASP-N (Roche). The proteolytic mixture was incubated overnight at 37°C.

Results

After cross-linking, the peptides generated by multi-enzymatic proteolysis are covering 81% of the total antigen sequence. After trypsin digestion of the antibody/antigen cross-linked complex with deuterated dO dl2, the nLC-orbitrap MS/MS analysis allowed to detect 35 peptides having monolinks. These monolinked peptides have been detected with both Xquest and Stavrox software.

After trypsin digestion of the antibody/antigen cross-linked complex with deuterated dO dl2, the nLC-orbitrap MS/MS analysis allowed to detect 4 cross-linked peptides between the antigen 168B and the antibody 5A1-1 heavy chain.

These cross-linked peptides have been detected with both Xquest and Stavrox software.
Trypsin Cross-link Identified: 168B/5A1-1

Protein 1/Protein 2
VDGVR-VKPGASVKL-a5-b2

Position Protein 1/Protein 2
5 168B 73-77 5A1-1 HC 11-19

Protein 1/Protein 2
VDGVR-VKPGASVKL-a5-b2

Position Protein 1/Protein 2
168B 73-77 5A1-1 HC 11-19

After chymotrypsin digestion of the antibody/antigen cross-linked complex with deuterated d0 d12, the nLC-orbitrap MS/MS analysis allowed to detect 10 cross-linked peptides between 168 B antigen and 5A1-1 antibody Heavy Chain and light chain. These cross-linked peptides have been detected with both Xquest and Stavrox software.

Chymotrypsin Cross-link Identified: 168B/5A1-1

Protein 1/Protein 2
HLFGYSWYK-VGDRVTITCRA-a6-b4

Position Protein 1/Protein 2
20 168B 61-69 5A1-1 LC 15-25

Protein 1/Protein 2
HLFGYSWYK-VGDRVTITCRA-a9-b 10

Position Protein 1/Protein 2
25 168B 61-69 5A1-1 LC 15-25

Protein 1/Protein 2
KGERVDRGVRQIIGY -TCRASQSISSYLN-al-b2

Position Protein 1/Protein 2
30 168B 69-82 5A1-1 LC 22-34

Protein 1/Protein 2
KGERVDRGVRQIIGY -TCRASQSISSYLN-a4-b 10

Position Protein 1/Protein 2
30 168B 69-82 5A1-1 LC 22-34
Protein 1/Protein 2
KGERVDGNRQIIGY-VKPGASVKLS-a9-b2

Position Protein 1 / Protein2
168B 69-82 5A1-1 HC 11-19

Protein 1/Protein 2
SGREIIYPNASL-ASVKLSC-al-b2

Position Protein 1 / Protein2
168B 96-107 5A1-1 HC 16-22

Protein 1/Protein 2
SGREIIYPNASL-ASVKLSC-a3-b4

Position Protein 1 / Protein2
168B 96-107 5A1-1 HC 16-22

Protein 1/Protein 2
SGREIIYPNASL-CKASGYTFTNYWINWVKP-al 1-b2

Position Protein 1 / Protein2
168B 96-107 5A1-1 HC 21-39

Protein 1/Protein 2
SGREIIYPNASL-CKASGYTFTNYWINWVKP-al 1-b17

Position Protein 1 / Protein2
168B 96-107 5A1-1 HC 21-39

After ASP-N digestion of the antibody/antigen cross-linked complex we did not detect any crosslinked peptides between 168B and 5A1-1.

Discussion
Using chemical cross-linking, High-Mass MALDI mass spectrometry and nLC-Orbitrap mass spectrometry we were able to characterize the interaction interface between the antigen 168B and the monoclonal antibody 5A1-1. Our analysis indicates that the epitope of this monoclonal antibody includes the following amino acids on 168B: 66; 69; 72; 77; 96; 98; 106.
On the antibody the paratope includes the following amino acids: Heavy chain: 15; 19; 21; 23; 25; 40 and CDRH1. Light Chain: 18; 24; 31 and CDRL1. These results are illustrated in figure 16.

**Example 10: Lateral flow test**

A "Dipstick" Lateral flow test (CAI 1-19) was designed and evaluations on the detection of the antigen in serum and/or urine-based test for discriminating between patients: was determined. Patients samples included serum and urine from: A) with Early Stage colorectal adenocarcinomas; B) with adenomas; C) normal individuals (confirmed by colonoscopy). The Lateral flow test utilized the same monoclonal cocktail on the nitrocellulose membrane (e.g. monoclonal 5A1-1 and 5E5-1) with the goat CAI 1-19 polyclonal antibody conjugated to colloidal gold on the conjugate pad. Preliminary testing showed excellent correlation of the serum and urine from normal individuals and colorectal cancer patients. See Figure 21. The lack of source of serum and "fresh" urine from the same cancer patients prevented further evaluation on the urine samples. CAI 1-19 antigen in serum or urine was seen with + line in colorectal (# 2) and lung (# 4 and 6) CA patients and with no + line in normal individuals (1,3,5 and 7) . See Figure 21.

These results show that an immunoassay utilizing lateral flow can readily detect the presence on 100 kDa CAI 1-19 antigen described herein in serum and urine from cancer patients.

It is apparent that many modifications and variations of this invention as set forth above may be made without departing from the spirit and scope. The specific embodiments described are given by way of example only, and the invention is limited only by the terms of the appended claims.
CLAIMS

What is claimed:

1. An immunoassay kit comprising:
   a) a first antibody that specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, but does not substantially bind to glycosylated 180kDa CEA and does not substantially bind to one or more linear peptides consisting of a linear peptide of 15 amino acids from amino acids 1-60 of SEQ ID NO:2, a linear peptide of amino acids 111 to 125 of SEQ ID NO:3, and/or a linear peptide of 15 amino acids from amino acids 150-701 of SEQ ID NO:4;
   b) an additional antibody that specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, labeled with a signal generating element, wherein the first antibody has a different epitope specificity than the second antibody; and
   c) a calibrator.

2. The immunoassay kit of claim 1, further comprising a second antibody or antigen binding fragment thereof, wherein the second antibody or antigen binding fragment thereof specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, but does not substantially bind to glycosylated 180kDa CEA and does not substantially bind to one or more linear peptides consisting of a linear peptide of 15 amino acids from amino acids 1-60 of SEQ ID NO:2, a linear peptide of amino acids 111 to 125 of SEQ ID NO:3, and/or a linear peptide of 15 amino acids from amino acids 150-701 of SEQ ID NO:4

3. The immunoassay kit of claim 1 or claims 2, wherein the first and/or second antibody are attached to a solid substrate.

4. The immunoassay kit of any one of claims 1-3, wherein the calibrator is an isolated 100 kDa glycoprotein comprising SEQ ID NO:1.
5. The immunoassay kit of any one of claims 1-4, further comprising instructions for identifying a subject as at risk for colorectal cancer if the amount of antigen detected in a sample from the subject is about 6.5 units/ml. or greater.

6. The immunoassay kit of any one of claims 1-5, wherein the first antibody is a monoclonal antibody is attached to a solid substrate, and the additional antibody is a polyclonal antibody that is detectably labelled.

7. The immunoassay kit of claim 6, wherein the second antibody is a monoclonal antibody attached to a solid substrate.

8. The immunoassay kit of any one of claims 1-7, wherein the first and second antibody or both bind to an epitope having an amino acid sequence selected from the group consisting of 61-69, 73-77, and 96-107 of SEQ ID NO:1 and amino acids 61-69, 78-98, 96-107, and 127-139 of SEQ ID NO:1.

9. The immunoassay kit of any one of claims 1-8, wherein the first or second antibody or both comprises heavy chain CDRs from a heavy chain variable region having an amino acid sequence of selected from the group consisting of SEQ ID NO:19 and SEQ ID NO:121.

10. The immunoassay kit of claim 9, wherein the first or second antibody or both comprises light chain CDRs from a light chain variable region having an amino acid sequence of selected from the group consisting of SEQ ID NO:120 and SEQ ID NO:122.

11. An isolated antibody or antigen-binding fragment thereof that specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, but does not substantially bind to glycosylated 180 kDa CEA and does not substantially bind to one or more linear peptides consisting of a linear peptide of 15 amino acids from amino acids 1-60 of SEQ ID NO:2, a linear peptide of amino acids 111 to 125 of SEQ ID NO:3, and/or a linear peptide of 15 amino acids from amino acids 150-701 of SEQ ID NO:4.
12. An isolated antibody or antigen-binding fragment thereof of claim 11, that specifically binds to or competes for binding to an epitope of 100 kDa glycoprotein comprising SEQ ID NO:1, wherein the epitope comprises amino acids 61-69, 73-77, and 96-107 of SEQ ID NO:1.

13. An isolated antibody or antigen-binding fragment thereof of claim 11, that specifically binds to or competes for binding to an epitope of 100 kDa glycoprotein comprising SEQ ID NO:1, wherein the epitope comprises amino acids 61-69, 78-98, 96-107, and 127-139 of SEQ ID NO:1.

14. An isolated antibody or antigen-binding fragment thereof that specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, but does not substantially bind to glycosylated 180kDa CEA, wherein the antibody or antigen binding fragment comprises: (a) heavy chain CDRs (HCDR1, HCDR2 and HCDR3) from a heavy chain variable region having an amino acid sequence selected from the group consisting of SEQ ID NO:119 and SEQ ID NO:121; and (b) light chain CDRs (LCDR1, LCDR2 and LCDR3) from a light region having an amino acid sequence selected from the group consisting of SEQ ID NO:120 and SEQ ID NO:122.

15. The isolated antibody or antigen-binding fragment thereof of claim 14, wherein the antibody or antigen binding fragment comprises: a) HCDR1, HCDR2, and HCDR3 selected from the group consisting of SEQ ID NOs: 15/16/17 and SEQ ID NOs:20/21/22; and b) LCDR1, LCDR2, and LCDR3 selected from the group consisting of SEQ ID NOs:26/27/28 and SEQ ID NOs:31/32/33.

16. The antibody or antigen binding fragment of claim 14, that comprises: a) a heavy chain variable region having an amino acid sequence SEQ ID NO: 119 and a light chain variable region having an amino acid sequence of SEQ ID NO: 120; or a heavy chain variable region having an amino acid sequence SEQ ID NO: 121 and a light chain variable region having an amino acid sequence of SEQ ID NO: 122.
17. An immunoassay for determining whether a subject is at risk of colorectal cancer comprising:
   a) combining an antibody or antigen binding fragment thereof with a human tissue or fluid sample from the subject; wherein the antibody or antigen binding fragment specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, but does not substantially bind to glycosylated 180 kDa CEA and does not substantially bind to one or more linear peptides consisting of a linear peptide of 15 amino acids from amino acids 1-60 of SEQ ID NO:2, a linear peptide of amino acids 111 to 125 of SEQ ID NO:3, and/ or a linear peptide of 15 amino acids from amino acids 150-701 of SEQ ID NO:4; and
   b) determining the amount of the antigen in the human tissue or fluid by determining the amount of the antibody antigen complex.

18. The immunoassay of claim 17, wherein the antibody or antigen binding fragment is a monoclonal antibody.

19. The immunoassay of claim 18, wherein the antibody is a capture antibody.

20. The immunoassay of any one of claims 17-19, further comprising a second antibody or antigen binding fragment thereof, wherein the antibody or antigen binding fragment specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1, but does not substantially bind to glycosylated 180 kDa CEA and does not substantially bind to one or more linear peptides consisting of a linear peptide of 15 amino acids from amino acids 1-60 of SEQ ID NO:2, a linear peptide of amino acids 111 to 125 of SEQ ID NO:3, and/ or a linear peptide of 15 amino acids from amino acids 150-701 of SEQ ID NO:4.

21. The immunoassay of claim 20, wherein the antibody or antigen binding fragment is a monoclonal antibody.

22. The immunoassay of claim 21, wherein the antibody is a capture antibody.
23. The immunoassay of any one of claims 1-6, further comprising an additional antibody that specifically binds to a 100 kDa glycoprotein comprising SEQ ID NO:1.

24. The immunoassay of claim 17-23, wherein said additional antibody or antigen binding fragment is labeled with a label selected from the group consisting of a radionuclide, an enzyme, a fluorescent agent, colloidal gold, and a chromophore.

25. The immunoassay of claim 23, wherein the additional antibody or antigen binding fragment thereof is a polyclonal antibody.

26. The immunoassay of any one of claims 17-25, further comprising identifying the subject as at risk for colorectal cancer if the amount of the antigen in the human tissue or fluid is about 6.5 units/ml or greater.

27. The immunoassay of any one of claims 17-26, wherein the first and second antibody or both bind to an epitope having an amino acid sequence selected from the group consisting of 61-69, 73-77, and 96-107 of SEQ ID NO:1 and amino acids 61-69, 78-98, 96-107, and 127-139 of SEQ ID NO:1.

28. The immunoassay kit of any one of claims 17-27, wherein the first or second antibody or both comprises heavy chain CDRs from a heavy chain variable region having an amino acid sequence of selected from the group consisting of SEQ ID NO:19 and SEQ ID NO:121.

29. The immunoassay kit of claim 28, wherein the first or second antibody or both comprises light chain CDRS from a light chain variable region having an amino acid sequence of selected from the group consisting of SEQ ID NO:120 and SEQ ID NO:122.
ROC Curve of CA11-19 Assay (2009-2010)

Sensitivity = 94%
Specificity = 84%
Starting material (CA11-19 168B)

SDS-PAGE

Stain protein bands with CBB
Excise target bands and cut the bands into smaller pieces
Destain CBB

Reduction
Carboxyamidomethylation

PNGase F digestion

In-gel

Extraction

C18 Sep pak

N-glycans

Permethylation

C18 Sep pak

Mass spectrometry
MALDI-MS

FIG. 4
Figure 6

A

100 KDa

B

45 KDa

SUBSTITUTE SHEET (RULE 26)
Starting material (CA11-19 168BSTK & CA11-19 168BNC)

SDS-PAGE

Stain protein bands with CBB
Excise target bands and cut the bands into smaller pieces
Destain CBB

Ethyl acetate wash

Reductive β-elimination In-gel

Desalt C18 Sep pak

O-glycans

Permethylation

C18 Sep pak

Mass spectrometry
MALDI-MS
NSI-LTQ-MS

FIG. 7

SUBSTITUTE SHEET (RULE 26)
Figure 8

Figure 2. Separation of CA11-19 Lot168B STK and NC by SDS-PAGE.
### A. CLASSIFICATION OF SUBJECT MATTER

| IPC(B) | CPC | A61K 39/395 (2015.01) | A61K 47/48384 (2015.07) |

According to International Patent Classification (IPC) or to both national classification and IPC.

### B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols):

| IPC(B) | CPC | A61K 31/337, 39/395, 47/48; A61P 35/00; C07D 498/18; C07K 15/10, 16/30; C12N 5/12 (2015.01) |
| CPC | A61K 47/48384, 47/48576; C07K 16/3007, 2317/24, 2317/33, 2317/34, 2317/73, 2317/77, 2317/92 (2015.07) |

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched:

| CPC | A61K 47/48384, 47/48576; C07K 16/3007, 2317/24, 2317/33, 2317/34, 2317/73, 2317/77, 2317/92 (2015.07) (keyword delimited) |

US Classes - 24/1 30. 1, 133. 1; 530/387. 1, 530/387.7

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used):

PatBase, Google Patents, Google, PubMed

Search terms used: CEACAM, CEACAM5, isoform 2, antibody, immunglobulin, binding, Carcinoembryonic antigen related cell adhesion molecule.

### C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>US 2012/0107347 A1 (HODGE et al) 03 May 2012 (03.05.2012) entire document</td>
<td>1-3, 11-22</td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of Box C. See patent family annex.

- Special categories of cited documents:
  - "A" document defining the general state of the art which is not considered to be of particular relevance
  - "E" earlier application or patent but published on or after the international filing date
  - "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
  - "O" document referring to an oral disclosure, use, exhibition or other means
  - "P" document published prior to the international filing date but later than the priority date claimed

- "I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "K" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search: 20 July 2015

Date of mailing of the international search report: 05 AUG 2015

Name and mailing address of the ISA:
- Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
- P.O. Box 1450, Alexandria, Virginia 22313-1450
- Facsimile No. 371-273-8300

Authorized officer: Blaine Copenheaver
- PCT Helpdesk: 371-272-4300
- PCT OSP: 371-272-7774

Form PCT/ISA/210 (second sheet) (January 2015)
1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
   a. ☒ forming part of the international application as filed:
      [☑] in the form of an Annex C/ST.2 five text file.
      [ ] on paper or in the form of an image file.
   b. [ ] furnished together with the international application under PCT Rule 13/ter.1(a) for the purposes of international search only in the form of an Annex C/ST.2S text file.
   c. [ ] furnished subsequent to the international filing date for the purposes of international search only:
      [ ] in the form of an Annex C/ST.2S text file (Rule 13/ter.1(a)).
      [ ] on paper or in the form of an image file (Rule 13/ter.1(b) and Administrative Instructions, Section 713).

2. [ ] In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

SEQ ID NOs: 1-4, 15-17, 20-22, 26-28, 31-33, and 119-122 were searched
International search report

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [ ] Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:

2. [ ] Claims Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [x] Claims Nos.: 4-10, 23-29 because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

The additional search fees were accompanied by the applicant’s protest and, where applicable, the payment of a protest fee.

The additional search fees were accompanied by the applicant’s protest but the applicable protest fee was not paid within the time limit specified in the invitation.

No protest accompanied the payment of additional search fees.