HETEROLOGOUS EXPRESSION AND PURIFICATION OF CELL JUNCTION COMPONENTS - AN EFFORT TOWARDS DEVELOPING AN ANTIGEN-CAPTURE ELISA DIAGNOSTICS FOR METASTATIC CANCERS

by

Michael Irvine

A Thesis Submitted to the Faculty of

The Charles E. Schmidt College of Biomedical Science

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

Florida Atlantic University

Boca Raton, Florida

December 2009

HETEROLOGOUS EXPRESSION AND PURIFICATION OF CELL JUNCTION COMPONENTS - AN EFFORT TOWARDS DEVELOPING AN ANTIGEN-CAPTURE ELISA DIAGNOSTICS FOR METASTATIC CANCERS

by

Michael Irvine

This thesis was prepared under the direction of the candidate's thesis advisor, Dr. Kasirajan Ayyanathan, Department of Biological Sciences, and has been approved by the members of his supervisory committee. It was submitted to the faculty of the Charles E. Schmidt College of Biomedical Science and was accepted in partial fulfillment of the requirements for the degree of Master of Science.

SUPERVISORY COMMITTEE:

Kasirajan Ayyanathan, Ph.D.
Thesis Advisor

Howard Prentice, Ph.D.

Michael Lu, Ph.D.

Keith Brew, Ph.D.
Chair, Department of Basic Science

Michael Friedland, M.D.
Dean, The Charles E. Schmidt
College of Biomedical Science

Barry T. Rosson, Ph.D.
Dean, Graduate College

Dec. 1, 2009
Date

ABSTRACT

Author:	Michael Irvine
Title:	Heterologous Expression And Purification Of Cell Junction Components - An Effort Towards Developing An Antigen-Capture ELISA Diagnostics For Metastatic Cancers
Institution:	Florida Atlantic University
Thesis Advisor:	Dr. Kasirajan Ayyanathan
Degree:	Master of Science
Year:	2009

Metastatic cancers are problematic because they spread throughout the body. A crucial step in cancer metastasis is the separation of the cancer cells from their surrounding normal cells. This occurs due to suppression or destruction of cell adhesion molecules such as E-cadherin, occludin, and various claudins. The Snail and Slug transcription factors play a direct role in suppressing these cell adhesion molecules through their SNAG repression domain. We explored the possibility of developing an ELISA diagnostics capable of detecting soluble E-cadherin, occludin, and claudin fragments in the serum of cancer patients. Using several bioinformatics tools, unique extracellular antigenic sequences were identified on claudins-1, 4, 16, occludin, and E-cadherin. These sequences were cloned as GST fusion proteins, expressed, and purified

in large quantities to raise antibodies. In parallel, expression profiling of metastatic cancer cell lines was carried out to derive a correlation between Snail-Slug expression and suppression of cell adhesion molecules.

HETEROLOGOUS EXPRESSION AND PURIFICATION OF CELL JUNCTION COMPONENTS - AN EFFORT TOWARDS DEVELOPING AN ANTIGEN-CAPTURE ELISA DIAGNOSTICS FOR METASTATIC CANCERS

List of Tables

LIST OF FIGURES

INTRODUCTION

Metastatic Cancers

Metastatic cancers are the most problematic form of cancer due to their ability to spread rapidly throughout the body. This process occurs when cancer cells are able to leave the tissue they originated in and enter the bloodstream. From the bloodstream, the cancer cells are able to enter a new tissue and divide there as well. Metastasis is triggered at the genetic level by various genes, some of which are typically expressed in embryonic development. Invasion, a prerequisite to metastasis, occurs due to a loss of adhesion between adjacent cells.

Desmosomes and Tight Junctions

Normally, cells in a tissue are held together by cell junctions such as desmosomes and tight junctions. Desmosomes are present on the basolateral surface of a cell. These protein complexes are the primary support for a cell to adhere to other cells or the extracellular matrix. In healthy tissue, a desmosome or hemidesmosome should always remain intact. Tight junctions are important for cell-to-cell adhesion, but they have additional roles in preventing migration of integral membrane proteins from apical to basolateral sides of the cell as well as regulation of small molecule and ion flow through the

space between cells.

Cadherins are the adhesion proteins that attach one desmosome to another. E-cadherin, also known as CDH1, is a calcium dependent glycoprotein that is a member of the Cadherin super family of cell-to-cell adhesion molecules. Mutations in E-cadherin have been associated with the development of cancer metastasis (Onder et al., 2008). E-cadherin can also be repressed as a result of snail expression (Batlle et al., 2000).

Snail Transcription Factor

Snail is a zinc-finger transcription factor that is involved in epithelial-mesenchymal transitions. Up regulation of snail has been shown to inhibit the production of E-cadherin. This inverse correlation between snail and E-cadherin expression can be seen in many metastatic cancers such as those of the breast, colon, and pancreas (Batlle et al., 2000). It has also been shown that snail directly binds to the E-box on the promoter region of E-cadherin, preventing transcription from occurring (Ikenouchi et al., 2003). Due to the role of snail in controlling epithelial to mesenchymal transitions (Cano et al., 2000), it is safe to assume that it would be similarly over expressed in all cancers that are epithelial in origin. All human carcinomas arise from epithelial tissue (Watson, 1987), so the dynamics between snail and E-cadherin expression is surely important.

Claudins

Claudins also seem to play a role in cancer aggressiveness. These proteins are located in the tight junctions discussed earlier. Similar to E-

cadherin, they play a role in cell-to-cell adhesion. Unlike E-cadherin, there is no set pattern of expression for a particular claudin in all cancers (Soini, 2005). Although expression of a particular claudin may change from one tissue to another, there are claudins associated with specific cancers. For example, there seems to be an increase in claudin-1 expression in colon cancer (Dhawan et al., 2005). Based on this finding, one would think the expression of claudin-1 is undesirable in adult epithelial tissue. However, other researchers have found that claudin-1 expression was lost in invasive breast cancer cell lines (Tokes et al., 2005b). It has also been shown that decreased expression of claudin-1 seems to increase the chances for breast cancer recurrence (Morohashi et al., 2007). There is also evidence of snail directly repressing claudins during epithelial-mesenchymal transitions by binding to the E-box in the promoter region of the claudin genes (Ikenouchi et al., 2003). Another study showed the repression of claudin-1 by snail in epithelial cells that up regulated the transcription factor (Martinez-Estrada et al., 2006). Claudin-4 seems to be particularly important in a variety of cancers. Increased expression of claudin-4 can be seen in cancers of the breast (Kulka et al., 2009), prostate (Landers et al., 2008; Szasz et al., 2009), pancreas (Michl et al., 2001), and ovaries (Santin et al., 2005).

Occludin

Occludin is another protein that is located in the tight junctions between cells. Unlike claudin, there is only one form of occludin known. Also unlike the claudins, occludin seems to have the same expression regardless of the tissue.

In human endometrial carcinomas, occludin levels decreased as the cancer progressed (Tobioka et al., 2004). Additionally, occludin has the same inverse correlation with snail expression that E-cadherin and some claudins have (Kurrey et al., 2005; Martinez-Estrada et al., 2006). It has also been shown that the down regulation of occludin via the activity of oncogenic Raf-1 leads to disruption of tight junction components. In the same study, it was also shown that restoration of occludin levels led to restoration of tight junction components as well as E-cadherin (Li and Mrsny, 2000).

Matrix Metalloproteinases

Matrix metalloproteinases (MMPs) are an important factor in cancer metastasis. MMPs are able to cleave extracellular matrix components. For example, MMP2 and MMP9 are both capable of cleaving type IV collagen. Type IV collagen is an important structural component of the basement membrane of a tissue. When the basement membrane is penetrated, it is possible for cancer cells to metastasize. Interestingly, it has been shown that soluble E-cadherin, cleaved from the cell membrane of human lung cancer cells, is capable of activating MMP2. It was also found that shedding E-cadherin seems to increase the invasiveness of the cancer cells (Nawrocki-Raby et al., 2003). Soluble E-cadherin has also been identified as a prognostic marker for gastric carcinoma patients (Chan et al., 2001).

Claudin-1 has also been associated with MMPs. It has been found that claudin-1 expression in oral squamous cell carcinoma causes MMP2 activation. In this case, MMP2 cleaves the γ2 chain of laminin-5, allowing epidermal

growth factor receptors to bind promoting cell growth and motility (Oku et al., 2006). This does not indicate claudin-1 expression is undesirable in all tissues, as it tends to vary depending on the location of tumor (Soini et al., 2006). As mentioned earlier, claudin-1 is lost in malignant breast cancer lines and the presence of claudin-1 helps prevent recurrence of breast cancer (Morohashi et al., 2007; Tokes et al., 2005a; Tokes et al., 2005b). Another claudin associated with MMPs is claudin-4. There is evidence that claudin-4 enhances invasive potential and is associated with an increase in MMP2 in ovarian cancers (Agarwal et al., 2005).

SNAG Repression Domain:

The SNAG repression domain is an important 21 amino acid sequence present on the snail transcription factor that allows interaction with the Ajuba co repressor (Ayyanathan et al., 2007). This sequence is also present on slug, a closely related transcription factor that is capable of repressing E-cadherin, occludin, and claudins similarly to snail (Martinez-Estrada et al., 2006). The name SNAG comes from the presence of the domain in the **Sna**il/**G**rowth factor independence-1 zinc-finger family of transcription factors (Ayyanathan et al., 2007).

The importance of the domain is its ability to interact with the LIM domain on the Ajuba co repressor protein (Ayyanathan et al., 2007). Without this interaction, the transcription factors snail and slug cannot repress target genes such as E-cadherin, occludin, and the claudins. Ajuba recruits an arginine methyltransferase 5 to methylate the target genes of snail. Methylation causes

the target gene to become compact, preventing transcription initiation (Hou et al., 2008). As mentioned above, there are other zinc-finger transcription factors besides snail and slug with the SNAG domain. Some of these other transcription factors have also been associated with repression of tumor suppressor genes. For example, GFI-1 represses Bax, preventing T-cell death and promoting cell growth (Grimes et al., 1996). GFI-1B represses $p21^{WAF1/CIP1}$, a p53 transcriptional target, and prevents myeloid cell differentiation (Tong et al., 1998).

Biomarkers for Prostate and Breast Cancers

Biomarkers are valuable tools in medicine because they allow for the identification of disease states in the body. An ideal biomarker is one that will allow a researcher to confirm without a doubt the presence or absence of a disease. Biomarkers can be antibodies, proteins, genes, cells, hormones, or virtually any alteration of the body's homeostasis. Indeed a rise in body temperature serves as a biomarker for fever. Similarly, a high blood pressure can be used as an indicator of a potential stroke.

Prostate cancer (PCa) has been studied extensively to identify possible indicators that would clearly reveal not only the presence of cancer, but the progression as well. Prostate-specific antigen (PSA) is a popular biomarker for PCa (De Angelis et al., 2007). The limitation of this marker seems to be inconsistent results that can lead to risky and often unnecessary treatments (De Angelis et al., 2007; Thompson et al., 2004). Although PSA is a widely used biomarker for PCa, there is controversy over the use of it as an indicator of

prostate cancer. A PSA score less than 4.0 ng/mL is generally considered normal. Despite this, one study found that out of 9,459 men, 2,950 men (62-91 years old) never had a PSA score greater than 4.0 ng/mL and 449 (15.2%) of these men were diagnosed with PCa. In addition, 67 of the 449 cancers (14.9%) had a Gleason score of 7 or higher (Thompson et al., 2004). The Gleason score rates cancer on a scale of 2-10, with 10 being the most invasive and disorganized.

The levels of PSA in the blood serum can vary due to other factors besides cancer. Inflammation of the prostate due to infection and benign prostate hyperplasia (BPH) are likely to increase PSA blood levels (Nadler et al., 1995). Ejaculation also raises the blood PSA levels for approximately 24 hours (Herschman et al., 1997). PSA is not useful as a therapeutic target because it is prostate organ specific, not cancer specific. If PSA were to be targeted with a drug, it could damage healthy tissue in the prostate.

Breast cancer (BrCa) has also been studied in order to locate reliable biomarkers. Currently, estrogen receptors (ER), progesterone receptors (PgR), and the epidermal growth factor Her2/neu receptors are used to target and treat breast cancer. These biomarkers are useful in determining the most effective treatment option for the patient (James et al., 2007). Estrogen and/or progesterone receptors are present in approximately 70% of metastatic breast cancers. Hormonal therapies such as tamoxifen are an effective treatment option for these types of cancer (James et al., 2007). However, there is always the possibility that the cancer will become resistant to the treatment (Riggins et

al., 2007). The Her2/neu epidermal growth factor receptors are amplified or overexpressed in approximately 30% of breast cancers. A human monoclonal antibody known as trastuzumab is a treatment option for these types of breast cancers. Treatment with trastuzumab typically yields a 20-30% response rate (James et al., 2007).

Potential for Claudin-4 as a Therapeutic Target

Claudin-4 is currently a protein of interest as not only a biomarker for primary and metastatic cancers (Kulka et al., 2009; Landers et al., 2008; Szasz et al., 2009), but also as a therapeutic target due to its ability to bind Clostridium perfringens enterotoxin (CPE) (Ebihara et al., 2006; Michl et al., 2001; Santin et al., 2005). As mentioned earlier, claudin-4 is known to be overexpressed in prostate cancer, breast cancer, pancreatic cancer, and ovarian cancer. CPE is an enterotoxin produced by the bacteria *Clostridium perfringens*. CPE is a polypeptide that is responsible for many cases of food poisoning. The N-terminus of CPE is a domain that causes cytotoxic effects while the C-terminus is a protein binding domain (Ebihara et al., 2006). The cytotoxic effects occur due to alterations in membrane permeability, which triggers either oncosis or apoptosis depending on high or low CPE concentrations respectively (McClane, 2000; McClane and Chakrabarti, 2004).

There is a concern of systemic toxicity when using CPE as a cancer treatment as claudin-4 expression is not unique to cancer cells. However, the right concentrations of CPE administered by intraperitoneal injection seem to eradicate human xenografted ovarian cancer in mice with little side effects. The

lack of normal epithelial tissue sensitivity to CPE was attributed to the significantly lower claudin-4 expression in non-cancerous epithelium (Santin et al., 2005). One potential option is creating a fusion drug with only the binding domain of CPE. A drug fused to the protein binding domain of CPE would be internalized by the cell following binding to claudin-4 (Ebihara et al., 2006; Matsuda et al., 2004). There is also evidence that CPE may be effective in treating brain metastases as claudin-4 is nearly non-existent in whole brain tissue and the blood-brain barrier may prevent the possibility of systemic toxicity (Kominsky et al., 2007).

Goals of the Present Study

The three main objectives of this thesis research are:

1) Analysis of claudins, occludin, and E-cadherin using bioinformatics tools in order to locate unique extracellular antigenic targets.

2) Cloning, expression of extracellular regions of claudin-1 (CLDN1), claudin-4 (CLDN4), claudin-16 (CLDN16), occludin (OCLN), and E-cadherin (CDH1) in *E.coli*, and large-scale purification of recombinant proteins to homogeneity in quantities amenable for raising antibodies.

3) Expression profiling of metastatic cancer cell lines MCF7, BT20, PC3, DU145, ZR75, and T47D in order to find a significant correlation between snail-slug expression and the cell adhesion molecules being studied.

MATERIALS AND METHODS

Bioinformatics

Amino acid reference sequences for were obtained from the NCBI PubMed database. These sequences were then analyzed with a Kyte-Doolittle GREASE hydropathy profile analysis tool provided by the UCSC Biology Workbench in order to find hydrophillic regions on each protein. Using existing information on the structure of each protein as well as our own hydropathy profile results, the extracellular regions of CLDN 1, CLDN 4, CLDN 16, OCLN, and CDH1 were located. These extracellular protein sequences were then selected as antigenic targets. The European Bioinformatics ClustalW tool (http://www.ebi.ac.uk) and the NCBI BLAST homology analysis tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi) allowed us to compare these antigenic target sequences with existing human proteins to identify any cross reactivity.

Processing of cDNA clones:

Full-length cDNA clones containing CLDN1, CLDN4, CLDN16, OCLN, and CDH1 genes were purchased from Thermo Fisher Scientific Open Biosystems Inc (Huntsville, AL). The plasmids were extracted using the alkaline lysis method (Sambrook and Russell, 2001). The bacteria containing

the abovementioned plasmids were grown overnight in Luria broth (LB) containing 100μg/ml of ampicillin. The cultures were centrifuged at 12,000g for 10 minutes at 4^0C. The supernatant was discarded and the pellets were resuspended in GTE (50 mM glucose; 25 mM Tris-Cl; 10 mM EDTA (pH8.0)) containing 10mg/ml lysozyme and kept on ice for 10 minutes. After 10 minutes lysis solution containing 1%SDS and 0.2N NaOH was added. The samples were left on ice for 5 minutes with intermittent mixing. Finally, 5M KOAc (potassium acetate) was added and the samples were left on ice for 20 minutes with intermittent mixing. Afterwards, the samples were spun at 23,400g for 15 minutes at 4^0C. The supernatant was transferred into new bottles containing equal volume of isopropanol. This mixture was left at room temperature (RT) for 20 minutes followed by another 20 minutes on ice. The samples were spun again at 23,400g for 10 minutes at 4^0C. The supernatant was discarded and the spin was repeated with 70% ethanol to wash the pellets. After the samples dried, TE buffer was used to dissolve the plasmid DNA.

The plasmid DNA samples were run on agarose gels to confirm the integrity. The DNA was treated with RNase and phenol:chloroform:isoamyl alcohol (PCI) purifications were performed on each sample to isolate the pure plasmid DNA. The plasmid DNA samples were run again on agarose gels to confirm that correct plasmids were obtained by comparing them to the λ HindIII-ϕX174 HaeIII marker on an agarose gel.

Restriction Enzyme Digestion of plasmids

The plasmids were digested with different enzymes according to the vector used to construct them. CLDN1, OCLN, and CDH1 were cloned in the pCMV-SPORT6 vector. All of these plasmids were double digested with EcoRI and Hind III restriction endonucleases. CLDN4 was cloned in the pOTB7 vector and was double digested with XhoI and EcoRI restriction endonucleases. CLDN16 was linked to the pPCR-Script Amp SK(+) vector and was sequentially digested with SacI and BamHI restriction endonucleases. The digested plasmid DNA samples were run on agarose gels to confirm that correct plasmids and cDNAs were present by comparing them to the λ HindIII-ϕX174 HaeIII marker. Upon confirmation, primers were designed to amplify specific regions from each of the genes that will be used to express as antigens for making antibodies.

PCR of Target Regions

The primer-pairs used to amplify the regions mentioned above from E-cadherin, claudin-1,4,16 and occludin are presented in **Table 1**. Convenient unique restriction sites were introduced into the primers for directional cloning. We performed PCR on the cDNA templates to amplify the regions that will be used as antigens for raising the respective antibodies. The PCR was not successful for the second region for CDH1. All other PCRs were determined to be successful based on the mobility of the bands on the agarose gel in relation to λ HindIII-ϕX174 HaeIII marker.

DNA Ligation

The *E.coli* expression vector, pGEX 4T1 vector was used for cloning CLDN1, CLDN4, CLDN16, and OCLN1 (loop 1 on OCLN) as BamHI-SalI fragments while pGEX 4T2 was used for cloning OCLN2 (loop 2 on OCLN) and CDH1 (E-cadherin region 1) as SmaI-XhoI fragments. The ligation reactions were transformed into *E.coli* DH5α competent cells.

Transformation of *E.coli* DH5α

E.coli DH5α competent cells can take up ligation reactions, but they cannot efficiently express the proteins since the promoter is not induced. 25μL of DH5α competent cells and 3μL of ligation cocktail were gently mixed and left on ice for 30 minutes. The samples were then heat shocked in a 37^0C water bath for 20 seconds. Samples were kept on ice for 2 minutes before addition of 500μL LB. The cultures were then incubated for 1 hour at 37^0C and 225 rpm. Following incubation, 200μL of each culture was transferred to LB plates. Glass beads were used to spread the bacteria and left at RT 10 minutes to allow the bacteria to absorb into the plates before discarding the beads. The LB plates were incubated at 37^0C overnight. Six colonies were selected for each expression construct and midi-plasmid preparations were made from 50mL cultures. These plasmids were digested with BamHI-SalI (for pGEX4T1 cloning) SmaI-XhoI (for pGEX4T2 cloning) restriction enzymes and analyzed on agarose gels to confirm the presence of correct size vector and inserts.

Transformation of *E.coli* BL21 Expression Host

The *E.coli* BL21 strain is also capable of expressing the recombinant protein. Supercoiled plasmids extracted from *E.coli* DH5α were used for transformation of BL21 cells. 0.5μL of recombinant plasmid DNA was gently mixed with 5μL of *E.coli* BL21 competent cells and kept on ice for 30 minutes. Six individual plasmids were transformed for each protein. The BL21 cells were then heat shocked at 37^0C for 15 seconds and put back on ice. 500μL of LB was then added to each tube and the samples were incubated at 37^0C 225 RPM for 1 hour. Following incubation, add 1.5mL of LB containing 100 μg/ml of ampicillin (amp) and 25 μg/ml of kanamycin (kan) to each culture. Cultures were then incubated overnight at 37^0C and 225 RPM.

Induction of Protein Expression in *E.coli* BL21 Transformants

The mini BL21 cultures were used to monitor protein expression. 750μL from each of the mini BL21 culture was used for the uninduced cultures. 60mL of LB containing amp and kan was prepared for use with uninduced cultures. 1.6mL of LB+amp+kan was added to the 750μL of BL21. Another 60mL of LB was prepared by adding 80μL of amp, kan, and IPTG. 1.25mL of LB+amp+kan+IPTG was added to 750μL of BL21 to make induced cultures. The presence of IPTG allows for induced expression of the GST-tagged recombinant proteins. The uninduced and induced cultures were incubated at 37^0C at 250 rpm for 2 hours. Six pairs of induced/uninduced cultures were prepared for each protein. Protein expression was analyzed on 12% SDS PAGE acrylamide gels. Two protein expression positive clones were selected

for each protein and grown as 500mL induced cultures as explained above. These large cultures would be used for purification of the GST-tagged recombinant proteins as seen below.

Induction of GST-CLDN1 Expression in E.coli BL21 Transformants

Glycerol stocks for the GST-CLDN1 fusion protein were used to inoculate 5mL LB cultures containing 5µL of amp and kan overnight. Overnight cultures were incubated at 30^0C at 250 rpm. The following day, the 5mL cultures were transferred to 125mL of LB with 125µL of amp and kan. 125mL cultures were incubated at 30^0C and a speed of 250 rpm. Upon reaching $A_{600} \approx$ 0.3, cultures were induced to express CLDN GST fusion protein with 125µL of 1M IPTG.

Sonication of BL21 Pellets

500mL BL21 cultures expressing our proteins were poured into bottles and spun at 8,000g 4^0C for 2 minutes. The pellets were stored at -80^0C until use. Pellets were resuspended in 30mL 1xPBS + 40mg lysozyme. PMSF was added to a final concentration of 1mM and the samples were left to gently mix on a rotator for 30 minutes at 4^0C. 600µL of 10% triton x 100 was added to a final concentration of 0.2%. Samples were sonicated for a total of 6 cycles. Each cycle consisted of pulsing for 1 minute followed by 1 minute of rest. Samples were centrifuged at 12,000g at 4^0C for 20 minutes. This supernatant was ultra centrifuged at 100,000g for 1 hour at 4^0C.

Sonication of BL21 Expressing GST-CLDN1

GST-CLDN pellets from 125mL cultures were re-suspended in 4mL 1xPBS + 4mg lysozyme. Following re-suspension, the solution was put on ice for 15 minutes. PMSF was added to 1mM final concentration, DTT was added to 5mM final concentration, and SLS (sodium lauryl sarcosine) was added to 3.5% final concentration. Sonication was done continuously for 30 seconds with a 1 minute rest in between cycles. Sonicated samples were then centrifuged at 12,000g at 4^0C for 20 minutes.

GSH Column Purification of Fusion Proteins

Columns were prepared for each protein with 300µL of GSH beads. The ultracentrifuged supernatant from each sample was then passed through the beads. When all of the flow through was collected, the column was washed with 5 mL 1xPBS. 500µL of wash was collected for each sample and the bottom of each tube was sealed with parafilm. 500µL of elution buffer was added to each column. The elution buffer was allowed to sit for 10 minutes before being collected in an eppendorf tube. This was repeated two more times for a total of 3 elutions. The elutions, flow-through, PBS wash and the sonicated extract were all analyzed on 14% SDS-PAGE gels.

Purification of GST-CLDN with GSH Beads

4mL of soluble protein was brought up to a volume of 15mL with 1xPBS. The soluble protein was then added to a 15mL conical tube with GSH beads. This tube was rotated at 4^0C for 1 hour. In order to collect the elutions, the soluble protein mixed with beads was poured back through the column and the

settled beads were washed with 5mL of 1xPBS. Elutions were collected as stated above.

Protein Refolding

Protein refolding was performed since some of the fusion proteins were expressed as inclusion bodies rather than in the supernatant we ran through the columns. Pellets from the centrifuged sonications were solubilized in 5mL of 2M urea/100mM Tris (pH 12.5) and kept on ice for 30 minutes. The solubilized pellets were then ultra centrifuged at 100,000g for 30 minutes at 4^0C. The refolding was carried out in the cold room. The ultracentrifuged supernatant was added to the refolding buffer (100 mM Tris Cl, 1 mM EDTA, 20% glycerol, 10% sucrose) containing 2M urea. The additions were done in a drop-wise manner at a rate of 1 drop every 5 seconds with continuous stirring. A 3-minute break was taken after every 1mL of supernatant was added. After the additions were complete, PMSF was added to each sample to a final concentration of 1mM.

Propagation of Cancer Cell Lines Used in Expression Profiling

The following human metastatic breast cancer cell lines (MCF7, BT20, ZR75, and T47D) and human metastatic prostate cancer cell lines (PC3 and DU145) were used in our studies. The cell lines were cultured in DMEM containing 10% fetal bovine serum supplemented with 1x penicillin/streptomycin, and 1x glutamine in a 5% CO_2 incubator at 37°C. Cultures were grown to ~90% confluency before removing the DMEM and

adding 2mL TRIzol to each dish to lyse the cells. The TRIzol lysates were evenly transferred into eppendorf microcentrifuge tubes.

RNA Preparation and First Strand cDNA Synthesis

200µL of chloroform was added to each lysate and shaken vigorously for 15 minutes. The samples were then centrifuged at 12,000 g for 15 minutes at 4°C. The aqueous phase was transferred to a new microcentrifuge tube using a micropipette and 500µL of isopropanol was added. After sitting for 10 minutes at room temperature, the tubes were centrifuged at 12,000 g for 10 minutes at 4°C. The supernatant was discarded and 1mL of 75% was added to each tube. Pellets were loosened from the wall of each tube and the tubes were shaken every 5 minutes for 20 minutes on ice. The tubes were then centrifuged at 7,500 g for 5 minutes at 4°C. Ethanol was decanted and the pellets were air-dried. After drying, 100µL of DEPC water was added in order to dissolve the pellets. 10µg of RNA was analyzed on 0.7% agarose gels. After verifying the integrity, 5µg of RNA was taken for first strand cDNA synthesis with 1µg of Oligo(dT) primer and the Ready-to-go You-Prime First-Strand beads kit obtained from Amersham Biosciences.

Semi-Quantitative PCR Analysis

The primers indicated in **Table 2** were used to PCR amplify the first strand cDNA templates synthesized from our RNA samples. PCR reactions were done using the Go Taq® Hot Start Green Master Mix and following the instructions provided. The PCR program consisted of an initial 2 minute 94°C denaturing step followed by a 2 minute 54°C annealing phase and then a 2

minute 72°C extension phase. The following 34 cycles were programmed as follows: 20 second 94°C denaturing step, a 30 second 54°C annealing step, and a 1 minute 72°C extension. After 34 cycles were complete, a final 10 minute 72°C extension phase. Aliquots were taken at 20, 26, 31, and 35 cycles. Aliquots of the PCR samples were electrophoresed on 1.2% agarose gels and photographed.

Alkaline Blotting

Biorad Zeta Probe membranes were used for alkaline blotting. Saran wrap was placed on a table with a dry stack of paper towels on top. Two dry Whatman 3 sheets were placed on top of the paper towels. A third Whatman sheet was soaked in 0.4 NaOH and placed on top of the dry sheets. The membrane was soaked in 0.4 NaOH prior to placing it on top of the Whatman sheets. Old film was cut into strips and placed on the edges of the membrane to ensure proper transfer. The agarose gel was placed on top of the membrane and bubbles between the two surfaces were removed. 3 remaining Whatman sheets were soaked in 0.4 NaOH and placed on top of the gel, taking care to remove any bubbles. A stack of paper towels soaked in 0.4 NaOH was placed on top. The transfer was allowed to take place over a period of 24 hours.

5' End Labeling of Internal Oligonucleotide Probes

The oligonucleotide sequences presented in **Table 3** were used in the kinasing reactions and used as internal probes in hybridization reactions. Each oligonucleotide was labeled with γ^{32}P-ATP. The 20µL reaction tubes contained 13µL of ddH_2O, 2µL of 10x PNK buffer, 2µL of T_4 polynucleotide kinase, and 1

µL of γ^{32}P-ATP (~166 µCi). Tubes were placed in a 37°C water bath for 2 hours. Following kinasing, 20µL of ddH_2O was added to each tube and the contents were added microcentrifuge columns containing 600µL of prespun G25 sephadex beads and centrifuged. 2µL was then taken to monitor the extent of labeling using a scintillation counter. The readings were typically ~1,000,000 counts per minute in a 2µL sample.

Prehybridization and Hybridization of Alkaline Blots

Alkaline blots were prepared for hybridization with radioactive oligonucleotides by incubation with a prehybridization solution. The prehybridization solution contained 5x SSC (0.75M NaCl, 0.075M sodium citrate), 20mM Na_2HPO_4 (pH 7.2), 7% SDS (sodium dodecyl sulfate), and 1x Denhardt's (0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin). Each blot was heat sealed in a bag with 20mL of prehybridization solution. The bags were then put into a 50°C water bath for 24 hours. nHybridization was carried out by cutting open each bag and adding 10µL of labeled oligonucleotide followed by resealing the bag. The bags were once again placed in a 50°C water bath for 24 hours with occasional mixing. Blots were washed the next day with a solution of 1xSSC and 1%SDS. A series of three 100mL 10 minute washes were done to remove excess label. Each blot was exposed to X-ray film for approximately 48 hours prior to development.

RESULTS

Bioinformatics

Hydropathy profiling allowed for identification of extracellular regions of claudin-1 (CLDN1), claudin-4 (CLDN4), and claudin-16 (CLDN16). These regions can be seen in **Figure 1**. Amino acids 28-80(CLDN1), 52-78(CLDN4), and 17-52(CLDN16) were selected as antigenic regions to be used in the GST fusion proteins. Both of the regions selected for occludin (OCLN) were located on one of two extracellular loops (**Figure 2A**). The sequence selected on loop 1 (OCLN1) of OCLN was amino acids 88-137 and the sequence on loop 2 (OCLN2) was amino acids 199-243. Two extracellular sequences were chosen for E-cadherin as antigenic regions (**Figure 2B**). The first region was amino acids 154-448 (CDH1) and the second was 489-644 (CDH2).

After locating the regions on each protein that were extracellular and soluble, we then needed to find a unique extracellular amino acid sequence that could be used to raise antibodies for an ELISA assay. ClustalW alignment of the extracellular regions of CLDN1 (**Figures 3** and **4**), CLDN4 (**Figures 5** and **6**), CLDN16 (**Figures 7** and **8**), and OCLN (**Figures 9** and **10**) with the extracellular regions of other claudins and E-cadherin allowed for identification of a more specific region to target. Two sequences were also

selected for E-cadherin by looking at ClustalW alignments with claudins, occludin, and other cadherins (**Figures 11** and **12**). All of these sequences can be found in **Figure 13.**

The ClustalW alignment results for the CLDN 1 loop 1 region in **Figure 3** indicate a highest possibility of cross reaction with CLDN 15 loop 1 where 5 out of 9 amino acids are an exact match. CLDN 1 loop 2 antigenic sequence in **Figure 4** is most closely related to CLDN 7 loop 2 where 5 out of 9 amino acids match exactly.

CLDN 4 loop 1 did not show as much promise as a unique antigenic target. The region that we found to be most unique on CLDN 4 loop 1 in **Figure 5** had 6 out of 12 or more amino acid matches with the loop 1 regions of claudins 1, 3, 5, 6, 7, 8, 9, 10A, 10B, 14, 15, 17, 19A, and 19B. The CLDN 4 loop 2 sequence in **Figure 6** looked more promising as its closest matches were CLDN 9 loop 2 with 6 out of 9 amino acids aligned perfectly and CLDN 6 loop 2 with 5 out of 9 matches.

The CLDN 16 loop 1 antigenic sequence in **Figure 7** had no real matches with respect to the claudins, OCLN, or E-cadherin in the ClustalW alignment. The strongest homology was found with the first extracellular E-cadherin sequence tested where 3 out of 12 amino acids matched. The closest match for the region selected on CLDN 16 loop 2 in **Figure 8** was found on CLDN 23 loop 2 where 4 out of 9 amino acids matched.

The region selected on OCLN loop 1 seen in **Figure 9** was very unique in the alignment with no significant cross reactivity. The closest matches were

two extracellular regions of E-cadherin we tested and CLDN 23 loop 1 where 3 out of 16 amino acids were an exact match. Our chosen OCLN loop 2 sequence in **Figure 10** also had little similarity to other proteins in the aligment with a maximum of 3 matching amino acids.

The first extracellular sequence selected for E-cadherin in **Figure 11** did not align at all with most of the proteins. Those that did not align are not presented. The closest matches were CLDN 5 loop 2, KSP-cadherin, R-cadherin, and M-cadherin which all aligned with only 2 out of 11 amino acids being the same. The second extracellular E-cadherin sequence we selected in **Figure 12** also aligned with few of the proteins tested in the alignment. LI-cadherin and type 2 vascular endothelium cadherins were the most similar with 3 out of 13 amino acids matches. R-cadherin matched with 4 out of 13 amino acids, but there was a gap in the alignment. The chosen polypeptide regions for raising peptide antibodies that resulted from extensive bioinformatics analysis are categorized in **Figure 13**.

Cloning, Expression, and Protein Purifications

The PCR amplification of selected antigenic regions is shown in **Figure 14**. All regions were successfully amplified with the exception of the second E-cadherin extracellular fragment selected (CDH2). Although the second loop fragment of OCLN (OCLN2) was successfully amplified, the GST fusion protein was not purified in large quantities.

Figure 15 shows the approximate relative sizes of fusion proteins as well as the exact location of each antigenic fragment on the original protein.

Coomassie staining of SDS-PAGE induction panels for CLDN1, CLDN4, CLDN16, OCLN1, OCLN2, and CDH1 GST fusion protein expression in recombinant clones are shown in **Figure 16**. All of the inductions were successful with equal expression of the fusion protein throughout the clones.

Initial purifications of GST-tagged recombinant protein is shown in **Figure 17**. Many of the initial purifications were not ideal, so the idea that most of the protein was present in inclusion bodies was explored. CDH1 was the only successful purification.

Protein refolding was performed to recover protein that went to inclusion bodies after centrifuging the sonicated extract. **Figure 18** shows the results of refolding. The CLDN16 protein was the only one successfully recovered as seen in **Figure 18A**. Analysis of base washes from column regeneration was also included in **Figure 18B** and it appears that CLDN16 was present once again. We also analyzed pellets obtained from centrifuging the refolded protein prior to running it through the GST columns. These pellets were compared to the unpurified refolded supernatant in **Figure 18C**.

Further purifications of our GST tagged proteins can be seen in **Figures 19** and **20**. CLDN 4 was prone to cleavage at times. This would result in a band similar to the GST tag. The OCLN (OCLN1) fusion protein always had significant cleavage of the GST tag. As a result, we acquired more than twice the amount of total protein needed for raising antibodies because our estimates of total fusion protein were most likely half of our estimation. The total amount

of fusion proteins purified after many rounds of purifications are presented in **Table 4**.

Expression Profiling

The RNA extractions from our cancer cell lines were successful, but multiple attempts were needed due to issues with contamination. The analysis of the RNA can be seen in **Figure 21**. The first extraction in **Figure 21A** was successful for the DU145, ZR75, and T47D cell lines. Another RNA extraction was performed on MCF7, BT20, and PC3 **(Figure 21B)**.

The semi-quantitative PCR-Southern analysis data for the DU145 cell line in **Figure 22** revealed high snail and slug expression with less E-cadherin and CLDN 4 expression. OCLN expression in the DU145 cell line appeared as a faint band in the terminal cycle, while CLDN1 and CLDN16 were absent entirely in all cell lines **(Fig 22A)** Analysis of the intensity of the bands at 20, 26, 31, and 35 cycles can be seen in **Figures 22B-E**. Snail expression was high intensity at 26 cycles, slug reached a high intensity at around 31 cycles, E-cadherin showed strong expression at 35 cycles, and CLDN 4 never peaked as it was lower intensity than the other detected genes.

A Southern blot analysis of CLDN 4 expression in MCF7, BT20, PC3, DU145, ZR75, and T47D cancer cell lines can be seen in **Figure 23**. The prostate cancer cell lines seemed to have the lowest expression of CLDN 4 while MCF7 had the highest. The intensity of the bands can be seen in **Figure 23**. MCF7 **(Figure 23B)** reached a high intensity at 31 cycles while BT20 **(Figure 23C)** and ZR75 **(Figure 23D)** reached a similar intensity at 35 cycles.

T47D **(Figure 23E)** did not reach the same level of expression as the other breast cancer cell lines but was still higher than PC3 **(Figure 23F)** and DU145 **(Figure 23G)**. PC3 had a higher expression level than DU145, which had the lowest expression of CLDN4.

DISCUSSION

Bioinformatics

The purpose of the bioinformatics work was to find an ideal extracellular sequence on each of our proteins that did not interact with the extracellular sequences of similar proteins as well as with any other human proteins that may be detectable in an ELISA. The extracellular regions of CLDN 1, CLDN 4, CLDN 16, OCLN, and CDH1 analyzed with ClustalW alignment and BLAST showed promising results for the goal of finding a unique antigenic target.

The OCLN loop 1 protein we identified would work well for our purposes. Similarly, the 2 sequences selected on E-cadherin were very distinct from any claudins or OCLN and in many cases did not match at all with the extracellular portions of these proteins. These E-cadherin sequences were also distinct with respect to the other cadherins, making them strong antibody targets. The region we selected for antibody production should be adequate.

It is worth noting that when comparing the entire loop 2 region of CLDN 4 to the whole loop 2 sequences of the other claudins, those closely related to CLDN 4 matched in the same regions. Loop 2 regions of CLDN 3, CLDN 6, CLDN 7, CLDN 8, CLDN 9, and CLDN 17 were not only closely related to

CLDN 4 loop 2, but also tended to match in the same locations. These similarities would suggest similar CPE binding due to the fact that the toxin binds to the second loop of CLDN 3 and CLDN 4 (Fujita et al., 2000; Ling et al., 2008; Takahashi et al., 2005). Indeed, one study found that CLDN 6 and CLDN7 have similar CPE affinities to that of CLDN 3 with CLDN 8 and CLDN 14 showing lesser but significant affinity (Fujita et al., 2000). CLDN 4 has been confirmed as a high affinity CPE receptor, while CLDN 3 has demonstrated lower but significant affinity (Fujita et al., 2000; Katahira et al., 1997). CLDN 3, CLDN 4, and CLDN 7 also share similar expression levels in cancer (Hewitt et al., 2006).

The ClustalW analysis and BLAST results revealed that while all of the regions we used for our fusion proteins had a unique sequence, there is also a chance for cross reactivity when it comes to the claudin antibodies. The CLDN 4 loop 1 protein region we chose would produce antibodies that could potentially react with several claudins. In this case, a monoclonal antibody to a sequence in the loop 2 region of CLDN 4 would be much more specific.

Protein Expression and Purifications

The PCR amplification of the CDH2 region we selected for E-cadherin did not seem to be successful. We are unsure of why the PCR did not work, but the CDH1 region should be sufficient to raise antibodies for E-cadherin that would suit our needs. The CDH1 region is a total of 294 amino acids while the CDH2 region is only 155 amino acids. The larger CDH1 sequence has a

greater potential to produce an antibody that will work for our future experiments.

Initially, we did not purify as much GST-tagged protein as we were expecting. We initially attempted protein refolding with the idea that there were inclusion bodies forming from our fusion proteins. However, this was most likely not the case as all of the proteins but CDH1 were similar in size to the 52 amino acid CLDN 1 loop 1 sequence which is too small to interfere with GST solubility. The refolding procedure seemed to work for CLDN 16, but it was not necessary as future expression and purification attempts yielded plenty of soluble protein.

GST-CLDN 1 was the most problematic in the beginning, but a significant alteration to the procedure allowed the protein to be collected in larger quantities. The first change to the procedure was the decrease in induction temperature from 37^0C to 30^0C. This temperature change led to more consistent fusion protein collection with the exception of GST-CLDN1. The second change to the GST-CLDN1 cultures was to the concentration of IPTG added. Using 250mL cultures, we induced one culture with 250µL of 100mM IPTG, another with 25µL of 1 M IPTG, and the last one with 250µL of 1M IPTG. The culture with 250µL of 1M IPTG yielded the most GST-CLDN1 fusion protein. Despite the success with the 250mL cultures, subsequent trials with 500mL cultures yielded little to no GST-CLDN1 protein.

The next change to the GST-CLDN1 purification procedure was the size of the cultures. It can be seen in the induction panels that all of the clones

produced a similar amount of fusion protein whether it was CLDN1, CLDN4, CLDN16, OCLN, or CDH1. Oddly enough, GST-CLDN1 expression was very strong in the induction panels, but hardly noticeable in the purifications from larger cultures. A change to 125 mL cultures induced with 125µL at 37^0C had some success.

The final alteration in the purification procedure of GST-CLDN1 involved having the GSH beads gently mix with the protein in a conical tube at 4^0C rather than pass through a GSH column at room temperature. With this final change, GST-CLDN1 was collected on a regular basis. The alterations to the procedure led to not only a higher yield of GST-CLDN1, but also to very clean elutions with all of the fusion protein intact.

In order to raise antibodies, we needed to collect 2mg of pure protein. If the proteins we needed were larger, this goal would have been fairly easy to meet. However, with the exception of CDH1, the proteins are all around 50 amino acids. This is much smaller than the 220 amino acid GST tag, so the difference in molecular weight needs to be considered when estimating the amount of protein collected. For example, in order to get 2.3mg of pure CLDN 4 protein, we had to collect a total of about 16.6g of fusion protein because CLDN 4 loop 1 is only 34 amino acids.

Expression Profiling

The semi-quantitative PCR-Southern blot analysis of the MCF-7, BT20, PC3, DU145, ZR75 and T47D cell lines provided some information we were looking for, but the data was not conclusive for all cell lines. DU145 was the

only cell line that showed a clear pattern of expression on the Southern blots expected of a metastatic cancer cell line.

Snail and slug were highly expressed in DU145, while E-cadherin was hardly present. In agreement with other studies, OCLN was hardly expressed in the cancer cell lines (Kurrey et al., 2005; Martinez-Estrada et al., 2006). CLDN 1 expression was not present or expected in the DU145 cell line based on existing data on prostate cancers (Szasz et al., 2009). There was no detectable CLDN 1 in any of the cancer cell lines, confirming previous studies that show a decrease in CLDN 1 expression in breast and prostate cancer (Martinez-Estrada et al., 2006; Morohashi et al., 2007; Szasz et al., 2009). CLDN 4, on the other hand, was present in all of the cell lines. The notable expression of CLDN 4 in all of the cell lines agrees with existing data on breast and prostate cancer (Kulka et al., 2009; Landers et al., 2008; Szasz et al., 2009). The amount of CLDN 4 detected for the prostate cancer cell lines DU145 and PC3 was not as significant as expected. It is also worth noting that the breast cancer cell lines all seemed to have higher CLDN 4 expression when compared to the prostate cell lines.

During the PCR of cancer cell line cDNA, aliquots were taken at 20, 26, 31, and 35 cycles. Analysis of these cycles in DU145 revealed that snail expression peaked at 26 cycles while E-cadherin was only in significant quantity at the terminal 35th cycle. Slug expression peaked at 31 cycles, indicating less protein in comparison to snail, but more than E-cadherin. While it is unclear when CLDN 4 expression would have peaked, the complete absence of CLDN

1 supports the inverse correlation between these proteins often seen in prostate cancers (Hewitt et al., 2006; Szasz et al., 2009).

CLDN 4 expression in MCF7 seemed to peak around 31 cycles, while the other breast cancer cell lines showed a steady increase in production. The data we obtained for the breast cancer cell lines seems to suggest a heavier expression of CLDN 4 in malignant breast tissue when compared to malignant prostate cancers.

FUTURE STUDIES

We will be raising both monoclonal and polyclonal antibodies against each of the proteins described earlier. Subsequent efforts will be made to develop a high-throughput ELISA-based assay to detect the presence of circulating cell junction components in a patient serum. The serum will be analyzed using the antibodies we raised towards E-cadherin, the claudins, and occludin. This would serve as a novel diagnostics for the detection of biomarkers in the serum of metastatic cancer patients.

PCR primers	
Primer Pair	**Sequence**
CDH_1-3'Xho	cacCTCGAGctgcttggcctcaaaatcc
CDH_1-5'Sma	cacCCCGGGagagactgggttattcc
CDH_2-3'Sal	gtgGTCGACagtcacccacctctaagg
CDH_2-5'Bam	gtgGGATCCcctgaaaagagagtggaag
CLDN_1-3'Sal	cacGTCGACggttgcttgcaatgtgc
CLDN_1-5'Bam	cacGGATCCccccagtggaggatttac
CLDN_4-3'Sal	cacGTCGACggcgcgggccgcctgc
CLDN_4-5'Bam	cacGGATCCaccatctgggagggcc
CLDN_16-3'Sal	cacGTCGACaagatccctcattgtggc
CLDN_16-5'Bam	cacGGATCCtactgcaactcaagacacc
Occludin_1-3'Sal	gtgGTCGACtgctgctcttgggtctg
Occludin_1-5'Bam	gtgGGATCCgcctgggacagaggctatg
Occludin_2-3'Xho	cacCTCGAGctcctggggatccacaac
Occludin_2-5'Sma	cacCCCGGGccaactgctcagtcttc

Table 1. Oligonucleotides used to create expression constructs. These oligonucleotides were used to create the respective GST fusion protein expression constructs.

PCR primers	
Primer Pair	**Sequence**
CLDN_1-5'-296	CCGTTGGCATGAAGTGTATG
CLDN_1-3'-503	CCAGTGAAGAGAGCCTGACC
CLDN_1-5'-330	CGATGAGGTGCAGAAGATGA
CLDN_1-3'-528	AAGGCAGAGAGAAGCAGCAG
CLDN_4-5'-6	CTCCATGGGGCTACAGGTAA
CLDN_4-3'-212	AGCAGCGAGTCGTACACCTT
CLDN_4-5'-134	CCATCTGGGAGGGCCTAT
CLDN_4-3'-331	CATCCTCCAGGCAGTTGG
CLDN_16-5'-112	TGCCAGGTACCAGAAACACA
CLDN_16-3'-310	CAGCATTCACCATCCAACAG
CLDN_16-5'-367	AATGCTTTTGATGGGATTCG
CLDN_16-3'-567	AACAAAGCAGATGCGGACTT
E-cad-5'-487	TGCCCAGAAAATGAAAAAGG
E-cad-3'-686	GTGTATGTGGCAATGCGTTC
E-cad-5'-1576	TGGAGAGACACTGCCAACTG
E-cad-3'-1776	GGCGTTGTCATTCACATCAG
OCLN-5'-1046	AAGTGGTTCAGGAGCTTCCA
OCLN-3'-1246	AGTCCTCCTCCAGCTCATCA
OCLN-5'-1239	GGAGGACTGGATCAGGGAAT
OCLN-3'-1436	TCAGCAGCAGCCATGTACTC
SCRT-5'-5	CCAGGTCCTTCCTGGTCA
SCRT-3'-217	CAGCTGCTGCGTACATGG
SCRT-5'-684	CAAGGTGTACGTGTCCATGC
SCRT-3'-921	CTTGCACTGGAAGTGCTTGA
Slug-5'-152	GAGCATACAGCCCCATCACT
Slug-3'-359	GGGTCTGAAAGCTTGGACTG
Slug-5'-452	GCGATGCCCAGTCTAGAAAA
Slug-3'-654	GCAGTGAGGGCAAGAAAAAG
Snail-5'-85	TTTACCTTCCAGCAGCCCTA
Snail-3'-283	CCTCATCTGACAGGGAGGTC
Snail-5'-169	CCAATGCTCATCTGGGACTC
Snail-3'-368	GCCTCCAAGGAAGAGACTGA

Table 2. Oligonucleotides used for expression profiling. Two pairs of primers were used for each gene. The cDNA from RT-PCR of extracted cancer cell line RNA was then amplified with these primers in order to analyze the expression levels of each gene.

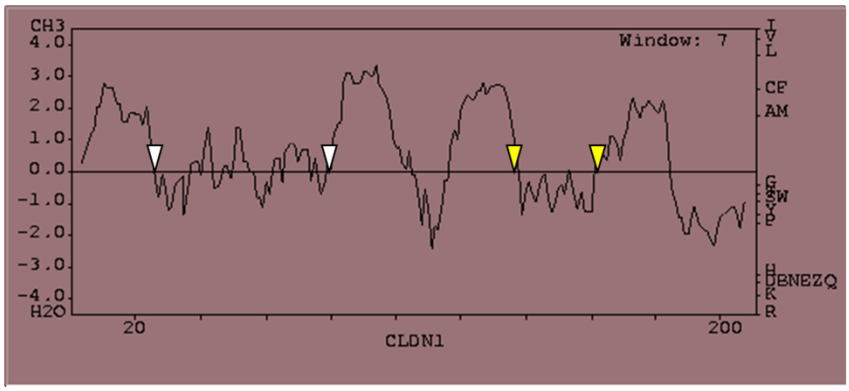
Oligonucleotide	Sequence
CLDN1_I&II	ggatggctgtcattggggg
CLDN4_I&II	ctgcgtggtgcagagcacc
CLDN16_I	ggctgcttgccacaatgagg
CLDN16_II	gcatcccttgaagctggtgg
E-cad-I	ggccaaggagctgacacacc
E-cad-II	gggctgagctggacaggg
OCLN_I&II	gagctggaggaggactgg
SCRT_I	cggccgacctggagagcgc
SCRT_II	catggccatgcacctgctc
SLUG_I	cctcctccatctgacacctc
SLUG_II	gtgagcctgggcgccctg
Snail_I&II	ctggcgccccaagcccagc

Table 3. Internal oligonucleotide probes used in Southern blotting. These oligonucleotide probes were labeled with ^{32}P radioactive isotope. Following labeling, the probes were used in a Southern blot procedure to specifically locate CLDN1, CLDN4, CLDN16, E-cadherin, OCLN, scratch, slug, and snail gene expression.

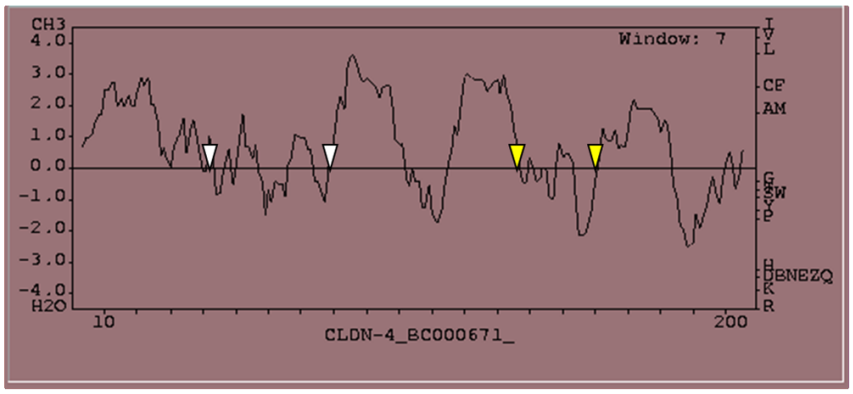
GST fusion protein quantities				
Protein	**Fusion prot. (kDa)**	**Antigenic prot. (kDa)**	**Total prot. (mg)**	**Antigenic prot. (mg)**
GST-CDH1	58.98	32.01	8.98	4.88
GST-CLDN1	32.88	5.91	9.36	1.68
GST-CLDN4	31.24	4.27	16.58	2.27
GST-CLDN16	33.46	6.49	25.32	4.91
GST-OCLN	32.06	5.09	26.60	4.22

Table 4. Quantities of purified proteins. We used the relative weights of the antigenic sequence and the GST fusion protein to determine what fraction of each fusion protein was our antigenic sequence. A protein estimation was done for all of the elutions and the amount of antigenic protein was calculated from that data.

A.



B.



C.

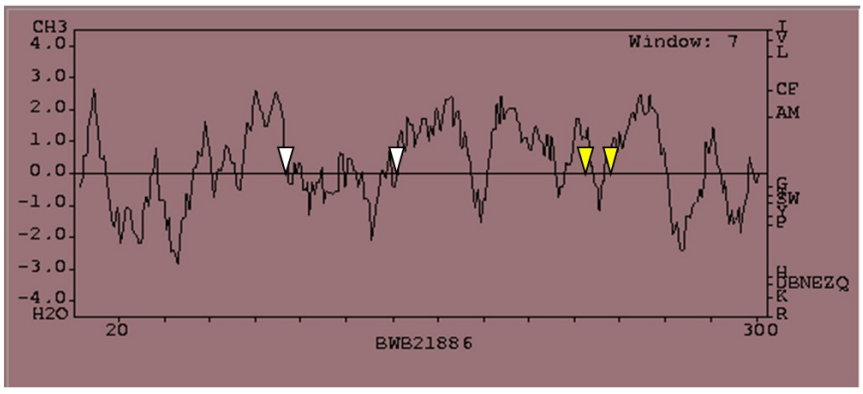
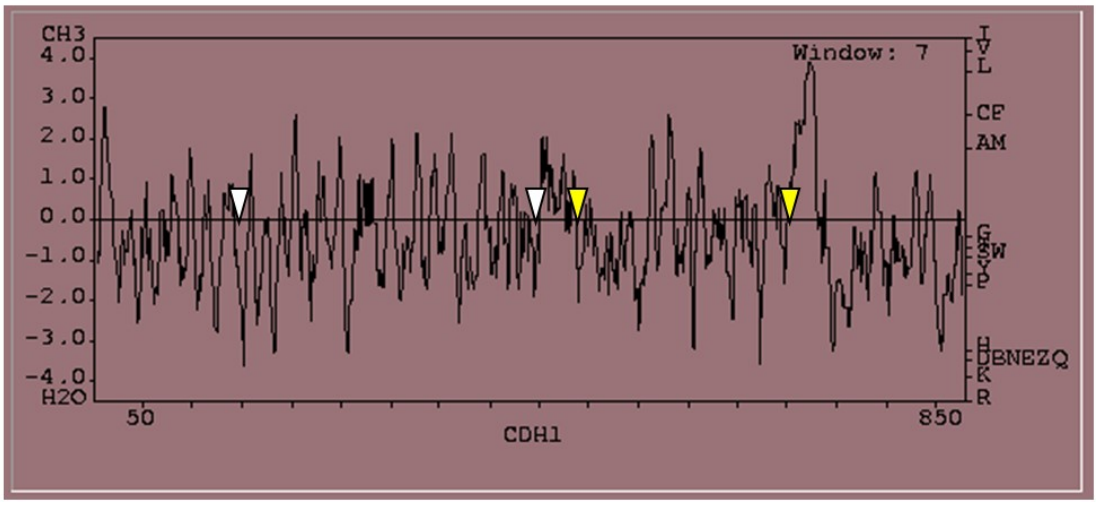


Figure 1. Hydropathy profiles of CLDN1 (A), CLDN4 (B), and CLDN16 (C). The line representing the amino acid sequence moves up if the region is hydrophobic and down if the region is hydrophilic. White arrows flank the first extracellular loop and the yellow arrows flank the second extracellular loop region. The first extracellular loop was used for creating all GST-claudin fusion proteins.

A.



B.

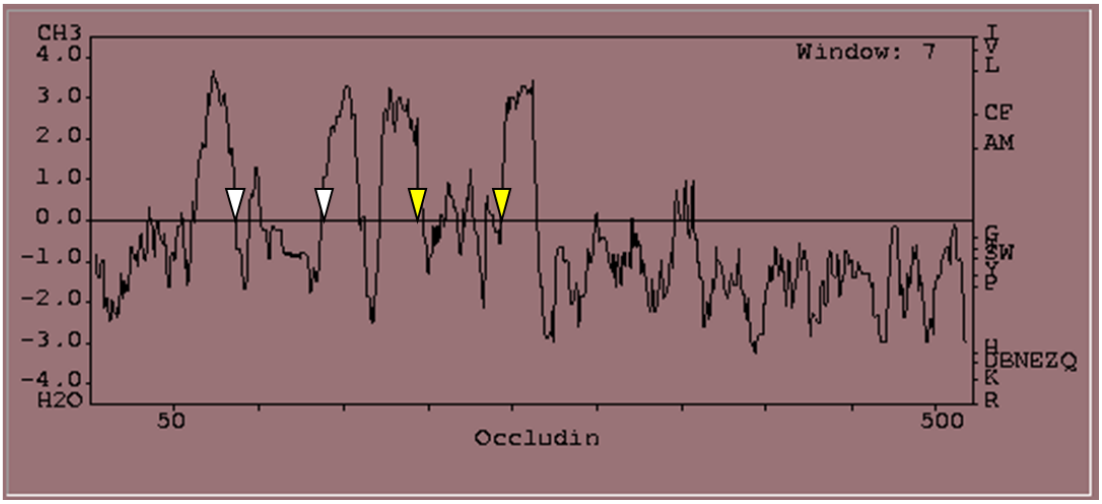


Figure 2. Hydropathy profiles of CDH1 and OCLN. The line representing the amino acid sequence moves up if the region is hydrophobic and down if the region is hydrophilic. A) White arrows indicate the extracellular region used for the GST fusion protein. The yellow arrows flank the other extracellular portion analyzed as a potential target for antibodies. B) The white arrows flank the first extracellular loop of OCLN. This first loop region was used for the GST fusion protein (GST-OCLN1). The yellow arrows flank the second extracellular loop used for the GST-OCLN2 fusion protein.

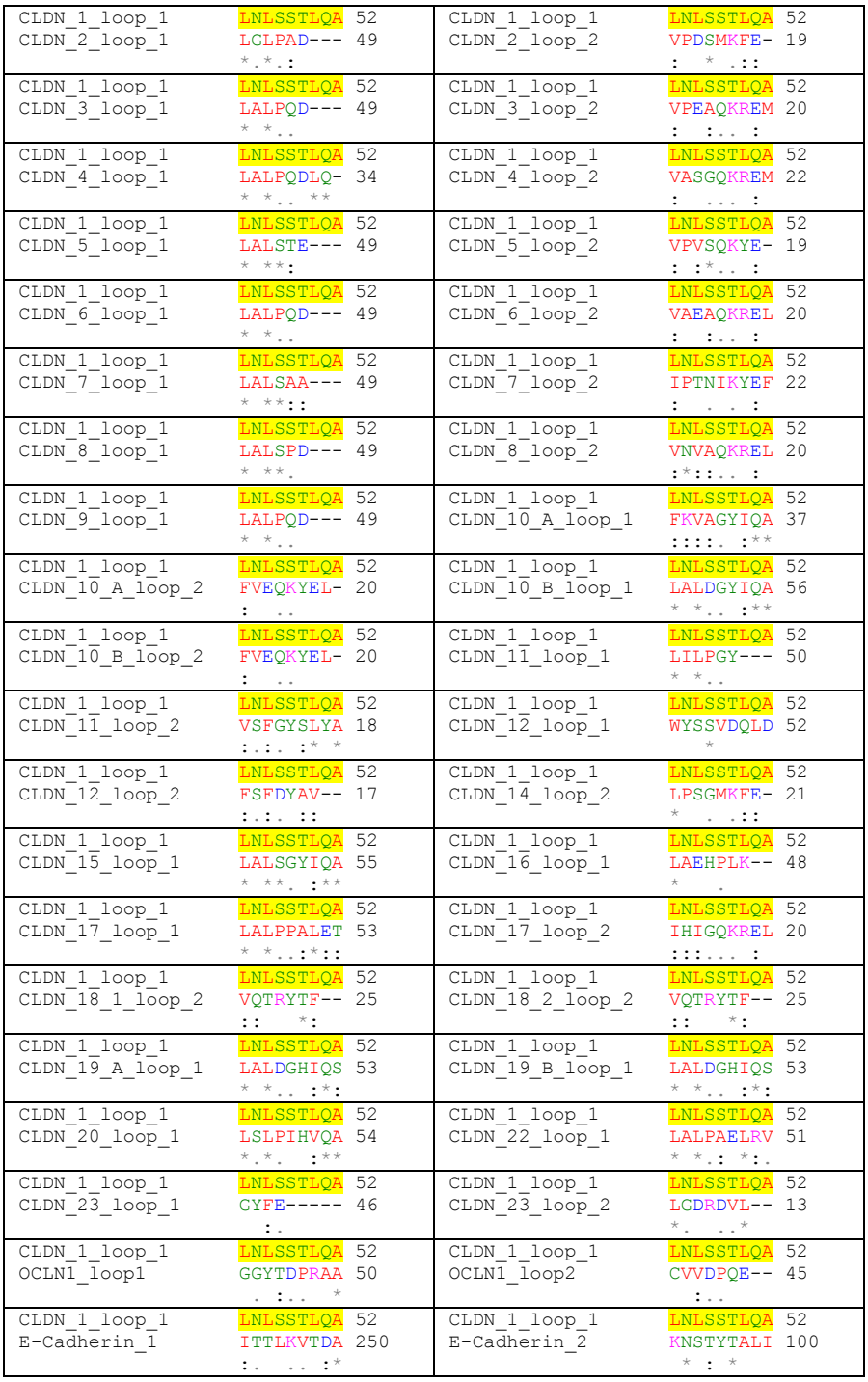


Figure 3. ClustalW alignment of the unique antigenic CLDN1 loop 1 sequence. The highlighted sequence was aligned with extracellular regions of claudins, occludin, and E-cadherin. This sequence on CLDN1 loop 1 is present on our GST-CLDN1 fusion protein.

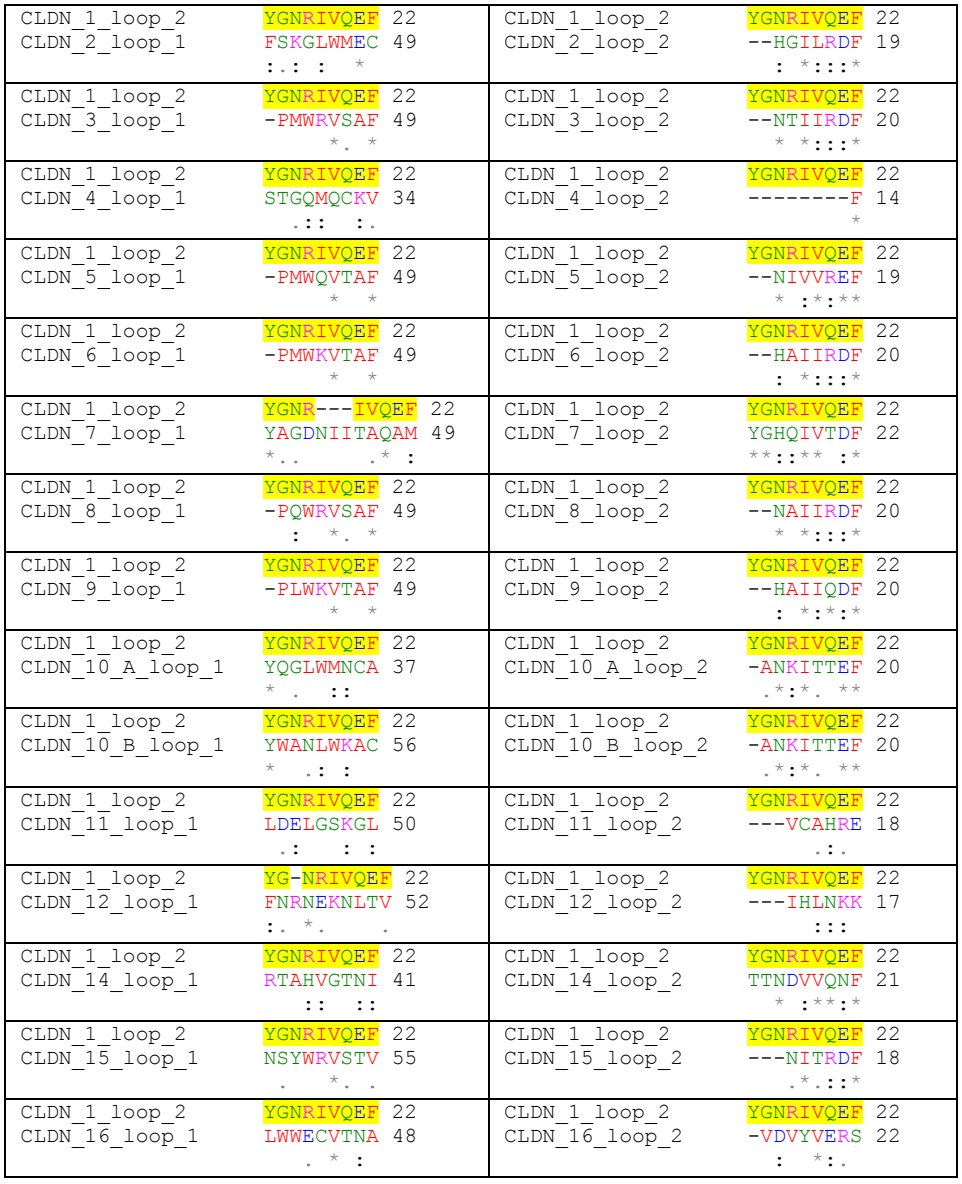


Figure 4. ClustalW alignment of the unique antigenic CLDN1 loop 2 sequence. Legend in next page

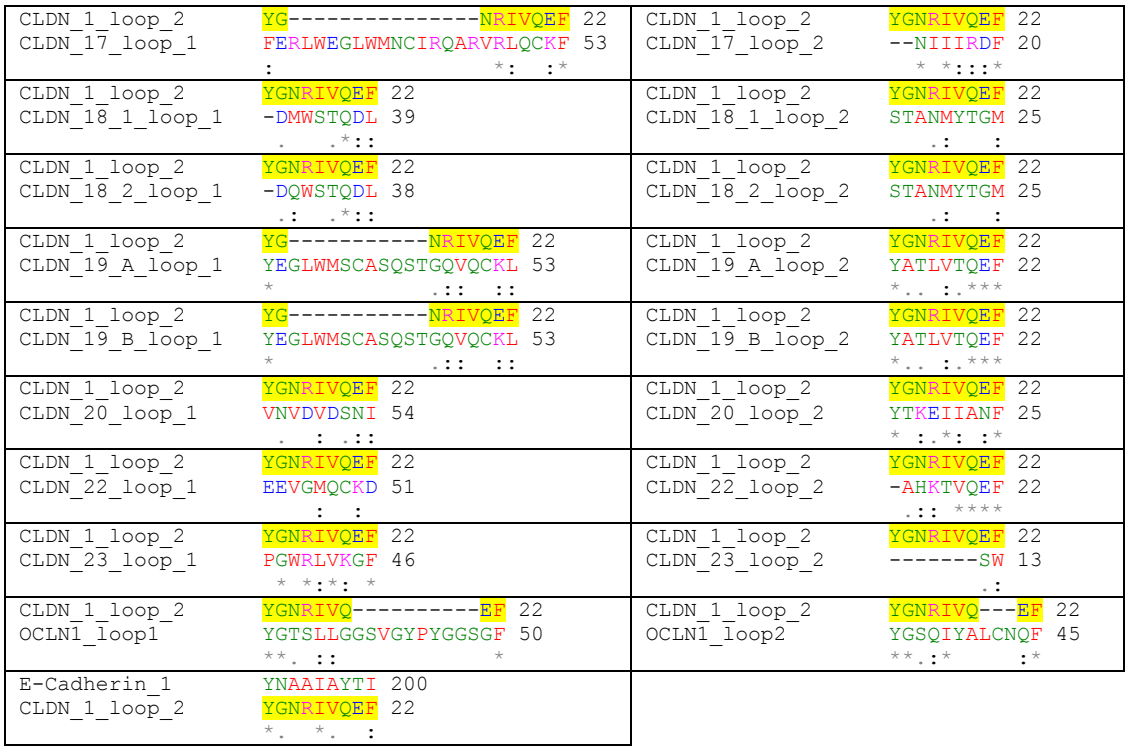


Figure 4. ClustalW alignment of the unique antigenic CLDN1 loop 2 sequence. The highlighted sequence was aligned with extracellular regions of claudins, occludin, and E-cadherin.

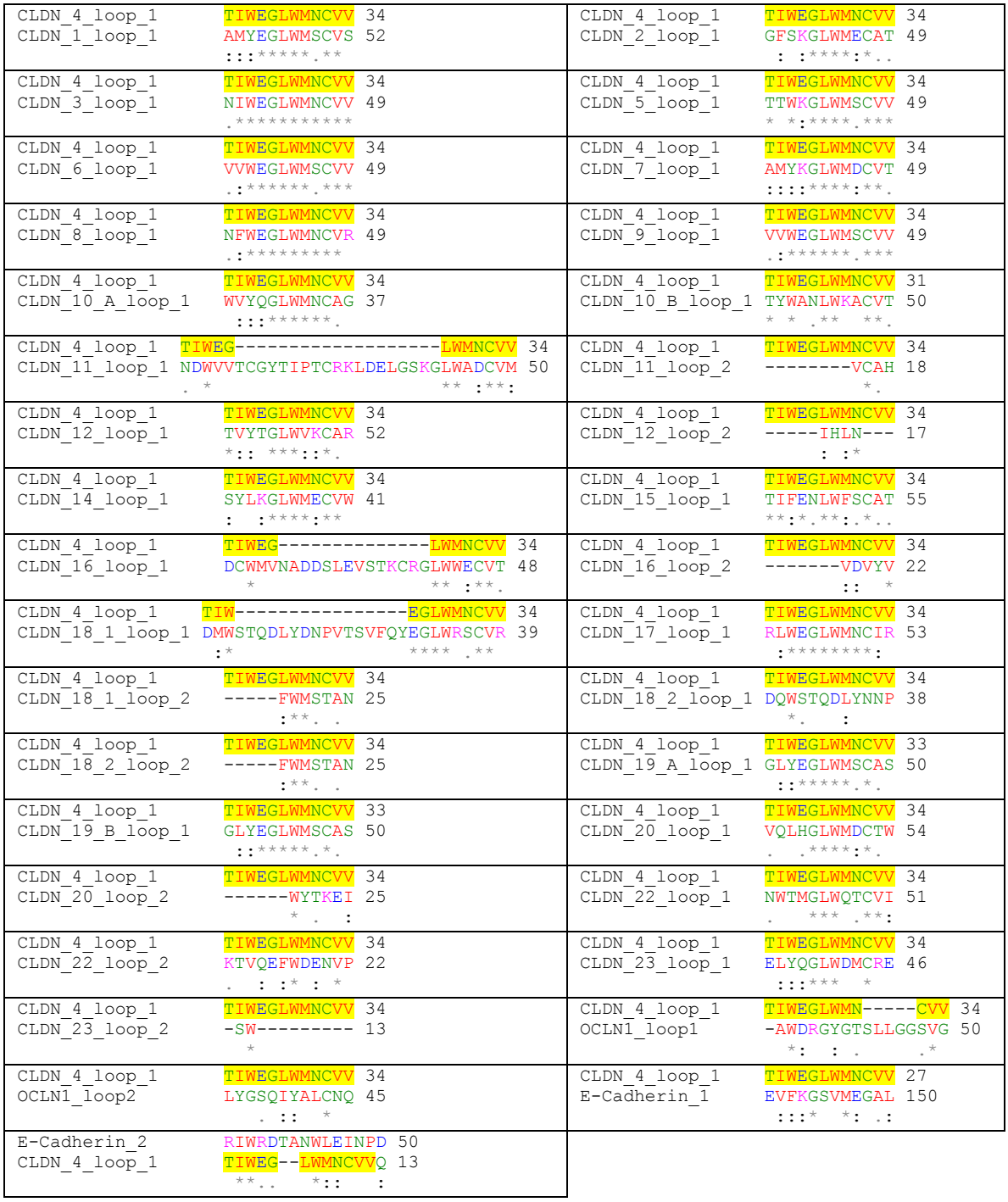


Figure 5. ClustalW alignment of the unique antigenic CLDN4 loop 1 sequence. The highlighted sequence was aligned with extracellular regions of claudins, occludin, and E-cadherin. This sequence is found on our GST-CLDN4 fusion protein.

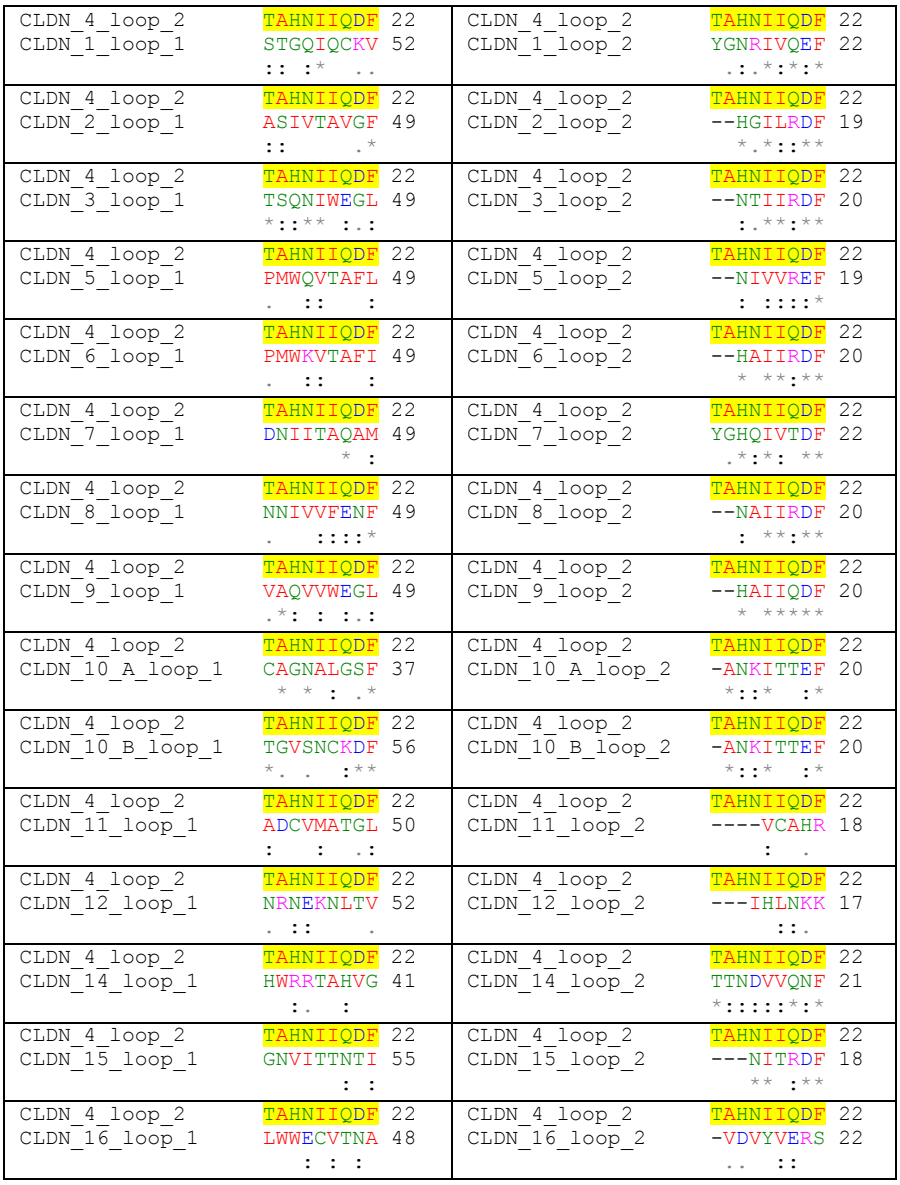


Figure 6. ClustalW alignment of the unique antigenic CLDN4 loop 2 sequence. Legend in next page

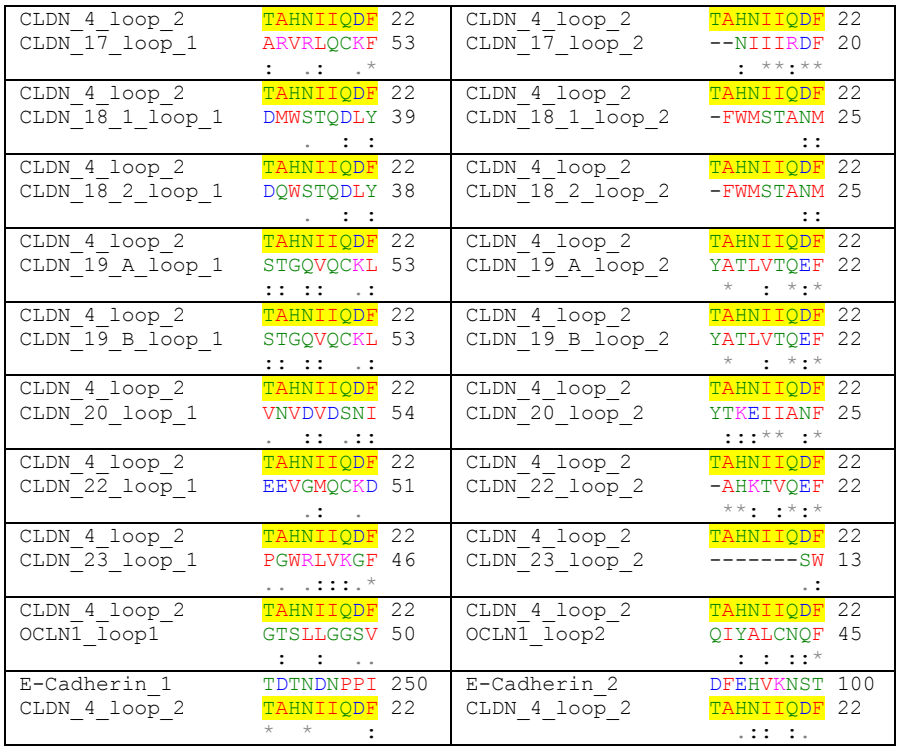


Figure 6. ClustalW alignment of the unique antigenic CLDN4 loop 2 sequence. The highlighted sequence was aligned with extracellular regions of claudins, occludin, and E-cadherin.

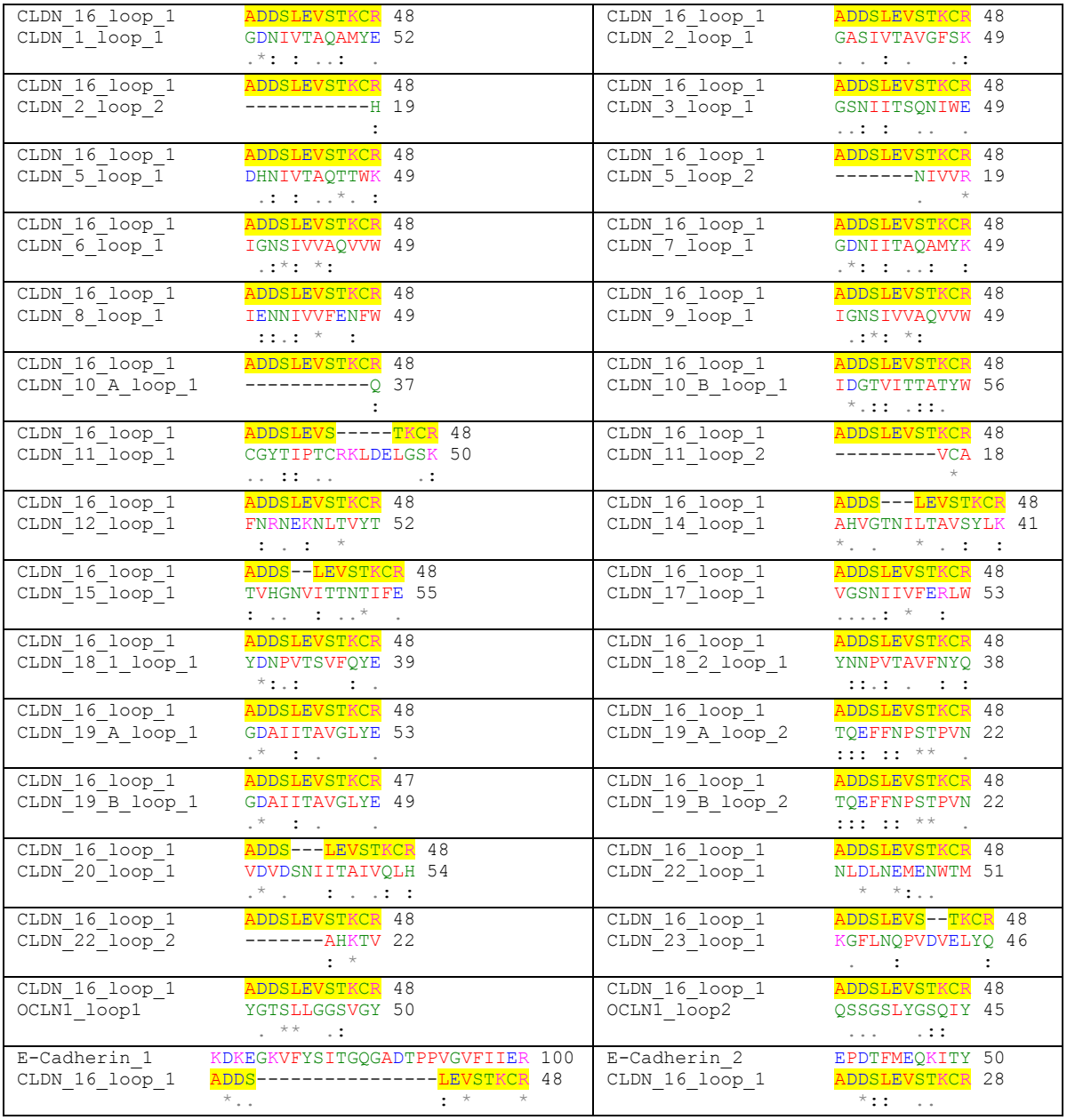


Figure 7. ClustalW alignment of a unique antigenic CLDN16 loop 1 sequence. The highlighted sequence was aligned with extracellular regions of claudins, occludin, and E-cadherin. This sequence is found on our GST-CLDN16 fusion protein.

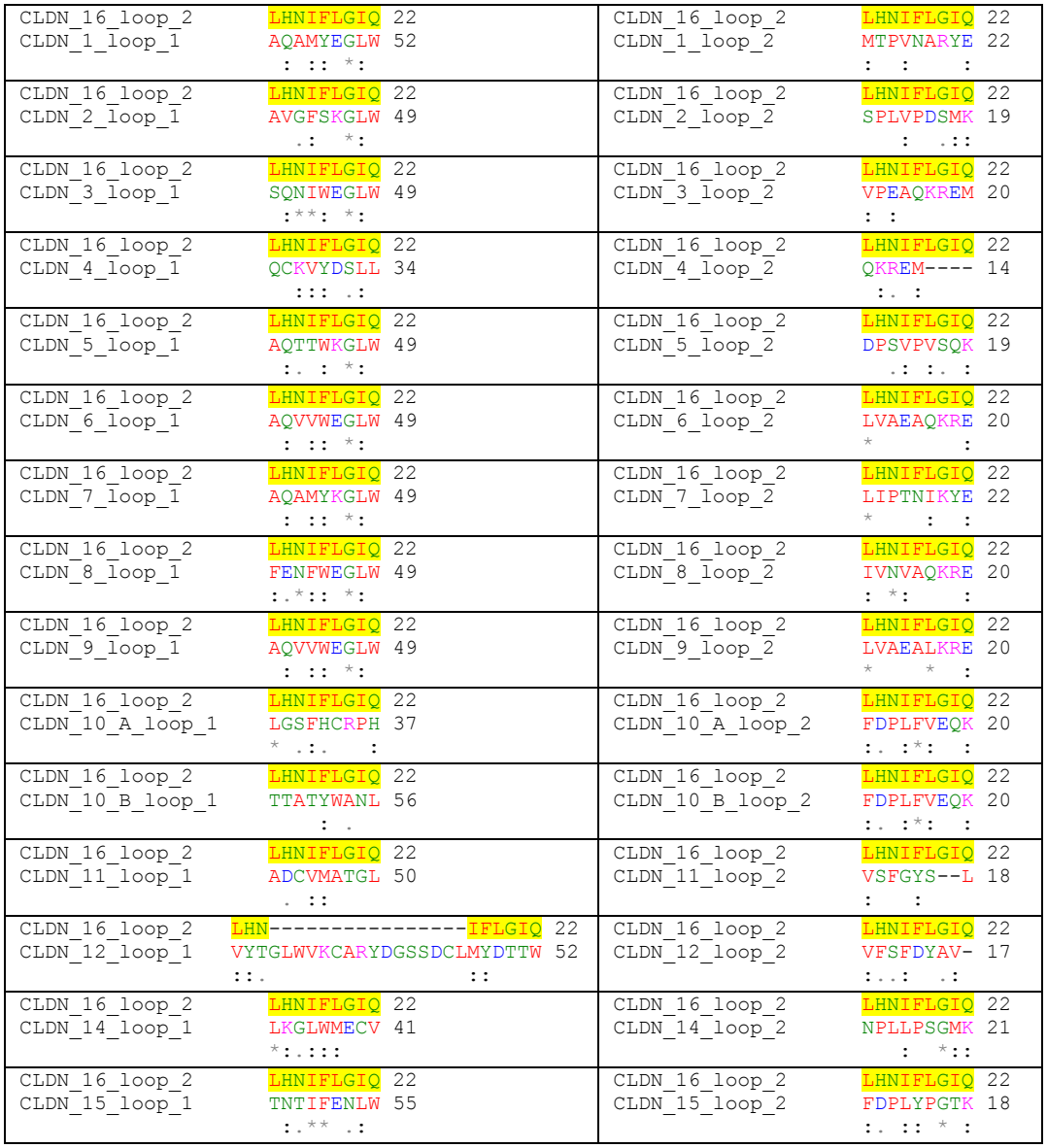


Figure 8. ClustalW alignment of a unique antigenic CLDN16 loop 2 sequence. Legend in next page

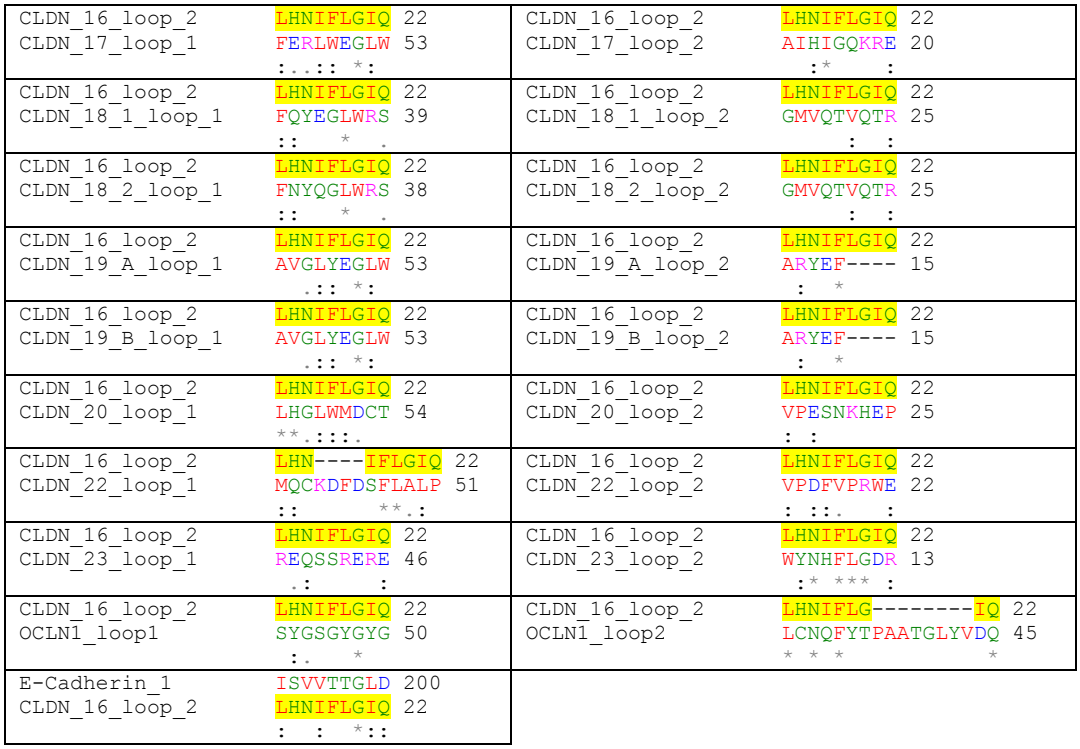


Figure 8. ClustalW alignment of a unique antigenic CLDN16 loop 2 sequence. The highlighted sequence was aligned with extracellular regions of claudins, occludin, and E-cadherin.

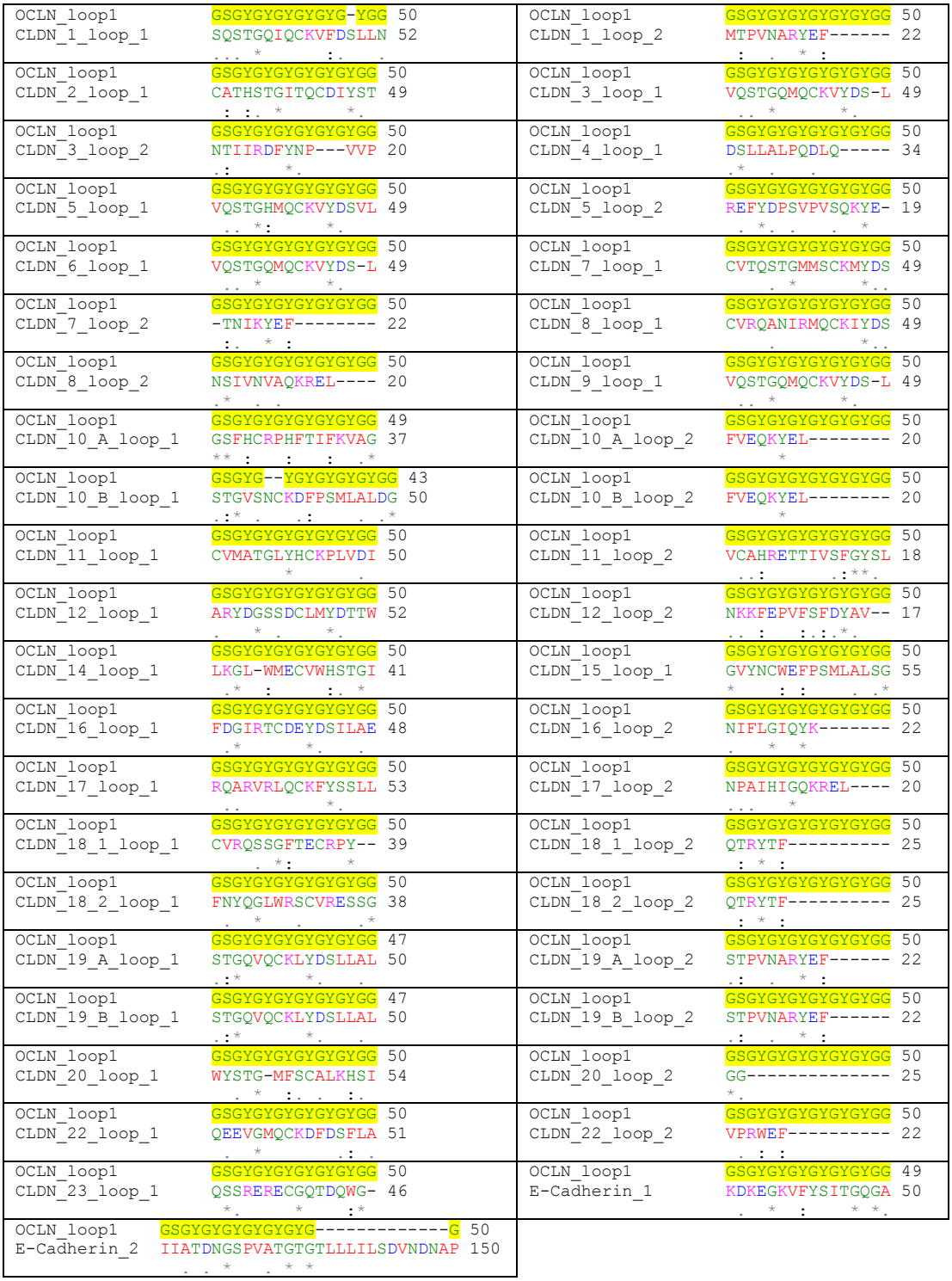


Figure 9. ClustalW alignment of a unique antigenic OCLN loop 1 sequence. The highlighted sequence was aligned with extracellular regions of claudins, occludin, and E-cadherin. This sequence is found on our GST-OCLN fusion protein.

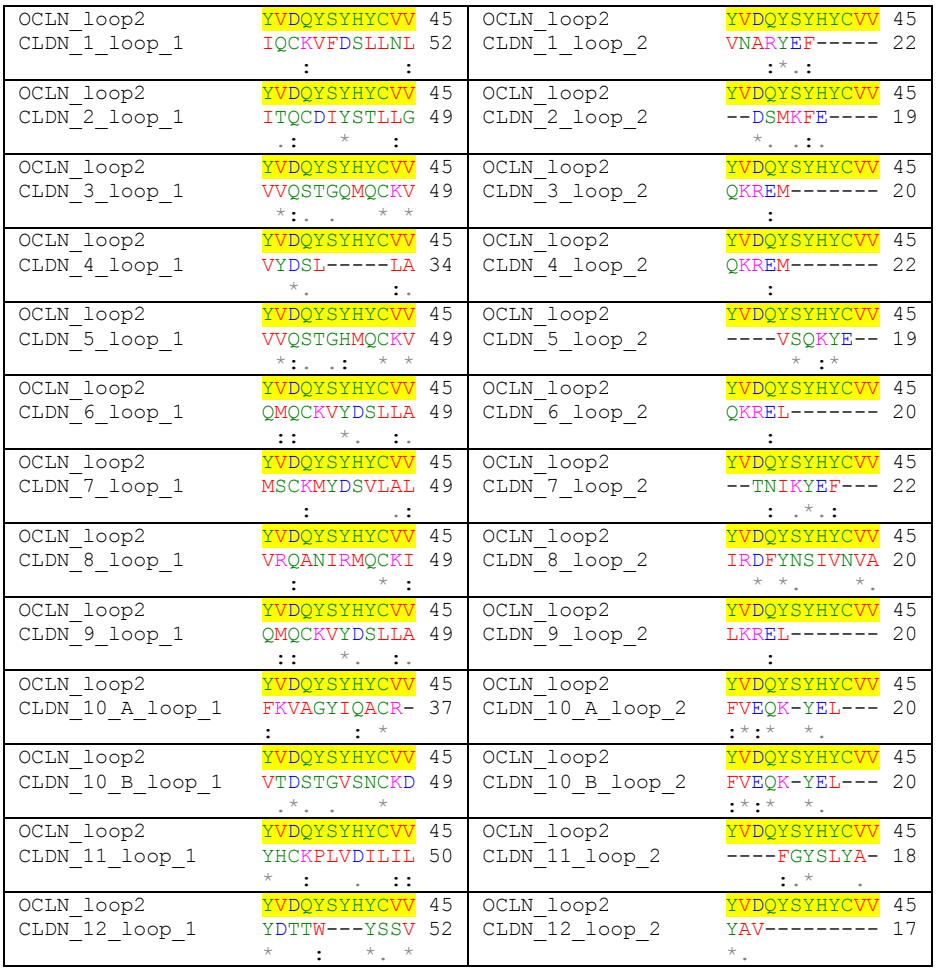


Figure 10. ClustalW alignment of a unique antigenic OCLN loop 2 sequence. Legend in next page.

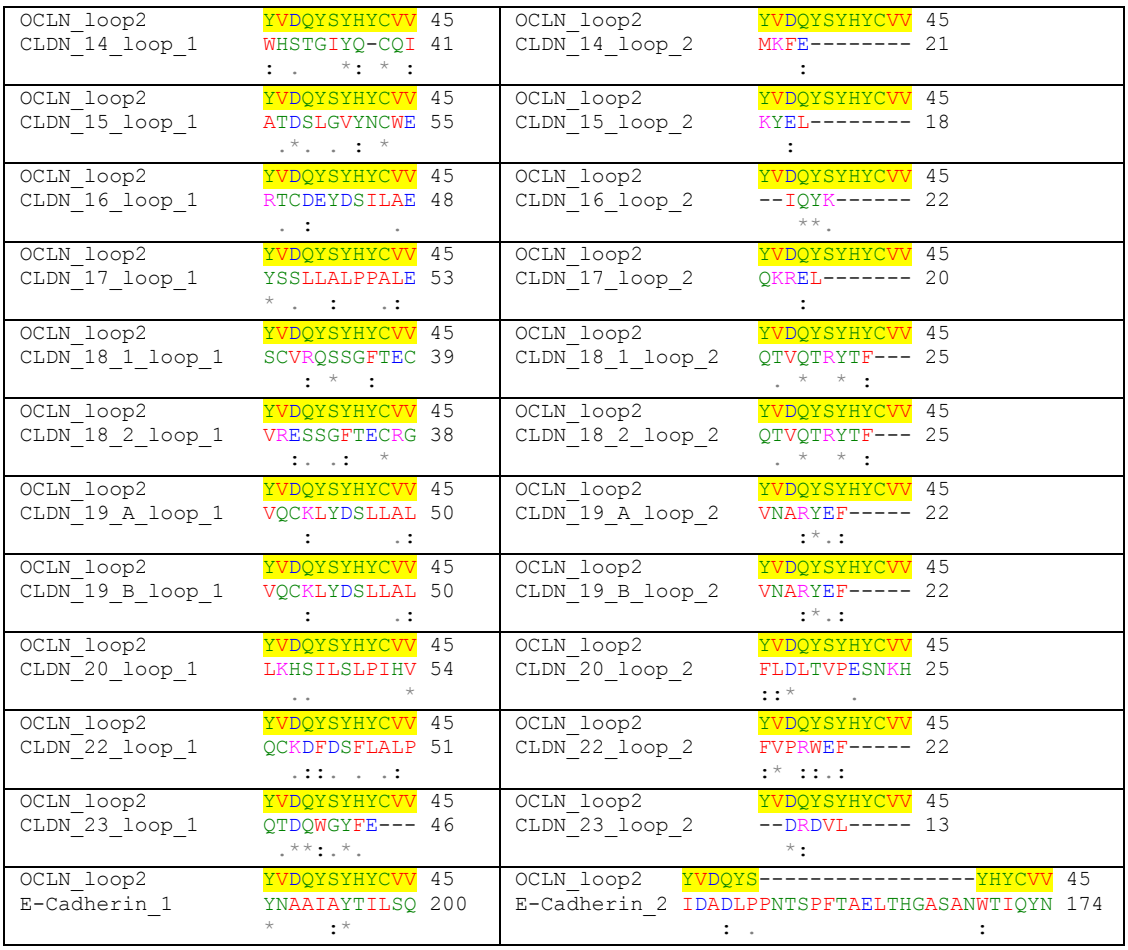


Figure 10. ClustalW alignment of a unique antigenic OCLN loop 2 sequence. The highlighted sequence was aligned with extracellular regions of claudins, occludin, and E-cadherin. A fusion protein was created for loop 2, but it was not purified in large amounts.

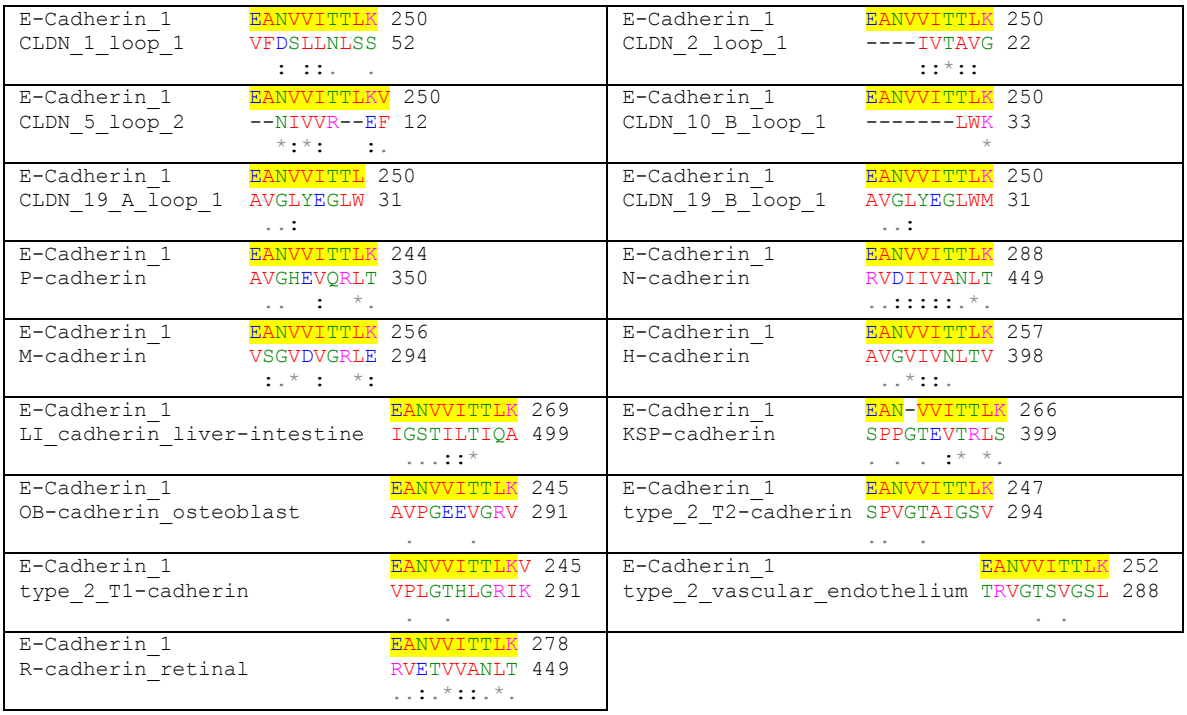


Figure 11. ClustalW alignment of a unique antigenic E-cadherin extracellular sequence. The highlighted sequence was aligned with extracellular regions of claudins, occludin, and other cadherins. This sequence is found on our GST-CDH1 fusion protein.

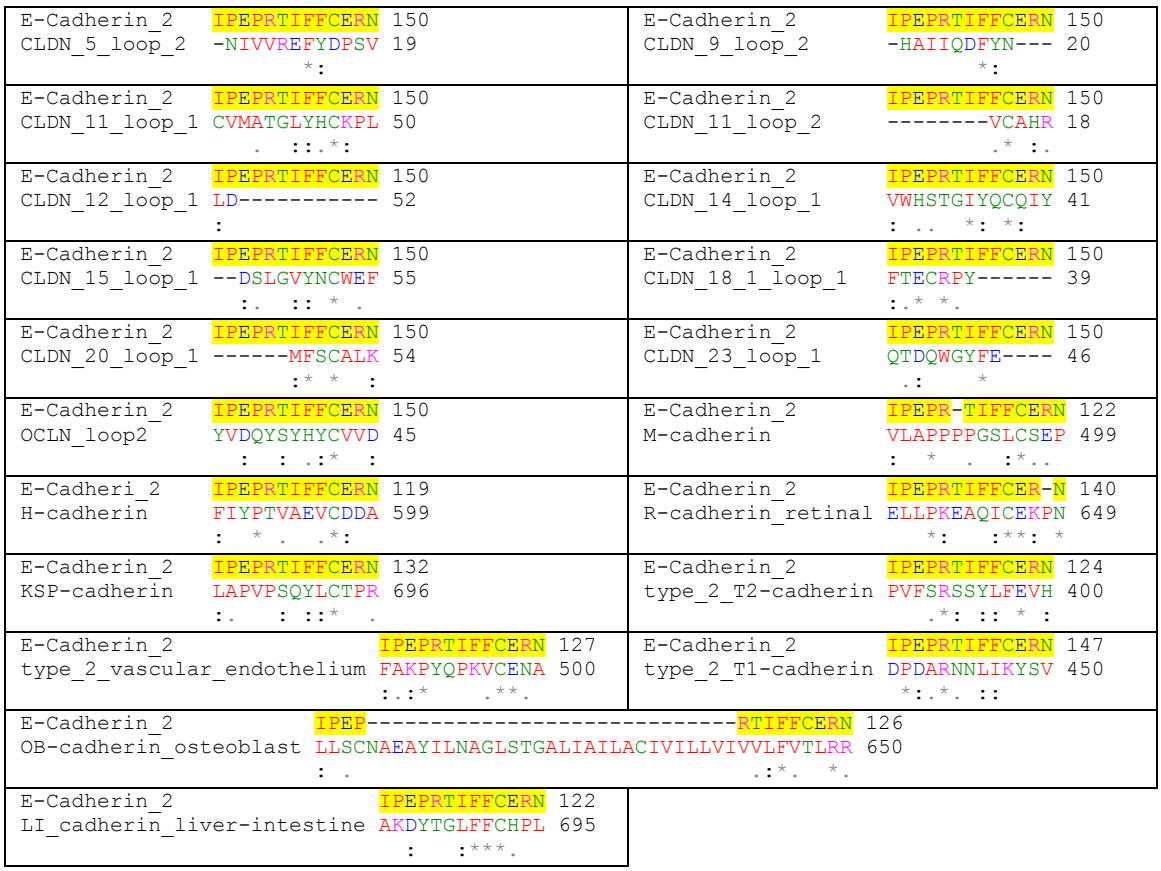


Figure 12. ClustalW alignment of a unique antigenic E-cadherin extracellular sequence. The highlighted sequence was aligned with extracellular regions of claudins, occludin, and other cadherins.

Peptide Targets	
Protein	**Sequence**
CLDN1 loop 1	LNLSSTLQA
CLDN1 loop 2	YGNRIVQEF
CLDN4 loop 1	TIWEGLWMNCVV
CLDN4 loop 2	TAHNIIQDF
CLDN16 loop 1	ADDSLEVSTKCR
CLDN16 loop 2	LHNIFLGIQ
OCLN loop 1	GSGYGYGYGYGYGYGG
OCLN loop 2	YVDQYSYHYCVV
CDH1 extracellular 1	EANVVITTLK
CDH1 extracellular 2	IPEPRTIFFCERN

Figure 13. **Unique extracellular antigenic targets for raising peptide antibodies.** Sequences were selected based on ClustalW alignments of extracellular regions with similar proteins as well as BLAST analysis with all human proteins.

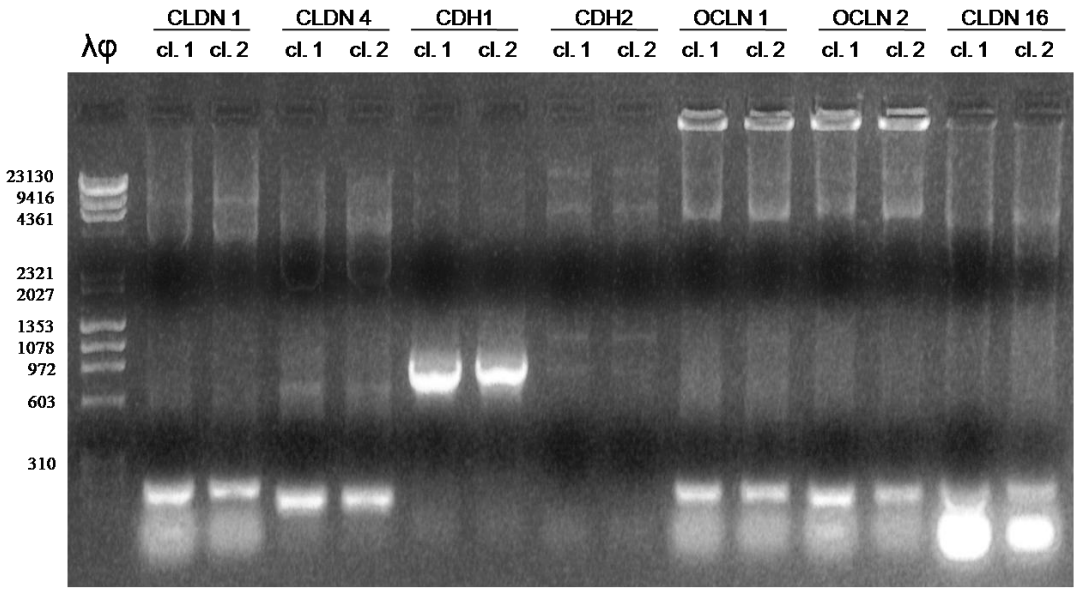


Figure 14. **PCR amplification of antigenic regions.** All of the regions were successfully amplified with the exception of CDH2.

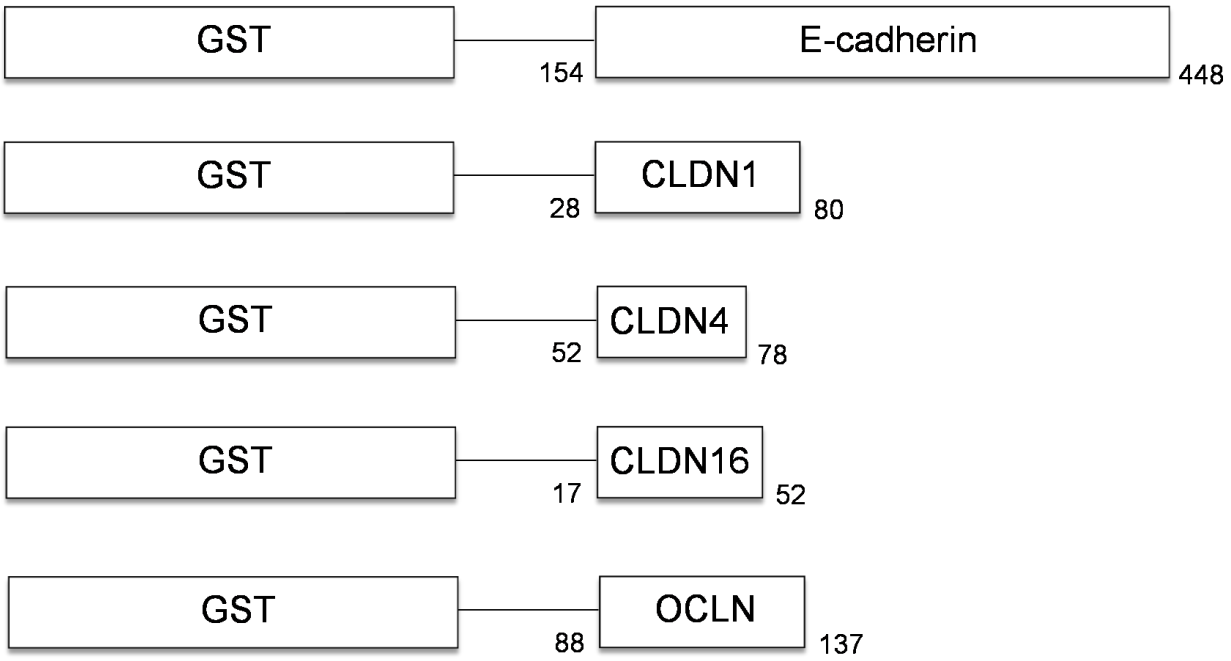


Figure 15. Diagrammatic representation of fusion proteins. These are the fusion proteins that were purified with the intent of raising antibodies. The approximate relative sizes of the fusion proteins are shown. The numbers represent the location of the antigenic fragment on the original protein.

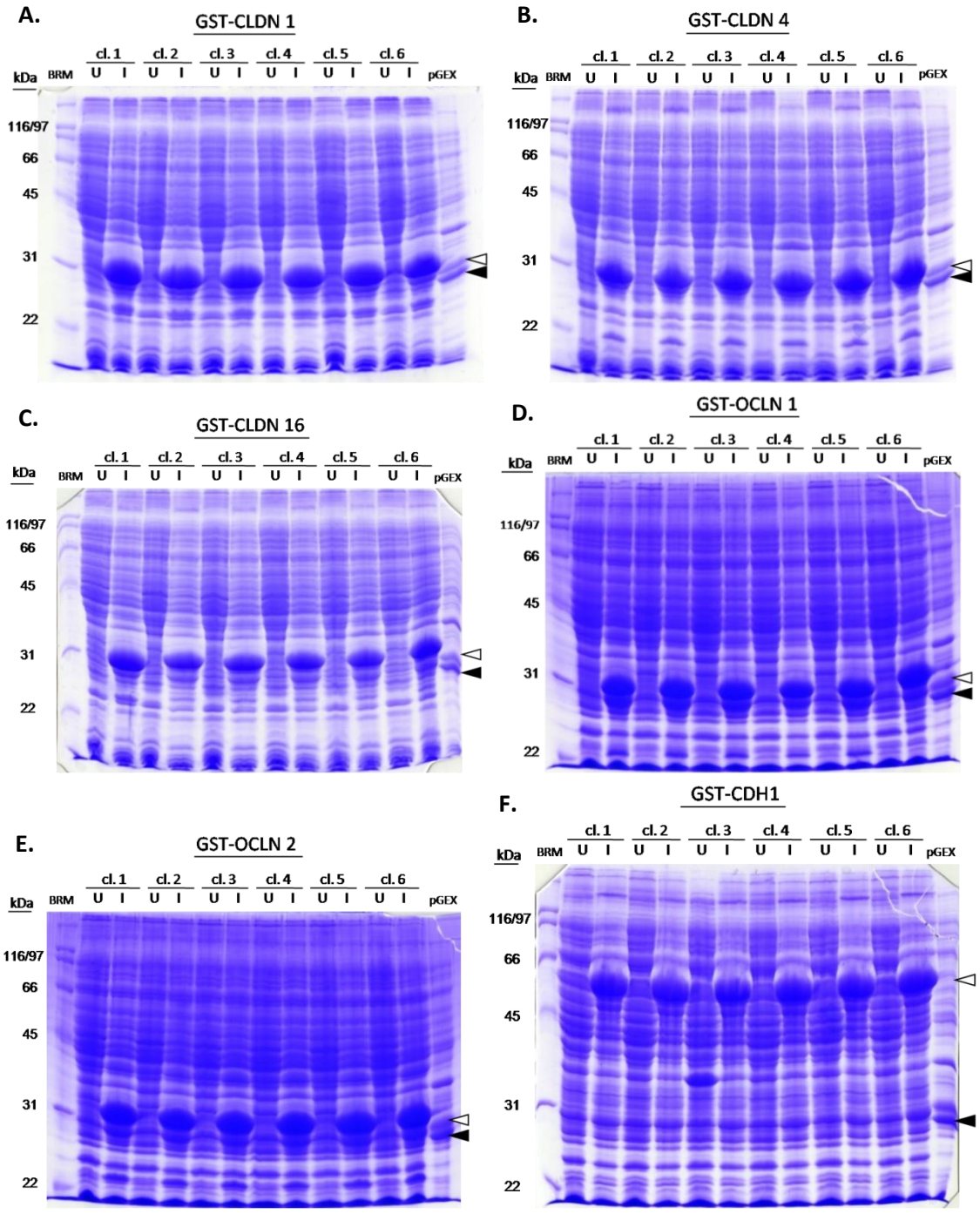


Figure 16. Induced expression of GST fusion proteins in BL21 cells. "U" stands for uninduced and "I" stands for induced. The white arrowhead points to the fusion protein and the black arrowhead points to the GST band. A) GST-CLDN1 inductions, B) GST-CLDN4 inductions, C) GST-CLDN16 inductions, D) GST-OCLN 1 inductions, E) GST-OCLN 2 inductions, F) GST-CDH1 inductions. All clones showed strong expression of the fusion proteins.

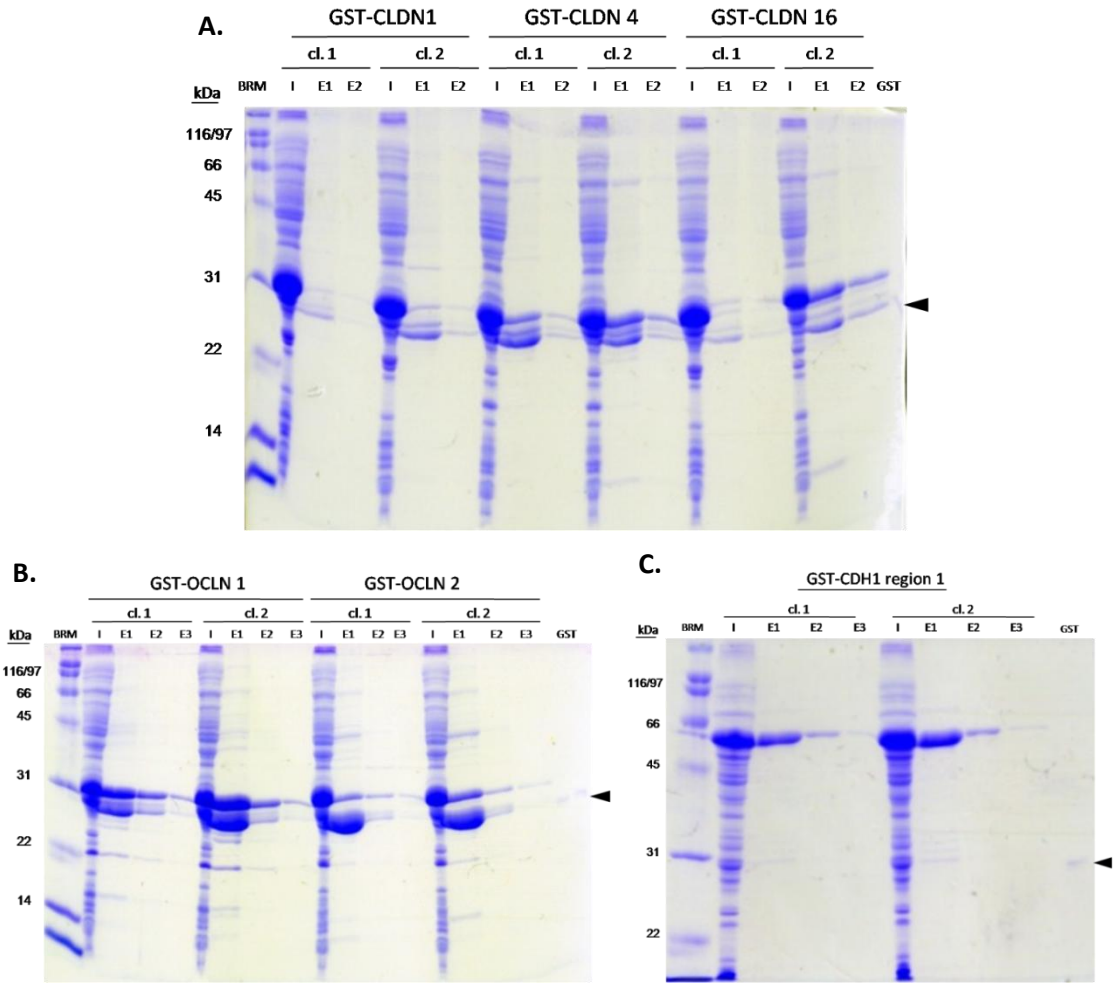


Figure 17. Initial purifications of GST fusion proteins. Samples were passed through a GSH affinity bead column. For each fusion protein, two clones from a successful induction were compared. "I" stands for the induced BL21 cells previously analyzed in figure#. The elutions are labeled as "E#". The black arrowhead indicates the GST band. A) GST-CLDN1, GST-CLDN4, and GST-CLDN16 purifications. Two clones were used for each protein. B) GST-OCLN1 and GST-OCLN2 purifications comparing two clones for each protein. C) GST-CDH1 purifications comparing two clones. With the exception of the GST-CDH1 purification, there was considerable cleavage of the GST tag from the fusion proteins in the initial purifications.

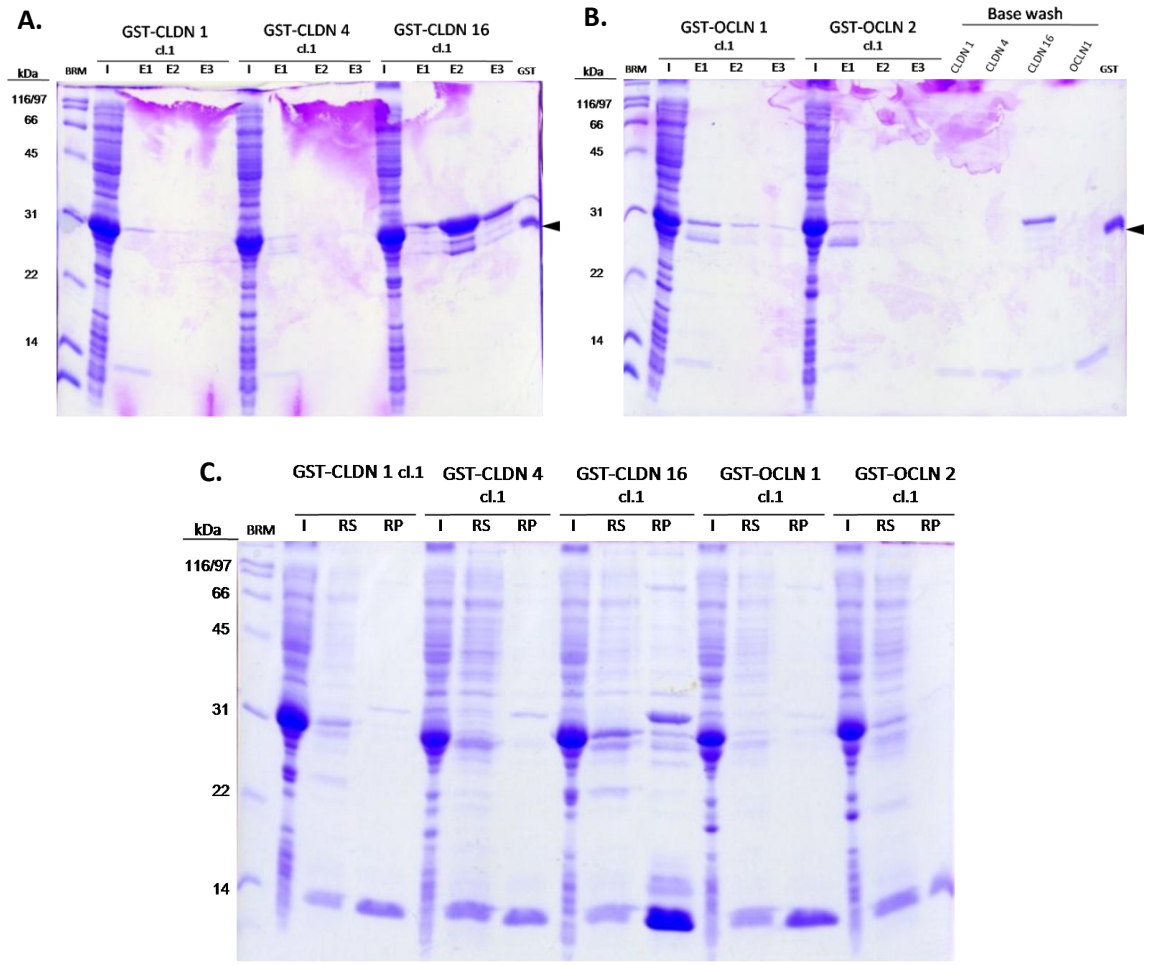


Figure 18. Protein refolding of sample previously passed through GSH affinity column. Refolding was performed in order to recover protein that went to inclusion bodies after centrifuging the sonicated extract. Following refolding, the protein was passed through a GSH affinity column and elutions (E1, E2, E3) were collected. Figure 17 shows the protein purifications prior to refolding. Black arrowheads indicate GST bands. "I" stands for the induced protein sample. A) CLDN1, CLDN4, and CLDN16 GST fusion protein purifications after refolding. B) GST-OCLN loop 1 and loop 2 fusion proteins after refolding. Base washes from column regeneration were also analyzed for protein. A considerable amount of CLDN16 was present in the base wash. C) Analysis of the unpurified supernatant (RS) of each sample following refolding as well as a pellet (RP) obtained from centrifuging the supernatant.

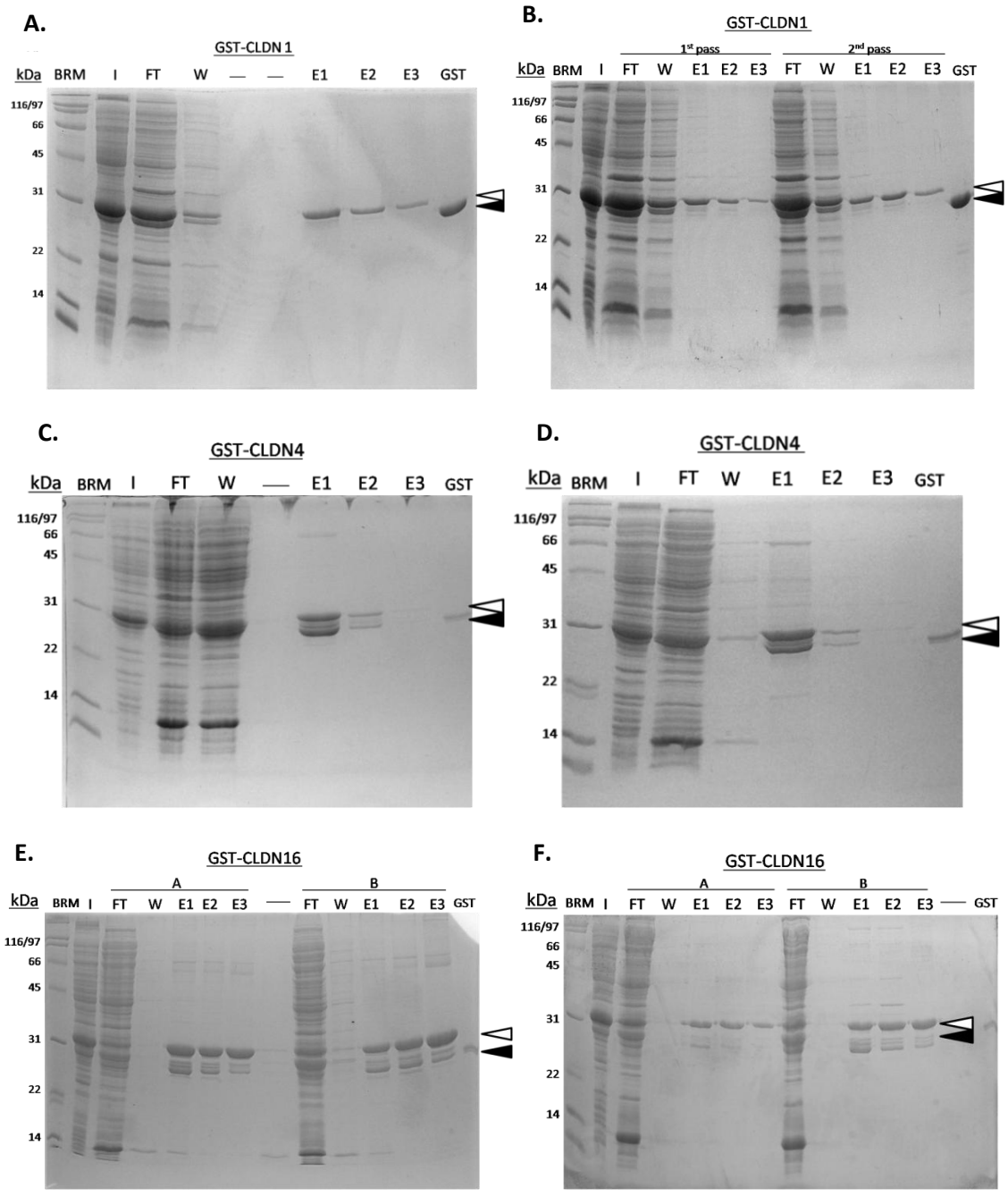


Figure 19. Purifications of GST-CLDN fusion proteins. "I" stands for the induced sample. "FT" is the sample after passing through the GSH affinity column. "W" is the wash transition prior to collecting elutions. Three elutions (E1, E2, E3) were collected for each protein. The black arrowhead points to the GST band while the white arrowhead points to the fusion protein. A) Purification of GST-CLDN1. B) GST-CLDN1 sample passed through the GSH affinity column twice. C, D) GST-CLDN4 purifications. E, F) GST-CLDN16 purifications of different samples. "A" and "B" represent proteins collected from separate cultures.

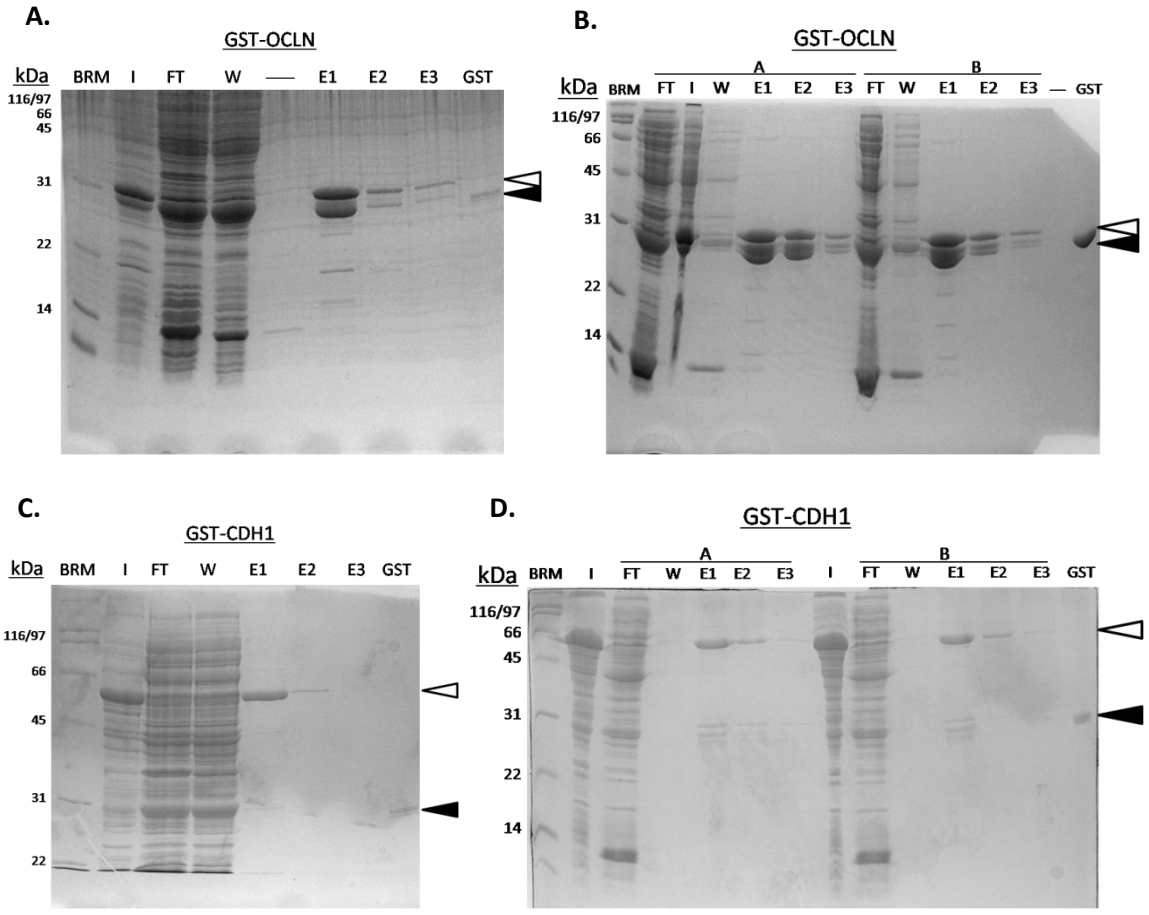


Figure 20. Purifications of GST-OCLN and GST-CDH1 fusion proteins. "I" stands for the induced sample. "FT" is the sample after passing through the GSH affinity column. "W" is the wash phase prior to collecting elutions. Three elutions (E1, E2, E3) were collected for each protein. The black arrowhead points to the GST band while the white arrowhead points to the fusion protein. A) GST-OCLN purification with noticeable free GST that was cleaved from the fusion protein. B) Two (A, B) different GST-OCLN purifications showing fusion proteins along with cleaved GST tag. C) Purification of GST-CDH1 protein. D) Two (A, B) different GST-CDH1 purifications.

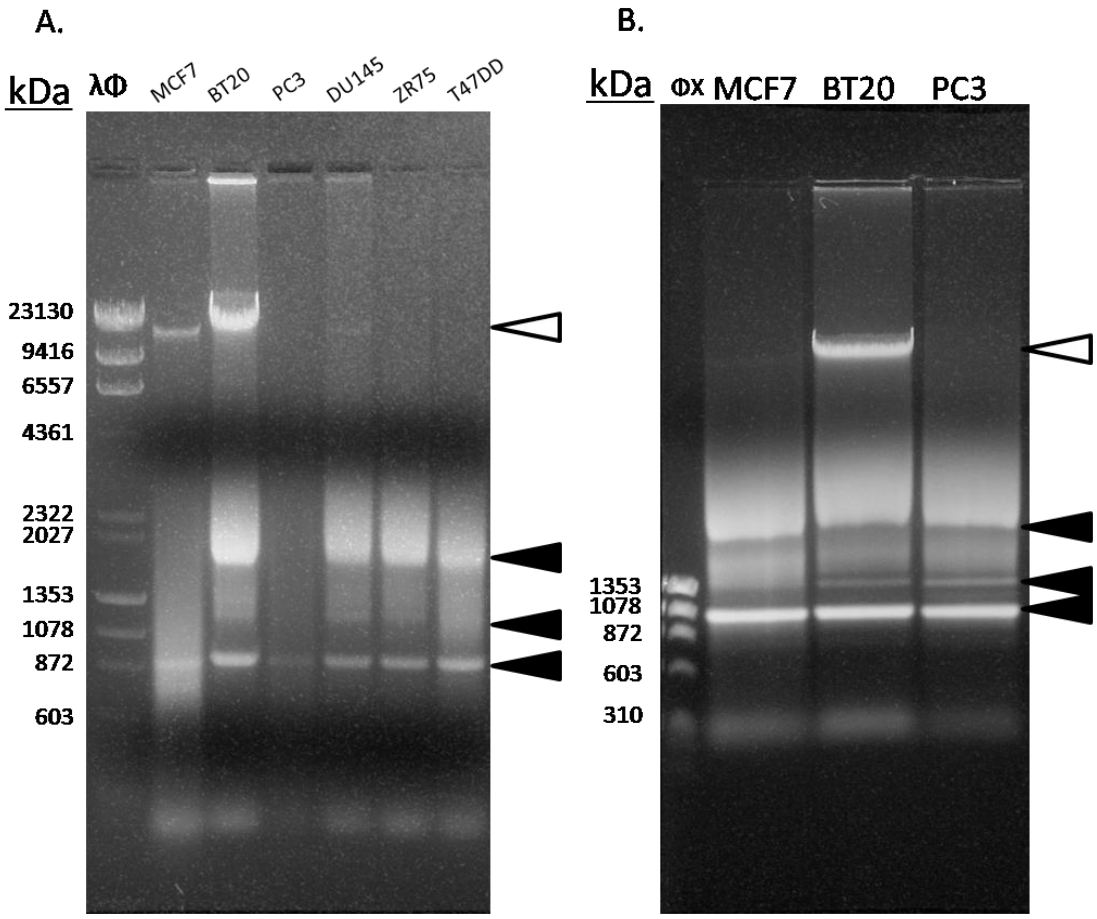


Figure 21. Analysis of RNA extracted from cancer cell lines. Intact clean RNA was used for cDNA synthesis. The black arrowheads point to 28S, 18S, and 5.8S subunits of the ribosomal RNA and the white arrowhead points to contamination. The smear is the mRNA extracted from the cancer cells. A) RNA extraction of MCF7, BT20, PC3, DU145, ZR75, T47DD. MCF7, BT20, and PC3 were contaminated. B) Second RNA extraction for MCF7, BT20, and PC3. There seemed to be an issue with contamination of the BT20 sample in all of the preparations.

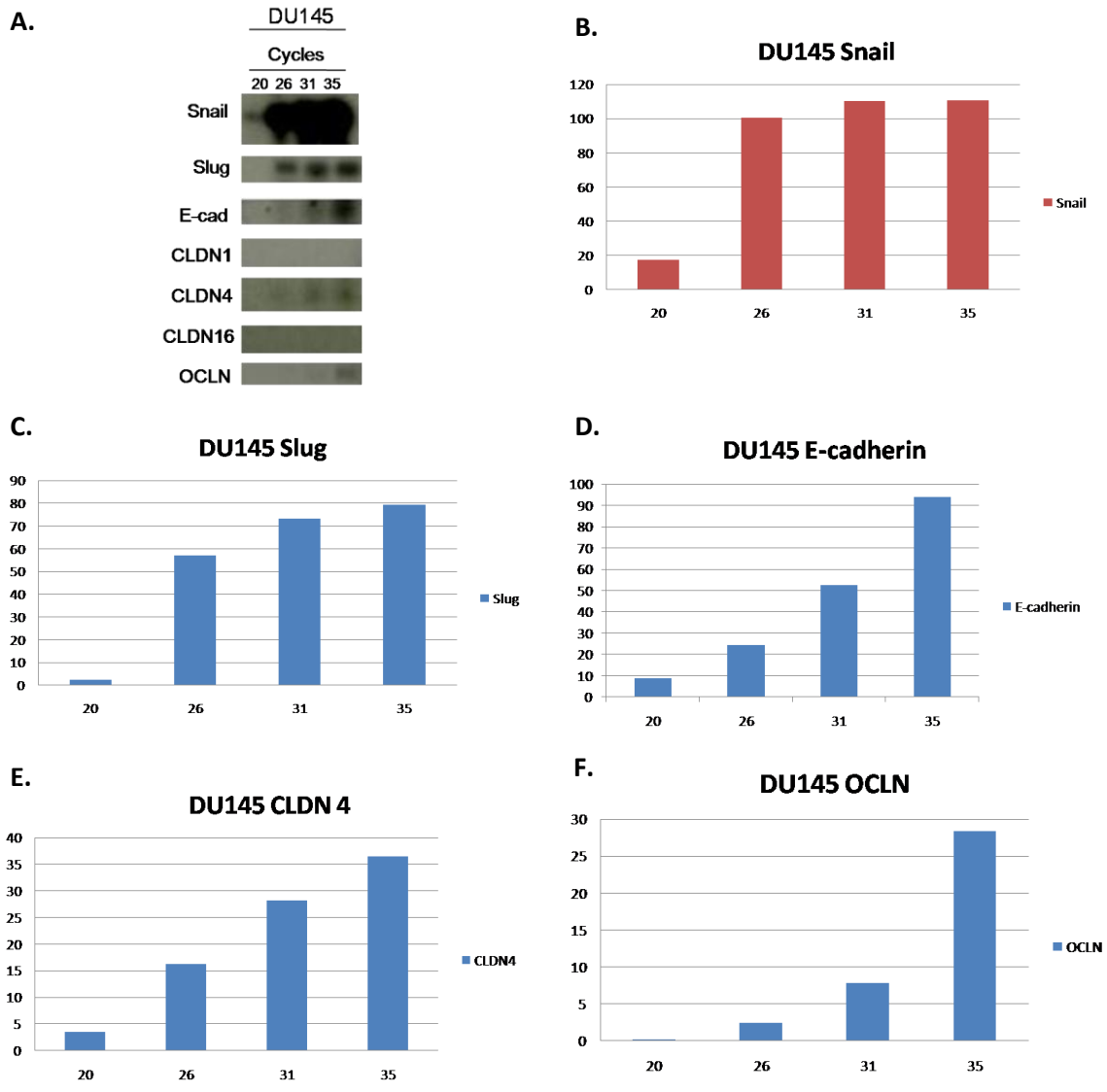


Figure 22. Expression profiling of DU145 metastatic cancer cell line. A) Aliquots from the PCR reaction were taken at 20, 26, 31, and 35 cycles and analyzed with a Southern blot. B) DU145 snail intensity. C) DU145 slug intensity. D) DU145 E-cadherin intensity. E) DU145 CLDN4 intensity. F) DU145 OCLN intensity. The x-axis shows an increase in cycles while the intensity of each band relative to the background is shown on the y-axis.

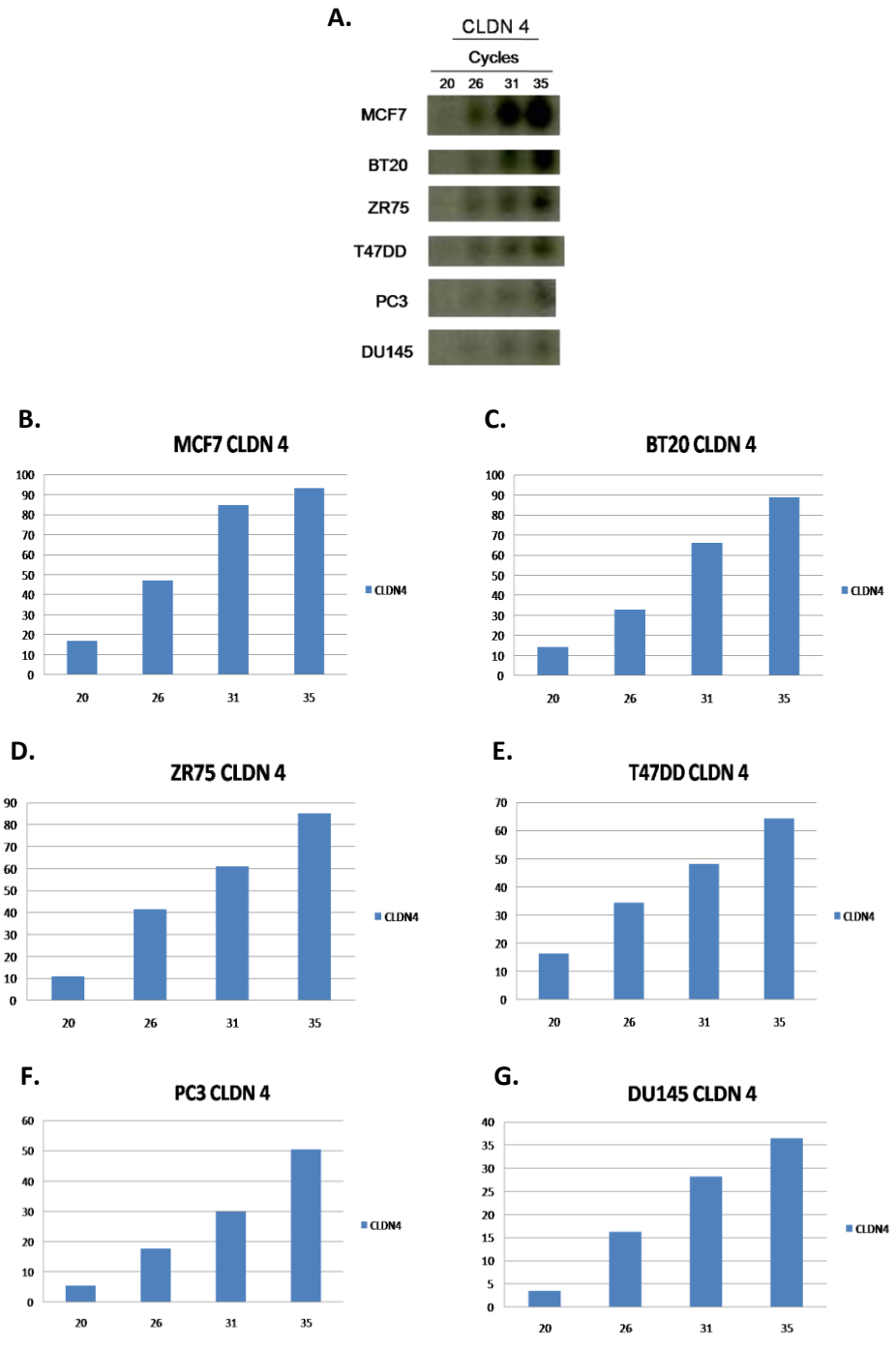


Figure 23. Expression profiling of CLDN4 in cancer cell lines. A) Aliquots from the PCR reaction were taken at 20, 26, 31, and 35 cycles and analyzed with a Southern blot. B) DU145 snail intensity. C) DU145 slug intensity. D) DU145 E-cadherin intensity. E) DU145 CLDN4 intensity. The x-axis shows an increase in cycles while the intensity of each band relative to the background is shown on the y-axis.

WORKS CITED

Agarwal, R., D'Souza, T., and Morin, P. J. (2005). Claudin-3 and claudin-4 expression in ovarian epithelial cells enhances invasion and is associated with increased matrix metalloproteinase-2 activity. *Cancer Res* **65,** 7378-85.

Ayyanathan, K., Peng, H., Hou, Z., Fredericks, W. J., Goyal, R. K., Langer, E. M., Longmore, G. D., and Rauscher, F. J., 3rd. (2007). The Ajuba LIM domain protein is a corepressor for SNAG domain mediated repression and participates in nucleocytoplasmic Shuttling. *Cancer Res* **67,** 9097-106.

Batlle, E., Sancho, E., Franci, C., Dominguez, D., Monfar, M., Baulida, J., and Garcia De Herreros, A. (2000). The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat Cell Biol* **2,** 84-9.

Cano, A., Perez-Moreno, M. A., Rodrigo, I., Locascio, A., Blanco, M. J., del Barrio, M. G., Portillo, F., and Nieto, M. A. (2000). The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat Cell Biol* **2,** 76-83.

Chan, A. O., Lam, S. K., Chu, K. M., Lam, C. M., Kwok, E., Leung, S. Y., Yuen, S. T., Law, S. Y., Hui, W. M., Lai, K. C., Wong, C. Y., Hu, H. C., Lai, C. L., and Wong, J. (2001). Soluble E-cadherin is a valid prognostic marker in gastric carcinoma. *Gut* **48,** 808-11.

De Angelis, G., Rittenhouse, H. G., Mikolajczyk, S. D., Blair Shamel, L., and Semjonow, A. (2007). Twenty Years of PSA: From Prostate Antigen to Tumor Marker. *Rev Urol* **9,** 113-23.

Dhawan, P., Singh, A. B., Deane, N. G., No, Y., Shiou, S. R., Schmidt, C., Neff, J., Washington, M. K., and Beauchamp, R. D. (2005). Claudin-1 regulates cellular transformation and metastatic behavior in colon cancer. *J Clin Invest* **115,** 1765-76.

Ebihara, C., Kondoh, M., Hasuike, N., Harada, M., Mizuguchi, H., Horiguchi, Y., Fujii, M., and Watanabe, Y. (2006). Preparation of a claudin-targeting molecule using a C-terminal fragment of Clostridium perfringens enterotoxin. *J Pharmacol Exp Ther* **316,** 255-60.

Fujita, K., Katahira, J., Horiguchi, Y., Sonoda, N., Furuse, M., and Tsukita, S. (2000). Clostridium perfringens enterotoxin binds to the second extracellular loop of claudin-3, a tight junction integral membrane protein. *FEBS Lett* **476,** 258-61.

Grimes, H. L., Gilks, C. B., Chan, T. O., Porter, S., and Tsichlis, P. N. (1996). The Gfi-1 protooncoprotein represses Bax expression and inhibits T-cell death. *Proc Natl Acad Sci U S A* **93,** 14569-73.

Herschman, J. D., Smith, D. S., and Catalona, W. J. (1997). Effect of ejaculation on serum total and free prostate-specific antigen concentrations. *Urology* **50,** 239-43.

Hewitt, K. J., Agarwal, R., and Morin, P. J. (2006). The claudin gene family: expression in normal and neoplastic tissues. *BMC Cancer* **6,** 186.

Hou, Z., Peng, H., Ayyanathan, K., Yan, K. P., Langer, E. M., Longmore, G. D., and Rauscher, F. J., 3rd. (2008). The LIM protein AJUBA recruits protein arginine methyltransferase 5 to mediate SNAIL-dependent transcriptional repression. *Mol Cell Biol* **28,** 3198-207.

Ikenouchi, J., Matsuda, M., Furuse, M., and Tsukita, S. (2003). Regulation of tight junctions during the epithelium-mesenchyme transition: direct repression of the gene expression of claudins/occludin by Snail. *J Cell Sci* **116,** 1959-67.

James, C. R., Quinn, J. E., Mullan, P. B., Johnston, P. G., and Harkin, D. P. (2007). BRCA1, a potential predictive biomarker in the treatment of breast cancer. *Oncologist* **12,** 142-50.

Katahira, J., Inoue, N., Horiguchi, Y., Matsuda, M., and Sugimoto, N. (1997). Molecular cloning and functional characterization of the receptor for Clostridium perfringens enterotoxin. *J Cell Biol* **136,** 1239-47.

Kominsky, S. L., Tyler, B., Sosnowski, J., Brady, K., Doucet, M., Nell, D., Smedley, J. G., 3rd, McClane, B., Brem, H., and Sukumar, S. (2007). Clostridium perfringens enterotoxin as a novel-targeted therapeutic for brain metastasis. *Cancer Res* **67,** 7977-82.

Kulka, J., Szasz, A. M., Nemeth, Z., Madaras, L., Schaff, Z., Molnar, I. A., and Tokes, A. M. (2009). Expression of tight junction protein claudin-4 in basal-like breast carcinomas. *Pathol Oncol Res* **15,** 59-64.

Kurrey, N. K., K, A., and Bapat, S. A. (2005). Snail and Slug are major determinants of ovarian cancer invasiveness at the transcription level. *Gynecol Oncol* **97,** 155-65.

Landers, K. A., Samaratunga, H., Teng, L., Buck, M., Burger, M. J., Scells, B., Lavin, M. F., and Gardiner, R. A. (2008). Identification of claudin-4 as a marker highly overexpressed in both primary and metastatic prostate cancer. *Br J Cancer* **99,** 491-501.

Li, D., and Mrsny, R. J. (2000). Oncogenic Raf-1 disrupts epithelial tight junctions via downregulation of occludin. *J Cell Biol* **148,** 791-800.

Ling, J., Liao, H., Clark, R., Wong, M. S., and Lo, D. D. (2008). Structural constraints for the binding of short peptides to claudin-4 revealed by surface plasmon resonance. *J Biol Chem* **283,** 30585-95.

Martinez-Estrada, O. M., Culleres, A., Soriano, F. X., Peinado, H., Bolos, V., Martinez, F. O., Reina, M., Cano, A., Fabre, M., and Vilaro, S. (2006). The transcription factors Slug and Snail act as repressors of Claudin-1 expression in epithelial cells. *Biochem J* **394,** 449-57.

Matsuda, M., Kubo, A., Furuse, M., and Tsukita, S. (2004). A peculiar internalization of claudins, tight junction-specific adhesion molecules, during the intercellular movement of epithelial cells. *J Cell Sci* **117,** 1247-57.

McClane, B. A. (2000). Clostridium perfringens enterotoxin and intestinal tight junctions. *Trends Microbiol* **8,** 145-6.

McClane, B. A., and Chakrabarti, G. (2004). New insights into the cytotoxic mechanisms of Clostridium perfringens enterotoxin. *Anaerobe* **10,** 107-14.

Michl, P., Buchholz, M., Rolke, M., Kunsch, S., Lohr, M., McClane, B., Tsukita, S., Leder, G., Adler, G., and Gress, T. M. (2001). Claudin-4: a new target for pancreatic cancer treatment using Clostridium perfringens enterotoxin. *Gastroenterology* **121,** 678-84.

Morohashi, S., Kusumi, T., Sato, F., Odagiri, H., Chiba, H., Yoshihara, S., Hakamada, K., Sasaki, M., and Kijima, H. (2007). Decreased expression of claudin-1 correlates with recurrence status in breast cancer. *Int J Mol Med* **20,** 139-43.

Nadler, R. B., Humphrey, P. A., Smith, D. S., Catalona, W. J., and Ratliff, T. L. (1995). Effect of inflammation and benign prostatic hyperplasia on elevated serum prostate specific antigen levels. *J Urol* **154,** 407-13.

Nawrocki-Raby, B., Gilles, C., Polette, M., Bruyneel, E., Laronze, J. Y., Bonnet, N., Foidart, J. M., Mareel, M., and Birembaut, P. (2003). Upregulation of MMPs by soluble E-cadherin in human lung tumor cells. *Int J Cancer* **105,** 790-5.

Oku, N., Sasabe, E., Ueta, E., Yamamoto, T., and Osaki, T. (2006). Tight junction protein claudin-1 enhances the invasive activity of oral squamous cell carcinoma cells by promoting cleavage of laminin-5 gamma2 chain via matrix metalloproteinase (MMP)-2 and membrane-type MMP-1. *Cancer Res* **66,** 5251-7.

Onder, T. T., Gupta, P. B., Mani, S. A., Yang, J., Lander, E. S., and Weinberg, R. A. (2008). Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. *Cancer Res* **68,** 3645-54.

Riggins, R. B., Schrecengost, R. S., Guerrero, M. S., and Bouton, A. H. (2007). Pathways to tamoxifen resistance. *Cancer Lett* **256,** 1-24.

Sambrook, J., and Russell, D. W. (2001). "Molecular cloning : a laboratory manual." Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Santin, A. D., Cane, S., Bellone, S., Palmieri, M., Siegel, E. R., Thomas, M., Roman, J. J., Burnett, A., Cannon, M. J., and Pecorelli, S. (2005). Treatment of chemotherapy-resistant human ovarian cancer xenografts in C.B-17/SCID mice by intraperitoneal administration of Clostridium perfringens enterotoxin. *Cancer Res* **65,** 4334-42.

Soini, Y. (2005). Expression of claudins 1, 2, 3, 4, 5 and 7 in various types of tumours. *Histopathology* **46,** 551-60.

Soini, Y., Tommola, S., Helin, H., and Martikainen, P. (2006). Claudins 1, 3, 4 and 5 in gastric carcinoma, loss of claudin expression associates with the diffuse subtype. *Virchows Arch* **448,** 52-8.

Szasz, A. M., Nyirady, P., Majoros, A., Szendroi, A., Szucs, M., Szekely, E., Tokes, A. M., Romics, I., and Kulka, J. (2009). beta-catenin expression and claudin expression pattern as prognostic factors of prostatic cancer progression. *BJU Int*.

Takahashi, A., Kondoh, M., Masuyama, A., Fujii, M., Mizuguchi, H., Horiguchi, Y., and Watanabe, Y. (2005). Role of C-terminal regions of the C-terminal fragment of Clostridium perfringens enterotoxin in its interaction with claudin-4. *J Control Release* **108,** 56-62.

Thompson, I. M., Pauler, D. K., Goodman, P. J., Tangen, C. M., Lucia, M. S., Parnes, H. L., Minasian, L. M., Ford, L. G., Lippman, S. M., Crawford, E. D., Crowley, J. J., and Coltman, C. A., Jr. (2004). Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter. *N Engl J Med* **350,** 2239-46.

Tobioka, H., Isomura, H., Kokai, Y., Tokunaga, Y., Yamaguchi, J., and Sawada, N. (2004). Occludin expression decreases with the progression of human endometrial carcinoma. *Hum Pathol* **35,** 159-64.

Tokes, A. M., Kulka, J., Paku, S., Mathe, M., Paska, C., Lodi, C., Kiss, A., and Schaff, Z. (2005a). The expression of five different claudins in invasive breast carcinomas: comparison of pT1pN1 and pT1pN0 tumors. *Pathol Res Pract* **201,** 537-44.

Tokes, A. M., Kulka, J., Paku, S., Szik, A., Paska, C., Novak, P. K., Szilak, L., Kiss, A., Bogi, K., and Schaff, Z. (2005b). Claudin-1, -3 and -4 proteins and mRNA expression in benign and malignant breast lesions: a research study. *Breast Cancer Res* **7,** R296-305.

Tong, B., Grimes, H. L., Yang, T. Y., Bear, S. E., Qin, Z., Du, K., El-Deiry, W. S., and Tsichlis, P. N. (1998). The Gfi-1B proto-oncoprotein represses p21WAF1 and inhibits myeloid cell differentiation. *Mol Cell Biol* **18,** 2462-73.

Watson, J. D. (1987). "Molecular biology of the gene." Benjamin/Cummings, Menlo Park, Calif.