Oncology Clinics & Research
ONCR, 1(1): 1-17
www.scitcentral.com

Review Article: Open Access

The Nature and Function of the Immunogenic Oncofetal Proteins and Monoclonals so derived, that Characterize Colorectal Carcinoma

Arlen M[1,2*], Arlen P[1,2], Tsang K[1,2], Coppa G[1,2], Conte C[1,2], Crawford J[1,2], Saric O[1,2] and Wang X[1,2]

*[*1]Department of Surgery, Division of Surgical Oncology, Northwell Health System and Hofstra College Medicine, Precision Biologics, USA*

[2]Department of Pathology, Northwell Health System and Hofstra College Medicine, Precision Biologics, USA.

Received June 15, 2018; Accepted July 07, 2018; Published April 03, 2019

ABSTRACT

Cancer as a cellular abnormality is caused by malfunction in the growth control system of the cell, leaving it with properties such as immortality and the capability of spreading beyond the site of origin. Such cells manipulate their growth pattern via a combination of proliferation/inhibition by tumor suppressor genes (e.g. Retinoblastoma protein (pRb), p53) and or proliferative activation by oncogenes (proto-oncogenes) (e.g. RAS, WNT, MYC, EKR and TRK).

Keywords: Immunogenic oncofetal proteins, Monoclonal antibodies, Cancer, Mutation, Immunohistocehmistry

INTRODUCTION

A mutation in these tumor suppressor genes and/or proto oncogenes in the cell, result in an unusually high rate of cell proliferation leading to the eventual appearance of the tumor cell [1]. Prior to this however, constant mutations are continually occurring in a wide array of superficial cells such as mucosa of bowel. This occurrence represents what one usually considers a field effect within the bowel mucosa probably induced by a virus such as the polyoma [2]. While most if not all of the cells within the field induced by such viral transformation appear phenotypically normal, should one employ immunohistochemistry (IHC) to define expression of existing tumor protein, those altered cells can now easily be defined. Such alterations represent the appearance of post translational modifications in the existing oncofetal proteins that characterize that specific type of malignancy. Here, one can see for the first time, sites of the expression of tumor protein within these normal appearing cells [3].

As one site of focus in the mucosa proliferates somewhat more rapidly than other adjacent sites, inhibitory factors are produced in the more advanced cell group keeping the other foci normal in their phenotypic appearance. This is probably the reason one rarely sees more than one clinically malignant growth (lesion) within a large segment of bowel. Removal of this particular growth is now the initiating factor for allowing any residual focus of atypia, after colectomy has been performed, to possibly reappear. Such as new growth usually expresses itself in the form of an anastomotic recurrence.

Knudson [4] described the resulting transforming cell as exhibiting early signs of aberrant growth such as altered morphology or unusually large size (hyperplasia). Developing tumor cells also tend to proliferate at a higher than usual rate to form that lesion characterized as either a benign tumor (dysplasia) or with the addition of further oncogene mutations, the malignant or cancerous growth. In later stages of cancer progression, such tumor cells proliferate at an unusually high rate, resulting in uncontrolled growth of the existing tumor.

In addition to progression and invasion of surrounding structures, the malignant lesion exhibits immortality as noted above. When we originally looked at the transcriptional protein FGARAT, it was apparent that this molecule, when overexpressed, is capable of activating the telomerase enzyme within the tumor system. Here, as the tumor grows, loss of a telomere at the time of cell division would normally shorten the chromosomal length until eventually the cell could no longer divide and as such reach its end state of

Corresponding author: Arlen M, Department of Surgery, Division of Surgical Oncology, Northwell Health System and Hofstra College Medicine, Precision Biologics, USA, Tel: +81-172-39-5004; E-mail: myronarlen@yahoo.com

Citation: Arlen M, Arlen P, Tsang K, Coppa G, Conte C, et al. (2019) The Nature and Function of the Immunogenic Oncofetal Proteins and Monoclonals so derived, that Characterize Colorectal Carcinoma. Oncol Clin Res, 1(1): 1-17.

development. In the presence of a functional enzyme system, induced by FGARAT, the chromosome cap is replaced and immortality of the cell is established. In looking for potential sites of telomerase replacement, a homologous sequence to FGARAT has been found in the capsule of the EBV virus. With this virus being almost universally found in most tissue of the body, it's not difficult to see how the virus helps the cancer cell achieve what is necessary for long term survival or immortality not seen in the benign lesion which is unable to maintain its telomere length.

AIM

In attempting to define which therapeutic approach can best be employed in managing the patient whose disease has progressed locally or has shown signs of metastasis, combination therapy appears ideal. Here, we have noted, that immunochemotherapy can surpass the results obtained with immunotherapy or chemotherapy alone. In order to define the molecular structure that characterizes the tumor and that can then be employed as a specific immunogen one must isolate and characterize that structure and the methods to employ it successfully. The purpose of this paper is to achieve this goal.

DISCUSSION

We have noted in experimental studies over the past several years that we can isolate and characterize the specific immunogen that essentially defines the tumor [5]. Such a molecule is represented by a post translational modification of an oncogene which in its altered state is extremely immunogenic [6]. Due to the suboptimal level of expression however, the cell is rarely recognized. This finding has allowed us to explain why a malignancy such as colon cancer is able to ward off or remain unrecognized by the host immune system. We have as such been able to show that in the transforming cell expression of this immunogenic molecule is less than 10% of the level needed for host recognition. Such a protein is almost always represented as noted, by a mutated oncofetal protein which acts as the immunogen but at too low a level to be able to induce a host immune response. This process of cellular transformation, wherein specific antigens appear, occurs at least 6 months prior to any atypical appearance of a mutated colonocyte that would suggest that the process of transformation is ongoing. Identification can only be achieved by immunohistochemistry (IHC). This process of expressing an altered or post translational modified oncofetal protein can be demonstrated in clusters of normal appearing cells adjacent to the malignant lesion. It suggests that the transformation process is not defined by one altered cell but rather occurs within a field in which the tumor transformation process had been initiated, again as noted above. This field appears to represent large clusters of transformed cells existing in a dormant state induced by the presence of an existing tumor that had progressed within the field of altered cells itself. Removal of the developing fully transformed malignancy appears to eliminate the presence of any localized immune suppressive activity arising from the

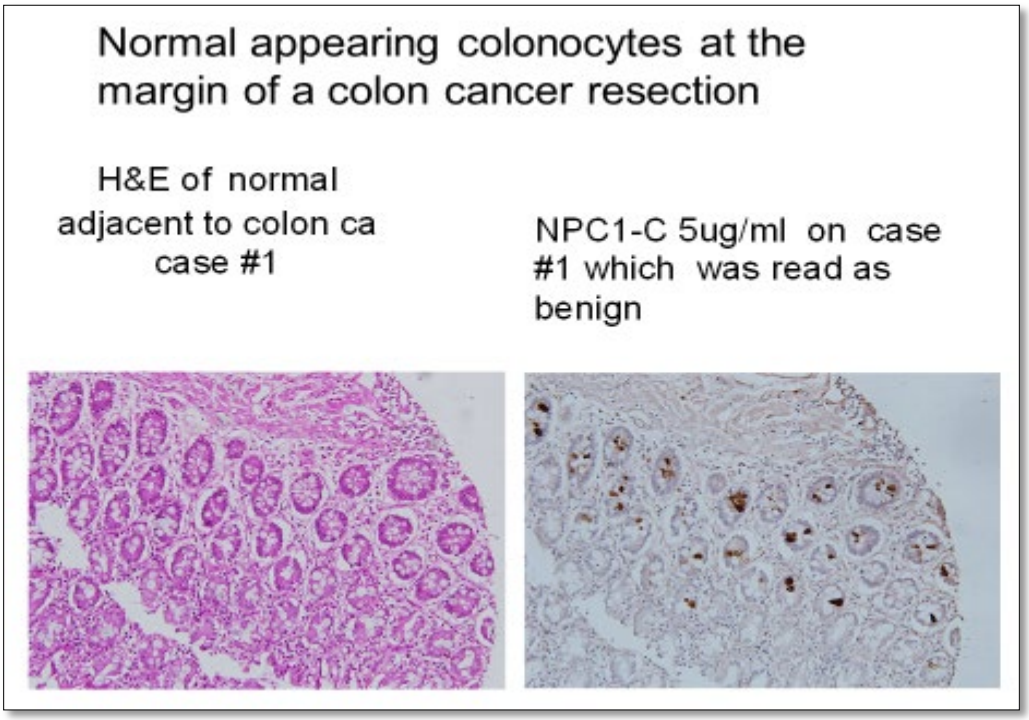

Figure 1. (a) Mucosa adjacent to the malignant lesion initially stained with H&E and secondarily with the antibody targeting the immunogenic oncofetal protein expressed by transforming cells as seen in (b).

primary lesion and allows further transformation within those atypical cells already existing within the field **(Figures 1a and 1b)**.

The slide section seen in **Figures 1a and 1b** was taken from normal appearing colon mucosa stained with H&E in an area a few cm. from a well-defined carcinomatous lesion. The margin of resection as shown here **(Figure 1a)** was considered to illustrate the appearance of normal colonocytes. As such an anastomosis was performed. Within one year, a recurrence appeared at the anastomotic site. When the same area of normal appearing tissue seen by the pathologist at the time of surgery, was stained with monoclonal antibody Neo102 which targets a post translational modification of the oncogene MUC5ac, it is apparent that many of the phenotypically normal appearing colonocytes had already converted to the premalignant state and were expressing tumor immunogenic protein **(Figure 1b)**. Should resection be performed for the primary malignancy and the pathologist is unable to identify transforming cells in the field around the malignancy, there is a possibility that these premalignant cells may be left behind, only to reappear in the anastomotic field within one to two years [7]. Such anastomotic recurrences can of course be prevented if a rapid immunohistochemistry procedure were performed by pathology in the OR and the surgeon notified of the need for further resection if it is felt that pre-malignancy will remain in the anastomosed colon.

As most tumors progress in their genetic transformation, they initiate the production of a number of factors including matrix metalloproteinase (MMP). This allows the cell producing MMP to invade adjacent vessels that have grown from the tumor stroma but also secondarily result in the breakdown of e-cadherin to its small form the s(small)-e cadherin which then initiates progression and metastasis from the primary lesion [8]. Many tumors that 'metastasize' or spread throughout the body reside in a specific location of implantation as a result of the surface glycoproteins expressed on the surface of the tumor, a seed and soil effect. As such, that lesion metastatic to liver may rely on the presence of hepatic ferritin (or an equivalent liver factor) to be present and can only grow in liver parenchyma and not spread to other organs. At best, such lesions can only proliferate and possibly spread within that organ in which it has implanted.

Should one be directed to identify a new colonic or other primary with a possible solitary metastatic lesion, resection for attempted cure is of course the main issue. One must be able to define heterogeneity compared to clonality of the metastatic site. As such whether histologic evaluation is performed with H&E, the addition of immunohistochemistry is essential to define the nature of the mutated oncofetal protein that is expressed in the lesion. In **Figure 2** we see that the colon lesion expresses the antigen for MUC5ac, but in an altered form in the tumor. This variant of the antigen of course is expressed only in the malignant component of the transformation process and not in adjacent normal mucosa. We now have determined that should immunotherapy be required as part of the therapeutic approach, when monoclonal antibodies are needed for any detected metastasis, that mAb Neo-102 would be selected in this particular situation. If the lesion proved to be localized but presenting features that might be considered of high risk for recurrence then a peptide vaccine construct of the immunogen would be utilized (**Figure 2**).

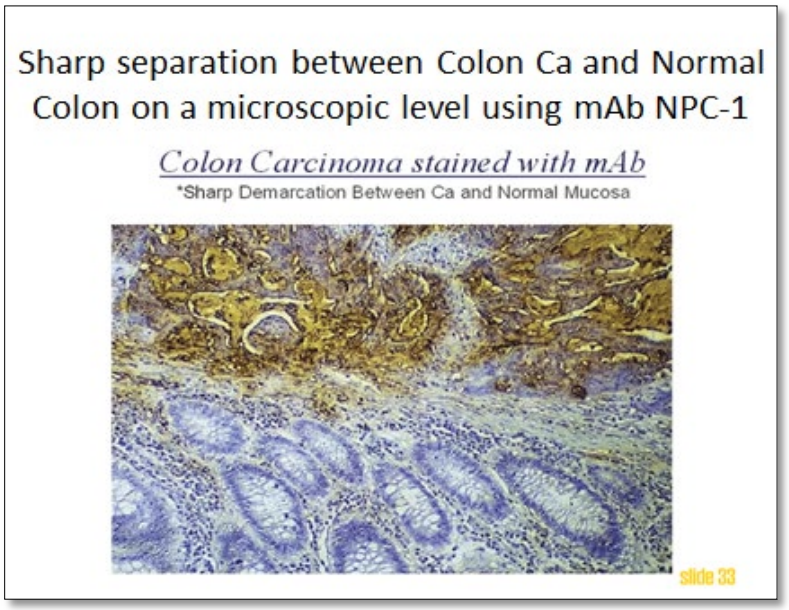

Figure 2. Colon carcinoma biopsy showing adjacent normal tissue. Tissue stained with Neo 102 identifies the mutated oncofetal protein expressed by the lesion (brown stain). No abnormalities are noted in adjacent normal bowel mucosa.

Because of the extremely low level of immunogen (Tumor Associated Antigen-TAA) expression within any tumor system however, the host virtually bypasses recognition of those cells expressing sub threshold levels of the immunogen. This allows for uninterrupted growth of the tumor. At best tumor surveillance may occur via the CD8 cells, wherein the host can identify and destroy several of the criminal cells in a population of thousands of such malignant cells. More frequently however it can be designed to recognize certain populations of malignant B cells. Prehn and Main [9] had originally suggested that in order to induce immune recognition, one must pool the tumor immunogen to reach its threshold level of activity needed and assure specificity in recognition of the system. Experiments performed in his lab confirmed the necessity of reaching so called threshold levels of immunogen before recognition could be established.

It is also of interest that in defining the presence of that oncogene characterizing the malignancy that the primary and metastatic lesion expresses the same unaltered immunogen. As such, if we were to have treated a large primary colon lesion by immuno chemotherapy, only to find at a later date that a metastatic lesion is now identified, the therapeutic mAb to be employed would remain unchanged from that having been expressed in the primary lesion.

The important antigens (immunogens) characterizing an array of malignant lesions was defined by Hollinshead [10] while present in their crude form as pooled allogeneic proteins. This pooled material was sub fractionated by molecular weight (MW) employing a Sephadex gel column. This was followed by discontinuous polyacrilamide gel electrophoresis. The components of the separation process were tested by delayed cutaneous skin hypersensitivity (DHR) in patients with the specific tumor of interest, patients with different malignancies and eventually normal volunteers. For colon cancer, an antigen of approximately 60-80 kDa was defined as the one producing the indicated immune response. In testing for the initiation of immune reactivity, doses of approximately 500-750 µg of partially purified antigen were tested and found to be the necessary therapeutic level needed to turn on an effective immune reactive response. Obviously these levels were not produced by the primary malignancy for host recognition.

FDA suspected that even though the final antigen preparations we were to use for therapy appeared to represent a single band on isoelectrophoresis, that the protein composition was more complex and probably represented several antigens migrating to the same region on the gel. As such it was felt that a genetically engineered preparation would be required for use in presumed to be/effective clinical trials. Monoclonal antibodies (mAbs) were developed to target the antigen used in the study. Obviously, several mAbs were obtained from hybridomas so produced. To clarify the antigen composition derived from the pooled allogeneic preparation, HPLC was performed and as suspected, several bands could be defined. These were matched against the antibodies being developed in our lab utilizing a crude antigen preparation **(Figure 3)**.

Antigens so obtained utilizing the Hollinshead [10] process of separation were employed in the initial clinical trials treating patients with recurrent disease or among those at high risk for recurrence. Peak 4, the more commonly expressed antigen, proved to be MUC5ac in a post translational modification from the innate oncofetal protein from which it was derived. Peaks 1 and 2 were the second most commonly identified antigens either seen alone or in combination with the other 2 immunogens. Peak 3 while less commonly identified in the colorectal lesion was a potent antigen with an antibody derived from it that had an ADCC in the 70% range. Our major goal at this point was to identify and sequence these immunogenic molecules and to define their levels of expression in the neoplastic disease process [11].

We were now able to demonstrate an effective host response to existing tumor. This occurred following the immunization process with the Hollinshead [10] pooled antigen, when the threshold level was reached for immune recognition.

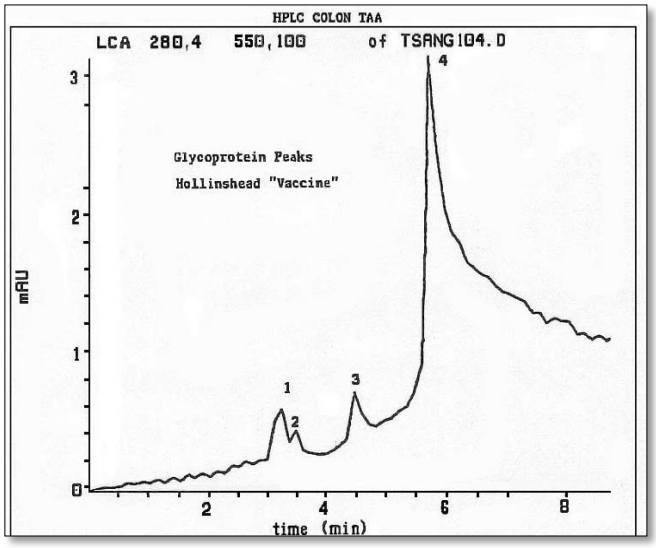

Figure 3. HPLC of the pooled allogeneic antigen that had later been purified by Sephadex G200 followed by isoelectric focusing.

The antigens as noted above were employed for induction of antibody expression thru hybridoma formation. Those antibodies specific for colon cancer and for what appeared to be selected GI malignancies including pancreatic cancer, demonstrated absence of cross reactivity to normal tissue. They showed an excellent capability for defining the selected tumor by immunohistochemistry (IHC) and serum ELISA. In addition they demonstrated strong (antibody dependent cell cytotoxicity) ADCC directed against the

neoplasm of interest [12]. Here, we were able to see for the first time, indications that the ideal tumor monoclonal antibody has two distinct functions, one that could diagnose the presence and nature of the tumor and at this point serve as a follow up therapeutic agent to destroy the tumor [13].

Antibodies directed against the tumor were now employed for identification of the specific antigen inducing the immune response as suggested by FDA. Affinity purification followed by mass spectroscopy was utilized in purification of the antigen. In each situation the antigen appeared to represent a modified oncofetal protein and not a so called surface molecule such as epidermal growth factor 1 or the vascular targets brought under temporary control by antibodies such as Avastin.

The most common of the immunogenic proteins isolated from malignancies of the bowel were related to mucin proteins expressed by those genes activated in the fetal state for the purpose of mucin production in the GI tract and pulmonary system. Of these MUC5ac was the most active of the proteins that were found. In the malignant state, while the first set of proteins were related to the MUC5ac protein, in actuality, in the tumor state they represented post translational modifications in structure of the original molecule [14]. As such, those commercial antibodies targeting MUC5ac protein found in cystic fibrosis did not cross react with tumor antigen modification and similarly those antibodies that we developed against the MUC5ac tumor target did not react with that protein expressed in cystic fibrosis.

THE ONCOFETAL PROTEINS THAT DEFINE COLORECTAL CARCINOMA

MUC5ac antigen

The MUC5AC mucins themselves are high molecular weight glycoproteins with O-linked oligosaccharides attached to serine or threonine residues of the apomucin protein backbone that is expressed in a cellular as well as a tissue-specific pattern in normal tissues. This mucin family includes proteins that contain tandem repeat structures with a high proportion of prolines, threonines and serines (which constitute the PTS domain) in the adenomatous lesion **(Figure 4)** [15].

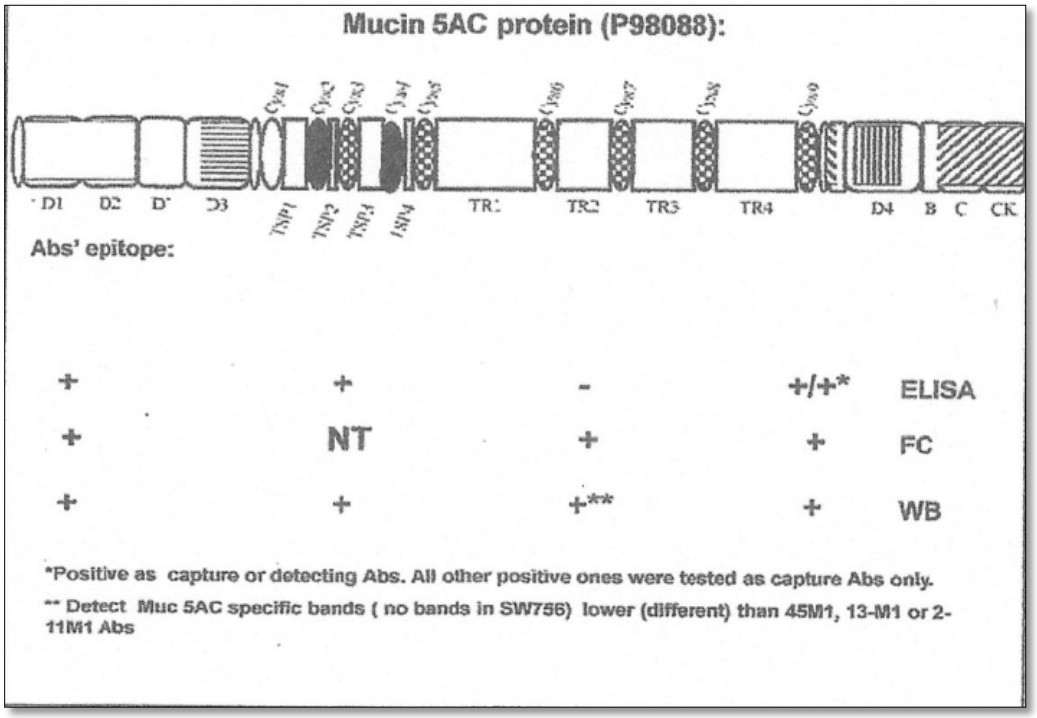

Figure 4. Reveals the various epitope regions defined by the corresponding/associated monoclonal antibody.

Additionally, the MUC5AC antigen in its modified form is expressed in tumors of gastrointestinal, pancreas-biliary and endocervical origin (e.g. colon, esophagus, liver, lung, pancreas, stomach and uterus). As the fetus matures prior to delivery, the gene for MUC5as is remethylated. Failure to suppress the function of this molecule with unrestricted expression of mucin will result in cystic fibrosis as the baby is born.

Later in adult life, either by viral or carcinogenic activity on the cellular nuclear components, the gene is reactivated but this time the process results in a post translational variant of the MUC5ac antigen. This new molecule, while present at small levels in the tumor cell, is extremely immunogenic. Antibodies targeting it in the tumor are diagnostically as well as therapeutically active. When challenging innate MUC5ac from fetal tissue with the tumor mAb, there is no

activity and antibodies derived from normally produced MUC5ac have no effect on the post translational modification of the newer antigen, now considered to be the tumor immunogen.

The study we performed, provided the peptidomimetics of an NPC-1 (commercial term for the mutated antigen) epitope derived from MUC5AC, including composition comprising the amino acid sequence of the epitope binding site defined by Phage display. Immunization of an animal with a fusion protein comprising a polypeptide of F(PHE) P(PRO) E(GLU) D(ASP) Y(TYR) F(PHE) R(ARG) Y(TYR) T(THR) N(ASN) Q((GLN) K(GLY) followed fusing a spleen cell with a myeloma cell, culturing these post-fusion cells and hybridoma selection indicated that the peptide vaccines could function to induce a needed immune response. Prior to use in patients post their surgical resection, the various polypeptides defined will be studied for optimum immune reactivity, i.e., FPEDYFRYTNQK or SLPDDWFRYINY.

While it is apparent that in treating cancer patients such as those colon or pancreatic cancers that have metastasized and have failed all known therapeutic modalities, they have a life span limited to weeks, we have found that an infusion of antibody can improve survival. This antitumor response can be seen in a matter of hours of the antibody delivery. When immunization with a vaccine occurs, it requires several months until the proper level of serum antibody results and is effective in controlling reappearance of tumor post-surgery. In those patients originally treated with the Hollinshead vaccine, significant levels of serum mAb could be detected 20 or more years post immunization with partially purified vaccine.

When Milstein began studying expression of fused lymphocytes, the latter derivative (that of the two fused lymphocytes) while possibly producing two different antibodies did not have long term survival in culture. Fusing the desired lymphocyte with a myeloma cell (malignant B cell) produced the needed response, a hybrid cell that was immortal (a hybridoma producing a monoclonal antibody). The oma suffix referred to the malignant heritage of the cell.

By removing said animal's spleen and preparing a single cell suspension followed by fusing of the spleen cell with a myeloma cell, the process resulted in a functional hybridoma. Culturing post-fusion cells in hybridoma selection medium and culturing the resultant hybridomas, is then followed by screening for specific activity.

NPC-IC expression system

The chimeric NPC-IC antibody was previously cloned into the pDHFR plasmid which contains the dhfr gene to allow for transgene amplification and selection of stable CHODG44 cells producing very high levels of the antibody. The vector contains expression elements from the CMV promoter and BGH polyadenylation signal to drive transgene expression and SV40 expression elements to drive dhfr gene expression. A diagram of the final constructs is shown below **(Figure 5)** and describes the two possible orientations for transgene expression from the two CMV promoters.

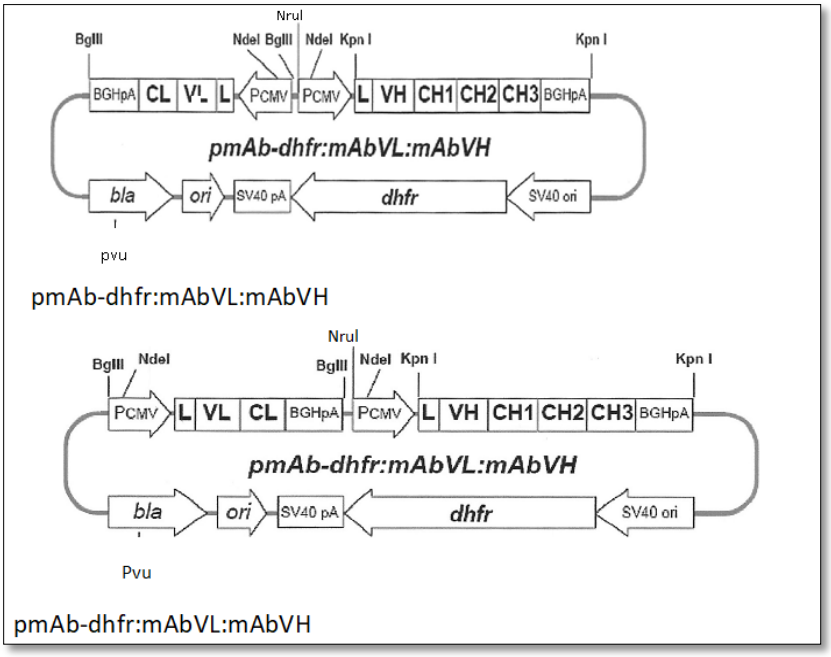

Figure 5. The pDI-IFR plasmid was purchased from BioFactura, the contract manufacturer employed for the development of the stable CHO cell clones and manufacturing of the clinical-grade Phase I-Il antibody material.

NPC-IC-CHO cell line development

An important strategic goal is the development of stable CHODG44 cells expressing very high levels of very high quality NPC-IC. These cells would represent the production cells for future manufacturing campaigns of the NPC-I C antibody for use in pre-clinical studies and clinical trials and potentially for commercialization. The **Figure 6** shows approximately the steps involved in achieving that goal. The development of such cell lines can take 45 weeks or more.

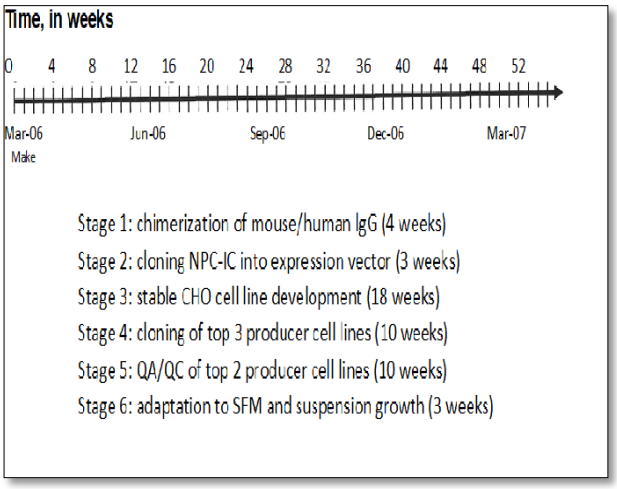

Figure 6. Steps involved in development of cell lines.

Currently, the NPC-IC expression plasmids have been stably transfected into CHODG44 cells using lipofectamine and the initial selection of cells in selection medium has been completed. Cells were chosen for further development based on cell growth characteristics and NPC-IC expression levels, as determined by specific productivity rate (SPR) assays.

After the first phase of selection (in stage 3 in the diagram above), there are approximately 60 cell cultures with SPR values of 0.3 to 1.4 pg/cell/day. Approximately 10 of these cell cultures will be chosen for the next phase of development, which includes amplification ofNPC-1C expression with increasing concentrations of methotrexate (Mtx). The concentration of Mtx in the cell growth medium currently ranges from 30100 nM and the final concentration of Mtx is expected to range from 1-10 uM following amplification. It is expected that NPC-IC expression may increase to 5-20 pg/cell/day by SPR the assay. This range of expression levels represents approximately 500-2000 mg/L/production run assuming a 10 day campaign, which is essential for the cost of antibody production **(Figure 7)**.

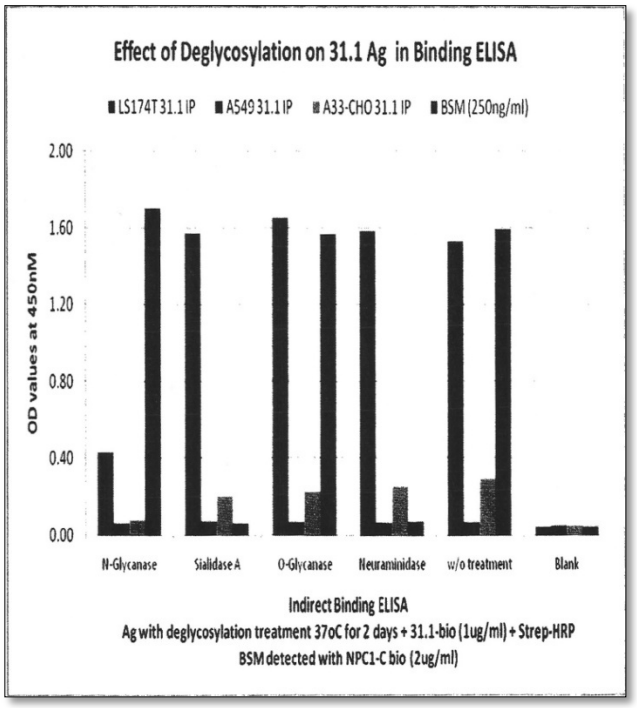

Figure 7. Antigen binding: This binding process for 31.1 holds true for the MUC5ac as well as other oncofetal proteins.

An initial study of the anti-tumor effect of NPC-1 using the antitumor model LS174T clearly shows the difference in effect on tumor growth when comparing saline IgG and NPC-1 **(Figure 8)**.

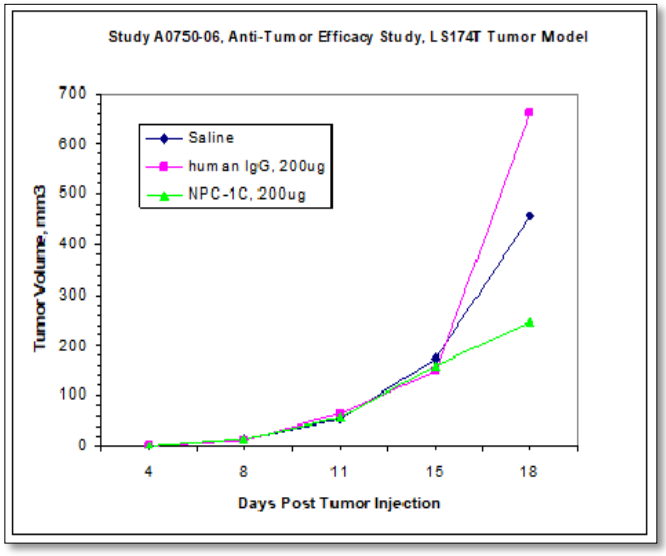

Figure 8: Shows toxicology.

Toxicology

In-life parameters To begin exploring its non-clinical toxicity, a preliminary non-GLP study using a research-

grade preparation of NPC-1C was conducted at an independent contractor, BioCon, Inc. Normal BALB/c mice were injected with a single IV dose of saline or 3, 10, 30 or 100 mg/kg of NPC-1C (n=3 female mice per group). Measures included body weights and clinical observations. Mice were humanely sacrificed 72 h following the injection and specimens were collected for analysis. Post-mortem parameters included macroscopic examination, blood cell counts, serum chemistries and histopathological evaluation of selected major organs and tissues. The results of the preliminary study demonstrated no significant changes in body weight, blood cell counts and histopathology of 7 major organs and tissues. The single remarkable finding was a mild elevation of serum aspartate transaminase (AST) observed in 2 out of 3 mice that received 100 mg/kg of NPC-1C. There were no other toxicities measured in the study, including by histopathological examination of the major organ systems in these mice. The biodistribution of the antibody is clearly seen in the comparison of different organs with that seen in the LS174T colorectal lesion. Clinically this was felt to be important in the antibody delivered IC, even going predominantly to the tumor was in essence not distributing any significant portion of the therapeutic pro-cut of other organs and as such diminishing what eventually resided within the tumor to initiate the process of ADCC. Apoptosis via Annexin V binding occurred to a much smaller extent **(Figure 9)**.

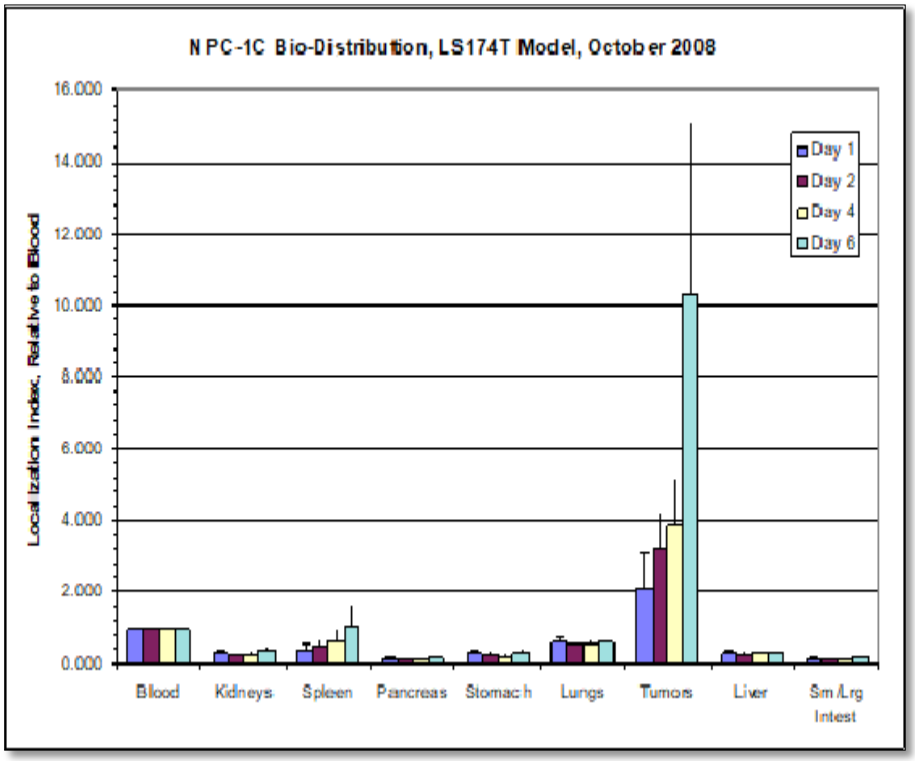

Figure 9. Considering that the level of Neo 101 needed for IV therapy in ongoing trials was in the hundreds of mg. it was decided to test different expression vectors for enhanced production of the monoclonals needed.

One such Vector licensed from Sellexis - more than quadrupled the level of antibody produced per liter of reactor fluid. When tested for activity, this new Biosimilar product appeared to produce better responses in terms of enhancing ADCC and eliminating the mild hemolysis found with mAb Neo 101. With this newer variant targeting MUC5ac, mAb Neo102 was further characterized and then scaled up for testing. The antibody was delivered in a Phase I trial. It indicated lack of any of the forms of toxicity previously encountered and indications of an improved therapeutic response. This resulted, not in varying the expression of antibody sequencing, but we believe in the way the antibody molecule was glycosylated. Phase II further indicated the efficacy of employing such an enhanced therapeutic agent in patients with recurrent and metastatic tumor having failed all forms of therapy, both standard and experimental [16-18].

In 2015 at the ASCO GI Symposium in San Francisco CA, results of the Phase 1 study of Ensituximab (Neo-102) in chemotherapy refractory metastatic colorectal Cancer Study were presented in a poster session. The study revealed a Maximal Tolerated Dose of 3.0 mg/kg IV every 2 weeks. The overall survival observed in this study demonstrated 10.4 months comparing favorably to the historical control for a similar population of patients with advanced colorectal Ca (5 months). This led to a larger Phase 2 multi-center colorectal cancer study using Ensituximab in the same

patient population where results were confirmed. Because tumors do shed an inhibitory molecule into the serum blocking many of the effects of the anti-tumor IgG1 in its ADCC response, new trials are being designed to initiate an updated study with low dose chemotherapy being added to minimize circulatory inhibitory antigen (mutated mitochondrial DNA) that could diminish the antitumor effect of the therapeutic mAb. We have considered the eventual addition of IL-15 which can enhance ADCC thru further activation of the NK cell effect on the tumor cell. When all factors are in place we anticipate that there will be a marked improvement in the management of patients with recurrent colon and other cancers.

In addition to immunogenic expression of antigens in colorectal cancer depicted by the presence of a post translational modification of MUC5ac as the primary immunogenic antigen, the initial HPLC of Hollinshead [10] antigen revealed that there were at least 2 additional and important immunogenic proteins to consider for possible vaccine therapy. These are 16C3 (oncofetal) antigen which proved to be a variant of CEAcam-5,6 and 31.1 which was derived from the antigen present in colon cancer first noted by Olds, that is A33.

Nature of the 16C3 antigen

As previously seen in the HPLC analysis of pooled colon tumor membrane antigen, the second most common antigen that we could identify in the malignant lesion was a combination of both CEAcam-5 and CEAcam-6 which comprised additional examples of cancer-specific antigens (Tumor Immunogens) expressed in colorectal and pancreatic cancer. This particular carcinoembryonic antigen (CEA) gene family is a member of the Ig Cam superfamily which includes 29 related genes and pseudogenes. These CEA proteins function as intercellular hemophilic and lipid rafts in the apical membrane of the polarized epithelial cells. CEAcam5 and CEAcam6 bind to a variety of gram-negative bacteria and mediate internalization/phagocytosis, participating in innate immune defense within the intestine [19,20].

CEAcam5 and CEAcam6 are overexpressed in many cancers (e.g. breast, ovarian, colon, pancreatic, lung and prostate). CEACAM5 and CEACAM6 are believed to be involved in cell adhesion, cellular invasiveness, resistance to anoikis (a form of programmed cell death) and metastatic behavior of tumor cells [21,22].

The 16C3 antibody (Neo-201) that we developed to target the corresponding colon immunogen is defined in US Patent Application No. 2009/0162931, but the nature of the antigen targeted by the mAb was not described at the time. To identify the 16C3 antigen sequence, several protein purification processes were utilized using either murine or humanized 16C3 mAb. The tumor antigen sources for these protein extracts were tumor cell lines, including LS174 (human colorectal tumor), CFPAC-1 (human pancreatic tumor) and cancer vaccine from the Hollinshead [23] library of cancer vaccines.

16C3 antigen, the second most commonly identified oncofetal protein, acts at a suboptimal level as one of the common tumor antigens in colon cancer. It was defined by coupling the antibody derived from the crude form of 16C3 to resins for needed antigen purification. This included magnetic beads for simple adsorption, washing and then elution from the beads. Proteins obtained from the beads were studied for determination of antigen presence, followed by their characterization and further identification.

Those proteins extracted from colon tumor tissue (operative specimens) or derived from the malignant cell lines LS-174T, HT-29, AsPC-l and CFPAC-I and their pellet extracts, were separated by SDS-PAGE, transferred to PVDF membrane and then stained with the 16C3 antibody. The results demonstrated two distinct molecular mass species cross-reactive with the 16C3 antibody. These were estimated to be 100 kDa and 200 kDa in size. The relative ratios of the two immunoreactive bands have however, generally been observed to be different among colorectal and pancreatic tumor cell lines. In colorectal tumor cells, the 200 kd band is found to be the predominant species, whereas in pancreatic tumor cells the 100 kd band is the predominant species.

The 16C3 antigen was prepared for identification by mass spectrometry by running immunopurified antigen preparations from several different tumor sources on SDS-PAGE and then excising the 16C3 immunoreactive bands from the polyacrylamide gel to define a protein with MW 300 kDa. A second band from HT-29 corresponded to a protein with MW-200 kDa, a third band from CFPAC-I corresponding to a MW-100 kDa. The proteins were then subjected to trypsin digestion followed by LC/MS/MS on a LTQ ORBITRAP@ XL mass spectrometer (Thermo Scientific).

Production data were searched against the forward and reverse International Protein Identification (IPI) human database using the Mascot search engine (Matrix Scientific, Ltd.). The database was appended with commonly observed background proteins to prevent false assignments of peptides derived from those proteins. Mascot output files were parsed into the Scaffold program for filtering to assess false discovery rates and allow only correct protein identifications.

Considered together, the three mass spectrometry experiments demonstrated the presence of CEAcam5- and/or CEAcam6-derived peptides in the 16C3 immunopurified preparations. These preparations were made from human colorectal (LS 174T, HT•29) and pancreatic (CFPAC-I) tumor cell lines. The CEAcam5 and CEAcam6 derived peptides appeared to be most relevant since the molecular mass of these CEACAM species are 100 kDa (CEAcam6)

and 200 kDa (CEAcam5) and are expressed in colon tissue and have been shown to be over-expressed in many colon cancer tissue samples. Thus, these experiments suggested that the tumor associated antigen recognized by 16C3 antibody is an epitope of the CEAcam6 glycoprotein. The identity of CEAcam5 and CEAcam6 as the target antigens of the 16C3 antibody was confirmed by comparing the immunoreactivity of 16C3 antibody with commercially-available antibodies against CEAcam5 and CEAcam6. The flow cytometry results demonstrate that 16C3 staining of LS-174T, CFPAC-I HT-29 and H226 cells is similar to that observed with other antibodies known to react with CEAcam5 and CEAcam6. The H226 cell line was included as a cell specificity control since these squamous lung tumor cells do not react with the 16C3 antibody.

Expression of the 16C3 antigen was successfully knocked down using siRNAs homologous to human CEAcam5 and CEAcam6 in cells known to express the 16C3-immunoreactive antigen. Several siRNA oligonucleotides were designed based upon the CEAcam5 and CEAcam6 sequences reported in public databases by methods known in the art. The sequences of the human CEAcam5 and CEAcam6 siRNA oligonucleotides are shown in **Figure 10**.

Sequence of CEACAM5 and CEACAM6 siRNA oligonucleotides

Oligonucleotide	Strand	Sequence
siRNA (CEACAM5)	Sense	AGAACUCAGUGAGUGCAAAtt (SEQ ID NO: 126)
	Anti-Sense	UUUGCACUCACUGAGUUCU (SE ID NO: 127)
		:CUAtt (SEQ ID NO: 128)
siRNA (CEACAM6)	Sense Anti-Sense	GGAACGAUGCAGGAUCCUAtt UAGGAUCCUGCAUCGUUCCtt (SEQ ID NO: 129

Figure 10. Oligonucleitide sequencing.

Similarly, siRNAs modeled from CEAcam6 sequence were transfected into CFPAC-I pancreatic tumor cells. The data demonstrate that siRNA specific for the CEAcam6 molecule reduced the amount of 16C3-reactive antigen expressed following transfection into CFPAC-1 tumor cells. Three commercially available anti-CEAcam5 and anti-CEAcam6 antibodies were used as controls in these experiments. They are clone CB30 against CEA/CD66e (#2383, Cell Signaling Inc), 9A6 against CEAcam6 (#ab78029. Abcam), MUS against CEAcam5/6 (#ab4539, Abcam). The reduction of CEAcam6 expression, as detected by both 16C3 and anti-CEAcam5 and anti-CEAcam6 antibodies, but not anti-CEAcam5 antibody alone, was dependent on the amount of CEAcam6 specific siRNA transfected into the cells. Approximately 75% of CEAcam6 expression was inhibited in CFPAC-I cell lines at 100 picomoles of the siRNA. These results confirmed that the 16C3 antigen is likely present in the CEAcam6 protein.

The identity of CEAcam5 and CEAcam6 as including the 16C3 antigens recognized by the 16C3 antibody was tested by cloning the genes encoding CEAcam5 and CEAcam6 into mammalian expression plasmids, transfected into human 293T cells (293T cells do not to express either CEAcam, 5 or CEAcam6.) After the DNA plasmids encoding CEAcam5 and CEAcam6 were transfected into 293T cells, the recombinant expression of the antigen targets was tested in western blots using 16C3 antibody and commercially available antibodies against CEAcam5 and CEAcam6. They included clone CB30 against CEA/CD66e (#2383, Cell Signaling Including 9A6 against CEAcam6 (#ab78029 Abcam) or MUS against CEAcam5/6 (#ab4539, Abcam). These results show that the 16C3 antibody detects 16C3 in both CEAcam5 and CEAcam6 expressed as recombinant antigenic molecules. The control anti-CEAcam5 and anti-CEAcam6 antibodies demonstrated that 16C3 detects proteins of the same approximate molecular mass.

The 16C3 antigen was defined by coupling the antibody to resins for needed antigen purification; this included magnetic beads for simple adsorption, washing and then elution from the beads. Proteins obtained from the beads were studied for determination of antigen presence, characterization and identification.

Proteins extracted from colon tumor tissue, or derived from LS 174T, HT-29, ASPC-l and CFPAC-I tumor cell pellet extracts, were separated by SDS-PAGE, transferred to PVDF membrane and then stained with the 16C3 antibody. The results demonstrate two distinct molecular mass species cross-reactive with the 16C3 antibody, estimated to be 100 kDa and 200 kDa. The relative ratios of the two immunoreactive bands have generally been observed to be different among colorectal and pancreatic tumor cell lines. In colorectal tumor cells, the 200 kd band is the predominant species whereas in pancreatic tumor cells the 100 kd band is predominant.

Production data were searched against the concatenated forward and reverse International Protein Identification (IPI) human database using the Mascot search engine (Matrix Scientific, Ltd.). The database was appended with commonly observed background proteins to prevent false assignments of peptides derived from those proteins. Mascot output files were parsed into the Scaffold program for filtering to assess false discovery rates and allow only correct protein identifications.

The siRNA ID#:S2885 was transfected into human colorectal LS 174T whereas siRNA ID#:S9285 was

transfected into human pancreatic CFPAC-I tumor cells. Following transfection of the siRNA into tumor cells, the CEAcam5 and CEAcam6 expressed by the cells was measured by specific PCR to measure the levels of CEAcam5 and CEAcam6 mRNAs. Quantitative Western blot analysis (CEAcam5) or quantitative flow cytometry (CEAcam6), each using a 6C3 antibody measured the levels of 16C3-immunoreactive protein.

The quantitative western blot data demonstrates that siRNA specific for the CEAcam5 molecule reduced the amount of 16C3-reactive protein following transfection into LSI 74T cells. A commercially available anti-CEAcam5 antibody was used as a positive control in these experiments. The reduction of CEAcam5 expression, as detected by both the commercial and J6C3 antibodies was dependent on the amount of siRNA transfected into the cells. Approximately 70% of CEAcam5 expression was inhibited in LSI 74T cells at 100 pmol of the siRNA. These results confirmed that CEAcam5 comprises an epitope bound by the 16C3 antibody.

These results confirmed that the 16C3 antigen is likely present in the CEAcam6 protein. This second target antigen, the immunogenic protein characterizing colorectal and pancreatic cancer has a wider spectrum for attacking tumors. The immunogen is also found in adenocarcinoma of ovary, lung and triple negative metastatic breast cancer. By itself it can play an important role in the managing a malignancy defined by this immunogen. Should this target protein exist in combination with the MUC5ac oncofetal protein, the combination of both mAbs may prove in experimental studies and later in the clinical situation that an improved efficacy has been characterized by this approach.

Target protein

The third antibody developed as a byproduct of the Hollinshead vaccine was mAb 31.1 [23]. Of the three, this appeared to demonostrate a higher level of ADCC than the other two mAbs. While we initially spent the most effort in producing and testing the mAb first isolated as Neo 101, the one targeting Mutated MUC5ac because of its activity in metastatic colon and pancreas cancer we then shifted to 16C3 because of initiate activity in targeting colon, lung, ovary and triple negative metastatic breast Ca. In the time alloted to 31.1 it appeared that this mAb was equivalent to a post translational modification of the colon cancer protein described by Old, that is A33 [24,25].

The mAb to A33 detects a membrane antigen that is expressed in normal human colonic and small bowel epithelium and in >95% of human colon cancers. It is absent from most other human tissues and tumor types. The murine A33 mAb has been shown to target colon cancer in clinical trials and the therapeutic potential of a humanized antibody is currently being evaluated. The true immunogen of course has been shown to be a post translational modification of A33 as has been shown to be true for MUC5ac and CEAcam5,6. Using detergent extracts of the human colon carcinoma cell lines LIM1215 and SW1222, in which the antigen is highly expressed, the molecule was purified, yielding a 43 kDa protein. The N-terminal sequence was determined and further an internal peptide sequence was obtained following enzymatic cleavage. Degenerate primers were used in PCRs to produce a probe to screen a LIM1215 cDNA library, yielding clones that enabled us to deduce the complete amino acid sequence of the A33 antigen and express the protein. The available data bases have been searched and reveal no overall sequence similarities with known proteins. Based on a hydrophilicity plot, the A33 protein has three distinct structural domains: an extracellular region of 213 amino acids (which, by sequence alignment of conserved residues, contains two putative immunoglobulin-like domains), a single hydrophobic transmembrane domain and a highly polar intracellular tail containing four consecutive cysteine residues. These data indicate that the A33 antigen is a novel cell surface receptor or cell adhesion molecule in the immunoglobulin superfamily. It may be expressed alone or with additional immunogenic targets **(Figure 11)**.

In developing products for the diagnosis and therapy of cancer colorectal cancer is among the most common malignancies of the Western world and is a leading cause of cancer deaths. Because of the high resistance of micro disseminated colorectal carcinoma cells to conventional therapies, new treatment methods are needed. Targeted antibody-based therapy is an active area of clinical investigation. The production of less immunogenic humanized or chimeric antibodies to reduce the immune reaction in patients, allowing antibodies to be repeatedly administered and the identification of more suitable antigenic targets offer promising new avenues for exploration. One potentially useful target for colon cancer is the surface molecule that interacts with the mAb A33. The expression of this antigen is highly restricted to the epithelial cells of the human small and large intestines and 95% of human colorectal tumors. Upon mAb binding to the A33 antigen, the antibody-antigen complex is internalized and sequestered in vesicles, a property that may underlie the marked retention of radioactivity in tumors for up to 6 weeks after administration of I- 125- or 131-labeled A33. Other factors contributing to its therapeutic potential are the high number of A33 binding sites per cell and the absence of detectable circulating A33 antigen. Preclinical evaluation of A33 mAb as a therapeutic reagent has been explored in a nude mouse model using xenografts of the human colorectal carcinoma cell line, SW1222. These transplanted tumors undergo regression after treatment with either 125I- or 131I-labeled A33. In phase I and II therapy trials, 131I- and 125I-labeled murine A33 mAb were shown to have anti-tumor effects without bowel toxicity. In view of these encouraging

data, the mAb A33 was genetically humanized and phase I trials commenced.

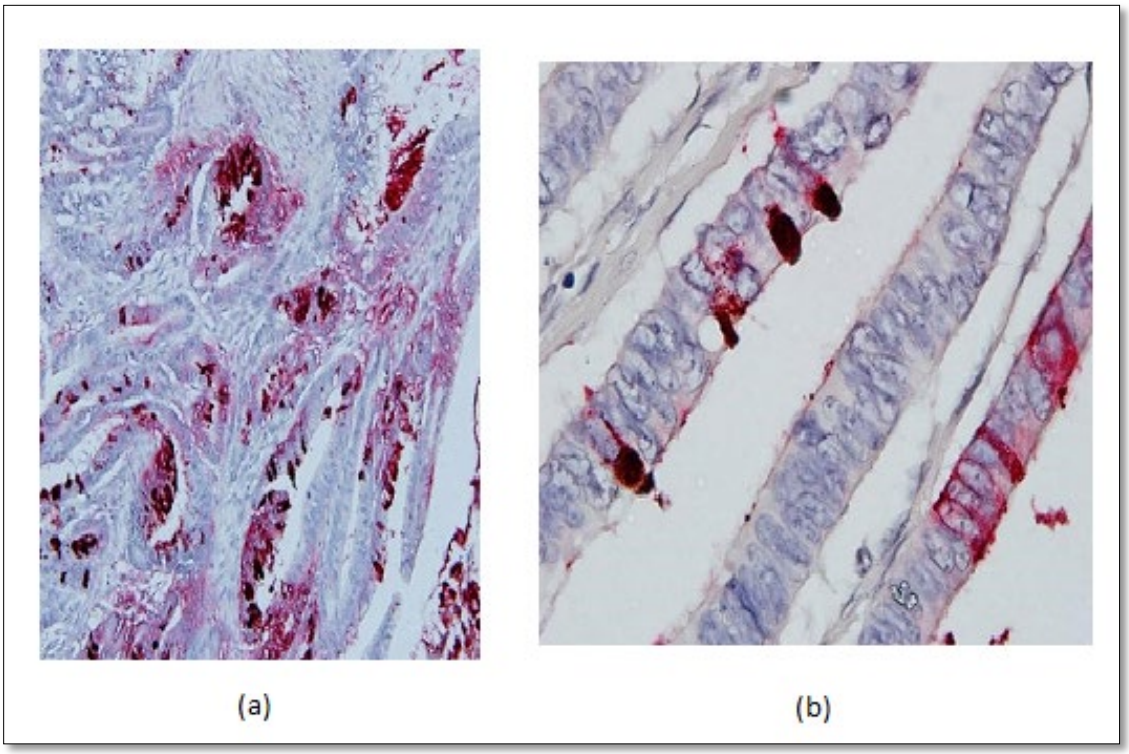

Figure 11. (a,b) Reveals use of a double stain technique to clarify colon cancer. The reddish stain specifically identifies colon cancer. NPC-1 is superimposed on the 31.1 stain. One can see in (b) that droplets of antigen are being shed into the lumen of the bowel where they can be detected by ELISA.

Using multidimensional high-resolution chromatography and biosensor detection, the A33 antigen was purified from Triton X-114 extracts of the human colon carcinoma cell line, LIM1215. This protocol yielded sufficient quantities of homogenous 40-45 kDa protein to permit N-terminal sequence analysis and identification of two tryptic peptides. The delineation of 33 N-terminal residues and subsequent analysis of available data bases suggested that the A33 antigen was a novel cell surface protein. Subsequently, affinity purification of the A33 antigen from SW1222 cells extended the N-terminal sequence characterization to 40 residues and identified an additional 101 residues of amino acid sequence from enzymatically cleaved peptide fragments (G.R., R.L.M., H.J., L.C., E.N., B.C., J.K.H., S.J.W., S.W., A.W.B., L.J.O. and R.J.S., unpublished data). The extended N-terminal amino acid sequence was used to design degenerate oligonucleotide primers for use in PCRs with a cDNA library constructed from LIM1215 poly(A)+ RNA. One of the amplified products encoded the N-terminal sequence of the A33 antigen and was used to isolate A33 antigen cDNAs from the LIM1215 cDNA library. In this paper we present the cDNA and deduced amino acid sequence of the A33 antigen, demonstrate expression of the protein on the surface of Cos cells and predict some of its structural characteristics. The A33 antigen sequence suggests a role for the A33 antigen in cell-cell recognition and signaling.

Cloning of A33 antigen cDNA

A LIM1215 cDNA library was custom-synthesized in the λZAPII expression vector by CLONTECH using poly(A)+ RNA prepared by us from confluent LIM1215 cells by two rounds of enrichment on columns of oligo(dT) cellulose. A cDNA probe encoding the A33 antigen N-terminal sequence was generated using PCR. Antisense primers were designed to hybridize to the carboxyl region of the N-terminal sequence obtained (unpublished data). Specifically, the six pools of 17-mer antisense oligonucleotides, each with 8-fold degeneracy, corresponding to amino acid residues 34-39 (LIQWDK) were as follows: primer 1477, 5′-A(AG)(CT)TT(AG)TCCCACTGAAT-3′; primer 1478, 5′-A(AG)(CT)TT(AG)TCCCATTGAAT-3′; primer 1479, 5′-A(AG)(CT)TT(AG)TCCCACTGGAT-3′; primer 1480, 5′-A(AG)(CT)TT(AG)TCCCATTGGAT-3′; primer 5915, 5′-A(AG)(CT)TT(AG)TCCCACTGTAT-3′; and primer 5916, 5′-A(AG)(CT)TT(AG)TCCCATTGTAT-3′.

These were paired with sense primers designed to hybridize to a primer sequence (KS; 5′-TCGAGGTCGACGGTATC) present in the backbone of the λZAPII vector. A 0.3 kb product obtained with primers KS and 1478 in a

"touchdown" PCR (11) with an initial primer-annealing temperature of 55°C was gel-purified, sequenced and found to encode a portion of the A33 N-terminal protein sequence. Precise primers for this A33 antigen cDNA sequence were then used (sense primer, 5′-CCTGTCTGGAGGCTGCCAGT; antisense primer, 5′-AGGTGCAGGGCAGGGTGACA) to amplify a 189-bp PCR product that was radiolabeled with [α-32P]ATP and [α-32P] CTP [both at 3000 Ci/mM (1 Ci=37 GBq); Bresatec, Adelaide, Australia] to a specific activity of $>10^7$ dpm/μg DNA using a Megaprime DNA labeling kit (Amersham). Clones (0.8×10^6) were screened according to the Stratagene instruction manual for λZAPII and 16 positive cDNA clones, ranging in size from 0.4 to 2.8 kb, were automatically excised from the λZAPII vector in the Bluescript plasmid, pBS SK(±), using the Lambda Zap Automatic Excision Process (Stratagene). Both strands of five independent clones were sequenced using an Applied Biosytems automated DNA sequencer. Initially, DNA sequence data were obtained using primers directed to the pBS backbone (KS, SK, T3, T7); the internal DNA sequence was then established by the specific primer-directed method. To establish the relationship of the A33 antigen sequence to known DNA sequences, four data bases [EMBL, GenBank, DDBJ (DNA Data Bank of Japan) and dbEST (data base of expressed sequence tags)] were searched using the data base similarity search algorithms, blast and fasta.

Northern blot analysis

For studies of A33 antigen mRNA expression, specimens of colorectal carcinoma (obtained during surgical resection) and confluent layers of colon carcinoma-derived cell lines (LIM1215, LIM1863, LIM1899, LIM2099, LIM2405 and LIM2437) [14] were directly solubilized in denaturing solution (4 M guanidinium isothiocyanate/0.5% Sarkosyl NL30 [BDH]/25 mM sodium citrate/0.1 M 2-mercaptoethanol) and total RNA was prepared according to the method of Kocer et al. [15]. Samples of normal human colon, counterparts of the tumor samples above and inflamed colon (from patients with Crohn disease) were first enriched for epithelial cells by incubating in PBS containing 3 mM EDTA and 0.5 mM DTT to release the crypts from the underlying stroma prior to solubilization in denaturing solution. Samples (20 μg) of total cytoplasmic RNA were electrophoresed in 0.4 M formaldehyde and 1% agarose gels and transferred overnight by capillary action to nylon filters (Hybond N; Amersham) and immobilized by exposure to UV light. Filters were prehybridized in a buffer containing 50% formamide, 5x Denhardt's solution (0.1% Ficoll/0.1% polyvinylpyrrolidone/0.1% BSA), 5x SSPE (1x SSPE=0.15 M NaCl/0.01 M NaH_2PO_4•$2H_2O$/1.2 mM EDTA, pH 7.4), 0.5% SDS and 1% (wt/vol) skimmed milk powder for at least 6 h at 42°C. The filters were then incubated at 42°C overnight in fresh hybridization solution with a 2.6 kb cDNA clone of A33 antigen, gel-purified and labeled with [α-32P]ATP (3000 Ci/mM; Bresatec) using a Megaprime DNA labeling kit. Filters were washed in 2x SSPE, 0.1% SDS followed by 1x SSPE and 0.1% SDS at 42°C and signals were visualized by autoradiography. To permit quantification of the mRNA signals, the filters were reprobed with an oligonucleotide specific for 18 S rRNA (5′-CGGCATGTATTAGCTCTAGAATTACCACAG), labeled with dATP[γ-32P] (3000 Ci/mmol; Bresatec) using T4 polynucleotide kinase.

Cos cell expression

A 2.6 kb putative A33 antigen cDNA clone was excised from the pBS (SK ±) plasmid using EcoRI and sub cloned in the sense orientation into the mammalian expression vector, pcDNA3 (InVitrogen, Leek, The Netherlands). Cos cells were seeded into 15 cm Petri dishes (Nunc) to achieve 50% confluency 24 h later. The cells were transfected over a 4 h incubation at 37°C with 15 μg of either pcDNA3/A33 or pcDNA3 (parental vector) using DEAE-dextran in the presence of chloroquine [17]. This was followed by dimethyl sulfoxide (DMSO) shock (10% DMSO in PBS) for 90 s; the cells were then returned to RPMI 1640 medium containing 10% fetal calf serum (FCS), 2 mM glutamine and 50 μg/ml gentamycin for 3-5 days. The cells were harvested and analyzed for A33 antigen expression by Western blot analysis, flow cytometry and immunocytochemistry.

Western blot analysis

Transfected cells were solubilized for 30 min at 4°C with 1% (vol/vol) Triton X-100 in 15 mM Tris•HCl (pH 7.4) containing 1 mM phenylmethylsulfonyl fluoride, 1 mM pepstatin, 0.1 mM leupeptin and 0.01 units per ml of aprotinin. The resulting extracts were centrifuged twice at 4°C for 20 min at 14,000x g and 2 μl aliquots were electrophoresed under nonreducing conditions in 8-25% SDS/PAGE Phastgels (Pharmacia) before they were transferred to poly(vinylidene difluoride) membranes and incubated with humanized mAb A33 (2 μg/ml). A33 signals were detected with anti-human IgG conjugated with horseradish peroxidase and visualized by enhanced chemiluminescence (Amersham).

Flow cytometry and immunochemistry

Transfected cells were detached using 10 mM EDTA in PBS for 10 min at 37°C and aspirated gently to produce a single cell suspension. They were then pelleted by centrifugation at 1500 rpm for 5 min and resuspended in 0.5 ml PBS containing 10 mM EDTA and 5% FCS and kept on ice for all remaining procedures to prevent internalization of antigen-antibody complexes. Murine A33 mAb was added to a final concentration of 20 μg/ml for 30 min. The cells were washed twice in PBS/EDTA/FCS and incubated on ice for a further 30 min in 0.5 ml PBS/EDTA/FCS containing fluorescein isothiocyanate-conjugated sheep anti-mouse IgG (Silenus, Hawthorn, Australia) diluted 1:50. Cells were then washed twice more and resuspended in 1 ml PBS/EDTA/FCS for fluorescence-activated cell sorting

(FACS) analysis in a Becton Dickinson FACScan and cytospin preparation. Cytospins were prepared in a Shandon cytospin 2 centrifuges (Shandon, Pittsburgh), allowed to air-dry, mounted in glycerol containing antifade and examined using a Nikon Fluorophot microscope.

RESULTS

Sixteen overlapping putative A33 antigen cDNA clones were obtained, the longest of which (clone 18) was 2793 bp prior to the addition of the poly(A)+ tail. The longest open reading frame in the cDNA sequence predicts a protein of 319 amino acids. Beginning at amino acid 22 of the predicted translation product, 40 contiguous residues are identical to the established amino-terminal sequence of native A33 antigen (GR, et al. unpublished data) [10]. The predicted sequence also contains regions (hatched boxes) identical with the amino acid sequences of several internal peptides released from the native molecule by enzymatic digestion

cDNA sequence and deduced amino acid sequence of the longest cDNA clone (clone 18; 2.8 kb) encoding the human A33 antigen. The longest open reading frame, encompassing nucleotides 345-1301, contains the known N-terminal sequence of the native A33 antigen and predicts a protein of 319 amino acids. The stop codon at 1302-1304 is boxed and the amino acid sequences of the internal peptides identified by digestion of the native molecule are shown (shaded areas). A putative signal sequence (bold underline), three potential N-linked glycosylation sites (overline) and a transmembrane domain (second bold underline) are indicated. Adjacent to the transmembrane domain, four consecutive cysteine residues are observed. The spans of the two putative Ig-like domains are enclosed by square brackets with the specific residues conserved in Ig superfamily members shown in circles. Other features of the DNA sequence include a tandem repeat of 25 bp in the 5′ untranslated region (bold overline) and a polyadenylylation signal (AATAAA) 11 bp upstream from the poly(A) tail. The asterisk above the C at position 294 denotes the fact that a C was found in this position in 2 out of 5 independent clones sequenced (including clone 18) and an A in the 3 other clones.

The predicted translation product of the human A33 antigen mRNA is initiated at the ATG positioned 345 bp from the 5′ end of clone 18. This ATG leads off the longest open reading frame and is in a favorable context for initiation of translation by reference to the Kozak consensus sequence, GCC(A/G)CCATGG. Following this, a sequence encoding 21 amino acids resembles a hydrophobic signal peptide. The putative cleavage site between alanine and isoleucine, which is required to produce the mature protein with the N-terminal sequence of the native molecule, is consistent with the (−3, −1) rule for signal peptide cleavage. The position of the first in-frame stop codon predicts a mature polypeptide chain comprising 298 amino acid residues, Mr 33,276. This Mr is not inconsistent with data demonstrating that the native A33 antigen is a glycoprotein of approximate Mr 40,000-45,000 because the sequence contains three potential N-linked glycosylation sites, one of which (N91) was strongly suspected from the initial sequence analysis of the peptide fragments. Each of these could be predicted to accommodate an average oligosaccharide chain length of 2.5-3 kDa. Indeed, in experiments where enzymes were used to remove sialic acid and both N- and O-linked glycosides from the A33 antigen polypeptide, there was a reduction of approximately 8000 in the Mr of the A33 antigen (unpublished data). Based on a Kyte-Doolittle hydrophilicity plot of the sequence, the molecule appears to have three structural domains: an extracellular region of 213 amino acids containing 6 cysteine residues, a hydrophobic transmembrane domain of 23 amino acids and a highly polar intracellular tail of 62 amino acids. Searches of available DNA and protein data bases revealed no direct sequence similarities with any known protein. However, relatively short spans of human A33 antigen nucleotide sequence could be matched with expressed sequence tags derived from the human colonic epithelial-derived cell line, T-84 (92% identity with GenBank accession no. AA055862; length, 344 nt) and the murine teratocarcinoma-derived cell line, F9 (74% identity with EMBL accession no. D28657D28657; length, 249 nt).

Kyte–Doolittle hydrophilicity plot of the deduced amino acid sequence of the A33 antigen. Several structural features are indicated: the presence of an N-terminal hydrophobic region likely to correspond to a signal sequence, a more distal highly hydrophobic sequence consistent with the presence of a single-span transmembrane domain, and a highly polar C-terminal region consistent with an intracellular location for this part of the molecule.

Manual inspection of the extracellular domain revealed the presence of several residues that are conserved in members of the Ig superfamily. Specifically, we noted a V-type Ig-like domain at the N terminus, in which the presence of a disulfide bond between the two conserved cysteines (C22, C96) was predicted from the microsequence analysis of the peptide fragments generated by enzymatic digestion. The V-type domain was followed by an Ig-like domain of the C2-type. This combination of a V-type and a C2-type domain is characteristic of the CD2 subgroup of Ig superfamily members. Sequence alignment to proteins in this subgroup with known three-dimensional structure revealed that the A33 antigen shares the highest sequence identity with the D1D2 fragment of human CD4 (16% over 208 residues) [22]. The V-type domain is most similar to the VL domains of antibodies (up to 22% sequence identity over 117 residues) and the C2-type domain to the D4 domain of rat CD4 (18% over 85 residues). The overall similarity with the CD2 group of molecules suggests that the A33 antigen may participate in cell-cell recognition processes, perhaps further implying the existence of a soluble or cell-associated

binding partner. The identification of non-antibody binding partners could be addressed using the same biosensor-based technology used recently to identify and purify both the A33 antigen and the ligand for the orphan receptor, HEK.

The N terminus of the predicted intracellular region of 62 amino acids begins with four consecutive cysteine residues. Data base analysis has identified 14 mature proteins containing the CCCC motif. Interestingly, three of these, G protein-coupled receptor 3, endothelial-1 receptor and the tachykinin-like peptide receptor, are members of the seven transmembrane G protein-coupled receptor family in which palmitoylation of the cysteine residues near the carboxyl terminus has been implicated in receptor coupling to G proteins and in the down-regulation of receptor activity by influencing receptor removal from the cell surface. Thus the presence of a CCCC motif adjacent to the transmembrane domain of the A33 antigen may indicate further membrane tethering of the molecule via palmitoylation and this could play a role in the trafficking of the protein to vesicles.

To verify that the clones we had isolated encoded the A33 antigen, we expressed a 2.6 kb cDNA transiently in Cos cells. The cells were then assayed for A33 antigen expression by Western blot analysis, flow cytometry and immunocytochemistry (data not shown). Only the Cos cells transfected with the expression vector containing a putative A33 antigen cDNA clone produced a protein that was recognized by the mAb A33. Furthermore, the Western blot analysis demonstrated that the recombinant A33 protein expressed by Cos cells was approximately the same Mr (40,000-45,000) as endogenously produced A33 antigen in LIM1215 cells. As expected, FACScan and immunocytochemical analysis of Cos cells transfected with the A33 antigen construct indicated that a large proportion of the A33 antigen was displayed on the cell surface.

Expression of recombinant A33 antigen by transfected Cos cells. Cos cells were transfected either with the parental vector, pcDNA3 or with pcDNA3 into which a 2.6 kb A33 antigen cDNA had been subcloned. Cells were harvested in this experiment 5 days after transfection and subjected to Western blot analysis (A) and flow cytometry (B). (A) Transfected cells were solubilized in Triton X-100 and electrophoresed into SDS/polyacrylamide gels without reduction and processed for Western blot analysis as described. The signal obtained with Cos cells transfected with pcDNA3 containing A33 antigen cDNA (lane 1) corresponds to approximately the same Mr (43,000) as the signal obtained with LIM1215 cells expressing high endogenous levels of A33 antigen (lane 3). Cells transfected with the parental vector gave no signal (lane 2). The positions of blue pre-stained standards (lane M) (NOVEX, San Diego) are indicated on the right. (B) FACScan profiles of control and A33 antigen-expressing Cos cells. The profile obtained with A33 antigen-expressing Cos cells (shown in bold) has been superimposed on the profile obtained with Cos cells transfected with pcDNA3 alone to show the shift to the right in fluorescence intensity of cells expressing A33 antigen.

Northern blot analysis demonstrated a strong A33 antigen signal (2.8 kb) in total RNA prepared from A33 antigen positive human colon carcinoma-derived cell lines (LIM1215, LIM1899 and LIM1863). No signal was obtained with total RNA from A33 antigen negative cell lines (LIM2099, LIM2405 and LIM2537; **Figure 4A**). Strong expression of A33 antigen mRNA in purified epithelial cells from normal human colon was always observed. In comparison, the A33 antigen mRNA signals obtained with RNA extracted from the adjacent tumor tissue were consistently weaker. The Northern blot analysis has been extended to more than 20 paired samples of normal and transformed colonic tissue with the same results. One explanation for this could be the higher fibroblast content of the tumor samples compared with the normal colonic crypt preparations, which are essentially pure epithelial cells.

Six colorectal carcinoma cell lines had been previously analyzed for A33 antigen expression using immunocytochemistry and flow cytometry (unpublished data). Half (LIM1215, LIM1863 and LIM1899) were found to be positive. The remaining half (LIM2099, LIM2405 and LIM2537) were negative. The pattern of A33 antigen mRNA expression was consistent with protein expression data. A33 antigen mRNA expression was found in samples of normal and diseased human colorectal tissue obtained from patients during surgical resection. In colonic mucosa, strong hybridization signals do correspond to relatively high expression of A33 antigen compared with that in the corresponding tumor preparations and the RNA extracted from the large polyp which all produced weaker A33 antigen mRNA signals. Crohn disease was found to produce an A33 antigen signal similar to that of its normal counterpart.

Immunohistochemical studies have demonstrated that the expression of the A33 antigen is essentially restricted to normal intestinal epithelium and 95% of colorectal tumors. Although several other mAbs capable of recognizing determinants on colon cancer cells exist, none of them matches the restricted tissue specificity of the A33 mAb. Ep-CAM (epithelial cell adhesion molecule) is a cell adhesion molecule expressed by most simple epithelia and its expression is maintained by a large proportion of colorectal, breast, lung, pancreatic and gastric tumors. Several mAbs to this tumor marker have been used in clinical trials and recently chimeric and single-chain versions of two of them, 17-1A and 323/A3, were developed and evaluated in cell-killing experiments. Similarly, members of the carcinoembryonic antigen family and the pancarcinoma antigen, TAG-72, have attracted attention as suitable targets for immunotherapy of colorectal, breast and lung tumors. Though their relatively wide expression offers the opportunity to develop antibodies that can target multiple

tumors, experience suggests that optimal localization of antibody to tumors may be impeded by the presence of antigen in a variety of normal tissues as well as by the presence of shed antigen in the circulation. In this respect, it might also have been expected that the strong expression of the A33 antigen in normal colon would have compromised its usefulness as a target in immunotherapy. However, these concerns have been allayed by phase I/II studies that have shown prolonged accumulation of 131I- and 125I-labeled A33 at sites of tumor metastasis for up to 6 weeks after administration, while normal colon was antibody-free by 7-8 days and minimal toxicity to the colonic mucosa was observed. The rapid clearance of the antibody from normal colon is not fully understood, but it may be due to rapid transcytosis of radiolabeled A33 mAb from colonocytes and/or the relatively rapid shedding of colonocytes from the top of the crypts. A basic understanding of the mechanisms underlying the tightly regulated, tissue-specific nature of A33 antigen expression as well as insights into its biological function is eagerly awaited. Such insights are likely to advance our understanding of cell-cell interactions and differentiation in the colon. In particular, knowledge of the characteristics of the A33 antigen gene promoter that confers such tissue-specific expression should allow us to develop transgenic mouse models of colon cancer, which could cast light on the role of certain oncogenes and tumor suppressor genes in the tumorigenic process. Indeed, it is conceivable that the tissue specificity of the A33 antigen promoter could also be exploited to target therapeutic genes to the colon.

Clearly the A33 antigen in its post translational form is an exciting target for immunotherapeutic approaches to colon cancer and our identification of the A33 antigen with its receptor-like structure points toward a new signaling system in the colon. The availability of the A33 antigen should facilitate the development of new immunotherapeutic and ligand-directed approaches to the treatment of metastatic colon cancer.

CONCLUSION

With all three of the oncofetal proteins defining colorectal cancer having been characterized, a more rational approach to the immunotherapy of advanced and metastatic colorectal cancer becomes defined. Depending on what is determined from immunohistochemistry studies of the primary lesion, a protocol of one mAb or combinations of those mAbs defining the 3 Immunogens can be devised **(Figure 11)**. As noted previously, when one defines the existence of a metastatic lesion at a time interval following treatment of the primary, it is not necessary to have a biopsy of the metastasis since the original primary contains unaltered immunogenic proteins expanding up the ladder from the original premalignant cell line, to the primary tumor in-situ, to the appearance of the functional lesion and under certain circumstances, to the metastatic lesion. Thus immunotherapy of the metastasis is based on the nature of the antigen expressed in the primary growth which can be employed as a vaccine [26]. The underlying immunogen that characterizes the primary lesion does not mutate as the lesion progresses. This is in contrast to many surface antigens that are eventually unstable and eventually result in a mutated variant.

Of importance to remember is that a good monoclonal antibody [27] is one that is extremely specific as well as effective in diagnosis and once the latter is established, the same monoclonal can demonstrate a strong immunotherapeutic approach.

REFERENCES

1. Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D, et al. (2000) Molecular Cell Biology. (4th Edn), W.H. Freeman, New York.
2. Lochhead P, Chan AT, Nishihara R, Fuch CS, Beck AH, et al. (2015) Etiologic field effect: Reappraisal of the field effect concept in cancer predisposition and progression. Mod Pathol 28: 14-29.
3. Wepsic HT (1983) Overview of oncofetal antigens in cancer. Ann Clin Lab Sci 13: 261-266.
4. Knudson AG (1971) Mutation and cancer: Statistical study of retinoblastoma. Proc Natl Acad Sci U S A 68: 820-823.
5. Arlen M, Hollinshead AC, Tsang KY (1993) Identification and characterization of a colon TAA. Ann N Y Acad Sci.
6. Arlen M, Molmenti E (2013) Identification of the oncofetal tumor antigens that characterize colorectal carcinoma for use in diagnosis and therapy. J Gastroenterol Oncol 10: 156-169.
7. Meighly MRB, Hall C (1987) Anastomotic recurrence of colorectal cancer - A biologic phenomenon or an avoidable calamity. Gut 28: 786-791.
8. Brouxhon S, Kyrkanides S, Teng M, Atha M, Ghazizadeh S, et al. (2014) Soluble E-cadherin: A critical oncogene modulating receptor tyrosine kinases, MAPK and PI3K/Akt/mTOR signaling. Oncogene 33: 225-235.
9. Prehn RT, Main JM (1957) Immunity to methylcholanthrene-induced sarcomas. J Natl Cancer Inst 18: 769-778.
10. Hollinshead AC, Elias EG, Arlen M, Buda B, Mosley M, et al. (1985) Specific active immunotherapy in patients with adenocarcinoma of the colon utilizing tumor-associated antigens (TAA) - A phase I clinical trial. Cancer 56: 480-489.
11. Elias EG, Hollinshead AC, Arlen M, Mosely MS (1985) Adjuvant-specific active immunotherapy in patients

with colon adenocarcinoma utilizing polypeptide tumor associated antigens (TAA): Proceedings. Am Soc Clin Oncol 4: 76.

12. Katz P, Simone CB, Pierre A, Faucci AS (1980) Mechanisms of antibody dependent cellular cytotoxicity. J Clin Invest 65: 55-83.
13. Arlen M, Arlen P, Bristol A, Luka J, Kantor J, et al. (2010) The dual functionality of monoclonal antibodies derived from tumor associated antigens. J Surg Oncol.
14. Bara J, Chastre E, Mahiou J, Singh RL, Forgue-Lafitte ME, et al. (1998) Gastric M1 mucin, an early oncofetal marker of colon carcinogenesis, is encoded by the MUC5AC gene. Int J Cancer 75: 767-773.
15. Kocer B, McKotanis J, Soran A (2006) Humoral immune response to MUC5AC in patients with colorectal polyps and colorectal carcinoma. BMC Gastroenterol 4.
16. Arlen M, Tsang KY (1993) Monoclonal antibodies and their role in modulation of the immume system. J Surg Oncol 54: 103-108.
17. Arlen M, Tsang KY, Bartal A, Wolf J, Saric O (1991) Monoclonal antibodies to immunoreactive tumor associated antigen (TAA) from human colon carcinoma. J Antibody Immunoconjugates Radiopharmaceutical 4: 895-905.
18. Arlen M, Arlen P, Bristol A, Wang X, Luka J, et al. (2012) The use of specific monoclonal antibodies to target immunogenic tumor proteins in patients with recurrent pancreatic and colorectal cancer. Curr Drug Deliv 9: 52-56.
19. Kolla V, Gonzales LW, Bailey NA, Wang P, Angampalli S, et al. (2009) Carcinoembryonic cell adhesion molecule 6 in human lung: Regulated expression of a multifunctional type II cell protein. Am J Physiol 296: 1019-1030.
20. Lund H (2003) Cancer. Gene Ther 10: 365-376.
21. Zhang S, Zhang HS, Cordon-Cardo C, Ragupathi G, Livingston PO (1998) Selection of tumor antigens as targets for immune attack using immunohistochemistry: Protein antigens1'2. Clin Cancer Res 4: 2669-2676.
22. Blumenthal RD, Hansen E, Goldenberg DM (2007) Expression patterns of CEACAM5 and CEACAM6 in primary and metastatic cancers. Cancer Res 65: 8809-8817.
23. Hollinshead AC, Elias G, Arlen M, Mosely MS, Scherrer J (1985) Specific active immunotherapy in patients with adenocarcinoma of the colon using TAA. Cancer 56: 480-489.
24. Old LJ, Boyse EA (1964) Immunology of experimental tumors. Annu Rev Med 15: 167-186.
25. Heath JK, White SJ, Johnstone CN, Catimel B, Simpson RJ, et al. (1997) The human A33 antigen is a transmembrane glycoprotein and a novel member of the immunoglobulin superfamily Proc Natl Acad Sci U S A 94: 469-474.
26. Arlen M, Arlen P, Tsang KY, Coppa G, Conte C, et al. (2017) The biologic nature and function of cancer vaccines. Vaccines 1: 23-30.
27. Arlen M, Arlen P, Coppa G, Tsang KY, Conte C, et al. (2017) Molecular mechanisms involved in the immune control of the malignant disease process. J Mol Immunol 2: 1-7.