(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
9 April 2009 (09.04.2009)

PCT

(10) International Publication Number
WO 2009/044208 A1

(51) International Patent Classification:
A61K 39/00 (2006.01) *C12N 15/11* (2006.01)

(21) International Application Number:
PCT/GB2008/050902

(22) International Filing Date: 3 October 2008 (03.10.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0719231.3 3 October 2007 (03.10.2007) GB
60/997,444 3 October 2007 (03.10.2007) US

(71) Applicant (*for all designated States except US*): OXFORD GENOME SCIENCES (UK) LIMITED [GB/GB]; James & Cowper, Buxton Court, Botley, Oxford Oxfordshire OX2 0JB (GB).

(72) Inventor; and
(75) Inventor/Applicant (*for US only*): ROHLFF, Christian [DE/GB]; Oxford Genome Sciences (UK) Ltd, The Forum, 86 Milton Park, Abingdon, Oxford Oxfordshire OX14 4RY (GB).

(74) Agents: TEUTEN, Andrew et al.; Sagittarius IPC Ltd, Taylor House, 39 High Street, Marlow Buckinghamshire SL7 1AF (GB).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

WO 2009/044208 A1

(54) Title: PROTEIN

(57) Abstract: The present invention provides methods and compositions for treatment, screening, diagnosis and prognosis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer, for monitoring the effectiveness of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer treatment, and for drug development.

PROTEIN

INTRODUCTION

The present invention relates to the identification of membrane protein associated with breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer, which has utility as a marker for breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer and breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer metastases and which also forms a biological target against which therapeutic antibodies (or other affinity reagents) or other pharmaceutical agents can be made, formulations/compositions comprising said protein/polypeptide, use of said protein/polypeptide or a composition comprising same in therapy, antibodies for use in therapy, compositions comprising a therapeutic antibody against a relevant polypeptide or a combination of antibodies and use of same in therapy. The invention also extends to use of the relevant protein, fragments thereof or antibodies directed against the same for diagnosis of one or more of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer and kits comprising said protein, fragments or antibodies and use of said kits in methods of diagnosis.

BACKGROUND OF THE INVENTION

Breast Cancer

Globally, breast cancer is both the most common cancer (10% of all cancer cases) and the leading cause of cancer death (6% of cancer deaths) in women. Global incidence of breast cancer is over 1 million cases per year, with about 400,000 deaths. Women in North America have the highest rate of breast cancer in the world (over 200,000 new cases per year, with about 40,000 deaths). The chance of developing invasive breast cancer at some time in a woman’s life is about 1 in 8. Breast cancer incidence increases with age, rising sharply after age 40. In the USA, about 77% of

invasive breast cancers occur in women over age 50. It has been estimated that approximately US$8.1 billion is spent in the USA each year on treating breast cancer.

Breast Cancer diagnosis:

Finding a breast cancer as early as possible improves the likelihood that treatment will be successful. Screening methods such as mammograms, clinical breast examinations and breast self-examinations are useful in detecting breast cancer. Current diagnostic methods include breast ultrasound, ductogram, full-field digital mammography (FFDM), scintimammography and MRI. A biopsy (fine needle aspiration biopsy, core biopsy or surgical biopsy) is then performed to confirm the presence of breast cancer. Imaging tests such as a chest x-ray, bone scan, CT, MRI and PET are used to detect if the breast cancer has spread.

Breast Cancer staging:

Breast cancer is staged using the American Joint Committee on Cancer (AJCC) TNM system - Stage 0 – Stage IV. Ductal carcinoma in situ (DCIS), a non-invasive cancer which accounts for 20% of new breast cancer cases is Stage 0. Nearly all women diagnosed at this early stage of breast cancer can be cured. Infiltrating (invasive) ductal carcinoma (IDC), which accounts for 80% of invasive breast cancer and infiltrating (invasive) lobular carcinoma (ILC), which accounts for 5% of invasive breast cancers are more severe Stage I-IV cancers and can metastasize.

Breast Cancer treatment:

Breast-conserving surgery (lumpectomy) or mastectomy are the usual treatments for breast cancer. For stage I or II breast cancer, breast-conserving surgery is as effective as mastectomy. Patients can then undergo reconstructive surgery. Axillary lymph node sampling and removal or sentinel lymph node biopsy (SLNB) is performed to see if the cancer has spread to the lymph nodes.

Neoadjuvant chemotherapy can be given before surgery to shrink large cancers. Adjuvant chemotherapy after surgery reduces the risk of breast cancer recurrence.

Chemotherapy can also be used as the main treatment for women whose cancer has spread outside the breast and underarm area. Chemotherapeutic agents used include anthracyclines (e.g. methotrexate, fluorouracil, doxorubicin, epirubicin), taxanes (e.g. paclitaxel, docetaxel, vinorelbine) and alkylating agents (e.g. cyclophosphamide).

5 Radiation therapy (usually external beam radiation but sometimes brachytherapy) is given once chemotherapy is complete.

Hormone therapy with selective estrogen receptor modulators (e.g. tamoxifen) can be given to women with estrogen receptor positive breast cancers. Taking tamoxifen after surgery for 5 years can reduce recurrence by about 50% in women with

10 early breast cancer. Aromatase inhibitors such as exemestane, letrozole or anastrozole can also be used.

Women with HER2 positive cancers (about 1/3 of breast cancers) can be given biological response modifiers such as trastuzumab (Herceptin). Clinical trials have shown that adding trastuzumab to chemotherapy lowers the recurrence rate and death

15 rate over chemotherapy alone after surgery in women with HER2 positive early breast cancers.

Breast Cancer Survival by Stage

This table shows survival by stage based on patients diagnosed between 1995 and1998. The survival rates now should be slightly higher.

20

Stage	5-year Relative Survival Rate
0	100%
I	100%
IIA	92%
IIB	81%
IIIA	67%
IIIB	54%
IV	20%

Colorectal Cancer

Colorectal cancer (CRC) is one of the leading causes of cancer-related morbidity and mortality, responsible for an estimated half a million deaths per year, mostly in Western, well developed countries. In these territories, CRC is the third most common malignancy (estimated number of new cases per annum in USA and EU is approximately 350,000 per year). Estimated healthcare costs related to treatment for colorectal cancer in the United States are more than $8 billion.

Colorectal Cancer diagnosis:

Today, the faecal occult blood test and colonoscopy, a highly invasive procedure, are the most frequently used screening and diagnostic methods for colorectal cancer.

Other diagnostic tools include Flexible Sigmoidoscopy (allowing the observation of only about half of the colon) and Double Contrast Barium Enema (DCBE, to obtain X-ray images).

Colorectal Cancer staging:

CRC has four distinct stages: patients with stage I disease have a five-year survival rate of >90%, while those with metastatic stage IV disease have a <5% survival rate according to the US National Institutes of Health (NIH).

Colorectal Cancer treatment:

Once CRC has been diagnosed, the correct treatment needs to be selected. Surgery is usually the main treatment for rectal cancer, although radiation and chemotherapy will often be given before surgery. Possible side effects of surgery include bleeding from the surgery, blood clots in the legs, and damage to nearby organs during the operation.

Currently, 60 percent of colorectal cancer patients receive chemotherapy to treat their disease; however, this form of treatment only benefits a few percent of the population, while carrying with it high risks of toxicity, thus demonstrating a need to better define the patient selection criteria.

Colorectal cancer has a 30 to 40 percent recurrence rate within an average of

18 months after primary diagnosis. As with all cancers, the earlier it is detected the more likely it can be cured, especially as pathologists have recognised that the majority of CRC tumours develop in a series of well-defined stages from benign adenomas.

Colon Cancer Survival by Stage

5

Stage	Survival Rate
I	93%
IIA	85%
IIB	72%
IIIA	83%
IIIB	64%
IIIC	44%
IV	8%

Gastric Cancer

Gastric cancer is the second-leading cause of cancer-related deaths in the world, 10 with about 700,000 deaths per year, mostly in less developed countries. In the USA, about 22,000 people are diagnosed with gastric cancer each year, with about 11,000 deaths. Two thirds of people diagnosed with gastric cancer are older than 65.

Gastric Cancer diagnosis:

Early stage gastric cancer rarely causes symptoms so only about 10-20% of 15 gastric cancers in the USA are found in the early stages, before they have spread to other areas of the body. Studies in the USA have not found mass screening for gastric cancer to be useful because the disease is not that common. Endoscopy followed by a biopsy is the main procedure used to diagnose gastric cancer. Other diagnostic methods include barium upper gastrointestinal radiographs, endoscopic ultrasound, CT scan, 20 PET scan, MRI scan, chest x-ray, laparoscopy, complete blood count (CBC) test and faecal occult blood test.

Gastric Cancer staging:

Gastric cancer is staged using the American Joint Commission on Cancer (AJCC) TNM system – Stage 0 – Stage IV. Patients with stage 0 disease have a 5-year survival rate of >90%, while there is usually no cure for patients with stage IV disease
5 and the 5-year survival rate is only 7%. The overall 5-year relative survival rate of people with gastric cancer in the USA is about 23%. The 5-year survival rate for cancers of the proximal stomach is lower than for cancers in the distal stomach.

Gastric Cancer treatment:

Surgery is the only way to cure gastric cancer. There are three types of surgery
10 used – endoscopic mucosal resection (only for early stage gastric cancer), subtotal gastrectomy or total gastrectomy. Gastric cancer often spreads to lymph nodes so these must also be removed. If the cancer has extended to the spleen, the spleen is also removed. Surgery for gastric cancer is difficult and complications can occur.

Chemotherapy may be given as the primary treatment for gastric cancer that has
15 spread to distant organs. Chemotherapy together with external beam radiation therapy may delay cancer recurrence and extend the life span of people with less advanced gastric cancer, especially when the cancer could not be removed completely by surgery. Chemotherapeutic agents used include fluorouracil, doxorubicin, methotrexate, etoposide and cisplatin.

20 ### *Gastric Cancer Survival by Stage*

Stage	Survival Rate
0	>90%
IA	80%
IB	60%
II	34%
IIIA	17%
IIIB	12%
IV	7%

Hepatocellular carcinoma (HCC)

Hepatocellular carcinoma (HCC) arises from the main cells of the liver (the hepatocytes) and accounts for around 80% of all cases of liver cancer. It is usually confined to the liver and is associated with cirrhosis in 50% to 80% of patients.
5 Hepatocellular carcinoma is about 3 times more common in males than in females. Chronic infection with hepatitis B virus (HBV) or hepatitis C virus (HCV) is a major cause of HCC and is responsible for making liver cancer the most common cancer in many parts of the world. In the United States, hepatitis C infection is responsible for about 50% to 60% of all liver cancers and hepatitis B is responsible for another 20%.
10 Exposure to Aflatoxins is also a cause of HCC, mostly in warmer and tropical countries.

Liver cancer accounts for about 5.8% of all cancer cases globally (about 626,000 cases) and 8.9% of deaths per year (about 598,000). It is the 3rd most common cause of cancer-related death in both men and women worldwide. HCC is
15 predominantly found in Asia and Africa, which account for 80% of cases. In the USA, there are approximately 18,500 new cases of HCC and 16,000 deaths per year.

About 85% of people diagnosed with liver cancer are between 45 and 85 years of age. About 4% are between 35 and 44 years of age and only 2.4% are younger than 35.

Hepatocellular carcinoma diagnosis:

Since symptoms of liver cancer often do not appear until the disease is advanced, only a small number of liver cancers are found in the early stages and can be removed with surgery. Many signs and symptoms of liver cancer are relatively nonspecific – that is, they can be caused by other cancers or by non-cancerous diseases. Imaging tests such as ultrasound, computed tomography (CT), magnetic resonance imaging (MRI) and angiography are commonly used to diagnose HCC. Other diagnostic tools include laparoscopy, biopsy, alpha-fetoprotein (AFP) blood test, liver function tests (LFTs), a prothrombin time (PT) and tests for hepatitis B and C.

Hepatocellular carcinoma staging:

HCC has four stages, stage I to stage IV according to the American Joint Committee on Cancer (AJCC) TNM system. HCC can be classified as localized resectable, localized unresectable or advanced. The overall 5-year relative survival rate from liver cancer is about 9%. One reason for this low survival rate is that most patients with liver cancer also have cirrhosis of the liver, which itself can be fatal (people with liver cancer and class C cirrhosis are generally too sick for any treatment and usually die in a few months). The 5 year survival for localized resectable HCC following surgery is between 40% and 70%. For advanced HCC there is no standard treatment and the 5 year survival rate is less than 5%. Survival continues to drop after diagnosis and treatment so that by 10 years it is half of what it was at 5 years.

Hepatocellular carcinoma treatment:

Treatment of liver cancer depends on the size of the tumour and whether the patient has cirrhosis. At this time, surgery, either by resection or liver transplantation, offers the only chance to cure a liver cancer. People without cirrhosis can do well with surgical removal of the tumour. However, in many cases, it might not be possible to safely remove a localized liver cancer. Less than 30% of the patients having explorative surgery are able to have their cancer completely removed by surgery. Partial

hepatectomy results in a 5-year survival of 30% to 40%. If there is cirrhosis, or a very large tumour, most experts recommend liver transplantation as the main treatment. The 5-year survival for liver transplantation patients is around 70% but the opportunities for liver transplantation are limited.

5 Other treatments include radiofrequency ablation (RFA), ethanol ablation, cryosurgery, hepatic artery embolization, chemoembolization or three-dimensional conformal radiation therapy (3DCRT). Chemotherapy can also be used but shrinks fewer than 1 in 5 tumours. This may be improved by hepatic artery infusion (HAI). Chemotherapeutic agents used include Adriamycin, VP-16, Cisplatinum, Mitomycin, 5-
10 FU and Leucovorin.

The prognosis for any treated primary liver cancer patient with progressing, recurring, or relapsing disease is poor. Treatment of liver cancer that returns after initial therapy depends on many factors, including the site of the recurrence, the type of initial treatment, and the functioning of the liver. Patients with localized resectable disease
15 that recurs in the same spot may be eligible for further surgery.

Lung Cancer

Lung cancer is the most common form of cancer worldwide (accounting for about 12% of cancer cases) and the main cause of death from cancer (accounting for
20 about 18% of deaths). Global incidence of lung cancer is over 1,300,000 per year, with the number of deaths over 1,100,000. In the USA, there are about 170,000 new cases per year (about 13% of all cancers), with about 160,000 deaths (about 28% of cancer deaths). Lung cancer is much more prevalent among men than women. Nearly 70% of people diagnosed with lung cancer are older than 65; fewer than 3% of all cases are
25 found in people under the age of 45. Around 15% of all lung cancers are small cell type (SCLC), which tend to spread widely through the body, while the remaining 85% are non-small cell (NSCLC). It has been estimated that approximately US$9.6 billion is spent in the USA each year on treating lung cancer.

Lung Cancer diagnosis:

Lung cancer is a life-threatening disease because it often metastasizes even before it can be detected on a chest x-ray. Usually symptoms of lung cancer do not appear until the disease is in an advanced stage. So far, there is no screening test that has been shown to improve a person's chance for a cure. Imaging tests such as a chest x-ray, CT scan, MRI scan or PET scan may be used to detect lung cancer. Tests to confirm the diagnosis are then performed and include sputum cytology, needle biopsy, bronchoscopy, endobronchial ultrasound and complete blood count (CBC).

Lung Cancer staging:

Nearly 60% of people diagnosed with lung cancer die within one year of diagnosis; 75% die within 2 years. The 5-year survival rate for people diagnosed with NSCLC is about 15%; for SCLC the 5-year survival rate is about 6%.

NSCLC is staged using the American Joint Committee on Cancer (AJCC) TNM system – Stage 0 – Stage IV. The 5-year survival rates by stage are as follows: stage I: 47%; stage II; 26%; stage III: 8% and stage IV: 2%.

SCLC has a 2-stage system – limited stage and extensive stage. About two thirds of SCLC patients have extensive disease at diagnosis. If SCLC is found very early and is localised to the lung alone, the 5-year survival rate is around 21%, but only 6% of patients fall into this category. Where the cancer has spread, the 5-year survival is around 11%. For patients with extensive disease, the 5-year survival is just 2%.

Lung Cancer treatment:

Surgery is the only reliable method to cure NSCLC. Types of surgery include lobectomy, pneumonectomy, segmentectomy and video-assisted thoracic surgery (for small tumours).

External beam radiation therapy is sometimes used as the primary treatment, especially if the patient's health is too poor to undergo surgery. Radiation therapy can also be used after surgery.

Chemotherapy may be given as the primary treatment or as an adjuvant to surgery.

Targeted therapy using epidermal growth factor receptor (EGFR) antagonists such as gefitinib or erlotinib can also be given after other treatments have failed. Antiangiogenesis drugs, such as bevacizumab, have been found to prolong survival of patients with advanced lung cancer. Photodynamic therapy is also being researched as a
5 treatment for lung cancer.

The main treatment for SCLC is chemotherapy, either alone or in combination with external beam radiation therapy and very rarely, surgery.

Chemotherapeutic agents used for NSCLC and SCLC include cisplatin, carboplatin, mitomycin C, ifosfamide, vinblastine, gemcitabine, etoposide, vinorelbine,
10 paclitaxel, docetaxel and irinotecan.

Pancreatic Cancer

Pancreatic cancer is a very difficult cancer to detect and the prognosis for patients is usually very poor. The number of new cases and deaths per year is almost
15 equal. Global incidence of pancreatic cancer is approximately 230,000 cases (about 2% of all cancer cases), with about 225,000 deaths (3.4% of cancer deaths) per year. It is much more prevalent in the developed world. In the USA, there are about 34,000 new cases per year, with about 32,000 deaths. It has been estimated that approximately US$1.5 billion is spent in the USA each year on treating pancreatic cancer.

20 ***Pancreatic Cancer diagnosis:***

Pancreatic cancer is very difficult to detect and very few pancreatic cancers are found early. Patients usually have no symptoms until the cancer has spread to other organs. There are currently no blood tests or easily available screening tests that can accurately detect early cancers of the pancreas. An endoscopic ultrasound followed by
25 a biopsy is the best way to diagnose pancreatic cancer. Other detection methods include CT, CT-guided needle biopsy, PET, ultrasonography and MRI. Blood levels of CA 19-9 and carcinoembryonic antigen (CEA) may be elevated but by the time blood levels are high enough to be detected, the cancer is no longer in its early stages.

Pancreatic Cancer staging:

Pancreatic cancer has four stages, stage I to stage IV according to the American Joint Committee on Cancer (AJCC) TNM system. Pancreatic cancer is also divided into resectable, locally advanced (unresectable) and metastatic cancer. For patients with advanced cancers, the overall survival rate is <1% at 5 years with most patients dying within 1 year.

Pancreatic Cancer treatment:

Surgery is the only method of curing pancreatic cancer. About 10% of pancreatic cancers are contained entirely within the pancreas at the time of diagnosis and attempts to remove the entire cancer by surgery may be successful in some of these patients. The 5-year survival for those undergoing surgery with the intent of completely removing the cancer is about 20%. Potentially curative surgery, usually by pancreaticoduodenectomy (Whipple procedure), is used when it may be possible to remove all of the cancer. Palliative surgery may be performed if the tumour is too widespread to be completely removed. Removing only part of the cancer does not allow patients to live longer. Pancreatic cancer surgery is difficult to perform and very hard for the patient to undergo with a high likelihood of complications.

External beam radiation therapy combined with chemotherapy can be given before or after surgery and can also be given to patients whose tumours are too widespread to be removed by surgery. The main chemotherapeutic agents which are used are gemcitabine and 5-fluorouracil.

Targeted therapy using drugs such as erlotinib and cetuximab may be of benefit to patients with advanced pancreatic cancer.

Therapeutic Challenges

The major challenges in treatment of the above mentioned cancers are to improve early detection rates, to find new non-invasive markers that can be used to follow disease progression and identify relapse, and to find improved and less toxic therapies, especially for more advanced disease where 5 year survival is still poor.

There is a great need to identify targets which are more specific to the cancer cells, e.g. ones which are expressed on the surface of the tumour cells so that they can be attacked by promising new approaches like immunotherapeutics and targeted toxins.

5 SUMMARY OF THE INVENTION

The present invention provides methods and compositions for screening, diagnosis, prognosis and therapy of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer, for breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic 10 cancer patients' stratification, for monitoring the effectiveness of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer treatment, and for drug development for treatment of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer.

We have used mass spectrometry to identify peptides generated by 1D gel 15 electrophoresis and tryptic digest of membrane proteins extracted from breast, colorectal, gastric epithelium, liver, lung and pancreatic cancer tissue samples. Peptide sequences were compared to existing protein and cDNA databases and the corresponding gene sequences identified. The protein of the invention has not been previously reported to originate from breast, colorectal, gastric epithelium, liver, lung 20 or pancreatic cancer cell membranes and represents a protein of new diagnostic and therapeutic value.

A first aspect of the invention is an agent capable of specific binding to Integrin beta 4, or a fragment thereof, or a hybridising agent capable of hybridizing to nucleic acid encoding Integrin beta 4 or an agent capable of detecting the activity of Integrin 25 beta 4 for use in treating, screening for, detecting and/or diagnosing disease, such as cancer, and especially breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

Another aspect of the invention is Integrin beta 4, or a fragment thereof for use in treating, screening for, detecting and/or diagnosing disease such as cancer, and

especially breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

Another aspect of the invention is an affinity reagent capable of specific binding to Integrin beta 4 or a fragment thereof, for example an affinity reagent which contains or is conjugated to a detectable label or contains or is conjugated to a therapeutic moiety such as a cytotoxic moiety. The affinity reagent may, for example, be an antibody.

In some embodiments, the antibody of the present invention is selected from the group consisting of: a whole antibody, an antibody fragment, a humanized antibody, a single chain antibody, an immunoconjugate, a defucosylated antibody, and a bispecific antibody. The antibody fragment may be selected from the group consisting of: a UniBody, a domain antibody, and a Nanobody. In some embodiments, the immunoconjugates of the invention comprise a therapeutic agent. In another aspect of the invention, the therapeutic agent is a cytotoxin or a radioactive isotope.

In some embodiments, the antibody of the present invention is selected from the group consisting of: an Affibody, a DARPin, an Anticalin, an Avimer, a Versabody, and a Duocalin.

Another aspect of the invention is a hybridizing agent capable of hybridizing to nucleic acid encoding Integrin beta 4, for example, a hybridizing agent which contains or is conjugated to a detectable label. One example of a hybridizing agent is an inhibitory RNA (RNAi). Other examples include anti-sense oligonucleotides and ribozymes.

The invention also provides a kit containing Integrin beta 4 and/or one or more fragments thereof or containing one or more aforementioned affinity reagents and/or hybridizing agents or containing one or more agents capable of detecting the activity of Integrin beta 4 together with instructions for their use in an aforementioned method. The kit may further contain reagents capable of detecting and reporting the binding of said affinity reagents and/or hybridizing agents to their binding partners.

Another aspect of the invention is a pharmaceutical composition comprising a therapeutically effective amount of an affinity reagent capable of specific binding to Integrin beta 4 or a fragment thereof.

Another aspect of the invention is a pharmaceutically acceptable diluent or
5 carrier and a pharmaceutical composition comprising one or more affinity reagents or hybridizing reagents as aforesaid and a pharmaceutically acceptable diluent or carrier.

In some embodiments, the present invention is a method for preparing an anti-Integrin beta 4 antibody, said method comprising the steps of: obtaining a host cell that contains one or more nucleic acid molecules encoding the antibody of the invention;
10 growing the host cell in a host cell culture; providing host cell culture conditions wherein the one or more nucleic acid molecules are expressed; and recovering the antibody from the host cell or from the host cell culture.

Other aspects of the invention are directed to methods of making the antibodies of the invention, comprising the steps of: immunizing a transgenic animal comprising human
15 immunoglobulin genes with a Integrin beta 4 peptide; recovering B-cells from said transgenic animal; making hybridomas from said B-cells; selecting hybridomas that express antibodies that bind Integrin beta 4; and recovering said antibodies that bind Integrin beta 4 from said selected hybridomas.

In other embodiments, the method of making anti-Integrin beta 4 antibodies,
20 comprises the steps of:

immunizing a transgenic animal comprising human immunoglobulin genes with a Integrin beta 4 peptide;

recovering mRNA from the B cells of said transgenic animal;

converting said mRNA to cDNA;

25 expressing said cDNA in phages such that anti-Integrin beta 4 antibodies encoded by said cDNA are presented on the surface of said phages;

selecting phages that present anti-Integrin beta 4 antibodies;

recovering nucleic acid molecules from said selected phages that encode said anti-Integrin beta 4 immunoglobulins;

expressing said recovered nucleic acid molecules in a host cell; and

recovering antibodies from said host cell that bind Integrin beta 4.

Another aspect of the invention provides use of a Integrin beta 4 polypeptide, one or more immunogenic fragments or derivatives thereof for the treatment or prophylaxis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

In another aspect, the invention provides methods of treating breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer, comprising administering to a patient a therapeutically effective amount of a compound that modulates (e.g., upregulates or downregulates) or complements the expression or the biological activity (or both) of the protein of the invention in patients having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer, in order to (a) prevent the onset or development of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer; (b) prevent the progression of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer; or (c) ameliorate the symptoms of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

According to another aspect of the invention we provide a method of detecting, diagnosing and/or screening for or monitoring the progression of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer or of monitoring the effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy in a subject which comprises detecting the presence or level of Integrin beta 4, or one or more fragments thereof, or the presence or level of nucleic acid encoding Integrin beta 4 or the presence or level of the activity of Integrin beta 4 or which comprises detecting a change in the level thereof in said subject.

According to another aspect of the invention we provide a method of detecting, diagnosing and/or screening for breast cancer, colorectal cancer, gastric cancer,

hepatocellular carcinoma, lung cancer or pancreatic cancer in a candidate subject which comprises detecting the presence of Integrin beta 4, or one or more fragments thereof, or the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of Integrin beta 4 in said candidate subject, in which either (a) the presence of an elevated level of Integrin beta 4 or said one or more fragments thereof or an elevated level of nucleic acid encoding Integrin beta 4 or the presence of an elevated level of Integrin beta 4 activity in the candidate subject as compared with the level in a healthy subject or (b) the presence of a detectable level of Integrin beta 4 or said one or more fragments thereof or a detectable level of nucleic acid encoding Integrin beta 4 or the presence of a detectable level of Integrin beta 4 activity in the candidate subject as compared with a corresponding undetectable level in a healthy subject indicates the presence of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in said subject.

According to another aspect of the invention we provide a method of monitoring the progression of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in a subject or of monitoring the effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy which comprises detecting the presence of Integrin beta 4, or one or more fragments thereof, or the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of Integrin beta 4 in said candidate subject at a first time point and at a later time point, the presence of an elevated or lowered level of Integrin beta 4 or said one or more fragments thereof or an elevated or lowered level of nucleic acid encoding Integrin beta 4 or the presence of an elevated or lowered level of Integrin beta 4 activity in the subject at the later time point as compared with the level in the subject at said first time point, indicating the progression or regression of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer or indicating the effect or non-effect of an anti-breast cancer, anti-colorectal

cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy in said subject.

The presence of Integrin beta 4, or one or more fragments thereof, or the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of
5 Integrin beta 4 may, for example, be detected by analysis of a biological sample obtained from said subject.

The method of invention may typically include the step of obtaining a biological sample for analysis from said subject.

The biological sample used can be from any source such as a serum sample or a
10 tissue sample, e.g. breast, colorectal, gastric epithelium, liver, lung or pancreatic tissue. For instance, when looking for evidence of metastatic breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer, one would look at major sites of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer metastasis, e.g. the liver, the lungs and
15 bones for breast cancer; the liver, the peritoneal cavity, the pelvis, the retroperitoneum and the lungs for colorectal cancer; the liver, the lungs, the brain and bones for gastric cancer; the lungs and bones for hepatocellular carcinoma; the brain, the liver, the bones and adrenal glands for lung cancer and the liver for pancreatic cancer.

Alternatively the presence of Integrin beta 4, or one or more fragments thereof,
20 or the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of Integrin beta 4 may be detected by analysis in situ.

In certain embodiments, methods of diagnosis described herein may be at least partly, or wholly, performed *in vitro*.

Suitably the presence of Integrin beta 4, or one or more fragments thereof, or
25 the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of Integrin beta 4 is detected quantitatively.

For example, quantitatively detecting may comprise:

(a) contacting a biological sample with an affinity reagent that is specific for Integrin beta 4, said affinity reagent optionally being conjugated to a

detectable label; and

(b) detecting whether binding has occurred between the affinity reagent and at least one species in the sample, said detection being performed either directly or indirectly.

5 Alternatively the presence of Integrin beta 4, or one or more fragments thereof, or the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of Integrin beta 4 may be detected quantitatively by means involving use of an imaging technology.

In another embodiment, the method of the invention involves use of 10 immunohistochemistry on breast, colorectal, gastric epithelium, liver, lung or pancreatic tissue sections in order to determine the presence of Integrin beta 4, or one or more fragments thereof, or the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of Integrin beta 4, and thereby to localise breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic 15 cancer cells.

In one embodiment the presence of Integrin beta 4 or one or more epitope-containing fragments thereof is detected, for example using an affinity reagent capable of specific binding to Integrin beta 4 or one or more fragments thereof, such as an antibody.

20 In another embodiment the activity of Integrin beta-4 is detected. Integrin beta-4 is a laminin-5 receptor which is phosphorylated on multiple tyrosines in the beta-4 cytoplasmic domain by a Src Family Kinase (SFK) after ligand binding. The phosphorylated complex activates the Ras to ERK cascade via the signaling adaptor protein She (Dans et al., 2001 J Biol. Chem. 276:1494-1502; Gagnoux-Palacios et al., 25 2003 J Cell Biol.162:1189-96; Mainiero et al.,1995 EMBO J. 14:4470-4481), as well as PI-3 kinase and Rac (Shaw, 2001; Shaw et al., 1997).

In vitro studies support a role for Integrin beta-4 in promoting tumor invasion, and a number of invasive carcinomas display elevated levels of Integrin beta-4 (Mercurio and Rabinovitz, 2001). Introduction of Integrin beta-4 in breast and colon carcinoma cells

activates PI-3K - Rac and confers a more invasive phenotype (Shaw et al. 1997 Cell 91:949-960). The activation of this pathway may also confer resistance to apoptosis.

Integrin beta-4 can also amplify signaling from certain receptor tyrosine kinases that have known involvement in tumour progression, such as c-met (Trusolino et al. 2001 Cell 107:643-54), erb-b2 and EGFR (Guo et al, 2006, Cell 126: 489-502).

Transfection of a dominant negative form of Integrin beta-4 impairs the survival of breast carcinoma cells phenotype (Weaver et al., 2002 Cancer Cell 2:205-16). Antibodies that inhibit the phosphorylation of Integrin beta-4 also blocked anchorage independent growth of MDA-MB-231 breast cancer cells, independently of the action of the EGFR-inhibitory antibody, erbitux (Gabarra et al., 2008 AACR Meeting Abstracts 2008: 5252). The same antibodies also blocked apoptosis induced by serum starvation of these and other Integrin beta-4-expressing cell lines (SW620, MCF7, SU159).

These results suggest that Integrin beta-4 promotes cell migration and invasion and confers resistance to apoptosis in carcinoma cells, functions that can be tested *in vitro* and used as an assay for inhibitors of Integrin beta-4 that could act as anti-cancer agents.

According to another aspect of the invention there is provided a method of detecting, diagnosing and/or screening for or monitoring the progression of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer or of monitoring the effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy in a subject which comprises detecting the presence or level of antibodies capable of immunospecific binding to Integrin beta 4, or one or more epitope-containing fragments thereof or which comprises detecting a change in the level thereof in said subject.

According to another aspect of the invention there is also provided a method of detecting, diagnosing and/or screening for breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in a subject which

comprises detecting the presence of antibodies capable of immunospecific binding to Integrin beta 4, or one or more epitope-containing fragments thereof in said subject, in which (a) the presence of an elevated level of antibodies capable of immunospecific binding to Integrin beta 4 or said one or more epitope-containing fragments thereof in 5 said subject as compared with the level in a healthy subject or (b) the presence of a detectable level of antibodies capable of immunospecific binding to Integrin beta 4 or said one or more epitope-containing fragments thereof in said subject as compared with a corresponding undetectable level in a healthy subject indicates the presence of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or 10 pancreatic cancer in said subject.

One particular method of detecting, diagnosing and/or screening for breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer comprises:

(a) bringing into contact with a biological sample to be tested Integrin beta 15 4, or one or more epitope-containing fragments thereof; and

(b) detecting the presence of antibodies in the subject capable of immunospecific binding to Integrin beta 4, or one or more epitope-containing fragments thereof

According to another aspect of the invention there is provided a method of 20 monitoring the progression of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer or of monitoring the effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy in a subject which comprises detecting the presence of antibodies capable of immunospecific binding to 25 Integrin beta 4, or one or more epitope-containing fragments thereof in said subject at a first time point and at a later time point, the presence of an elevated or lowered level of antibodies capable of immunospecific binding to Integrin beta 4, or one or more epitope-containing fragments thereof in said subject at the later time point as compared with the level in said subject at said first time point, indicating the progression or

regression of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer or the effect or non-effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy in said subject.

5 The presence of antibodies capable of immunospecific binding to Integrin beta 4, or one or more epitope-containing fragments thereof is typically detected by analysis of a biological sample obtained from said subject (exemplary biological samples are mentioned above, e.g. the sample is a sample of breast, colorectal, gastric epithelium, liver, lung or pancreatic tissue, or else a sample of blood or saliva).

10 The method typically includes the step of obtaining said biological sample for analysis from said subject.

The antibodies that may be detected include IgA, IgM and IgG antibodies.

In any of the above methods, the level that may be detected in the candidate subject who has breast cancer, colorectal cancer, gastric cancer, hepatocellular 15 carcinoma, lung cancer or pancreatic cancer is 2 or more fold higher than the level in the healthy subject.

In one embodiment the cancer to be detected, prevented or treated is breast cancer.

In another embodiment the cancer to be detected, prevented or treated is 20 colorectal cancer.

In another embodiment the cancer to be detected, prevented or treated is gastric cancer.

In another embodiment the cancer to be detected, prevented or treated is hepatocellular carcinoma.

25 In another embodiment the cancer to be detected, prevented or treated is lung cancer.

In another embodiment the cancer to be detected, prevented or treated is pancreatic cancer.

Other aspects of the present invention are set out below and in the claims

herein.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows the amino acid sequences of the five splice variants of the
5 protein of the invention. The tryptics detected experimentally by mass spectrometry are highlighted – mass match peptides are shown in bold, tandem peptides are underlined. Recombinant protein is shown in italics.

Figure 2 shows the Protein Index for the protein of the invention.

10 DETAILED DESCRIPTION OF THE INVENTION

The invention described in detail below encompasses the administration of therapeutic compositions to a mammalian subject to treat or prevent breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer. The invention also provides methods and compositions for clinical screening,
15 diagnosis and prognosis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer in a mammalian subject for identifying patients most likely to respond to a particular therapeutic treatment, for monitoring the results of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer therapy, for drug screening and drug
20 development.

In one aspect the invention provides an agent capable of specific binding to Integrin beta 4, or a fragment thereof, or a hybridising agent capable of hybridizing to nucleic acid encoding Integrin beta 4 or an agent capable of detecting the activity of Integrin beta 4 for use in treating, screening for, detecting and/or diagnosing disease,
25 such as cancer, and especially breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

Another aspect of the invention is an affinity reagent capable of specific binding to Integrin beta 4 or a fragment thereof, for example an affinity reagent which contains or is conjugated to a detectable label or contains or is conjugated to a therapeutic

moiety such as a cytotoxic moiety. The affinity reagent may, for example, be an antibody.

Another aspect of the invention is a pharmaceutical composition comprising a therapeutically effective amount of an affinity reagent capable of specific binding to
5 Integrin beta 4 or a fragment thereof.

In another aspect the invention provides use of a Integrin beta 4 polypeptide, or one or more fragments or derivatives thereof, for the treatment or prophylaxis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

10 The invention also provides use of a Integrin beta 4 polypeptide, one or more fragments or derivatives thereof in the manufacture of a medicament for the treatment or prophylaxis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

In one aspect there is provided a method of treatment comprising administering
15 a therapeutically effective amount of a Integrin beta 4 polypeptide, one or more fragments or derivatives thereof, or one or more fragments or derivatives thereof, for the treatment or prophylaxis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

The invention further provides a method for the treatment or prophylaxis of
20 breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in a subject, or of vaccinating a subject against breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer, which comprises the step of administering to the subject an effective amount of a Integrin beta 4 polypeptide and/or one or more antigenic or immunogenic fragments
25 thereof, for example as a vaccine.

The mammalian subject may be a non-human mammal, but is preferably human, more preferably a human adult, i.e. a human subject at least 21 (more preferably at least 35, at least 50, at least 60, at least 70, or at least 80) years old.

In one aspect there is provided a composition capable of eliciting an immune

response in a subject, which composition comprises a Integrin beta 4 polypeptide and/or one or more antigenic or immunogenic fragments thereof, and one or more suitable adjuvants (suitable adjuvants are discussed below).

The composition capable of eliciting an immune response may for example be provided as a vaccine comprising a Integrin beta 4 polypeptide or derivatives thereof, and/or one or more antigenic or immunogenic fragments thereof.

For clarity of disclosure, and not by way of limitation, the invention will be described with respect to the analysis of breast, colorectal, gastric epithelium, liver, lung and pancreatic tissue. However, as one skilled in the art will appreciate, the assays and techniques described below can be applied to other types of patient samples, including body fluids (*e.g.* blood, urine or saliva), a tissue sample from a patient at risk of having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer (*e.g.* a biopsy such as a breast, liver, stomach, lung or pancreatic biopsy) or homogenate thereof. The methods and compositions of the present invention are specially suited for screening, diagnosis and prognosis of a living subject, but may also be used for postmortem diagnosis in a subject, for example, to identify family members at risk of developing the same disease.

Integrin beta 4

In one aspect of the invention, one-dimensional electrophoresis or another appropriate method is used to analyze breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer tissue samples from a subject, preferably a living subject, in order to measure the expression of the protein of the invention for screening or diagnosis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer, to determine the prognosis of a breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer patient, to monitor the effectiveness of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer therapy, or for drug development.

As used herein, the term "Protein of the invention", or "Integrin beta 4", refers to the protein illustrated in Figure 1 in all its splice variants, in particular in its five different splice variants detected experimentally by 1D electrophoresis of breast, colorectal, gastric epithelium, liver, lung and pancreatic cancer tissue samples (Integrin beta 4a to Integrin beta 4e). Protein derivatives of these sequences may also be useful for the same purposes as described herein.

This protein has been identified in membrane protein extracts of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer tissue samples from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer patients, through the methods and apparatus of the Preferred Technology (1D gel electrophoresis and tryptic digest of membrane protein extracts). Peptide sequences were compared to the SWISS-PROT and trEMBL databases (held by the Swiss Institute of Bioinformatics (SIB) and the European Bioinformatics Institute (EBI) which are available at www.expasy.com), and the following entry: P16144, Integrin beta-4, was identified.

According to SWISS-PROT, integrin beta-4 is predominantly expressed by epithelia. In addition, Integrin beta 4d is also expressed in colon and placenta while Integrin beta 4e is also expressed in epidermis, lung, duodenum, heart, spleen and stomach. The function of Integrin alpha-6/beta-4 is to be a receptor for laminin. It plays a critical structural role in the hemidesmosome of epithelial cells.

The expression of Integrin beta-4 has been reported to be associated with basal-like cancers and it is hypothesized that Integrin beta-4 may function in concert with a discrete set of proteins to facilitate the aggressive behaviour of a subset of tumours (see, e.g., Lu et al., Analysis of Integrin β4 Expression in Human Breast Cancer: Association with Basal-like Tumors and Prognostic Significance, Clin Cancer Res 2008;14(4):1050-58, which is incorporated herein by reference in its entirety).

Integrin beta-4 has also been identified as a candidate biomarker for the clinical outcome of tongue squamous cell carcinoma (see, e.g., Kurokawa et al., Diagnostic Value of Integrin α3, β4, and β5 Gene Expression Levels for the Clinical Outcome of

Tongue Squamouos Cell Carcinoma, Cancer 2008; 112:1272-81, which is incorporated herein by reference in its entirety).

Integrin beta-4 has been reported to promote the migratory and invasive phenotype of pancreatic carcinoma cells through the Tiam1-Rac1 pathway in part
5 through the upregulation of Tiam1 (see, e.g., Cruz-Monserrate and O'Connor, Integrin α6β4 Promotes Migration, Invasion through Tiam1 Upregulation, and Subsequent Rac Activation, Neoplasia, 2008; 10(5):408-417, which is incorporated herein by reference in its entirety).

Integrin beta-4 contains two pairs of fibronectin type III (FNIII) repeats in the
10 cytoplasmic domain (see Figure 1). Deletion of the first pair of FNIII repeats has been associated with Carmi syndrome, a rare autosomal recessive disorder (see, e.g., Birnbaum et al., Deletion of the First Pair of Fibronectin Type III Repeats of the Integrin β-4 Gene Is Associated With Epidermolysis Bullosa, Pyloric Atresia and Aplasia Cutis Congenita in the Original Carmi Syndrome Patients, 2008, Am J Med
15 Genet Part A 146A:1063-66, which is incorporated herein by reference in its entirety).

For further discussion of the function of Integrin beta-4 see Wilhelmsen et al., Multiple Functions of the Integrin α6β4 in Epidermal Homeostasis and Tumorigenesis, Molecular and Cellular Biology, 2006; 26(8);2877–86, which is incorporated herein by reference in its entirety.

20 The protein of the invention is useful as are fragments particularly epitope containing fragments e.g. antigenic or immunogenic fragments thereof and derivatives thereof. Epitope containing fragments including antigenic or immunogenic fragments will typically be of length 12 amino acids or more e.g. 20 amino acids or more e.g. 50 or 100 amino acids or more. Fragments may be 95% or more of the length of the full
25 protein e.g. 90% or more e.g. 75% or 50% or 25% or 10% or more of the length of the full protein.

Alternatively, the protein/polypeptide employed or referred to herein may be limited to those specifically recited/described in the present specification or a moiety 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98 or 99% identical or similar thereto.

Epitope containing fragments including antigenic or immunogenic fragments will be capable of eliciting a relevant immune response in a patient. DNA encoding the protein of the invention is also useful as are fragments thereof e.g. DNA encoding fragments of the protein of the invention such as immunogenic fragments thereof. Fragments of nucleic acid (e.g. DNA) encoding the protein of the invention may be 95% or more of the length of the full coding region e.g. 90% or more e.g. 75% or 50% or 25% or 10% or more of the length of the full coding region. Fragments of nucleic acid (e.g. DNA) may be 36 nucleotides or more e.g. 60 nucleotides or more e.g. 150 or 300 nucleotides or more in length.

Derivatives of the protein of the invention include variants on the sequence in which one or more (e.g. 1-20 such as 15 amino acids, or up to 20% such as up to 10% or 5% or 1% by number of amino acids based on the total length of the protein) deletions, insertions or substitutions have been made. Substitutions may typically be conservative substitutions. Derivatives will typically have essentially the same biological function as the protein from which they are derived. Derivatives will typically be comparably antigenic or immunogenic to the protein from which they are derived. Derivatives will typically have either the ligand-binding activity, or the active receptor-complex forming ability, or preferably both, of the protein from which they are derived.

Derivatives of proteins also include chemically treated protein such as carboxymethylated, carboxyamidated, acetylated proteins, for example treated during purification.

Tables 1a to 1f below illustrate the different occurrences of Integrin beta 4 as detected by 1D gel electrophoresis and mass spectrometry of membrane protein extracts of breast, colorectal, gastric epithelium, liver, lung and pancreatic tissue samples from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer patients respectively. The first column provides the molecular weight, the second column gives information on the subfractionation

protocol used, if any (see Example 1 below), the third column gives information on the preferred splice variants, based on tandem peptides detected experimentally by mass spectrometry (see Figure 1), and the last column provides a list of the sequences observed by mass spectrometry and the corresponding SEQ ID Nos.

5 Table 1a – Breast cancer

MW (Da)	Subfractionation	Preferred splice variants	Tryptics identified [SEQ ID No]
153448		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AEEVVVR [7], DVVSFEQPEFSVSR [13], GMVEFQEGVELVDVR [27], HVTQEFVSR [30], MTTTSAAAYGTHLSPHVPHR [45], NVISLTEDVDEFR [48], RCNTQAELLAAGCQR [56], SEHSHSTTLPR [64], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83], YWIQGDSESEAHLLDSK [85]
162825		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	ACLALLPCCNR [6], AQSQEGWGR [10], DVVSFEQPEFSVSR [13], DYSTLTSVSSHDSR [16], EDHYMLR [19], EGIITIESQDGGPFPQLGSR [21], LVFSALGPTSLR [39], MDFAFPGSTNSLHR [42], MLLIENLR [44], NGAGWGPER [47], QDHTIVDTVLMAPR [50], QEVEENLNEVYR [51], QLLVEAIDVPAGTATLGR [53], RAEEVVVR [55], RCNTQAELLAAGCQR [56], SEHSHSTTLPR [64], SFTSQMLSSQPPPHGDLGAPQNPNAK [65], THQEVPSEPGR [72], VAPGYYTLTADQDAR [76], VLSTSSTLTR [79], VPLFIRPEDDDEK [81]
164389		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	GMVEFQEGVELVDVR [27], IHFNWLPPSGKPMGYR [31], ISGNLDAPEGGFDAILQTAVCTR [32], MDAGIICDVCTCELQK [41], MDFAFPGSTNSLHR [42], VDGDSPESR [78], VSVPQTDMRPEK [83]
165165		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AEEVVVR [7], DVVSFEQPEFSVSR [13], EGIITIESQDGGPFPQLGSR [21], GMVEFQEGVELVDVR [27], ISGNLDAPEGGFDAILQTAVCTR [32], MTTTSAAAYGTHLSPHVPHR [45], NVISLTEDVDEFR [48], SEHSHSTTLPR [64], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83], YWIQGDSESEAHLLDSK [85]

Table 1b – Colorectal Cancer

MW (Da)	Subfractionation	Preferred splice variants	Tryptics identified [SEQ ID No]
88905		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], GMVEFQEGVELVDVR [27], IHFNWLPPSGKPMGYR [31], ISGNLDAPEGGFDAILQTAVCTR [32], RAEEVVVR [55], TLTTSGTLSTHMDQQFFQT [73], TTEGFGPER [75], VPSVELTNLYPYCDYEMK [82]
104028		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], ENLMASDHLDTPMLR [22], HNIIPIFAVTNYSYSYYEK [29], ISGNLDAPEGGFDAILQTAVCTR [32], LLELQEVDSLLR [35], MGQNLAR [43], MTTTSAAAYGTHLSPHVPHR [45], NVISLTEDVDEFR [48], RFHVQLSNPK [57], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81]
107091		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AQSQEGWGR [10], DVVSFEQPEFSVSR [13], GEVGIYQVQLR [26], MGQNLAR [43], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83]
113823		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], ENLMASDHLDTPMLR [22], GMVEFQEGVELVDVR [27], HNIIPIFAVTNYSYSYYEK [29], KIHFNWLPPSGKPMGYR [33] [31], NVISLTEDVDEFR [48], VAPGYYTLTADQDAR [76], VSWQEPR [84], YWIQGDSESEAHLLDSK [85]
117533		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], ENLMASDHLDTPMLR [22], GMVEFQEGVELVDVR [27], HNIIPIFAVTNYSYSYYEK [29], ISGNLDAPEGGFDAILQTAVCTR [32], RCNTQAELLAAGCQR [56], VAPGYYTLTADQDAR [76], VPSVELTNLYPYCDYEMK [82], VSVPQTDMRPEK [83], YWIQGDSESEAHLLDSK [85]
121506		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], DYIPVEGELLFQPGEAWK [14], ENLMASDHLDTPMLR [22], GEVGIYQVQLR [26], GMVEFQEGVELVDVR [27], MDFAFPGSTNSLHR [42], NVISLTEDVDEFR [48], TGSFHIRR [71], VAPGYYTLTADQDAR [76]

125772		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	LVFSALGPTSLR [39]
130363		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	LVFSALGPTSLR [39]
140688		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	ALEHVDGTHVCQLPEDQK [9], AQSQEGWGR [10], DVVSFEQPEFSVSR [13], DYSTLTSVSSHDSR [16], EGIITIESQDGGPFPQLGSR [21], ENLMASDHLDTPMLR [22], GEVGIYQVQLR [26], GMVEFQEGVELVDVR [27], GNIHLKPSFSDGLK [28], MDFAFPGSTNSLHR [42], MTTTSAAAYGTHLSPHVPHR [45], NGAGWGPER [47], NVISLTEDVDEFR [48], QEVEENLNEVYR [51], RCNTQAELLAAGCQR [56], RFHVQLSNPK [57], RGEVGIYQVQLR [58], TTEGFGPER [75], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], YWIQGDSESEAHLLDSK [85]
146522		Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AQSQEGWGR [10], DVVSFEQPEFSVSR [13], DYNSLTR [15], DYSTLTSVSSHDSR [16], EDHYMLR [19], EGIITIESQDGGPFPQLGSR [21], ENLMASDHLDTPMLR [22], GEVGIYQVQLR [26], GMVEFQEGVELVDVR [27], MAGPRPSPWAR [40], MDFAFPGSTNSLHR [42], MLLIENLR [44], MTTTSAAAYGTHLSPHVPHR [45], NGAGWGPER [47], NVISLTEDVDEFR [48], PSVSDDTEHLVNGR [49], QDHTIVDTVLMAPR [50], QEVEENLNEVYR [51], QLLVEAIDVPAGTATLGR [53], RCNTQAELLAAGCQR [56], RGEVGIYQVQLR [58], SEHSHSTTLPR [64], SQMSPQGLR [67], TTEGFGPER [75], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSWQEPR [84], YWIQGDSESEAHLLDSK [85]
159861		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AQSQEGWGR [10], DVVSFEQPEFSVSR [13], EGIITIESQDGGPFPQLGSR [21], GMVEFQEGVELVDVR [27], RCNTQAELLAAGCQR [56], VAPGYYTLTADQDAR [76], YWIQGDSESEAHLLDSK [85]

	Heparin Binding	Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], EGIITIESQDGGPFPQLGSR [21], GMVEFQEGVELVDVR [27], ISGNLDAPEGGFDAILQTAVCTR [32], LVFSALGPTSLR [39], SQMSPQGLR [67], VAPGYYTLTADQDAR [76]
	Heparin Binding	Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	EGIITIESQDGGPFPQLGSR [21], LVFSALGPTSLR [39], MVDELK [46], QISGVHK [52], SQMSPQGLR [67], VLVDNPK [80]

Table 1c – Gastric cancer

MW (Da)	Subfractionation	Preferred splice variants	Tryptics identified [SEQ ID No]
84210		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], ECAQLR [18], EDHYMLR [19], RFHVQLSNPK [57], SNLDIR [66]
85955		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DPDELDR [12]
87777		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	GMVEFQEGVELVDVR [27]
98257		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d, Integrin beta 4e	NVISLTEDVDEFR [48]
100678		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], FHVQLSNPK [25], RFHVQLSNPK [57]
114938		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	LTAGVPDTPTR [37]
121910		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], ECAQLR [18], LTAGVPDTPTR [37], LTVPGLSENVPYK [38], NVISLTEDVDEFR [48], RAEEVVVR [55]
125740		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], LTAGVPDTPTR [37], MAGPRPSPWAR [40], SNLDIR [66], VSWQEPR [84]

129831		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AFHDLK [8], EGEDKPCSGR [20], FEPLLGEELDLR [23], LCTENLLKPDTR [34], MVDELK [46], SNLDIR [66], SQMSPQGLR [67], VSVPQTDMRPEK [83], VSWQEPR [84]

Table 1d – Hepatocellular carcinoma

MW (Da)	Subfractionation	Preferred splice variants	Tryptics identified [SEQ ID No]
82761		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	LLELQEVDSLLR [35]
102119		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AEEVVVR [7], AQSQEGWGR [10], GEVGIYQVQLR [26], HVTQEFVSR [30], LVFSALGPTSLR [39], MDFAFPGSTNSLHR [42], QEVEENLNEVYR [51], QISGVHK [52], SEHSHSTTLPR [64], THQEVPSEPGR [72], TQDYPSVPTLVR [74], TTEGFGPER [75], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83]
106345		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AEEVVVR [7], AQSQEGWGR [10], DVVSFEQPEFSVSR [13], GEVGIYQVQLR [26], GMVEFQEGVELVDVR [27], HVTQEFVSR [30], LTAGVPDTPTR [37], NGAGWGPER [47], QEVEENLNEVYR [51], QISGVHK [52], SEHSHSTTLPR [64], SQMSPQGLR [67], THQEVPSEPGR [72], TQDYPSVPTLVR [74], TTEGFGPER [75], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83], VSWQEPR [84]

114377		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AQSQEGWGR [10], DVVSFEQPEFSVSR [13], DYSTLTSVSSHDSR [16], EGIITIESQDGGPFPQLGSR [21], ENLMASDHLDTPMLR [22], GEVGIYQVQLR [26], GMVEFQEGVELVDVR [27], HVTQEFVSR [30], MAGPRPSPWAR [40], MDFAFPGSTNSLHR [42], MTTTSAAAYGTHLSPHVPHR [45], NGAGWGPER [47], NVISLTEDVDEFR [48], QEVEENLNEVYR [51], SEHSHSTTLPR [64], SQMSPQGLR [67], SQVSYR [68], THQEVPSEPGR [72], TQDYPSVPTLVR [74], TTEGFGPER [75], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VPSVELTNLYPYCDYEMK [82], VSVPQTDMRPEK [83], VSWQEPR [84], YWIQGDSESEAHLLDSK [85]
178322		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	LLELQEVDSLLR [35]

Table 1e – Lung cancer

MW (Da)	Subfractionation	Preferred splice variants	Tryptics identified [SEQ ID No]
181240		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AQSQEGWGR [10], DVVSFEQPEFSVSR [13], DYNSLTR [15], DYSTLTSVSSHDSR [16], GEVGIYQVQLR [26], GMVEFQEGVELVDVR [27], GNIHLKPSFSDGLK [28], HVTQEFVSR [30], LLELQEVDSLLR [35], LVFSALGPTSLR [39], MGQNLAR [43], MLLIENLR [44], NGAGWGPER [47], QEVEENLNEVYR [51], RFHVQLSNPK [57], RGEVGIYQVQLR [58], RVTWR [61], SCVQCQAWGTGEK [63], THQEVPSEPGR [72], TTEGFGPER [75], VAPGYYTLTADQDAR [76], VLSTSSTLTR [79], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83], VSWQEPR [84]

Table 1f – Pancreatic cancer

MW (Da)	Subfractionation	Preferred splice variants	Tryptics identified [SEQ ID No]

31738		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d, Integrin beta 4e	FEPLLGEELDLRR [24], FHVQLSNPK [25], NVISLTEDVDEFR [48]
105212		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], LLELQEVDSLLR [35], SATPGPPGEHLVNGR [62], VPLFIRPEDDDEK [81]
107788		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	RAEEVVVR [55], THQEVPSEPGR [72], VCAYGAQGEGPYSSLVSCR [77]
122734		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], GMVEFQEGVELVDVR [27], QEVEENLNEVYR [51], RCNTQAELLAAGCQR [56], SNLDIR [66], TGSFHIRR [71], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSWQEPR [84]
126215		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], ENLMASDHLDTPMLR [22], LNIPNPAQTSVVVEDLLPNHSYVFR [36], MAGPRPSPWAR [40], MDFAFPGSTNSLHR [42], QEVEENLNEVYR [51], VAPGYYTLTADQDAR [76], VCAYGAQGEGPYSSLVSCR [77]
134980		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], ECAQLR [18], GMVEFQEGVELVDVR [27], IHFNWLPPSGKPMGYR [31], KIHFNWLPPSGKPMGYR [33], LCTENLLKPDTR [34], LNIPNPAQTSVVVEDLLPNHSYVFR [36], MAGPRPSPWAR [40], MTTTSAAAYGTHLSPHVPHR [45], QEVEENLNEVYR [51], QLLVEAIDVPAGTATLGR [53], RGEVGIYQVQLR [58], SQMSPQGLR [67], TQDYPSVPTLVR [74], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83], VSWQEPR [84]
142690		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], ECAQLR [18], EGIITIESQDGGPFPQLGSR [21], GMVEFQEGVELVDVR [27], IHFNWLPPSGKPMGYR [31], MAGPRPSPWAR [40], NVISLTEDVDEFR [48], QDHTIVDTVLMAPR [50], QEVEENLNEVYR [51], RAEEVVVR [55], SQVSYR [68], VSVPQTDMRPEK [83], VSWQEPR [84]

148609		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AQSQEGWGR [10], DVVSFEQPEFSVSR [13], DYNSLTR [15], EGIITIESQDGGPFPQLGSR [21], GEVGIYQVQLR [26], HVTQEFVSR [30], LCTENLLKPDTR [34], NVISLTEDVDEFR [48], QEVEENLNEVYR [51], QQPNAGK [54], SEHSHSTTLPR [64], TGSFHIR [70], THQEVPSEPGR [72], TTEGFGPER [75], VAPGYYTLTADQDAR [76], VLVDNPK [80], VPLFIRPEDDDEK [81], YWIQGDSESEAHLLDSK [85]
154726		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AEEVVVR [7], AQSQEGWGR [10], DVVSFEQPEFSVSR [13], ECAQLR [18], EGIITIESQDGGPFPQLGSR [21], ENLMASDHLDTPMLR [22], GEVGIYQVQLR [26], GMVEFQEGVELVDVR [27], HVTQEFVSR [30], MDFAFPGSTNSLHR [42], MGQNLAR [43], MLLIENLR [44], NGAGWGPER [47], NVISLTEDVDEFR [48], QEVEENLNEVYR [51], RGEVGIYQVQLR [58] [26], SEHSHSTTLPR [64], SQMSPQGLR [67], TGSFHIR [70], THQEVPSEPGR [72], TQDYPSVPTLVR [74], TTEGFGPER [75], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83], VSWQEPR [84], YWIQGDSESEAHLLDSK [85]
161314		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AQSQEGWGR [10], DVVSFEQPEFSVSR [13], EGIITIESQDGGPFPQLGSR [21], ENLMASDHLDTPMLR [22], FHVQLSNPK [25], GEVGIYQVQLR [26], GNIHLKPSFSDGLK [28], HVTQEFVSR [30], LVFSALGPTSLR [39], MDFAFPGSTNSLHR [42], MGQNLAR [43], NGAGWGPER [47], NVISLTEDVDEFR [48], QEVEENLNEVYR [51], RGEVGIYQVQLR [58], RSQMSPQGLR [67], SEHSHSTTLPR [64], TGSFHIR [70], THQEVPSEPGR [72], TQDYPSVPTLVR [74], TTEGFGPER [75], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83], VSWQEPR [84], YWIQGDSESEAHLLDSK [85]

167857		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AQSQEGWGR [10], DVVSFEQPEFSVSR [13], EGIITIESQDGGPFPQLGSR [21], GEVGIYQVQLR [26], GMVEFQEGVELVDVR [27], HVTQEFVSR [30], IHFNWLPPSGKPMGYR [31], ISGNLDAPEGGFDAILQTAVCTR [32], MTTTSAAAYGTHLSPHVPHR [45], MVDELK [46], NGAGWGPER [47], NVISLTEDVDEFR [48], QLLVEAIDVPAGTATLGR [53], RCNTQAELLAAGCQR [56], RPNGDIVGYLVTCEMAQGGGPATAFR [59], SATPGPPGEHLVNGR [62], SEHSHSTTLPR [64], SFTSQMLSSQPPPHGDLGAPQNPNAK [65], SQMSPQGLR [67], TQDYPSVPTLVR [74], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83], YWIQGDSESEAHLLDSK [85]
168430		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AQSQEGWGR [10], DVVSFEQPEFSVSR [13], GMVEFQEGVELVDVR [27], MGQNLAR [43], NVISLTEDVDEFR [48], QEVEENLNEVYR [51], SEHSHSTTLPR [64], SSDAEAPHGPPDDGGAGGK [69], TGSFHIR [70], THQEVPSEPGR [72], TTEGFGPER [75], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83], YWIQGDSESEAHLLDSK [85]
174222		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	DVVSFEQPEFSVSR [13], DYIPVEGELLFQPGEAWK [14], ECAQLR [18], EGIITIESQDGGPFPQLGSR [21], GEVGIYQVQLR [26], GMVEFQEGVELVDVR [27], LCTENLLKPDTR [34], MDAGIICDVCTCELQK [41], MTTTSAAAYGTHLSPHVPHR [45], NVISLTEDVDEFR [48], QEVEENLNEVYR [51], SFTSQMLSSQPPPHGDLGAPQNPNAK [65], VAPGYYTLTADQDAR [76], VPSVELTNLYPYCDYEMK [82], VSVPQTDMRPEK [83], VSWQEPR [84], YWIQGDSESEAHLLDSK [85]

179222		Integrin beta 4a, Integrin beta 4b, Integrin beta 4d	AQSQEGWGR [10], DVVSFEQPEFSVSR [13], DYSTLTSVSSHDSR [16], ECAQLR [18], EGIITIESQDGGPFPQLGSR [21], ENLMASDHLDTPMLR [22], GEVGIYQVQLR [26], GMVEFQEGVELVDVR [27], GNIHLKPSFSDGLK [28], HVTQEFVSR [30], MDFAFPGSTNSLHR [42], NGAGWGPER [47], NVISLTEDVDEFR [48], QEVEENLNEVYR [51], QLLVEAIDVPAGTATLGR [53], RGEVGIYQVQLR [58], SEHSHSTTLPR [64], SQMSPQGLR [67], SQVSYR [68], TQDYPSVPTLVR [74], TTEGFGPER [75], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83], VSWQEPR [84], YWIQGDSESEAHLLDSK [85]
192682		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AQSQEGWGR [10], CERPLQGYSVEYQLLNGGELHR [11], DVVSFEQPEFSVSR [13], DYSTLTSVSSHDSR [16], ECAQLR [18], EGIITIESQDGGPFPQLGSR [21], GEVGIYQVQLR [26], GMVEFQEGVELVDVR [27], IHFNWLPPSGKPMGYR [31], ISGNLDAPEGGFDAILQTAVCTR [32], MTTTSAAAYGTHLSPHVPHR [45], NGAGWGPER [47], NVISLTEDVDEFR [48], QEVEENLNEVYR [51], RGEVGIYQVQLR [58], RPNGDIVGYLVTCEMAQGGGPATAFR [59], SEHSHSTTLPR [64], SFTSQMLSSQPPPHGDLGAPQNPNAK [65], TQDYPSVPTLVR [74], TTEGFGPER [75], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VPSVELTNLYPYCDYEMK [82], VSVPQTDMRPEK [83], VSWQEPR [84], YWIQGDSESEAHLLDSK [85]
198088		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	EAIINLATQPK [17]

207419		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AQSQEGWGR [10], DVVSFEQPEFSVSR [13], DYSTLTSVSSHDSR [16], ECAQLR [18], EGIITIESQDGGPFPQLGSR [21], ENLMASDHLDTPMLR [22], GEVGIYQVQLR [26], GMVEFQEGVELVDVR [27], HVTQEFVSR [30], LLELQEVDSLLR [35], MTTTSAAAYGTHLSPHVPHR [45], NGAGWGPER [47], NVISLTEDVDEFR [48], QEVEENLNEVYR [51], QLLVEAIDVPAGTATLGR [53], RFHVQLSNPK [57], RGEVGIYQVQLR [58], SEHSHSTTLPR [64], SQMSPQGLR [67], THQEVPSEPGR [72], TQDYPSVPTLVR [74], TTEGFGPER [75], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83], VSWQEPR [84], YWIQGDSESEAHLLDSK [85]
208929		Integrin beta 4a, Integrin beta 4b, Integrin beta 4c, Integrin beta 4d	AQSQEGWGR [10], DVVSFEQPEFSVSR [13], EGIITIESQDGGPFPQLGSR [21], GEVGIYQVQLR [26], GMVEFQEGVELVDVR [27], MDFAFPGSTNSLHR [42], MTTTSAAAYGTHLSPHVPHR [45], NVISLTEDVDEFR [48], QEVEENLNEVYR [51], SEHSHSTTLPR [64], TQDYPSVPTLVR [74], VAPGYYTLTADQDAR [76], VPLFIRPEDDDEK [81], VSVPQTDMRPEK [83], VSWQEPR [84]

For Integrin beta 4, the detected level obtained upon analyzing tissue from subjects having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer relative to the detected level obtained upon analyzing tissue from subjects free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer will depend upon the particular analytical protocol and detection technique that is used. Accordingly, the present invention contemplates that each laboratory will establish a reference range in subjects free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer according to the analytical protocol and detection technique in use, as is conventional in the diagnostic art. Preferably, at least one control positive tissue sample from a subject known to have breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic

cancer or at least one control negative tissue sample from a subject known to be free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer (and more preferably both positive and negative control samples) are included in each batch of test samples analysed.

5 Integrin beta 4 can be used for detection, prognosis, diagnosis, or monitoring of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer or for drug development. In one embodiment of the invention, tissue from a subject (*e.g.*, a subject suspected of having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer) is
10 analysed by 1D electrophoresis for detection of Integrin beta 4. An increased abundance of Integrin beta 4 in the tissue from the subject relative to tissue from a subject or subjects free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer (*e.g.*, a control sample) or a previously determined reference range indicates the presence of breast cancer,
15 colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

The sequences shown in Table 1 may be employed in any relevant aspect of the invention.

In a particular embodiment, specific splice variants of Integrin beta 4 are used.
20 As indicated in Table 1, for colorectal cancer the preferred splice variants are Integrin beta 4b, Integrin beta 4c and Integrin beta 4d; for breast cancer the preferred splice variants are Integrin beta 4a, Integrin beta 4b, Integrin beta 4c and Integrin beta 4d; for gastric cancer the preferred splice variants are Integrin beta 4a, Integrin beta 4b, Integrin beta 4c and Integrin beta 4d; for hepatocellular carcinoma the preferred splice
25 variants are Integrin beta 4a, Integrin beta 4b, Integrin beta 4c and Integrin beta 4d; for lung cancer the preferred splice variants are Integrin beta 4a, Integrin beta 4b, Integrin beta 4c and Integrin beta 4d and for pancreatic cancer the preferred splice variants are Integrin beta 4a, Integrin beta 4b and Integrin beta 4d.

Integrin beta 4 may, in particular, be characterized as an isoform having a MW

substantially as recited (e.g. +/- 10%, particularly +/-5% of the value) in column 1 of any of the rows of Tables 1a-1f.

In relation to variants, fragments, epitope-containing fragments, immunogenic fragments or antigenic fragments of Integrin beta 4:

5 -for breast cancer applications: in one aspect of the invention these comprise one or more of the sequences identified as tryptic sequences in the 4th column of Table 1a;

-for colorectal cancer applications: in one aspect of the invention these comprise one or more of the sequences identified as tryptic sequences in the 4th column of Table 1b;

-for gastric cancer applications: in one aspect of the invention these comprise one or 10 more of the sequences identified as tryptic sequences in the 4th column of Table 1c;

-for hepatocellular cancer applications: in one aspect of the invention these comprise one or more of the sequences identified as tryptic sequences in the 4th column of Table 1d;

-for lung cancer applications: in one aspect of the invention these comprise one or more 15 of the sequences identified as tryptic sequences in the 4th column of Table 1e;

-for pancreatic cancer applications: in one aspect of the invention these comprise one or more of the sequences identified as tryptic sequences in the 4th column of Table 1f.

As used herein, Integrin beta 4 is "isolated" when it is present in a preparation that is substantially free of contaminating proteins, *i.e.*, a preparation in which less than 20 10% (preferably less than 5%, more preferably less than 1%) of the total protein present is contaminating protein(s). A contaminating protein is a protein having a significantly different amino acid sequence from that of isolated Integrin beta 4, as determined by mass spectral analysis. As used herein, a "significantly different" sequence is one that permits the contaminating protein to be resolved from Integrin 25 beta 4 by mass spectral analysis, performed according to the Reference Protocol.

Thus in one aspect the invention provides a pharmaceutical composition for the treatment of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer comprising a therapeutically effective amount of a Integrin beta 4 polypeptide (particularly those defined above) or an immunogenic

fragment thereof and an adjuvant.

Integrin beta 4 can be assayed by any method known to those skilled in the art, including but not limited to, the Preferred Technology described herein, kinase assays, enzyme assays, binding assays and other functional assays, immunoassays, and western blotting. In one embodiment, Integrin beta 4 is separated on a 1-D gel by virtue of its MW and visualized by staining the gel. In one embodiment, Integrin beta 4 is stained with a fluorescent dye and imaged with a fluorescence scanner. Sypro Red (Molecular Probes, Inc., Eugene, Oregon) is a suitable dye for this purpose. A preferred fluorescent dye is disclosed in U.S. Application No. 09/412,168, filed on October 5, 1999, which is incorporated herein by reference in its entirety.

Alternatively, Integrin beta 4 can be detected in an immunoassay. In one embodiment, an immunoassay is performed by contacting a sample from a subject to be tested with an anti-Integrin beta 4 antibody (or other affinity reagent) under conditions such that binding (e.g. immunospecific binding) can occur if Integrin beta 4 is present, and detecting or measuring the amount of any binding (e.g. immunospecific binding) by the affinity reagent. Integrin beta 4 binding agents can be produced by the methods and techniques taught herein.

Integrin beta 4 may be detected by virtue of the detection of a fragment thereof e.g. an epitope containing (e.g. an immunogenic or antigenic) fragment thereof. Fragments may have a length of at least 10, more typically at least 20 amino acids e.g. at least 50 or 100 amino acids e.g. at least 200 or 500 amino acids e.g. at least 1000 amino acids e.g. at least 1500 amino acids.

In one embodiment, binding of an affinity reagent (e.g. an antibody) in tissue sections can be used to detect aberrant Integrin beta 4 localization or an aberrant level of Integrin beta 4. In a specific embodiment, an antibody (or other affinity reagent) to Integrin beta 4 can be used to assay a patient tissue (*e.g.*, breast, colorectal, gastric epithelium, liver, lung or pancreatic tissue) for the level of Integrin beta 4 where an aberrant level of Integrin beta 4 is indicative of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. As used herein, an

"aberrant level" means a level that is increased compared with the level in a subject free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer or a reference level.

Any suitable immunoassay can be used, including, without limitation, competitive and non-competitive assay systems using techniques such as western blots, radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoprecipitation assays, precipitin reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays, complement-fixation assays, immunoradiometric assays, fluorescent immunoassays and protein A immunoassays.

For example, Integrin beta 4 can be detected in a fluid sample (e.g., blood, urine, or saliva) by means of a two-step sandwich assay. In the first step, a capture reagent (e.g., an anti-Integrin beta 4 antibody or other affinity reagent) is used to capture Integrin beta 4. The capture reagent can optionally be immobilized on a solid phase. In the second step, a directly or indirectly labeled detection reagent is used to detect the captured Integrin beta 4. In one embodiment, the detection reagent is a lectin. Any lectin can be used for this purpose that preferentially binds to Integrin beta 4 rather than to other isoforms that have the same core protein as Integrin beta 4 or to other proteins that share the antigenic determinant recognized by the antibody. In a preferred embodiment, the chosen lectin binds Integrin beta 4 with at least 2-fold greater affinity, more preferably at least 5-fold greater affinity, still more preferably at least 10-fold greater affinity, than to said other isoforms that have the same core protein as Integrin beta 4 or to said other proteins that share the antigenic determinant recognized by the affinity reagent. Based on the present description, a lectin that is suitable for detecting Integrin beta 4 can readily be identified by methods well known in the art, for instance upon testing one or more lectins enumerated in Table I on pages 158-159 of Sumar et al., Lectins as Indicators of Disease-Associated Glycoforms, *In*: Gabius H-J & Gabius S (eds.), 1993, Lectins and Glycobiology, at pp. 158-174 (which is incorporated herein by reference in its entirety). In an alternative embodiment, the detection reagent is an antibody (or other affinity reagent), e.g., an antibody that

specifically (e.g. immunospecifically) detects other post-translational modifications, such as an antibody that immunospecifically binds to phosphorylated amino acids. Examples of such antibodies include those that bind to phosphotyrosine (BD Transduction Laboratories, catalog nos.: P11230-050/P11230-150; P11120; P38820; P39020), those that bind to phosphoserine (Zymed Laboratories Inc., South San Francisco, CA, catalog no. 61-8100) and those that bind to phosphothreonine (Zymed Laboratories Inc., South San Francisco, CA, catalogue nos. 71-8200, 13-9200).

If desired, a gene encoding Integrin beta 4, a related gene, or related nucleic acid sequences or subsequences, including complementary sequences, can also be used in hybridization assays. A nucleotide encoding Integrin beta 4, or subsequences thereof comprising at least 8 nucleotides, preferably at least 12 nucleotides, and most preferably at least 15 nucleotides can be used as a hybridization probe. Hybridization assays can be used for detection, prognosis, diagnosis, or monitoring of conditions, disorders, or disease states, associated with aberrant expression of the gene encoding Integrin beta 4, or for differential diagnosis of subjects with signs or symptoms suggestive of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. In particular, such a hybridization assay can be carried out by a method comprising contacting a subject's sample containing nucleic acid with a nucleic acid probe capable of hybridizing to a DNA or RNA that encodes Integrin beta 4, under conditions such that hybridization can occur, and detecting or measuring any resulting hybridization.

Hence nucleic acid encoding Integrin beta 4 (e.g. DNA or more suitably RNA) may be detected, for example, using a hybridizing agent capable of hybridizing to nucleic acid encoding Integrin beta 4.

One such exemplary method comprises:

(a) contacting one or more oligonucleotide probes comprising 10 or more consecutive nucleotides complementary to a nucleotide sequence encoding Integrin beta 4, with an RNA obtained from a biological sample from the subject or with cDNA copied from the RNA, wherein

said contacting occurs under conditions that permit hybridization of the probe to the nucleotide sequence if present;

(b) detecting hybridization, if any, between the probe and the nucleotide sequence; and

5 (c) comparing the hybridization, if any, detected in step (b) with the hybridization detected in a control sample, or with a previously determined reference range.

The invention also provides diagnostic kits, comprising an anti-Integrin beta 4 antibody (or other affinity reagent). In addition, such a kit may optionally comprise one 10 or more of the following: (1) instructions for using the anti-Integrin beta 4 affinity reagent for diagnosis, prognosis, therapeutic monitoring or any combination of these applications; (2) a labeled binding partner to the affinity reagent; (3) a solid phase (such as a reagent strip) upon which the anti-Integrin beta 4 affinity reagent is immobilized; and (4) a label or insert indicating regulatory approval for diagnostic, prognostic or 15 therapeutic use or any combination thereof. If no labeled binding partner to the affinity reagent is provided, the anti-Integrin beta 4 affinity reagent itself can be labeled with a detectable marker, *e.g.*, a chemiluminescent, enzymatic, fluorescent, or radioactive moiety.

The invention also provides a kit comprising a nucleic acid probe capable of 20 hybridizing to nucleic acid, suitably RNA encoding Integrin beta 4. In a specific embodiment, a kit comprises in one or more containers a pair of primers (*e.g.*, each in the size range of 6-30 nucleotides, more preferably 10-30 nucleotides and still more preferably 10-20 nucleotides) that under appropriate reaction conditions can prime amplification of at least a portion of a nucleic acid encoding Integrin beta 4, such as by 25 polymerase chain reaction (see, *e.g.*, Innis et al., 1990, PCR Protocols, Academic Press, Inc., San Diego, CA), ligase chain reaction (see EP 320,308) use of Qβ replicase, cyclic probe reaction, or other methods known in the art.

A kit can optionally further comprise a predetermined amount of Integrin beta 4 or a nucleic acid encoding Integrin beta 4, *e.g.*, for use as a standard or control.

Use in Clinical Studies

The diagnostic methods and compositions of the present invention can assist in monitoring a clinical study, *e.g.* to evaluate drugs for therapy of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. In one embodiment, candidate molecules are tested for their ability to restore Integrin beta 4 levels in a subject having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer to levels found in subjects free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer or, in a treated subject, to preserve Integrin beta 4 levels at or near non-breast cancer, non-colorectal cancer, non-gastric cancer, non-hepatocellular carcinoma, non-lung cancer or non-pancreatic cancer values.

In another embodiment, the methods and compositions of the present invention are used to screen candidates for a clinical study to identify individuals having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer; such individuals can then be excluded from the study or can be placed in a separate cohort for treatment or analysis.

Production of Protein of the Invention and Corresponding Nucleic Acid

In one aspect the invention provides a method of treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer, comprising administering to a subject in need of such treatment or prevention a therapeutically effective amount of nucleic acid encoding Integrin beta 4 or one or more fragments or derivatives thereof, for example in the form of a vaccine.

In another aspect there is provided a method of treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer comprising administering to a subject in need of such treatment or prevention a therapeutically effective amount of nucleic acid that inhibits the function

or expression of Integrin beta 4.

The methods (and/or other DNA aspects disclosed herein) of the invention may, for example include wherein the nucleic acid is a Integrin beta 4 anti-sense nucleic acid or ribozyme.

5 Thus the invention includes the use of nucleic acid encoding Integrin beta 4 or one or more fragments or derivatives thereof, in the manufacture of a medicament for treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

There is also provided the use of nucleic acid that inhibits the function or
10 expression of Integrin beta 4 in the manufacture of a medicament for treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

A DNA employed in the present invention can be obtained by isolation as a cDNA fragment from cDNA libraries using as starter materials commercial mRNAs and
15 determining and identifying the nucleotide sequences thereof. That is, specifically, clones are randomly isolated from cDNA libraries, which are prepared according to Ohara et al.'s method (DNA Research Vol.4, 53-59 (1997)). Next, through hybridization, duplicated clones (which appear repeatedly) are removed and then *in vitro* transcription and translation are carried out. Nucleotide sequences of both termini
20 of clones, for which products of 50 kDa or more are confirmed, are determined.

Furthermore, databases of known genes are searched for homology using the thus obtained terminal nucleotide sequences as queries.

In addition to the above screening method, the 5' and 3' terminal sequences of cDNA are related to a human genome sequence. Then an unknown long-chain gene is
25 confirmed in a region between the sequences, and the full-length of the cDNA is analyzed. In this way, an unknown gene that is unable to be obtained by a conventional cloning method that depends on known genes can be systematically cloned.

Moreover, all of the regions of a human-derived gene containing a DNA of the present invention can also be prepared using a PCR method such as RACE while

paying sufficient attention to prevent artificial errors from taking place in short fragments or obtained sequences. As described above, clones having DNA of the present invention can be obtained.

In another means for cloning DNA of the present invention, a synthetic DNA primer having an appropriate nucleotide sequence of a portion of a polypeptide of the present invention is produced, followed by amplification by the PCR method using an appropriate library. Alternatively, selection can be carried out by hybridization of the DNA of the present invention with a DNA that has been incorporated into an appropriate vector and labeled with a DNA fragment or a synthetic DNA encoding some or all of the regions of the polypeptide of the present invention. Hybridization can be carried out by, for example, the method described in Current Protocols in Molecular Biology (edited by Frederick M. Ausubel et al., 1987). DNA of the present invention may be any DNA, as long as they contain nucleotide sequences encoding the polypeptides of the present invention as described above. Such a DNA may be a cDNA identified and isolated from cDNA libraries or the like that are derived from breast, colorectal, gastric epithelium, liver, lung or pancreatic tissue. Such a DNA may also be a synthetic DNA or the like. Vectors for use in library construction may be any of bacteriophages, plasmids, cosmids, phargemids, or the like. Furthermore, by the use of a total RNA fraction or a mRNA fraction prepared from the above cells and/or tissues, amplification can be carried out by a direct reverse transcription coupled polymerase chain reaction (hereinafter abbreviated as "RT-PCR method").

DNA encoding the above polypeptide consisting of an amino acid sequence that is substantially identical to the amino acid sequence of Integrin beta 4 or DNA encoding the above polypeptide consisting of an amino acid sequence derived from the amino acid sequence of Integrin beta 4 by deletion, substitution, or addition of one or more amino acids composing a portion of the amino acid sequence can be easily produced by an appropriate combination of, for example, a site-directed mutagenesis method, a gene homologous recombination method, a primer elongation method, and the PCR method known by persons skilled in the art. In addition, at this time, a possible

method for causing a polypeptide to have substantially equivalent biological activity is substitution of homologous amino acids (e.g. polar and nonpolar amino acids, hydrophobic and hydrophilic amino acids, positively-charged and negatively charged amino acids, and aromatic amino acids) among amino acids composing the polypeptide. Furthermore, to maintain substantially equivalent biological activity, amino acids within functional domains contained in the polypeptide of the present invention are preferably conserved.

Furthermore, examples of DNA of the present invention include DNA comprising a nucleotide sequence that encodes the amino acid sequence of Integrin beta 4 and DNA hybridizing under stringent conditions to the DNA and encoding a polypeptide (protein) having biological activity (function) equivalent to the function of the polypeptide consisting of the amino acid sequence of Integrin beta 4. Under such conditions, an example of such DNA capable of hybridizing to DNA comprising the nucleotide sequence that encodes the amino acid sequence of Integrin beta 4 is DNA comprising a nucleotide sequence that has a degree of overall mean homology with the entire nucleotide sequence of the DNA, such as approximately 80% or more, preferably approximately 90% or more, and more preferably approximately 95% or more. Hybridization can be carried out according to a method known in the art such as a method described in Current Protocols in Molecular Biology (edited by Frederick M. Ausubel et al., 1987) or a method according thereto. Here, "stringent conditions" are, for example, conditions of approximately "1*SSC, 0.1% SDS, and 37°C, more stringent conditions of approximately "0.5*SSC, 0.1% SDS, and 42°C, or even more stringent conditions of approximately "0.2*SSC, 0.1% SDS, and 65°C. With more stringent hybridization conditions, the isolation of a DNA having high homology with a probe sequence can be expected. The above combinations of SSC, SDS, and temperature conditions are given for illustrative purposes. Stringency similar to the above can be achieved by persons skilled in the art using an appropriate combination of the above factors or other factors (for example, probe concentration, probe length, and reaction time for hybridization) for determination of hybridization stringency.

A cloned DNA of the present invention can be directly used or used, if desired, after digestion with a restriction enzyme or addition of a linker, depending on purposes. The DNA may have ATG as a translation initiation codon at the 5' terminal side and have TAA, TGA, or TAG as a translation termination codon at the 3' terminal side. These translation initiation and translation termination codons can also be added using an appropriate synthetic DNA adapter.

In the methods/uses of the invention Integrin beta 4 may, for example, be provided in isolated form, such as where the Integrin beta 4 polypeptide has been purified at least to some extent. Integrin beta 4 polypeptide may be provided in substantially pure form, that is to say free, to a substantial extent, from other proteins. Integrin beta 4 polypeptide can also be produced using recombinant methods, synthetically produced or produced by a combination of these methods. Integrin beta 4 can be easily prepared by any method known by persons skilled in the art, which involves producing an expression vector containing a DNA of the present invention or a gene containing a DNA of the present invention, culturing a transformant transformed using the expression vector, generating and accumulating a polypeptide of the present invention or a recombinant protein containing the polypeptide, and then collecting the resultant.

Recombinant Integrin beta 4 polypeptide may be prepared by processes well known in the art from genetically engineered host cells comprising expression systems. Accordingly, the present invention also relates to expression systems which comprise a Integrin beta 4 polypeptide or nucleic acid, to host cells which are genetically engineered with such expression systems and to the production of Integrin beta 4 polypeptide by recombinant techniques. For recombinant Integrin beta 4 polypeptide production, host cells can be genetically engineered to incorporate expression systems or portions thereof for nucleic acids. Such incorporation can be performed using methods well known in the art, such as, calcium phosphate transfection, DEAD-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection,

electroporation, transduction, scrape loading, ballistic introduction or infection (see e.g. Davis et al., Basic Methods in Molecular Biology, 1986 and Sambrook et al. , Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbour laboratory Press, Cold Spring Harbour, NY, 1989).

5 As host cells, for example, bacteria of the genus Escherichia, Streptococci, Staphylococci, Streptomyces, bacteria of the genus Bacillus, yeast, Aspergillus cells, insect cells, insects, and animal cells are used. Specific examples of bacteria of the genus Escherichia, which are used herein, include Escherichia coli K12 and DH1 (Proc. Natl. Acad. Sci. U.S.A., Vol. 60, 160 (1968)), JM103 (Nucleic Acids Research, Vol. 9,
10 309 (1981)), JA221 (Journal of Molecular Biology, Vol. 120, 517 (1978)), and HB101 (Journal of Molecular Biology, Vol. 41, 459 (1969)). As bacteria of the genus Bacillus, for example, Bacillus subtilis MI114 (Gene, Vol. 24, 255 (1983)) and 207-21 (Journal of Biochemistry, Vol. 95, 87 (1984)) are used. As yeast, for example, Saccaromyces cerevisiae AH22, AH22R-, NA87-11A, DKD-5D, and 20B-12, Schizosaccaromyces
15 pombe NCYC1913 and NCYC2036, and Pichia pastoris are used. As insect cells, for example, Drosophila S2 and Spodoptera Sf9 cells are used. As animal cells, for example, COS-7 and Vero monkey cells, CHO Chinese hamster cells (hereinafter abbreviated as CHO cells), dhfr-gene-deficient CHO cells, mouse L cells, mouse AtT-20 cells, mouse myeloma cells, rat GH3 cells, human FL cells, COS, HeLa, C127,3T3,
20 HEK 293, BHK and Bowes melanoma cells are used.

Cell-free translation systems can also be employed to produce recombinant polypeptides (e.g. rabbit reticulocyte lysate, wheat germ lysate, SP6/T7 in vitro T&T and RTS 100 E. Coli HY transcription and translation kits from Roche Diagnostics Ltd., Lewes, UK and the TNT Quick coupled Transcription/Translation System from
25 Promega UK, Southampton, UK).

The expression vector can be produced according to a method known in the art. For example, the vector can be produced by (1) excising a DNA fragment containing a DNA of the present invention or a gene containing a DNA of the present invention and (2) ligating the DNA fragment downstream of the promoter in an appropriate

expression vector. A wide variety of expression systems can be used, such as and without limitation, chromosomal, episomal and virus-derived systems, e.g. plasmids derived from Escherichia coli (e.g. pBR322, pBR325, pUC18, and pUC118), plasmids derived from Bacillus subtilis (e.g. pUB110, pTP5, and pC194), from bacteriophage, from transposons, from yeast episomes (e.g. pSH19 and pSH15), from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage (such as [lambda] phage) genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Promoters to be used in the present invention may be any promoters as long as they are appropriate for hosts to be used for gene expression. For example, when a host is Escherichia coli, a trp promoter, a lac promoter, a recA promoter, a pL promoter, an lpp promoter, and the like are preferred. When a host is Bacillus subtilis, an SPO1 promoter, an SPO2 promoter, a penP promoter, and the like are preferred. When a host is yeast, a PHO5 promoter, a PGK promoter, a GAP promoter, an ADH promoter, and the like are preferred. When an animal cell is used as a host, examples of promoters for use in this case include an SRa promoter, an SV40 promoter, an LTR promoter, a CMV promoter, and an HSV-TK promoter. Generally, any system or vector that is able to maintain, propagate or express a nucleic acid to produce a polypeptide in a host may be used.

The appropriate nucleic acid sequence may be inserted into an expression system by any variety of well known and routine techniques, such as those set forth in Sambrook et al., supra. Appropriate secretion signals may be incorporated into the Integrin beta 4 polypeptide to allow secretion of the translated protein into the lumen of the endoplasmic reticulum, the periplasmic space or the extracellular environment. These signals may be endogenous to the Integrin beta 4 polypeptide or they may be heterologous signals. Transformation of the host cells can be carried out according to methods known in the art. For example, the following documents can be referred to:

Proc. Natl. Acad. Sci. U.S.A., Vol. 69, 2110 (1972); Gene, Vol. 17, 107 (1982); Molecular & General Genetics, Vol. 168, 111 (1979); Methods in Enzymology, Vol. 194, 182-187 (1991); Proc. Natl. Acad. Sci. U.S.A.), Vol. 75, 1929 (1978); Cell Technology, separate volume 8, New Cell Technology, Experimental Protocol. 263-267 (1995) (issued by Shujunsha); and Virology, Vol. 52, 456 (1973). The thus obtained transformant transformed with an expression vector containing a DNA of the present invention or a gene containing a DNA of the present invention can be cultured according to a method known in the art. For example, when hosts are bacteria of the genus Escherichia, the bacteria are generally cultured at approximately 15°C to 43°C for approximately 3 to 24 hours. If necessary, aeration or agitation can also be added. When hosts are bacteria of the genus Bacillus, the bacteria are generally cultured at approximately 30°C to 40°C for approximately 6 to 24 hours. If necessary, aeration or agitation can also be added. When transformants whose hosts are yeast are cultured, culture is generally carried out at approximately 20°C to 35°C for approximately 24 to 72 hours using media with pH adjusted to be approximately 5 to 8. If necessary, aeration or agitation can also be added. When transformants whose hosts are animal cells are cultured, the cells are generally cultured at approximately 30°C to 40°C for approximately 15 to 60 hours using media with the pH adjusted to be approximately 6 to 8. If necessary, aeration or agitation can also be added.

If a Integrin beta 4 polypeptide is to be expressed for use in cell-based screening assays, it is preferred that the polypeptide be produced at the cell surface. In this event, the cells may be harvested prior to use in the screening assay. If the Integrin beta 4 polypeptide is secreted into the medium, the medium can be recovered in order to isolate said polypeptide. If produced intracellularly, the cells must first be lysed before the Integrin beta 4 polypeptide is recovered.

Integrin beta 4 polypeptide can be recovered and purified from recombinant cell cultures or from other biological sources by well known methods including, ammonium sulphate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, affinity chromatography,

hydrophobic interaction chromatography, hydroxylapatite chromatography, molecular sieving chromatography, centrifugation methods, electrophoresis methods and lectin chromatography. In one embodiment, a combination of these methods is used. In another embodiment, high performance liquid chromatography is used. In a further 5 embodiment, an antibody which specifically binds to a Integrin beta 4 polypeptide can be used to deplete a sample comprising a Integrin beta 4 polypeptide of said polypeptide or to purify said polypeptide.

To separate and purify a polypeptide or a protein of the present invention from the culture products, for example, after culture, microbial bodies or cells are collected 10 by a known method, they are suspended in an appropriate buffer, the microbial bodies or the cells are disrupted by, for example, ultrasonic waves, lysozymes, and/or freeze-thawing, the resultant is then subjected to centrifugation or filtration, and then a crude extract of the protein can be obtained. The buffer may also contain a protein denaturation agent such as urea or guanidine hydrochloride or a surfactant such as 15 Triton X-100(TM). When the protein is secreted in a culture solution, microbial bodies or cells and a supernatant are separated by a known method after the completion of culture and then the supernatant is collected. The protein contained in the thus obtained culture supernatant or the extract can be purified by an appropriate combination of known separation and purification methods. The thus obtained polypeptide (protein) of 20 the present invention can be converted into a salt by a known method or a method according thereto. Conversely, when the polypeptide (protein) of the present invention is obtained in the form of a salt, it can be converted into a free protein or peptide or another salt by a known method or a method according thereto. Moreover, an appropriate protein modification enzyme such as trypsin or chymotrypsin is caused to 25 act on a protein produced by a recombinant before or after purification, so that modification can be arbitrarily added or a polypeptide can be partially removed. The presence of a polypeptide (protein) of the present invention or a salt thereof can be measured by various binding assays, enzyme immunoassays using specific antibodies, and the like.

Techniques well known in the art may be used for refolding to regenerate native or active conformations of the Integrin beta 4 polypeptide when the polypeptide has been denatured during isolation and or purification. In the context of the present invention, Integrin beta 4 polypeptide can be obtained from a biological sample from 5 any source, such as and without limitation, a blood sample or tissue sample, e.g. a breast, colorectal, gastric epithelium, liver, lung or pancreatic tissue sample.

Integrin beta 4 polypeptide may be in the form of a "mature protein" or may be part of a larger protein such as a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, a pre-, 10 pro- or prepro- protein sequence, or a sequence which aids in purification such as an affinity tag, for example, but without limitation, multiple histidine residues, a FLAG tag, HA tag or myc tag.

Integrin beta 4 may, for example, be fused with a heterologous fusion partner such as the surface protein, known as protein D from Haemophilus Influenza B, a non- 15 structural protein from influenzae virus such as NS1, the S antigen from Hepatitis B or a protein known as LYTA such as the C terminal thereof.

An additional sequence that may provide stability during recombinant production may also be used. Such sequences may be optionally removed as required by incorporating a cleavable sequence as an additional sequence or part thereof. Thus, a 20 Integrin beta 4 polypeptide may be fused to other moieties including other polypeptides or proteins (for example, glutathione S-transferase and protein A). Such a fusion protein can be cleaved using an appropriate protease, and then separated into each protein. Such additional sequences and affinity tags are well known in the art. In addition to the above, features known in the art, such as an enhancer, a splicing signal, 25 a polyA addition signal, a selection marker, and an SV40 replication origin can be added to an expression vector, if desired.

Production of Affinity Reagents to Integrin beta 4

According to those in the art, there are three main types of immunoaffinity

reagent – monoclonal antibodies, phage display antibodies and smaller antibody-derived molecules such as Affibodies, Domain Antibodies (dAbs), Nanobodies, Unibodies, DARPins, Anticalins, Duocalins, Avimers or Versabodies. In general in applications according to the present invention where the use of antibodies is stated, other affinity
5 reagents (e.g. Affibodies, Domain Antibodies, Nanobodies, Unibodies, DARPins, Anticalins, Duocalins, Avimers or Versabodies) may be employed. Such substances may be said to be capable of immunospecific binding to Integrin beta 4. Where appropriate the term “affinity agent” shall be construed to embrace immunoaffinity reagents and other substances capable of specific binding to Integrin beta 4 including
10 but not limited to ligands, lectins, streptavidins, antibody mimetics and synthetic binding agents.

Production of Antibodies to Integrin beta 4

According to the invention Integrin beta 4, a Integrin beta 4 analogue, a
15 Integrin beta 4-related protein or a fragment or derivative of any of the foregoing may be used as an immunogen to generate antibodies which immunospecifically bind such an immunogen. Such immunogens can be isolated by any convenient means, including the methods described above. The term "antibody" as used herein refers to a peptide or polypeptide derived from, modelled after or substantially encoded by an
20 immunoglobulin gene or immunoglobulin genes, or fragments thereof, capable of specifically binding an antigen or epitope. See, e.g. *Fundamental Immunology*, 3rd Edition, W.E. Paul, ed., Raven Press, N.Y. (1993); Wilson (1994) *J. Immunol. Methods* 175:267-273; Yarmush (1992) *J. Biochem. Biophys. Methods* 25:85-97. The term antibody includes antigen-binding portions, i.e., “antigen binding sites,” (e.g.,
25 fragments, subsequences, complementarity determining regions (CDRs)) that retain capacity to bind antigen, including (i) a Fab fragment, a monovalent fragment consisting of the VL, VH, CL and CH1 domains; (ii) a $F(ab')_2$ fragment, a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; (iii) a Fd fragment consisting of the VH and CH1 domains; (iv) a Fv fragment

consisting of the VL and VH domains of a single arm of an antibody, (v) a dAb fragment (Ward et al., (1989) *Nature* 341:544-546), which consists of a VH domain; and (vi) an isolated complementarity determining region (CDR). Single chain antibodies are also included by reference in the term "antibody." Antibodies of the
5 invention include, but are not limited to polyclonal, monoclonal, bispecific, humanized or chimeric antibodies, single chain antibodies, Fab fragments and $F(ab')_2$ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. The immunoglobulin molecules of the invention can be of any class (*e.g.*, IgG, IgE, IgM, IgD and IgA) or subclass of
10 immunoglobulin molecule.

The term “specifically binds” (or “immunospecifically binds”) is not intended to indicate that an antibody binds exclusively to its intended target. Rather, an antibody “specifically binds” if its affinity for its intended target is about 5-fold greater when compared to its affinity for a non-target molecule. Suitably there is no significant cross-
15 reaction or cross-binding with undesired substances, especially naturally occurring proteins or tissues of a healthy person or animal. The affinity of the antibody will, for example, be at least about 5 fold, such as 10 fold, such as 25-fold, especially 50-fold, and particularly 100-fold or more, greater for a target molecule than its affinity for a non-target molecule. In some embodiments, specific binding between an antibody or
20 other binding agent and an antigen means a binding affinity of at least 10^6 M^{-1}. Antibodies may, for example, bind with affinities of at least about 10^7 M^{-1}, such as between about 10^8 M^{-1} to about 10^9 M^{-1}, about 10^9 M^{-1} to about 10^{10} M^{-1}, or about 10^{10} M^{-1} to about 10^{11} M^{-1}.

Affinity is calculated as $K_d = k_{off} / k_{on}$ (k_{off} is the dissociation rate constant, k_{on} is
25 the association rate constant and K_d is the equilibrium constant. Affinity can be determined at equilibrium by measuring the fraction bound (r) of labelled ligand at various concentrations (c). The data are graphed using the Scatchard equation: r/c = K(n-r):

where

r = moles of bound ligand/mole of receptor at equilibrium;

c = free ligand concentration at equilibrium;

K = equilibrium association constant; and

n = number of ligand binding sites per receptor molecule

5 By graphical analysis, r/c is plotted on the Y-axis versus r on the X-axis thus producing a Scatchard plot. The affinity is the negative slope of the line. k_{off} can be determined by competing bound labelled ligand with unlabeled excess ligand (see, e.g., U.S. Pat No. 6,316,409). The affinity of a targeting agent for its target molecule is, for example, at least about 1 x 10^{-6} moles/litre, such as at least about 1 x 10^{-7} moles/litre, such as at
10 least about 1 x 10^{-8} moles/litre, especially at least about 1 x 10^{-9} moles/litre, and particularly at least about 1 x 10^{-10} moles/litre. Antibody affinity measurement by Scatchard analysis is well known in the art. *See, e.g.*, van Erp *et al., J. Immunoassay* 12: 425-43, 1991; Nelson and Griswold, *Comput. Methods Programs Biomed*. 27: 65-8, 1988.

15 In one embodiment, antibodies that recognize gene products of genes encoding Integrin beta 4 are publicly available. In another embodiment, methods known to those skilled in the art are used to produce antibodies that recognize Integrin beta 4, a Integrin beta 4 analogue, a Integrin beta 4-related polypeptide, or a fragment or derivative of any of the foregoing. One skilled in the art will recognize that many
20 procedures are available for the production of antibodies, for example, as described in Antibodies, A Laboratory Manual, Ed Harlow and David Lane, Cold Spring Harbor Laboratory (1988), Cold Spring Harbor, N.Y. One skilled in the art will also appreciate that binding fragments or Fab fragments which mimic antibodies can also be prepared from genetic information by various procedures (Antibody Engineering: A Practical
25 Approach (Borrebaeck, C., ed.), 1995, Oxford University Press, Oxford; J. Immunol. 149, 3914-3920 (1992)).

In one embodiment of the invention, antibodies to a specific domain of Integrin beta 4 are produced. In a specific embodiment, hydrophilic fragments of Integrin beta 4 are used as immunogens for antibody production.

In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, *e.g.* ELISA (enzyme-linked immunosorbent assay). For example, to select antibodies which recognize a specific domain of Integrin beta 4, one may assay generated hybridomas for a product which binds to a Integrin beta 4 fragment containing such domain. For selection of an antibody that specifically binds a first Integrin beta 4 homolog but which does not specifically bind to (or binds less avidly to) a second Integrin beta 4 homolog, one can select on the basis of positive binding to the first Integrin beta 4 homolog and a lack of binding to (or reduced binding to) the second Integrin beta 4 homolog. Similarly, for selection of an antibody that specifically binds Integrin beta 4 but which does not specifically bind to (or binds less avidly to) a different isoform of the same protein (such as a different glycoform having the same core peptide as Integrin beta 4), one can select on the basis of positive binding to Integrin beta 4 and a lack of binding to (or reduced binding to) the different isoform (*e.g.*, a different glycoform). Thus, the present invention provides an antibody (such as a monoclonal antibody) that binds with greater affinity (for example at least 2-fold, such as at least 5-fold, particularly at least 10-fold greater affinity) to Integrin beta 4 than to a different isoform or isoforms (*e.g.*, glycoforms) of Integrin beta 4.

Polyclonal antibodies which may be used in the methods of the invention are heterogeneous populations of antibody molecules derived from the sera of immunized animals. Unfractionated immune serum can also be used. Various procedures known in the art may be used for the production of polyclonal antibodies to Integrin beta 4, a fragment of Integrin beta 4, a Integrin beta 4-related polypeptide, or a fragment of a Integrin beta 4-related polypeptide. For example, one way is to purify polypeptides of interest or to synthesize the polypeptides of interest using, e.g., solid phase peptide synthesis methods well known in the art. See, e.g., *Guide to Protein Purification*, Murray P. Deutcher, ed., *Meth. Enzymol.* Vol 182 (1990); Solid Phase Peptide Synthesis, Greg B. Fields ed., *Meth. Enzymol.* Vol 289 (1997); Kiso *et al., Chem. Pharm. Bull.* (Tokyo) 38: 1192-99, 1990; Mostafavi *et al., Biomed. Pept. Proteins*

Nucleic Acids 1: 255-60, 1995; Fujiwara *et al., Chem. Pharm. Bull.* (Tokyo) 44: 1326-31, 1996. The selected polypeptides may then be used to immunize by injection various host animals, including but not limited to rabbits, mice, rats, etc., to generate polyclonal or monoclonal antibodies. The Preferred Technology described herein provides isolated Integrin beta 4 suitable for such immunization. If Integrin beta 4 is purified by gel electrophoresis, Integrin beta 4 can be used for immunization with or without prior extraction from the polyacrylamide gel. Various adjuvants (i.e. immunostimulants) may be used to enhance the immunological response, depending on the host species, including, but not limited to, complete or incomplete Freund's adjuvant, a mineral gel such as aluminium hydroxide, surface active substance such as lysolecithin, pluronic polyol, a polyanion, a peptide, an oil emulsion, keyhole limpet hemocyanin, dinitrophenol, and an adjuvant such as BCG (bacille Calmette-Guerin) or corynebacterium parvum. Additional adjuvants are also well known in the art.

For preparation of monoclonal antibodies (mAbs) directed toward Integrin beta 4, a fragment of Integrin beta 4, a Integrin beta 4-related polypeptide, or a fragment of a Integrin beta 4-related polypeptide, any technique which provides for the production of antibody molecules by continuous cell lines in culture may be used. For example, the hybridoma technique originally developed by Kohler and Milstein (1975, Nature 256:495-497), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAbs of the invention may be cultivated *in vitro* or *in vivo*. In an additional embodiment of the invention, monoclonal antibodies can be produced in germ-free animals utilizing known technology (PCT/US90/02545, incorporated herein by reference).

The monoclonal antibodies include but are not limited to human monoclonal antibodies and chimeric monoclonal antibodies (*e.g.*, human-mouse chimeras). A

chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a human immunoglobulin constant region and a variable region derived from a murine mAb. (See, *e.g.*, Cabilly et al., U.S. Patent No. 4,816,567; and Boss et al., U.S. Patent No. 4,816397, which are incorporated herein by reference in their entirety.) Humanized antibodies are antibody molecules from non-human species having one or more complementarity determining regions (CDRs) from the non-human species and a framework region from a human immunoglobulin molecule. (See, *e.g.*, Queen, U.S. Patent No. 5,585,089, which is incorporated herein by reference in its entirety.)

Chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT Publication No. WO 87/02671; European Patent Application 184,187; European Patent Application 171,496; European Patent Application 173,494; PCT Publication No. WO 86/01533; U.S. Patent No. 4,816,567; European Patent Application 125,023; Better et al., 1988, Science 240:1041-1043; Liu et al., 1987, Proc. Natl. Acad. Sci. USA 84:3439-3443; Liu et al., 1987, J. Immunol. 139:3521-3526; Sun et al., 1987, Proc. Natl. Acad. Sci. USA 84:214-218; Nishimura et al., 1987, Canc. Res. 47:999-1005; Wood et al., 1985, Nature 314:446-449; and Shaw et al., 1988, J. Natl. Cancer Inst. 80:1553-1559; Morrison, 1985, Science 229:1202-1207; Oi et al., 1986, Bio/Techniques 4:214; U.S. Patent 5,225,539; Jones et al., 1986, Nature 321:552-525; Verhoeyan et al. (1988) Science 239:1534; and Beidler et al., 1988, J. Immunol. 141:4053-4060.

Completely human antibodies are particularly desirable for therapeutic treatment of human subjects. Such antibodies can be produced using transgenic mice which are incapable of expressing endogenous immunoglobulin heavy and light chain genes, but which can express human heavy and light chain genes. The transgenic mice are immunized in the normal fashion with a selected antigen, *e.g.*, all or a portion of Integrin beta 4. Monoclonal antibodies directed against the antigen can be obtained using conventional hybridoma technology. The human immunoglobulin transgenes

harboured by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA, IgM and IgE antibodies. For an overview of this technology for producing human antibodies, see Lonberg and Huszar (1995, Int. Rev. Immunol. 13:65-93). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, *see, e.g.*, U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806. In addition, companies such as Abgenix, Inc. (Freemont, CA) and Genpharm (San Jose, CA) can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as “guided selection.” In this approach a selected non-human monoclonal antibody, *e.g.*, a mouse antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. (Jespers et al. (1994) Bio/technology 12:899-903).

The antibodies of the present invention can also be generated by the use of phage display technology to produce and screen libraries of polypeptides for binding to a selected target. See, e.g., Cwirla et al., *Proc. Natl. Acad. Sci. USA* 87, 6378-82, 1990; Devlin et al., *Science* 249, 404-6, 1990, Scott and Smith, *Science* 249, 386-88, 1990; and Ladner et al., U.S. Pat. No. 5,571,698. A basic concept of phage display methods is the establishment of a physical association between DNA encoding a polypeptide to be screened and the polypeptide. This physical association is provided by the phage particle, which displays a polypeptide as part of a capsid enclosing the phage genome which encodes the polypeptide. The establishment of a physical association between polypeptides and their genetic material allows simultaneous mass screening of very large numbers of phage bearing different polypeptides. Phage displaying a polypeptide with affinity to a target bind to the target and these phage are enriched by affinity screening to the target. The identity of polypeptides displayed from

these phage can be determined from their respective genomes. Using these methods a polypeptide identified as having a binding affinity for a desired target can then be synthesized in bulk by conventional means. See, e.g., U.S. Patent No. 6,057,098, which is hereby incorporated in its entirety, including all tables, figures, and claims. In particular, such phage can be utilized to display antigen binding domains expressed from a repertoire or combinatorial antibody library (*e.g.*, human or murine). Phage expressing an antigen binding domain that binds the antigen of interest can be selected or identified with antigen, *e.g.,* using labelled antigen or antigen bound or captured to a solid surface or bead. Phage used in these methods are typically filamentous phage including fd and M13 binding domains expressed from phage with Fab, Fv or disulfide stabilized Fv antibody domains recombinantly fused to either the phage gene III or gene VIII protein. Phage display methods that can be used to make the antibodies of the present invention include those disclosed in Brinkman et al., J. Immunol. Methods 182:41-50 (1995); Ames et al., J. Immunol. Methods 184:177-186 (1995); Kettleborough et al., Eur. J. Immunol. 24:952-958 (1994); Persic et al., Gene 187 9-18 (1997); Burton et al., Advances in Immunology 57:191-280 (1994); PCT Application No. PCT/GB91/01134; PCT Publications WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and U.S. Patent Nos. 5,698,426; 5,223,409; 5,403,484; 5,580,717; 5,427,908; 5,750,753; 5,821,047; 5,571,698; 5,427,908; 5,516,637; 5,780,225; 5,658,727; 5,733,743 and 5,969,108; each of which is incorporated herein by reference in its entirety.

As described in the above references, after phage selection, the antibody coding regions from the phage can be isolated and used to generate whole antibodies, including human antibodies, or any other desired antigen binding fragment, and expressed in any desired host, including mammalian cells, insect cells, plant cells, yeast, and bacteria, e.g., as described in detail below. For example, techniques to recombinantly produce Fab, Fab' and $F(ab')_2$ fragments can also be employed using methods known in the art such as those disclosed in PCT publication WO 92/22324; Mullinax et al., BioTechniques 12(6):864-869 (1992); and Sawai et al., AJRI 34:26-34 (1995); and

Better et al., Science 240:1041-1043 (1988) (said references incorporated by reference in their entireties).

Examples of techniques which can be used to produce single-chain Fvs and antibodies include those described in U.S. Patents 4,946,778 and 5,258,498; Huston *et al., Methods in Enzymology* 203:46-88 (1991); Shu *et al., PNAS* 90:7995-7999 (1993); and Skerra *et al., Science* 240:1038-1040 (1988).

The invention further provides for the use of bispecific antibodies, which can be made by methods known in the art. Traditional production of full length bispecific antibodies is based on the coexpression of two immunoglobulin heavy chain-light chain pairs, where the two chains have different specificities (Milstein et al., 1983, Nature 305:537-539). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of 10 different antibody molecules, of which only one has the correct bispecific structure. Purification of the correct molecule, which is usually done by affinity chromatography steps, is rather cumbersome, and the product yields are low. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., 1991, EMBO J. 10:3655-3659.

According to a different and more preferred approach, antibody variable domains with the desired binding specificities (antibody-antigen combining sites) are fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light chain binding, present in at least one of the fusions. DNAs encoding the immunoglobulin heavy chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. This provides for great flexibility in adjusting the mutual proportions of the three polypeptide fragments in embodiments when unequal ratios of the three polypeptide chains used in the construction provide the optimum yields. It is, however, possible to insert the coding sequences for two or all

three polypeptide chains in one expression vector when the expression of at least two polypeptide chains in equal ratios results in high yields or when the ratios are of no particular significance.

In a preferred embodiment of this approach, the bispecific antibodies are composed of a hybrid immunoglobulin heavy chain with a first binding specificity in one arm, and a hybrid immunoglobulin heavy chain-light chain pair (providing a second binding specificity) in the other arm. It was found that this asymmetric structure facilitates the separation of the desired bispecific compound from unwanted immunoglobulin chain combinations, as the presence of an immunoglobulin light chain in only one half of the bispecific molecule provides for a facile way of separation. This approach is disclosed in WO 94/04690 published March 3, 1994. For further details for generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 1986, 121:210.

The invention provides functionally active fragments, derivatives or analogues of the anti-Integrin beta 4 immunoglobulin molecules. Functionally active means that the fragment, derivative or analogue is able to elicit anti-anti-idiotype antibodies (*i.e.*, tertiary antibodies) that recognize the same antigen that is recognized by the antibody from which the fragment, derivative or analogue is derived. Specifically, in a particular embodiment the antigenicity of the idiotype of the immunoglobulin molecule may be enhanced by deletion of framework and CDR sequences that are C-terminal to the CDR sequence that specifically recognizes the antigen. To determine which CDR sequences bind the antigen, synthetic peptides containing the CDR sequences can be used in binding assays with the antigen by any binding assay method known in the art.

The present invention provides antibody fragments such as, but not limited to, $F(ab')_2$ fragments and Fab fragments. Antibody fragments which recognize specific epitopes may be generated by known techniques. $F(ab')_2$ fragments consist of the variable region, the light chain constant region and the CH1 domain of the heavy chain and are generated by pepsin digestion of the antibody molecule. Fab fragments are generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. The invention

also provides heavy chain and light chain dimers of the antibodies of the invention, or any minimal fragment thereof such as Fvs or single chain antibodies (SCAs) (*e.g.*, as described in U.S. Patent 4,946,778; Bird, 1988, Science 242:423-42; Huston et al., 1988, Proc. Natl. Acad. Sci. USA 85:5879-5883; and Ward et al., 1989, Nature 5 334:544-54), or any other molecule with the same specificity as the antibody of the invention. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide. Techniques for the assembly of functional Fv fragments in *E. coli* may be used (Skerra et al., 1988, Science 242:1038-1041).

10 In other embodiments, the invention provides fusion proteins of the immunoglobulins of the invention (or functionally active fragments thereof), for example in which the immunoglobulin is fused via a covalent bond (*e.g.*, a peptide bond), at either the N-terminus or the C-terminus to an amino acid sequence of another protein (or portion thereof, preferably at least 10, 20 or 50 amino acid portion of the 15 protein) that is not the immunoglobulin. Preferably the immunoglobulin, or fragment thereof, is covalently linked to the other protein at the N-terminus of the constant domain. As stated above, such fusion proteins may facilitate purification, increase half-life *in vivo*, and enhance the delivery of an antigen across an epithelial barrier to the immune system.

20 The immunoglobulins of the invention include analogues and derivatives that are modified, i.e., by the covalent attachment of any type of molecule as long as such covalent attachment does not impair immunospecific binding. For example, but not by way of limitation, the derivatives and analogues of the immunoglobulins include those that have been further modified, e.g., by glycosylation, acetylation, pegylation, 25 phosphylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc. Any of numerous chemical modifications may be carried out by known techniques, including, but not limited to specific chemical cleavage, acetylation, formylation, etc. Additionally, the analogue or derivative may contain one or more non-classical amino acids.

The foregoing antibodies can be used in methods known in the art relating to the localization and activity of Integrin beta 4, e.g., for imaging this protein, measuring levels thereof in appropriate physiological samples, in diagnostic methods, etc.

5 Production of Affibodies to Integrin beta 4

Affibody molecules represent a new class of affinity proteins based on a 58-amino acid residue protein domain, derived from one of the IgG-binding domains of staphylococcal protein A. This three helix bundle domain has been used as a scaffold for the construction of combinatorial phagemid libraries, from which Affibody variants 10 that target the desired molecules can be selected using phage display technology (Nord K, Gunneriusson E, Ringdahl J, Stahl S, Uhlen M, Nygren PA, Binding proteins selected from combinatorial libraries of an α-helical bacterial receptor domain, Nat Biotechnol 1997;15:772-7. Ronmark J, Gronlund H, Uhlen M, Nygren PA, Human immunoglobulin A (IgA)-specific ligands from combinatorial engineering of protein A, 15 Eur J Biochem 2002;269:2647-55.). The simple, robust structure of Affibody molecules in combination with their low molecular weight (6 kDa), make them suitable for a wide variety of applications, for instance, as detection reagents (Ronmark J, Hansson M, Nguyen T, et al, Construction and characterization of affibody-Fc chimeras produced in *Escherichia coli*, J Immunol Methods 2002;261:199-211) and to inhibit 20 receptor interactions (Sandstorm K, Xu Z, Forsberg G, Nygren PA, Inhibition of the CD28-CD80 co-stimulation signal by a CD28-binding Affibody ligand developed by combinatorial protein engineering, Protein Eng 2003;16:691-7). Further details of Affibodies and methods of production thereof may be obtained by reference to US Patent No 5831012 which is herein incorporated by reference in its entirety.

25 Labelled Affibodies may also be useful in imaging applications for determining abundance of Isoforms.

Production of Domain Antibodies to Integrin beta 4

References to antibodies herein embrace references to Domain Antibodies.

Domain Antibodies (dAbs) are the smallest functional binding units of antibodies, corresponding to the variable regions of either the heavy (V_H) or light (V_L) chains of human antibodies. Domain Antibodies have a molecular weight of approximately 13 kDa. Domantis has developed a series of large and highly functional libraries of fully
5 human V_H and V_L dAbs (more than ten billion different sequences in each library), and uses these libraries to select dAbs that are specific to therapeutic targets. In contrast to many conventional antibodies, Domain Antibodies are well expressed in bacterial, yeast, and mammalian cell systems. Further details of domain antibodies and methods of production thereof may be obtained by reference to US Patent 6,291,158; 6,582,915;
10 6,593,081; 6,172,197; 6,696,245; US Serial No. 2004/0110941; European patent application No. 1433846 and European Patents 0368684 & 0616640; WO05/035572, WO04/101790, WO04/081026, WO04/058821, WO04/003019 and WO03/002609, each of which is herein incorporated by reference in its entirety.

15 Production of Nanobodies to Integrin beta 4

Nanobodies are antibody-derived therapeutic proteins that contain the unique structural and functional properties of naturally-occurring heavy-chain antibodies. These heavy-chain antibodies contain a single variable domain (VHH) and two constant domains (C_H2 and C_H3). Importantly, the cloned and isolated VHH domain is a
20 perfectly stable polypeptide harbouring the full antigen-binding capacity of the original heavy-chain antibody. Nanobodies have a high homology with the VH domains of human antibodies and can be further humanised without any loss of activity. Importantly, Nanobodies have a low immunogenic potential, which has been confirmed in primate studies with Nanobody lead compounds.

25 Nanobodies combine the advantages of conventional antibodies with important features of small molecule drugs. Like conventional antibodies, Nanobodies show high target specificity, high affinity for their target and low inherent toxicity. However, like small molecule drugs they can inhibit enzymes and readily access receptor clefts. Furthermore, Nanobodies are extremely stable, can be administered by means other

than injection (see e.g. WO 04/041867, which is herein incorporated by reference in its entirety) and are easy to manufacture. Other advantages of Nanobodies include recognising uncommon or hidden epitopes as a result of their small size, binding into cavities or active sites of protein targets with high affinity and selectivity due to their unique 3-dimensional, drug format flexibility, tailoring of half-life and ease and speed of drug discovery.

Nanobodies are encoded by single genes and are efficiently produced in almost all prokaryotic and eukaryotic hosts e.g. *E. coli* (see e.g. US 6,765,087, which is herein incorporated by reference in its entirety), moulds (for example *Aspergillus* or *Trichoderma*) and yeast (for example *Saccharomyces*, *Kluyveromyces*, *Hansenula* or *Pichia*) (see e.g. US 6,838,254, which is herein incorporated by reference in its entirety). The production process is scalable and multi-kilogram quantities of Nanobodies have been produced. Because Nanobodies exhibit a superior stability compared with conventional antibodies, they can be formulated as a long shelf-life, ready-to-use solution.

The Nanoclone method (see e.g. WO 06/079372, which is herein incorporated by reference in its entirety) is a proprietary method for generating Nanobodies against a desired target, based on automated high-throughput selection of B-cells.

Production of Unibodies to Integrin beta 4

UniBodies are another antibody fragment technology; however this one is based upon the removal of the hinge region of IgG4 antibodies. The deletion of the hinge region results in a molecule that is essentially half the size of traditional IgG4 antibodies and has a univalent binding region rather than the bivalent binding region of IgG4 antibodies. It is also well known that IgG4 antibodies are inert and thus do not interact with the immune system, which may be advantageous for the treatment of diseases where an immune response is not desired, and this advantage is passed onto UniBodies. For example, UniBodies may function to inhibit or silence, but not kill, the cells to which they are bound. Additionally, UniBody binding to cancer cells do not stimulate them to

proliferate. Furthermore, because UniBodies are about half the size of traditional IgG4 antibodies, they may show better distribution over larger solid tumours with potentially advantageous efficacy. UniBodies are cleared from the body at a similar rate to whole IgG4 antibodies and are able to bind with a similar affinity for their antigens as whole
5 antibodies. Further details of UniBodies may be obtained by reference to patent WO2007/059782, which is herein incorporated by reference in its entirety.

Production of DARPins to Integrin beta 4

DARPins (Designed Ankyrin Repeat Proteins) are one example of an antibody
10 mimetic DRP (Designed Repeat Protein) technology that has been developed to exploit the binding abilities of non-antibody polypeptides. Repeat proteins such as ankyrin or leucine-rich repeat proteins, are ubiquitous binding molecules, which occur, unlike antibodies, intra- and extracellularly. Their unique modular architecture features repeating structural units (repeats), which stack together to form elongated repeat domains displaying variable
15 and modular target-binding surfaces. Based on this modularity, combinatorial libraries of polypeptides with highly diversified binding specificities can be generated. This strategy includes the consensus design of self-compatible repeats displaying variable surface residues and their random assembly into repeat domains.

DARPins can be produced in bacterial expression systems at very high yields and
20 they belong to the most stable proteins known. Highly specific, high-affinity DARPins to a broad range of target proteins, including human receptors, cytokines, kinases, human proteases, viruses and membrane proteins, have been selected. DARPins having affinities in the single-digit nanomolar to picomolar range can be obtained.

DARPins have been used in a wide range of applications, including ELISA,
25 sandwich ELISA, flow cytometric analysis (FACS), immunohistochemistry (IHC), chip applications, affinity purification or Western blotting. DARPins also proved to be highly active in the intracellular compartment for example as intracellular marker proteins fused to green fluorescent protein (GFP). DARPins were further used to inhibit viral entry with IC50 in the pM range. DARPins are not only ideal to block protein-protein interactions,

but also to inhibit enzymes. Proteases, kinases and transporters have been successfully inhibited, most often an allosteric inhibition mode. Very fast and specific enrichments on the tumour and very favourable tumour to blood ratios make DARPins well suited for in vivo diagnostics or therapeutic approaches.

5 Additional information regarding DARPins and other DRP technologies can be found in US Patent Application Publication No. 2004/0132028, and International Patent Application Publication No. WO 02/20565, both of which are hereby incorporated by reference in their entirety.

10 Production of Anticalins to Integrin beta 4

Anticalins are an additional antibody mimetic technology, however in this case the binding specificity is derived from lipocalins, a family of low molecular weight proteins that are naturally and abundantly expressed in human tissues and body fluids. Lipocalins have evolved to perform a range of functions in vivo associated with the physiological transport 15 and storage of chemically sensitive or insoluble compounds. Lipocalins have a robust intrinsic structure comprising a highly conserved ß-barrel which supports four loops at one terminus of the protein. These loops form the entrance to a binding pocket and conformational differences in this part of the molecule account for the variation in binding specificity between individual lipocalins.

20 While the overall structure of hypervariable loops supported by a conserved ß-sheet framework is reminiscent of immunoglobulins, lipocalins differ considerably from antibodies in terms of size, being composed of a single polypeptide chain of 160-180 amino acids which is marginally larger than a single immunoglobulin domain.

Lipocalins are cloned and their loops are subjected to engineering in order to create 25 Anticalins. Libraries of structurally diverse Anticalins have been generated and Anticalin display allows the selection and screening of binding function, followed by the expression and production of soluble protein for further analysis in prokaryotic or eukaryotic systems. Studies have successfully demonstrated that Anticalins can be developed that are specific for virtually any human target protein; they can be isolated and binding affinities in the

nanomolar or higher range can be obtained.

Anticalins can also be formatted as dual targeting proteins, so-called Duocalins. A Duocalin binds two separate therapeutic targets in one easily produced monomeric protein using standard manufacturing processes while retaining target specificity and affinity
5 regardless of the structural orientation of its two binding domains.

Modulation of multiple targets through a single molecule is particularly advantageous in diseases known to involve more than a single causative factor. Moreover, bi- or multivalent binding formats such as Duocalins have significant potential in targeting cell surface molecules in disease, mediating agonistic effects on signal transduction
10 pathways or inducing enhanced internalization effects via binding and clustering of cell surface receptors. Furthermore, the high intrinsic stability of Duocalins is comparable to monomeric Anticalins, offering flexible formulation and delivery potential for Duocalins.

Additional information regarding Anticalins can be found in US Patent No. 7,250,297 and International Patent Application Publication No. WO 99/16873, both of
15 which are hereby incorporated by reference in their entirety.

Production of Avimers to Integrin beta 4

Avimers are evolved from a large family of human extracellular receptor domains by in vitro exon shuffling and phage display, generating multidomain proteins with binding
20 and inhibitory properties. Linking multiple independent binding domains has been shown to create avidity and results in improved affinity and specificity compared with conventional single-epitope binding proteins. Other potential advantages include simple and efficient production of multitarget-specific molecules in Escherichia coli, improved thermostability and resistance to proteases. Avimers with sub-nanomolar affinities have been obtained
25 against a variety of targets.

Additional information regarding Avimers can be found in US Patent Application Publication Nos. 2006/0286603, 2006/0234299, 2006/0223114, 2006/0177831, 2006/0008844, 2005/0221384, 2005/0164301, 2005/0089932, 2005/0053973, 2005/0048512, 2004/0175756, all of which are hereby incorporated by

reference in their entirety.

Production of Versabodies to Integrin beta 4

Versabodies are small proteins of 3-5 kDa with >15% cysteines, which form a high disulfide density scaffold, replacing the hydrophobic core that typical proteins have. The replacement of a large number of hydrophobic amino acids, comprising the hydrophobic core, with a small number of disulfides results in a protein that is smaller, more hydrophilic (less aggregation and non-specific binding), more resistant to proteases and heat, and has a lower density of T-cell epitopes, because the residues that contribute most to MHC presentation are hydrophobic. All four of these properties are well-known to affect immunogenicity, and together they are expected to cause a large decrease in immunogenicity.

The inspiration for Versabodies comes from the natural injectable biopharmaceuticals produced by leeches, snakes, spiders, scorpions, snails, and anemones, which are known to exhibit unexpectedly low immunogenicity. Starting with selected natural protein families, by design and by screening the size, hydrophobicity, proteolytic antigen processing, and epitope density are minimized to levels far below the average for natural injectable proteins.

Given the structure of Versabodies, these antibody mimetics offer a versatile format that includes multi-valency, multi-specificity, a diversity of half-life mechanisms, tissue targeting modules and the absence of the antibody Fc region. Furthermore, Versabodies are manufactured in E. coli at high yields, and because of their hydrophilicity and small size, Versabodies are highly soluble and can be formulated to high concentrations. Versabodies are exceptionally heat stable (they can be boiled) and offer extended shelf-life.

Additional information regarding Versabodies can be found in US Patent Application Publication No. 2007/0191272 which is hereby incorporated by reference in its entirety.

Expression of Affinity Reagents

Expression of Antibodies

The antibodies of the invention can be produced by any method known in the art for the synthesis of antibodies, in particular, by chemical synthesis or by recombinant expression, and are preferably produced by recombinant expression techniques.

Recombinant expression of antibodies, or fragments, derivatives or analogues thereof, requires construction of a nucleic acid that encodes the antibody. If the nucleotide sequence of the antibody is known, a nucleic acid encoding the antibody may be assembled from chemically synthesized oligonucleotides (*e.g.*, as described in Kutmeier et al., 1994, BioTechniques 17:242), which, briefly, involves the synthesis of overlapping oligonucleotides containing portions of the sequence encoding antibody, annealing and ligation of those oligonucleotides, and then amplification of the ligated oligonucleotides by PCR.

Alternatively, the nucleic acid encoding the antibody may be obtained by cloning the antibody. If a clone containing the nucleic acid encoding the particular antibody is not available, but the sequence of the antibody molecule is known, a nucleic acid encoding the antibody may be obtained from a suitable source (*e.g.*, an antibody cDNA library, or cDNA library generated from any tissue or cells expressing the antibody) by PCR amplification using synthetic primers hybridizable to the 3' and 5' ends of the sequence or by cloning using an oligonucleotide probe specific for the particular gene sequence.

If an antibody molecule that specifically recognizes a particular antigen is not available (or a source for a cDNA library for cloning a nucleic acid encoding such an antibody), antibodies specific for a particular antigen may be generated by any method known in the art, for example, by immunizing an animal, such as a rabbit, to generate polyclonal antibodies or, more preferably, by generating monoclonal antibodies. Alternatively, a clone encoding at least the Fab portion of the antibody may be obtained by screening Fab expression libraries (*e.g.*, as described in Huse et al., 1989, *Science* 246:1275-1281) for clones of Fab fragments that bind the specific antigen or by

screening antibody libraries (See, *e.g.,* Clackson et al., 1991, Nature 352:624; Hane et al., 1997 Proc. Natl. Acad. Sci. USA 94:4937).

Once a nucleic acid encoding at least the variable domain of the antibody molecule is obtained, it may be introduced into a vector containing the nucleotide
5 sequence encoding the constant region of the antibody molecule (see, *e.g.*, PCT Publication WO 86/05807; PCT Publication WO 89/01036; and U.S. Patent No. 5,122,464). Vectors containing the complete light or heavy chain for co-expression with the nucleic acid to allow the expression of a complete antibody molecule are also available. Then, the nucleic acid encoding the antibody can be used to introduce the
10 nucleotide substitution(s) or deletion(s) necessary to substitute (or delete) the one or more variable region cysteine residues participating in an intrachain disulfide bond with an amino acid residue that does not contain a sulfhydyl group. Such modifications can be carried out by any method known in the art for the introduction of specific mutations or deletions in a nucleotide sequence, for example, but not limited to, chemical
15 mutagenesis, in vitro site directed mutagenesis (Hutchinson et al., 1978, J. Biol. Chem. 253:6551), PCT based methods, etc.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., 1984, Proc. Natl. Acad. Sci. 81:851-855; Neuberger et al., 1984, *Nature* 312:604-608; Takeda et al., 1985, Nature 314:452-454) by splicing genes from
20 a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. As described *supra*, a chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human antibody constant region, *e.g.*, humanized antibodies.

25 Once a nucleic acid encoding an antibody molecule of the invention has been obtained, the vector for the production of the antibody molecule may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing the protein of the invention by expressing nucleic acid containing the antibody molecule sequences are described herein. Methods which are well known to

those skilled in the art can be used to construct expression vectors containing an antibody molecule coding sequences and appropriate transcriptional and translational control signals. These methods include, for example, *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. See, for example, the techniques described in Sambrook et al. (1990, Molecular Cloning, A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) and Ausubel et al. (eds., 1998, Current Protocols in Molecular Biology, John Wiley & Sons, NY).

The expression vector is transferred to a host cell by conventional techniques and the transfected cells are then cultured by conventional techniques to produce an antibody of the invention.

The host cells used to express a recombinant antibody of the invention may be either bacterial cells such as *Escherichia coli,* or, preferably, eukaryotic cells, especially for the expression of whole recombinant antibody molecule. In particular, mammalian cells such as Chinese hamster ovary cells (CHO), in conjunction with a vector such as the major intermediate early gene promoter element from human cytomegalovirus are an effective expression system for antibodies (Foecking et al., 1986, Gene 45:101; Cockett et al., 1990, Bio/Technology 8:2).

A variety of host-expression vector systems may be utilized to express an antibody molecule of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, express the antibody molecule of the invention *in situ.* These include but are not limited to microorganisms such as bacteria (*e.g.*, *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing antibody coding sequences; yeast (*e.g.*, *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing antibody coding sequences; insect cell systems infected with recombinant virus expression vectors (*e.g.*, baculovirus) containing the antibody coding sequences;

plant cell systems infected with recombinant virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (*e.g.*, Ti plasmid) containing antibody coding sequences; or mammalian cell systems (*e.g.*, COS, CHO, BHK, 293, 3T3 cells) harbouring
5 recombinant expression constructs containing promoters derived from the genome of mammalian cells (*e.g.*, metallothionein promoter) or from mammalian viruses (*e.g.*, the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the antibody molecule being expressed.
10 For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions comprising an antibody molecule, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther et al., 1983, EMBO J. 2:1791), in which the
15 antibody coding sequence may be ligated individually into the vector in frame with the *lac Z* coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, *Nucleic Acids Res.* 13:3101-3109; Van Heeke & Schuster, 1989, J. Biol. Chem. 24:5503-5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such
20 fusion proteins are soluble and can easily be purified from lysed cells by adsorption and binding to a matrix glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.
25 In an insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The antibody coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). In mammalian

host cells, a number of viral-based expression systems (*e.g.*, an adenovirus expression system) may be utilized.

As discussed above, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (*e.g.*, glycosylation) and processing (*e.g.*, cleavage) of protein products may be important for the function of the protein.

For long-term, high-yield production of recombinant antibodies, stable expression is preferred. For example, cell lines that stably express an antibody of interest can be produced by transfecting the cells with an expression vector comprising the nucleotide sequence of the antibody and the nucleotide sequence of a selectable (*e.g.*, neomycin or hygromycin), and selecting for expression of the selectable marker. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that interact directly or indirectly with the antibody molecule.

The expression levels of the antibody molecule can be increased by vector amplification (for a review, see Bebbington and Hentschel, The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells in DNA cloning, Vol.3. (Academic Press, New York, 1987)). When a marker in the vector system expressing antibody is amplifiable, increase in the level of inhibitor present in culture of host cell will increase the number of copies of the marker gene. Since the amplified region is associated with the antibody gene, production of the antibody will also increase (Crouse et al., 1983, Mol. Cell. Biol. 3:257).

The host cell may be co-transfected with two expression vectors of the invention, the first vector encoding a heavy chain derived polypeptide and the second vector encoding a light chain derived polypeptide. The two vectors may contain identical selectable markers which enable equal expression of heavy and light chain polypeptides. Alternatively, a single vector may be used which encodes both heavy and light chain polypeptides. In such situations, the light chain should be placed before the heavy chain to avoid an excess of toxic free heavy chain (Proudfoot, 1986, Nature 322:52; Kohler, 1980, Proc. Natl. Acad. Sci. USA 77:2197). The coding sequences for

the heavy and light chains may comprise cDNA or genomic DNA.

Once the antibody molecule of the invention has been recombinantly expressed, it may be purified by any method known in the art for purification of an antibody molecule, for example, by chromatography (*e.g.*, ion exchange chromatography, affinity chromatography such as with protein A or specific antigen, and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of proteins.

Alternatively, any fusion protein may be readily purified by utilizing an antibody specific for the fusion protein being expressed. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht et al., 1991, Proc. Natl. Acad. Sci. USA 88:8972-897). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the open reading frame of the gene is translationally fused to an amino-terminal tag consisting of six histidine residues. The tag serves as a matrix binding domain for the fusion protein. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni^{2+} nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

The antibodies that are generated by these methods may then be selected by first screening for affinity and specificity with the purified polypeptide of interest and, if required, comparing the results to the affinity and specificity of the antibodies with polypeptides that are desired to be excluded from binding. The screening procedure can involve immobilization of the purified polypeptides in separate wells of microtiter plates. The solution containing a potential antibody or groups of antibodies is then placed into the respective microtiter wells and incubated for about 30 min to 2 h. The microtiter wells are then washed and a labelled secondary antibody (for example, an anti-mouse antibody conjugated to alkaline phosphatase if the raised antibodies are mouse antibodies) is added to the wells and incubated for about 30 min and then washed. Substrate is added to the wells and a colour reaction will appear where antibody to the immobilized polypeptide(s) is present.

The antibodies so identified may then be further analyzed for affinity and specificity in the assay design selected. In the development of immunoassays for a target protein, the purified target protein acts as a standard with which to judge the sensitivity and specificity of the immunoassay using the antibodies that have been selected. Because the binding affinity of various antibodies may differ; certain antibody pairs (*e.g.*, in sandwich assays) may interfere with one another sterically, etc., assay performance of an antibody may be a more important measure than absolute affinity and specificity of an antibody.

Those skilled in the art will recognize that many approaches can be taken in producing antibodies or binding fragments and screening and selecting for affinity and specificity for the various polypeptides, but these approaches do not change the scope of the invention.

For therapeutic applications, antibodies (particularly monoclonal antibodies) may suitably be human or humanized animal (e.g. mouse) antibodies. Animal antibodies may be raised in animals using the human protein (e.g. Integrin beta 4) as immunogen. Humanisation typically involves grafting CDRs identified thereby into human framework regions. Normally some subsequent retromutation to optimize the conformation of chains is required. Such processes are known to persons skilled in the art.

Expression of Affibodies

The construction of Affibodies has been described elsewhere (Ronnmark J, Gronlund H, Uhle´ n, M., Nygren P.A˚, Human immunoglobulin A (IgA)-specific ligands from combinatorial engineering of protein A, 2002, Eur. J. Biochem. 269, 2647–2655.), including the construction of affibody phage display libraries (Nord, K., Nilsson, J., Nilsson, B., Uhle´ n, M. & Nygren, P.A˚, A combinatorial library of an a-helical bacterial receptor domain, 1995, Protein Eng. 8, 601–608. Nord, K., Gunneriusson, E., Ringdahl, J., Sta˚ hl, S., Uhle´ n, M. & Nygren, P.A˚, Binding proteins selected from combinatorial libraries of an a-helical bacterial receptor domain,

1997, Nat. Biotechnol.15, 772–777.)

The biosensor analyses to investigate the optimal Affibody variants using biosensor binding studies has also been described elsewhere (Ronnmark J, Gronlund H, Uhle´ n, M., Nygren P.A°, Human immunoglobulin A (IgA)-specific ligands from combinatorial engineering of protein A, 2002, Eur. J. Biochem. 269, 2647–2655.).

Affinity Reagent Modifications

In a preferred embodiment, anti-Integrin beta 4 affinity reagents such as antibodies or fragments thereof are conjugated to a diagnostic moiety (such as a detectable label) or a therapeutic moiety. The antibodies can be used for diagnosis or to determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling the antibody to a detectable substance (label). Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, radioactive nuclides, positron emitting metals (for use in positron emission tomography), and nonradioactive paramagnetic metal ions. See generally U.S. Patent No. 4,741,900 for metal ions which can be conjugated to antibodies for use as diagnostics according to the present invention. Suitable enzymes include horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; suitable prosthetic groups include streptavidin, avidin and biotin; suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride and phycoerythrin; suitable luminescent materials include luminol; suitable bioluminescent materials include luciferase, luciferin, and aequorin; and suitable radioactive nuclides include ^{125}I, ^{131}I, ^{111}In and ^{99}Tc. ^{68}Ga may also be employed.

Anti-Integrin beta 4 antibodies or fragments thereof as well as other affinity reagents can be conjugated to a therapeutic agent or drug moiety to modify a given biological response. An exemplary therapeutic agent to which the affinity reagent may be conjugated is a cytotoxic moiety. The therapeutic agent or drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug

moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumour necrosis factor, α-interferon, β-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator, a thrombotic agent or an anti-angiogenic agent, *e.g.*, angiostatin or endostatin; or, a biological response modifier such as a lymphokine, interleukin-1 (IL-1), interleukin-2 (IL-2), interleukin-6 (IL-6), granulocyte macrophage colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF), nerve growth factor (NGF) or other growth factor.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld et al. (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom et al., "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson et al. (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera et al. (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin et al. (eds.), pp. 303-16 (Academic Press 1985), and Thorpe et al., "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev., 62:119-58 (1982).

Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

An antibody with or without a therapeutic moiety conjugated to it can be used as a therapeutic that is administered alone or in combination with cytotoxic factor(s) and/or cytokine(s).

The invention also provides for fully human, or humanised antibodies that induce antibody-directed cell-mediated cytotoxicity (ADCC). A fully human antibody is one in which the protein sequences are encoded by naturally occurring human

immunoglobulin sequences, either from isolated antibody-producing human B-lymphocytes, or from transgenic murine B-lymphocytes of mice in which the murine immunoglobulin coding chromosomal regions have been replaced by orthologous human sequences. Transgenic antibodies of the latter type include, but are not restricted to, HuMab (Medarex, Inc., CA) and Xenomouse (Abgenix Inc., CA). A humanised antibody is one in which the constant region of a non-human antibody molecule of appropriate antigen specificity, is replaced by the constant region of a human antibody, preferably of the IgG subtype, with appropriate effector functions (Morrison et al., 1984, Proc. Natl. Acad. Sci. 81:851-855; Neuberger et al., 1984, *Nature* 312:604-608; Takeda et al., 1985, Nature 314:452-454). Appropriate effector functions include ADCC, which is a natural process by which fully-human antibodies or humanized antibodies, when bound to targets on the surface of cancer cells, switch on the cell killing properties of lymphocytes that are part of the normal immune system. These active lymphocytes, called Natural Killer (NK) cells, use a cytotoxic process to destroy living cells to which the antibodies are bound. ADCC activity may be detected and quantified by measuring release of Europium (Eu3+) from Eu3+ labelled, living cells in the presence of an antigen-specific antibody and peripheral blood mononuclear cells extracted from an immunocompetent, living human subject. The ADCC process is described in detail in Janeway Jr. C.A. et al., Immunobiology, 5th ed., 2001, Garland Publishing, ISBN 0-8153-3642-X; Pier G.B. et al., Immunology, Infection, and Immunity, 2004, p246-5; Albanell J. et al., Advances in Experimental Medicine and Biology, 2003, 532:p2153-68 and Weng, W.-K. et al., Journal of Clinical Oncology, 2003, 21:p 3940-3947. Suitable methods for the detection and quantification of ADCC can be found in Blomberg et al., Journal of Immunological Methods. 1986, 86:p225-9; Blomberg et al., Journal of Immunological Methods. 1986, 21;92:p117-23 and Patel & Boyd, Journal of Immunological Methods. 1995, 184:p29-38.

ADCC typically involves activation of NK cells and is dependent on the recognition of antibody-coated cells by Fc receptors on the surface of the NK cell. The Fc receptors recognize the Fc (crystalline) portion of antibodies such as IgG, bound

specifically to the surface of a target cell. The Fc receptor that triggers activation of the NK cell is called CD16 or FcγRIIIa. Once the FcγRIIIa receptor is bound to the IgG Fc, the NK cell releases cytokines such as IFN-γ, and cytotoxic granules containing perforin and granzymes that enter the target cell and promote cell death by triggering
5 apoptosis.

The induction of antibody-dependent cellular cytotoxicity (ADCC) by an antibody can be enhanced by modifications that alter interactions between the antibody constant region (Fc) and various receptors that are present on the surface of cells of the immune system. Such modifications include the reduction or absence of alpha1,6-linked
10 fucose moieties in the complex oligosaccharide chains that are normally added to the Fc of antibodies during natural or recombinant synthesis in mammalian cells. In a particular embodiment, non-fucosylated anti-Integrin beta 4 affinity reagents such as antibodies or fragments thereof are produced for the purpose of enhancing their ability to induce the ADCC response.

15 Techniques for reducing or ablating alpha1,6-linked fucose moieties in the oligosaccharide chains of the Fc are well established. In one example, the recombinant antibody is synthesized in a cell line that is impaired in its ability to add fucose in an alpha 1,6 linkage to the innermost N-acetylglucosamine of the N-linked biantennary complex-type Fc oligosaccharides. Such cell lines include, but are not limited to, the rat
20 hybridoma YB2/0, which expresses a reduced level of the alpha 1,6-fucosyltransferase gene, FUT8. Preferably, the antibody is synthesized in a cell line that is incapable of adding alpha 1,6-linked fucosyl moieties to complex oligosaccharide chains, due to the deletion of both copies of the FUT8 gene. Such cell lines include, but are not limited to, FUT8-/- CHO/DG44 cell lines. Techniques for synthesizing partially fucosylated, or
25 non-fucosylated antibodies and affinity reagents are described in Shinkawa et al., J. Biol. Chem. 278:3466–34735 (2003); Yamane-Ohnuki et al., Biotechnology and Bioengineering 87: 614-22 (2004) and in WO00/61739 A1, WO02/31140 A1 and WO03/085107 A1. In a second example, the fucosylation of a recombinant antibody is reduced or abolished by synthesis in a cell line that has been genetically engineered to

overexpress a glycoprotein-modifying glycosyl transferase at a level that maximizes the production of complex N-linked oligosaccharides carrying bisecting N-acetylglucosamine. For example, the antibody is synthesized in a Chinese Hamster Ovary cell line expressing the enzyme N-acetyl glucosamine transferase III (GnT III).
5 Cell lines stably transfected with suitable glycoprotein-modifying glycosyl transferases, and methods of synthesizing antibodies using these cells are described in WO9954342.

A non-fucosylated antibody or affinity reagent can be used as a therapeutic that is administered alone or in combination with cytotoxic factor(s) and/or cytokine(s).

In a further modification, the amino acid sequences of the antibody Fc are
10 altered in a way that enhances ADCC activation, without affecting ligand affinity. Examples of such modifications are described in Lazar et al., Proceedings of the National Academy of Sciences 2006, 103:p4005-4010; WO03074679 and WO2007039818. In these examples, substitution of amino acids in the antibody Fc, such as aspartate for serine at position 239, and isoleucine for glutamate at position
15 332, altered the binding affinity of an antibody for Fc receptors, leading to an increase in ADCC activation.

An antibody reagent with enhanced ADCC activation due to amino acid substitutions can be used as a therapeutic that is administered alone or in combination with cytotoxic factor(s) and/or cytokine(s).

20

Diagnosis of Breast Cancer, Colorectal Cancer, Gastric Cancer, Hepatocellular Carcinoma, Lung Cancer and Pancreatic Cancer

In accordance with the present invention, test samples of breast, colorectal, gastric epithelium, liver, lung or pancreatic tissue, serum, plasma or urine obtained from
25 a subject suspected of having or known to have breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer can be used for diagnosis or monitoring. In one embodiment, a change in the abundance of Integrin beta 4 in a test sample relative to a control sample (from a subject or subjects free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer

and pancreatic cancer) or a previously determined reference range indicates the presence of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. In another embodiment, the relative abundance of Integrin beta 4 in a test sample compared to a control sample or a previously determined reference range indicates a subtype of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer (*e.g.* inflammatory breast cancer, familial or sporadic colorectal cancer, gastrointestinal stromal tumours, fibrolamellar hepatocellular carcinoma, squamous cell lung carcinoma or endocrine tumours of the pancreas). In yet another embodiment, the relative abundance of Integrin beta 4 in a test sample relative to a control sample or a previously determined reference range indicates the degree or severity of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer (e.g. the likelihood for metastasis). In any of the aforesaid methods, detection of Integrin beta 4 may optionally be combined with detection of one or more of additional biomarkers for breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. Any suitable method in the art can be employed to measure the level of Integrin beta 4, including but not limited to the Preferred Technology described herein, kinase assays, immunoassays to detect and/or visualize Integrin beta 4 (*e.g.*, Western blot, immunoprecipitation followed by sodium dodecyl sulfate polyacrylamide gel electrophoresis, immunocytochemistry, etc.). In a further embodiment, a change in the abundance of mRNA encoding Integrin beta 4 in a test sample relative to a control sample or a previously determined reference range indicates the presence of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. Any suitable hybridization assay can be used to detect Integrin beta 4 expression by detecting and/or visualizing mRNA encoding the Integrin beta 4 (*e.g.*, Northern assays, dot blots, *in situ* hybridization, etc.).

In another embodiment of the invention, labelled antibodies (or other affinity reagents), derivatives and analogues thereof, which specifically bind to Integrin beta 4

can be used for diagnostic purposes to detect, diagnose, or monitor breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer. For example, breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer may be detected in an animal, such as in a mammal and particularly in a human.

Screening Assays

The invention provides methods for identifying agents (e.g., candidate compounds or test compounds) that bind to Integrin beta 4 or have a stimulatory or inhibitory effect on the expression or activity of Integrin beta 4. The invention also provides methods of identifying agents, candidate compounds or test compounds that bind to a Integrin beta 4-related polypeptide or a Integrin beta 4 fusion protein or have a stimulatory or inhibitory effect on the expression or activity of a Integrin beta 4-related polypeptide or a Integrin beta 4 fusion protein. Examples of agents, candidate compounds or test compounds include, but are not limited to, nucleic acids (*e.g.*, DNA and RNA), carbohydrates, lipids, proteins, peptides, peptidomimetics, small molecules and other drugs. Agents can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, 1997, Anticancer Drug Des. 12:145; U.S. Patent No. 5,738,996; and U.S. Patent No.5,807,683, each of which is incorporated herein in its entirety by reference).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt et al., 1993, Proc. Natl. Acad. Sci. USA 90:6909; Erb et al., 1994, Proc. Natl. Acad. Sci. USA 91:11422; Zuckermann et al., 1994, J. Med.

Chem. 37:2678; Cho et al., 1993, Science 261:1303; Carrell et al., 1994, Angew. Chem. Int. Ed. Engl. 33:2059; Carell et al., 1994, Angew. Chem. Int. Ed. Engl. 33:2061; and Gallop et al., 1994, J. Med. Chem. 37:1233, each of which is incorporated herein in its entirety by reference.

5 Libraries of compounds may be presented, e.g., presented in solution (*e.g.*, Houghten, 1992, Bio/Techniques 13:412-421), or on beads (Lam, 1991, Nature 354:82-84), chips (Fodor, 1993, Nature 364:555-556), bacteria (U.S. Patent No. 5,223,409), spores (Patent Nos. 5,571,698; 5,403,484; and 5,223,409), plasmids (Cull et al., 1992, Proc. Natl. Acad. Sci. USA 89:1865-1869) or phage (Scott and Smith, 10 1990, Science 249:386-390; Devlin, 1990, Science 249:404-406; Cwirla et al., 1990, Proc. Natl. Acad. Sci. USA 87:6378-6382; and Felici, 1991, J. Mol. Biol. 222:301-310), each of which is incorporated herein in its entirety by reference.

In one embodiment, agents that interact with (*i.e.*, bind to) Integrin beta 4, a Integrin beta 4 fragment (e.g. a functionally active fragment), a Integrin beta 4-related 15 polypeptide, a fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein are identified in a cell-based assay system. In accordance with this embodiment, cells expressing Integrin beta 4, a fragment of Integrin beta 4, a Integrin beta 4-related polypeptide, a fragment of the Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein are contacted with a candidate compound or a control 20 compound and the ability of the candidate compound to interact with Integrin beta 4 is determined. If desired, this assay may be used to screen a plurality (e.g. a library) of candidate compounds. The cell, for example, can be of prokaryotic origin (*e.g.*, *E. coli*) or eukaryotic origin (*e.g.*, yeast or mammalian). Further, the cells can express Integrin beta 4, fragment of Integrin beta 4, Integrin beta 4-related polypeptide, a fragment of 25 the Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein endogenously or be genetically engineered to express Integrin beta 4, fragment of Integrin beta 4, Integrin beta 4-related polypeptide, a fragment of the Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein. In certain instances, Integrin beta 4, fragment of Integrin beta 4, Integrin beta 4-related polypeptide, a fragment of the

Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein or the candidate compound is labelled, for example with a radioactive label (such as ^{32}P, ^{35}S, and ^{125}I) or a fluorescent label (such as fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde or fluorescamine) to enable detection of
5 an interaction between Integrin beta 4 and a candidate compound. The ability of the candidate compound to interact directly or indirectly with Integrin beta 4, a fragment of Integrin beta 4, a Integrin beta 4-related polypeptide, a fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein can be determined by methods known to those of skill in the art. For example, the interaction between a candidate
10 compound and Integrin beta 4, a Integrin beta 4-related polypeptide, a fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein can be determined by flow cytometry, a scintillation assay, immunoprecipitation or western blot analysis.

In another embodiment, agents that interact with (*i.e.*, bind to) Integrin beta 4, a
15 Integrin beta 4 fragment (e.g., a functionally active fragment), a Integrin beta 4-related polypeptide, a fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein are identified in a cell-free assay system. In accordance with this embodiment, a native or recombinant Integrin beta 4 or fragment thereof, or a native or recombinant Integrin beta 4-related polypeptide or fragment thereof, or a Integrin beta
20 4-fusion protein or fragment thereof, is contacted with a candidate compound or a control compound and the ability of the candidate compound to interact with Integrin beta 4 or Integrin beta 4-related polypeptide, or Integrin beta 4 fusion protein is determined. If desired, this assay may be used to screen a plurality (e.g. a library) of candidate compounds. Preferably, Integrin beta 4, Integrin beta 4 fragment, Integrin
25 beta 4-related polypeptide, a fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4-fusion protein is first immobilized, by, for example, contacting Integrin beta 4, Integrin beta 4 fragment, Integrin beta 4-related polypeptide, a fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein with an immobilized antibody (or other affinity reagent) which specifically recognizes and binds

it, or by contacting a purified preparation of Integrin beta 4, Integrin beta 4 fragment, Integrin beta 4-related polypeptide, fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein with a surface designed to bind proteins. Integrin beta 4, Integrin beta 4 fragment, Integrin beta 4-related polypeptide, a fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein may be partially or completely purified (*e.g.*, partially or completely free of other polypeptides) or part of a cell lysate. Further, Integrin beta 4, Integrin beta 4 fragment, Integrin beta 4-related polypeptide, fragment of a Integrin beta 4-related polypeptide may be a fusion protein comprising the Integrin beta 4 or a biologically active portion thereof, or Integrin beta 4-related polypeptide and a domain such as glutathionine-S-transferase. Alternatively, Integrin beta 4, the Integrin beta 4 fragment, the Integrin beta 4-related polypeptide, the fragment of a Integrin beta 4-related polypeptide or the Integrin beta 4 fusion protein can be biotinylated using techniques well known to those of skill in the art (*e.g.*, biotinylation kit, Pierce Chemicals; Rockford, IL). The ability of the candidate compound to interact with Integrin beta 4, a Integrin beta 4 fragment, a Integrin beta 4-related polypeptide, a fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein can be determined by methods known to those of skill in the art.

In another embodiment, a cell-based assay system is used to identify agents that bind to or modulate the activity of a protein, such as an enzyme, or a biologically active portion thereof, which is responsible for the production or degradation of Integrin beta 4 or is responsible for the post-translational modification of Integrin beta 4. In a primary screen, a plurality (e.g., a library) of compounds are contacted with cells that naturally or recombinantly express: (i) Integrin beta 4, an isoform of Integrin beta 4, a Integrin beta 4 homolog, a Integrin beta 4-related polypeptide, a Integrin beta 4 fusion protein, or a biologically active fragment of any of the foregoing; and (ii) a protein that is responsible for processing of Integrin beta 4, the Integrin beta 4 isoform, the Integrin beta 4 homolog, the Integrin beta 4-related polypeptide, the Integrin beta 4 fusion protein, or fragment in order to identify compounds that modulate the production,

degradation, or post-translational modification of Integrin beta 4, the Integrin beta 4 isoform, the Integrin beta 4 homolog, the Integrin beta 4-related polypeptide, the Integrin beta 4 fusion protein or fragment. If desired, compounds identified in the primary screen can then be assayed in a secondary screen against cells naturally or 5 recombinantly expressing Integrin beta 4. The ability of the candidate compound to modulate the production, degradation or post-translational modification of Integrin beta 4, isoform, homolog, Integrin beta 4-related polypeptide, or Integrin beta 4 fusion protein can be determined by methods known to those of skill in the art, including without limitation, flow cytometry, a scintillation assay, immunoprecipitation and 10 western blot analysis.

In another embodiment, agents that competitively interact with (*i.e.*, bind to) Integrin beta 4, a Integrin beta 4 fragment, a Integrin beta 4-related polypeptide, a fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein are identified in a competitive binding assay. In accordance with this embodiment, cells 15 expressing Integrin beta 4, a Integrin beta 4 fragment, a Integrin beta 4-related polypeptide, a fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein are contacted with a candidate compound and a compound known to interact with Integrin beta 4, the Integrin beta 4 fragment, the Integrin beta 4-related polypeptide, a fragment of a Integrin beta 4-related polypeptide or a Integrin beta 4 20 fusion protein; the ability of the candidate compound to preferentially interact with Integrin beta 4, the Integrin beta 4 fragment, the Integrin beta 4-related polypeptide, the fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein is then determined. Alternatively, agents that preferentially interact with (*i.e.*, bind to) Integrin beta 4, a Integrin beta 4 fragment, a Integrin beta 4-related polypeptide or 25 fragment of a Integrin beta 4-related polypeptide are identified in a cell-free assay system by contacting Integrin beta 4, a Integrin beta 4 fragment, a Integrin beta 4-related polypeptide, a fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein with a candidate compound and a compound known to interact with Integrin beta 4, the Integrin beta 4-related polypeptide or the Integrin beta 4

fusion protein. As stated above, the ability of the candidate compound to interact with Integrin beta 4, a Integrin beta 4 fragment, a Integrin beta 4-related polypeptide, a fragment of a Integrin beta 4-related polypeptide, or a Integrin beta 4 fusion protein can be determined by methods known to those of skill in the art. These assays, whether
5 cell-based or cell-free, can be used to screen a plurality (e.g., a library) of candidate compounds.

In another embodiment, agents that modulate (*i.e.*, upregulate or downregulate) the expression or activity of Integrin beta 4, or a Integrin beta 4-related polypeptide are identified by contacting cells (*e.g.*, cells of prokaryotic origin or eukaryotic origin)
10 expressing Integrin beta 4, or Integrin beta 4-related polypeptide with a candidate compound or a control compound (*e.g.*, phosphate buffered saline (PBS)) and determining the expression of Integrin beta 4, Integrin beta 4-related polypeptide, or Integrin beta 4 fusion protein, mRNA encoding Integrin beta 4, or mRNA encoding the Integrin beta 4-related polypeptide. The level of expression of Integrin beta 4, Integrin
15 beta 4-related polypeptide, mRNA encoding Integrin beta 4, or mRNA encoding the Integrin beta 4-related polypeptide in the presence of the candidate compound is compared to the level of expression of Integrin beta 4, Integrin beta 4-related polypeptide, mRNA encoding Integrin beta 4, or mRNA encoding the Integrin beta 4-related polypeptide in the absence of the candidate compound (*e.g.*, in the presence
20 of a control compound). The candidate compound can then be identified as a modulator of the expression of Integrin beta 4, or the Integrin beta 4-related polypeptide based on this comparison. For example, when expression of Integrin beta 4 or mRNA is significantly greater in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of expression of
25 Integrin beta 4 or mRNA. Alternatively, when expression of Integrin beta 4 or mRNA is significantly less in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of the expression of the Integrin beta 4 or mRNA. The level of expression of Integrin beta 4 or the mRNA that encodes it can be determined by methods known to those of skill in the art. For example, mRNA

expression can be assessed by Northern blot analysis or RT-PCR, and protein levels can be assessed by western blot analysis.

In another embodiment, agents that modulate the activity of Integrin beta 4 or a Integrin beta 4-related polypeptide are identified by contacting a preparation containing Integrin beta 4 or Integrin beta 4-related polypeptide or cells (*e.g.*, prokaryotic or eukaryotic cells) expressing Integrin beta 4 or Integrin beta 4-related polypeptide with a test compound or a control compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of Integrin beta 4 or Integrin beta 4-related polypeptide. The activity of Integrin beta 4 or a Integrin beta 4-related polypeptide can be assessed by detecting induction of a cellular signal transduction pathway of Integrin beta 4 or Integrin beta 4-related polypeptide (*e.g.*, intracellular Ca^{2+}, diacylglycerol, IP3, etc.), detecting catalytic or enzymatic activity of the target on a suitable substrate, detecting the induction of a reporter gene (*e.g.*, a regulatory element that is responsive to Integrin beta 4 or a Integrin beta 4-related polypeptide and is operably linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cellular differentiation, or cell proliferation. Based on the present description, techniques known to those of skill in the art can be used for measuring these activities (see, *e.g.*, U.S. Patent No. 5,401,639, which is incorporated herein by reference). The candidate compound can then be identified as a modulator of the activity of Integrin beta 4 or a Integrin beta 4-related polypeptide by comparing the effects of the candidate compound to the control compound. Suitable control compounds include phosphate buffered saline (PBS) and normal saline (NS).

In another embodiment, agents that modulate (i.e., upregulate or downregulate) the expression, activity or both the expression and activity of Integrin beta 4 or a Integrin beta 4-related polypeptide are identified in an animal model. Examples of suitable animals include, but are not limited to, mice, rats, rabbits, monkeys, guinea pigs, dogs and cats. Preferably, the animal used represents a model of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer (*e.g.* xenografts of breast cell lines such as MCF-7 (Ozzello L, Sordat M., Eur J

Cancer. 1980;16:553–559) and MCF10AT (Miller et al., J Natl Cancer Inst. 1993;85:1725–1732) in nude or SCID mice; xenografts of colorectal cell lines such as MDA-MB-345 in oestrogen-deprived Severe Combined Immunodeficient (SCID) mice, Eccles et al. 1994 Cell Biophysics 24/25, 279; xenografts of gastric cell lines such as AZ-521 in nude mice; xenografts of hepatocellular carcinoma cell lines such as MHCC97 in nude mice, Tian et al., Br J Cancer 1999 Nov;81(5):814-21; xenografts of non small cell lung cancer cell lines such as A549 and H460 and xenografts of small cell lung cancer cell lines such as NCI-H345 or xenografts of pancreatic cancer cell lines such as MIA PaCa-2 in nude mice, Marincola et al., J Surg Res 1989 Dec;47(6):520-9.) These can be utilized to test compounds that modulate Integrin beta 4 levels, since the pathology exhibited in these models is similar to that of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer. In accordance with this embodiment, the test compound or a control compound is administered (*e.g.*, orally, rectally or parenterally such as intraperitoneally or intravenously) to a suitable animal and the effect on the expression, activity or both expression and activity of Integrin beta 4 or Integrin beta 4-related polypeptide is determined. Changes in the expression of Integrin beta 4 or a Integrin beta 4-related polypeptide can be assessed by the methods outlined above.

In yet another embodiment, Integrin beta 4 or a Integrin beta 4-related polypeptide is used as a "bait protein" in a two-hybrid assay or three hybrid assay to identify other proteins that bind to or interact with Integrin beta 4 or a Integrin beta 4-related polypeptide (*see*, *e.g.*, U.S. Patent No. 5,283,317; Zervos et al. (1993) Cell 72:223-232; Madura et al. (1993) J. Biol. Chem. 268:12046-12054; Bartel et al. (1993) Bio/Techniques 14:920-924; Iwabuchi et al. (1993) Oncogene 8:1693-1696; and PCT Publication No. WO 94/10300). As those skilled in the art will appreciate, such binding proteins are also likely to be involved in the propagation of signals by Integrin beta 4 as, for example, upstream or downstream elements of a signalling pathway involving Integrin beta 4.

This invention further provides novel agents identified by the above-described

screening assays and uses thereof for treatments as described herein. In addition, the invention also provides the use of an agent which interacts with, or modulates the activity of, Integrin beta 4 in the manufacture of a medicament for the treatment of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer
5 or pancreatic cancer.

Therapeutic Use of Integrin beta 4

The invention provides for treatment or prevention of various diseases and disorders by administration of a therapeutic compound. Such compounds include but
10 are not limited to: Integrin beta 4, Integrin beta 4 analogues, Integrin beta 4-related polypeptides and derivatives (including fragments) thereof; antibodies (or other affinity reagents) to the foregoing; nucleic acids encoding Integrin beta 4, Integrin beta 4 analogues, Integrin beta 4-related polypeptides and fragments thereof; antisense nucleic acids to a gene encoding Integrin beta 4 or a Integrin beta 4-related polypeptide; and
15 modulator (e.g., agonists and antagonists) of a gene encoding Integrin beta 4 or a Integrin beta 4-related polypeptide. An important feature of the present invention is the identification of genes encoding Integrin beta 4 involved in breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. Breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer
20 or pancreatic cancer can be treated (e.g. to ameliorate symptoms or to retard onset or progression) or prevented by administration of a therapeutic compound that reduces function or expression of Integrin beta 4 in the serum or tissue of subjects having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.
25 In one embodiment, one or more antibodies (or other affinity reagents) each specifically binding to Integrin beta 4 are administered alone or in combination with one or more additional therapeutic compounds or treatments.

A biological product such as an antibody (or other affinity reagent) is, for example, allogeneic to the subject to which it is administered. In one embodiment, a

human Integrin beta 4 or a human Integrin beta 4-related polypeptide, a nucleotide sequence encoding a human Integrin beta 4 or a human Integrin beta 4-related polypeptide, or an antibody (or other affinity reagent) to a human Integrin beta 4 or a human Integrin beta 4-related polypeptide, is administered to a human subject for therapy (e.g. to ameliorate symptoms or to retard onset or progression) or prophylaxis.

Without being limited by theory, it is conceived that the therapeutic activity of antibodies (or other affinity reagents) which specifically bind to Integrin beta 4 may be achieved through the phenomenon of Antibody –Dependent Cell-mediated Cytotoxicity (ADCC) (see e.g. Janeway Jr. C.A. et al., Immunobiology, 5th ed., 2001, Garland Publishing, ISBN 0-8153-3642-X; Pier G.B. et al., Immunology, Infection, and Immunity, 2004, p246-5; Albanell J. et al., Advances in Experimental Medicine and Biology, 2003, 532:p2153-68 and Weng, W.-K. et al., Journal of Clinical Oncology, 2003, 21:p 3940-3947).

Treatment and Prevention of Breast Cancer, Colorectal Cancer, Gastric Cancer, Hepatocellular Carcinoma, Lung Cancer and Pancreatic Cancer

Breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer is treated or prevented by administration to a subject suspected of having or known to have breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer or to be at risk of developing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer of a compound that modulates (*i.e.*, increases or decreases) the level or activity (*i.e.*, function) of Integrin beta 4 that is differentially present in the serum or tissue of subjects having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer compared with serum or tissue of subjects free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer. In one embodiment, breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer is treated or prevented by administering to a subject

suspected of having or known to have breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer or to be at risk of developing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer a compound that upregulates (*i.e.*, increases) the level
5 or activity (*i.e.*, function) of Integrin beta 4 that are decreased in the serum or tissue of subjects having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. Examples of such a compound include, but are not limited to, Integrin beta 4 antisense oligonucleotides, ribozymes, antibodies (or other affinity reagents) directed against Integrin beta 4, and compounds that inhibit
10 the enzymatic activity of Integrin beta 4. Other useful compounds *e.g.*, Integrin beta 4 antagonists and small molecule Integrin beta 4 antagonists, can be identified using *in vitro* assays.

Breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer is also treated or prevented by administration to a subject
15 suspected of having or known to have breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer or to be at risk of developing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer of a compound that downregulates the level or activity (i.e. function) of Integrin beta 4 that are increased in the serum or tissue of subjects
20 having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. Examples of such a compound include but are not limited to: Integrin beta 4, Integrin beta 4 fragments and Integrin beta 4-related polypeptides; nucleic acids encoding Integrin beta 4, a Integrin beta 4 fragment and a Integrin beta 4-related polypeptide (*e.g.*, for use in gene therapy); and, for those Integrin beta 4 or
25 Integrin beta 4-related polypeptides with enzymatic activity, compounds or molecules known to modulate that enzymatic activity. Other compounds that can be used, *e.g.*, Integrin beta 4 agonists, can be identified using in *in vitro* assays.

In another embodiment, therapy or prophylaxis is tailored to the needs of an individual subject. Thus, in specific embodiments, compounds that promote the level

or function of Integrin beta 4 are therapeutically or prophylactically administered to a subject suspected of having or known to have breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer, in whom the levels or functions of said Integrin beta 4 are absent or are decreased relative to a control or normal reference range. In further embodiments, compounds that promote the level or function of Integrin beta 4 are therapeutically or prophylactically administered to a subject suspected of having or known to have breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in whom the levels or functions of said Integrin beta 4 are increased relative to a control or to a reference range. In further embodiments, compounds that decrease the level or function of Integrin beta 4 are therapeutically or prophylactically administered to a subject suspected of having or known to have breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in whom the levels or functions of said Integrin beta 4 are increased relative to a control or to a reference range. In further embodiments, compounds that decrease the level or function of Integrin beta 4 are therapeutically or prophylactically administered to a subject suspected of having or known to have breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in whom the levels or functions of said Integrin beta 4 are decreased relative to a control or to a reference range. The change in Integrin beta 4 function or level due to the administration of such compounds can be readily detected, *e.g.*, by obtaining a sample (*e.g.*, blood or urine) and assaying *in vitro* the levels or activities of said Integrin beta 4, or the levels of mRNAs encoding said Integrin beta 4, or any combination of the foregoing. Such assays can be performed before and after the administration of the compound as described herein.

The compounds of the invention include but are not limited to any compound, *e.g.*, a small organic molecule, protein, peptide, antibody (or other affinity reagent), nucleic acid, etc. that restores the Integrin beta 4 profile towards normal. The compounds of the invention may be given in combination with any other chemotherapy

drugs.

Vaccine Therapy

Another aspect of the invention is an immunogenic composition, suitably a vaccine composition, comprising Integrin beta 4 or an epitope containing fragment thereof, or nucleic acid encoding Integrin beta 4 or a fragment thereof optionally together with an immunostimulant.

There is also provided a method of raising an immune response which comprises administering to a subject such compositions and a method for treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer which comprises administering to a subject in need thereof a therapeutically effective amount of such compositions and such compositions for use in preventing or treating breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

Thus, Integrin beta 4 may be useful as antigenic material, and may be used in the production of vaccines for treatment or prophylaxis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer. Such material can be “antigenic” and/or “immunogenic”. Generally, “antigenic” is taken to mean that the protein is capable of being used to raise antibodies (or other affinity reagents) or indeed is capable of inducing an antibody response in a subject or experimental animal. “Immunogenic” is taken to mean that the protein is capable of eliciting a protective immune response in a subject or experimental animal. Thus, in the latter case, the protein may be capable of not only generating an antibody response but, in addition, non-antibody based immune responses. “Immunogenic” also embraces whether the protein may elicit an immune-like response in an in-vitro setting e.g. a T-cell proliferation assay. The generation of an appropriate immune response may require the presence of one or more adjuvants and/or appropriate presentation of an antigen.

The skilled person will appreciate that homologues or derivatives of Integrin beta 4 will also find use as antigenic/immunogenic material. Thus, for instance proteins

which include one or more additions, deletions, substitutions or the like are encompassed by the present invention. In addition, it may be possible to replace one amino acid with another of similar "type". For instance, replacing one hydrophobic amino acid with another. One can use a program such as the CLUSTAL program to compare amino acid sequences. This program compares amino acid sequences and finds the optimal alignment by inserting spaces in either sequence as appropriate. It is possible to calculate amino acid identity or similarity (identity plus conservation of amino acid type) for an optimal alignment. A program like BLASTx will align the longest stretch of similar sequences and assign a value to the fit. It is thus possible to obtain a comparison where several regions of similarity are found, each having a different score. Both types of analysis are contemplated in the present invention.

In the case of homologues and derivatives, the degree of identity with a protein as described herein is less important than that the homologue or derivative should retain its antigenicity and/or immunogenicity. However, suitably, homologues or derivatives having at least 60% similarity (as discussed above) with the proteins or polypeptides described herein are provided, for example, homologues or derivatives having at least 70% similarity, such as at least 80% similarity. Particularly, homologues or derivatives having at least 90% or even 95% similarity are provided. Suitably, homologues or derivatives have at least 60% sequence identity with the proteins or polypeptides described herein, for example, homologues or derivatives have at least 70% identity, such as at least 80% identity. Particularly, homologues or derivatives have at least 90% or even 95% identity.

In an alternative approach, the homologues or derivatives could be fusion proteins, incorporating moieties which render purification easier, for example by effectively tagging the desired protein or polypeptide. It may be necessary to remove the "tag" or it may be the case that the fusion protein itself retains sufficient antigenicity to be useful.

It is well known that it is possible to screen an antigenic protein or polypeptide to identify epitopic regions, i.e. those regions which are responsible for the protein or

polypeptide's antigenicity or immunogenicity. Methods well known to the skilled person can be used to test fragments and/or homologues and/or derivatives for antigenicity. Thus, the fragments of the present invention should include one or more such epitopic regions or be sufficiently similar to such regions to retain their
5 antigenic/immunogenic properties. Thus, for fragments according to the present invention the degree of identity is perhaps irrelevant, since they may be 100% identical to a particular part of a protein or polypeptide, homologue or derivative as described herein. The key issue, once again, is that the fragment retains the antigenic/immunogenic properties of the protein from which it is derived.

10 What is important for homologues, derivatives and fragments is that they possess at least a degree of the antigenicity/immunogenicity of the protein or polypeptide from which they are derived. Thus, in an additional aspect of the invention, there is provided antigenic/or immunogenic fragments of Integrin beta 4, or of homologues or derivatives thereof.

15 Integrin beta 4, or antigenic fragments thereof, can be provided alone, as a purified or isolated preparation. They may be provided as part of a mixture with one or more other proteins of the invention, or antigenic fragments thereof. In a further aspect, therefore, the invention provides an antigen composition comprising Integrin beta 4 and/or one or more antigenic fragments thereof. Such a composition can be
20 used for the detection and/or diagnosis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

Vaccine compositions according to the invention may be either a prophylactic or therapeutic vaccine composition.

The vaccine compositions of the invention can include one or more adjuvants
25 (immunostimulants). Examples well known in the art include inorganic gels, such as aluminium hydroxide, and water-in-oil emulsions, such as incomplete Freund's adjuvant. Other useful adjuvants will be well known to the skilled person.

Suitable adjuvants for use in vaccine compositions for the treatment of cancer include: 3De-O-acylated monophosphoryl lipid A (known as 3D-MPL or simply MPL

see WO92/116556), a saponin, for example QS21 or QS7, and TLR4 agonists such as a CpG containing molecule, for example as disclosed in WO95/26204.

The adjuvants employed may be a combination of components, for example MPL and QS21 or MPL, QS21 and a CpG containing moiety.

5 Adjuvants may be formulated as oil-in-water emulsions or liposomal formulations.

Such preparations may include other vehicles.

In another embodiment, a preparation of oligonucleotides comprising 10 or more consecutive nucleotides complementary to a nucleotide sequence encoding 10 Integrin beta 4 or a Integrin beta 4 peptide fragment is used as a vaccine for the treatment of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. Such preparations may include adjuvants or other vehicles.

15 Inhibition of Integrin beta 4 to Treat Breast Cancer, Colorectal Cancer, Gastric Cancer, Hepatocellular Carcinoma, Lung Cancer and Pancreatic Cancer

In one embodiment of the invention, breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer is treated or prevented by administration of a compound that antagonizes (inhibits) the level and/or 20 function of Integrin beta 4 which are elevated in the serum or tissue of subjects having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer as compared with serum or tissue of subjects free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer.

25 Compounds useful for this purpose include but are not limited to anti-Integrin beta 4 antibodies (or other affinity reagents, and fragments and derivatives containing the binding region thereof), Integrin beta 4 antisense or ribozyme nucleic acids, and nucleic acids encoding dysfunctional Integrin beta 4 that are used to "knockout" endogenous Integrin beta 4 function by homologous recombination (see, *e.g.*,

Capecchi, 1989, *Science* 244:1288-1292). Other compounds that inhibit Integrin beta 4 function can be identified by use of known *in vitro* assays, *e.g.*, assays for the ability of a test compound to inhibit binding of Integrin beta 4 to another protein or a binding partner, or to inhibit a known Integrin beta 4 function.

5 Such inhibition is, for example, assayed *in vitro* or in cell culture, but genetic assays may also be employed. The Preferred Technology described herein can also be used to detect levels of Integrin beta 4 before and after the administration of the compound. Suitable *in vitro* or *in vivo* assays are utilized to determine the effect of a specific compound and whether its administration is indicated for treatment of the 10 affected tissue, as described in more detail below.

In a specific embodiment, a compound that inhibits Integrin beta 4 function (activity) is administered therapeutically or prophylactically to a subject in whom an increased serum or tissue level or functional activity of Integrin beta 4 (*e.g.*, greater than the normal level or desired level) is detected as compared with serum or tissue of 15 subjects with breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer who do not receive treatment according to the invention or to bring the level or activity to that found in subjects free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer or a predetermined reference range. Methods standard in the art can 20 be employed to measure the increase in Integrin beta 4 level or function, as outlined above. Suitable Integrin beta 4 inhibitor compositions may, for example, include small molecules, i.e., molecules of 1000 Daltons or less. Such small molecules can be identified by the screening methods described herein.

25 Assays for Therapeutic or Prophylactic Compounds

The present invention also provides assays for use in drug discovery in order to identify or verify the efficacy of compounds for treatment or prevention of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

Thus there is provided a method of screening for compounds that modulate the activity of Integrin beta 4, the method comprising: (a) contacting Integrin beta 4 or a biologically active portion thereof with a candidate compound; and (b) determining whether activity of Integrin beta 4 is thereby modulated. Such a process may comprise (a) contacting Integrin beta 4 or a biologically active portion thereof with a candidate compound in a sample; and (b) comparing the activity of Integrin beta 4 or a biologically active portion thereof in said sample after contact with said candidate compound with the activity of Integrin beta 4 or a biologically active portion thereof in said sample before contact with said candidate compound, or with a reference level of activity.

The method of screening may be a method of screening for compounds that inhibit activity of Integrin beta 4.

Integrin beta 4 or a biologically active portion thereof may, for example be expressed on or by a cell. Integrin beta 4 or a biologically active portion thereof may, for example, be isolated from cells which express it. Integrin beta 4 or a biologically active portion thereof may, for example, be immobilised onto a solid phase.

There is also provided a method of screening for compounds that modulate the expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4, the method comprising: (a) contacting cells expressing Integrin beta 4 or nucleic acid encoding Integrin beta 4 with a candidate compound; and (b) determining whether expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4 is thereby modulated. Such a process may comprise (a) contacting cells expressing Integrin beta 4 or nucleic acid encoding Integrin beta 4 with a candidate compound in a sample; and (b) comparing the expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4 by cells in said sample after contact with said candidate compound with the expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4 of cells in said sample before contact with said candidate compound, or with a reference level of expression.

The method may be a method of screening for compounds that inhibit expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4.

Other aspects of the invention include: a compound obtainable by an aforementioned screening method, a compound which modulates the activity or expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4, for example a compound which inhibits the activity or expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4.

Such a compound is provided for use in treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. There is also provided a method for treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer which comprises administering to a subject in need thereof a therapeutically effective amount of such a compound.

Test compounds can be assayed for their ability to restore Integrin beta 4 levels in a subject having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer towards levels found in subjects free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer or to produce similar changes in experimental animal models of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. Compounds able to restore Integrin beta 4 levels in a subject having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer towards levels found in subjects free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer or to produce similar changes in experimental animal models of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer can be used as lead compounds for further drug discovery, or used therapeutically. Integrin beta 4 expression can be assayed by the Preferred Technology described herein, immunoassays, gel electrophoresis followed by visualization, detection of Integrin beta 4 activity, or any other method taught herein or known to those skilled in the art. Such assays can be used to screen candidate drugs, in clinical monitoring or in drug development, where abundance of Integrin beta 4 can serve as a

surrogate marker for clinical disease.

In various specific embodiments, *in vitro* assays can be carried out with cells representative of cell types involved in a subject's disorder, to determine if a compound has a desired effect upon such cell types.

5 Compounds for use in therapy can be tested in suitable animal model systems prior to testing in humans, including but not limited to rats, mice, chicken, cows, monkeys, rabbits, etc. For *in vivo* testing, prior to administration to humans, any animal model system known in the art may be used. Examples of animal models of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer 10 and pancreatic cancer include, but are not limited to xenografts of breast cell lines such as MCF-7 (Ozzello L, Sordat M., Eur J Cancer. 1980;16:553–559) and MCF10AT (Miller et al., J Natl Cancer Inst. 1993;85:1725–1732) in nude or SCID mice; xenografts of colorectal cancer cell lines such as MDA-MB-435 in oestrogen-deprived Severe Combined Immunodeficient (SCID) mice (Eccles et al., 1994 Cell Biophysics 15 24/25, 279); xenografts of gastric cell lines such as AZ-521 in nude mice; xenografts of hepatocellular carcinoma cell lines such as MHCC97 in nude mice (Tian et al., Br J Cancer 1999 Nov;81(5):814-21); xenografts of non small cell lung cancer cell lines such as A549 and H460 and xenografts of small cell lung cancer cell lines such as NCI-H345 and xenografts of pancreatic cancer cell lines such as MIA PaCa-2 in nude mice 20 (Marincola et al., J Surg Res 1989 Dec;47(6):520-9). These can be utilized to test compounds that modulate Integrin beta 4 levels, since the pathology exhibited in these models is similar to that of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer. It is also apparent to the skilled artisan that based upon the present disclosure, transgenic animals can be 25 produced with "knock-out" mutations of the gene or genes encoding Integrin beta 4. A "knock-out" mutation of a gene is a mutation that causes the mutated gene to not be expressed, or expressed in an aberrant form or at a low level, such that the activity associated with the gene product is nearly or entirely absent. The transgenic animal is, for example, a mammal; such as a mouse.

In one embodiment, test compounds that modulate the expression of Integrin beta 4 are identified in non-human animals (*e.g.*, mice, rats, monkeys, rabbits, and guinea pigs), preferably non-human animal models for breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer, expressing Integrin beta 4. In accordance with this embodiment, a test compound or a control compound is administered to the animals, and the effect of the test compound on expression of Integrin beta 4 is determined. A test compound that alters the expression of Integrin beta 4 can be identified by comparing the level of Integrin beta 4 (or mRNA encoding the same) in an animal or group of animals treated with a test compound with the level of Integrin beta 4 or mRNA in an animal or group of animals treated with a control compound. Techniques known to those of skill in the art can be used to determine the mRNA and protein levels, for example, *in situ* hybridization. The animals may or may not be sacrificed to assay the effects of a test compound.

In another embodiment, test compounds that modulate the activity of Integrin beta 4 or a biologically active portion thereof are identified in non-human animals (*e.g.*, mice, rats, monkeys, rabbits, and guinea pigs), preferably non-human animal models for breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer, expressing Integrin beta 4. In accordance with this embodiment, a test compound or a control compound is administered to the animals, and the effect of a test compound on the activity of Integrin beta 4 is determined. A test compound that alters the activity of Integrin beta 4 can be identified by assaying animals treated with a control compound and animals treated with the test compound. The activity of Integrin beta 4 can be assessed by detecting induction of a cellular second messenger of Integrin beta 4 (*e.g.*, intracellular Ca^{2+}, diacylglycerol, IP3, etc.), detecting catalytic or enzymatic activity of Integrin beta 4 or binding partner thereof, detecting the induction of a reporter gene (*e.g.*, a regulatory element that is responsive to Integrin beta 4 operably linked to a nucleic acid encoding a detectable marker, such as luciferase or green fluorescent protein), or detecting a cellular response (*e.g.*, cellular differentiation or cell proliferation). Techniques known to those of skill in the art can be utilized to

detect changes in the activity of Integrin beta 4 (see, *e.g.*, U.S. Patent No. 5,401,639, which is incorporated herein by reference).

In yet another embodiment, test compounds that modulate the level or expression of Integrin beta 4 are identified in human subjects having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer, particularly those having severe breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. In accordance with this embodiment, a test compound or a control compound is administered to the human subject, and the effect of a test compound on Integrin beta 4 expression is determined by analyzing the expression of Integrin beta 4 or the mRNA encoding the same in a biological sample (*e.g.*, serum, plasma, or urine). A test compound that alters the expression of Integrin beta 4 can be identified by comparing the level of Integrin beta 4 or mRNA encoding the same in a subject or group of subjects treated with a control compound to that in a subject or group of subjects treated with a test compound. Alternatively, alterations in the expression of Integrin beta 4 can be identified by comparing the level of Integrin beta 4 or mRNA encoding the same in a subject or group of subjects before and after the administration of a test compound. Techniques known to those of skill in the art can be used to obtain the biological sample and analyze the mRNA or protein expression. For example, the Preferred Technology described herein can be used to assess changes in the level of Integrin beta 4.

In another embodiment, test compounds that modulate the activity of Integrin beta 4 are identified in human subjects having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer, (particularly those with severe breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer). In this embodiment, a test compound or a control compound is administered to the human subject, and the effect of a test compound on the activity of Integrin beta 4 is determined. A test compound that alters the activity of Integrin beta 4 can be identified by comparing biological samples from subjects treated with a control compound to samples from subjects treated with the test compound.

Alternatively, alterations in the activity of Integrin beta 4 can be identified by comparing the activity of Integrin beta 4 in a subject or group of subjects before and after the administration of a test compound. The activity of Integrin beta 4 can be assessed by detecting in a biological sample (*e.g.*, serum, plasma, or urine) induction of a cellular signal transduction pathway of Integrin beta 4 (*e.g.*, intracellular Ca^{2+}, diacylglycerol, IP3, etc.), catalytic or enzymatic activity of Integrin beta 4 or a binding partner thereof, or a cellular response, for example, cellular differentiation, or cell proliferation. Techniques known to those of skill in the art can be used to detect changes in the induction of a second messenger of Integrin beta 4 or changes in a cellular response. For example, RT-PCR can be used to detect changes in the induction of a cellular second messenger.

In another embodiment, a test compound that changes the level or expression of Integrin beta 4 towards levels detected in control subjects (*e.g.*, humans free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer) is selected for further testing or therapeutic use. In another embodiment, a test compound that changes the activity of Integrin beta 4 towards the activity found in control subjects (*e.g.*, humans free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer) is selected for further testing or therapeutic use.

In another embodiment, test compounds that reduce the severity of one or more symptoms associated with breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer are identified in human subjects having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer, particularly subjects with severe breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. In accordance with this embodiment, a test compound or a control compound is administered to the subjects, and the effect of a test compound on one or more symptoms of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer is determined. A test compound that

reduces one or more symptoms can be identified by comparing the subjects treated with a control compound to the subjects treated with the test compound. Techniques known to physicians familiar with breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer can be used to determine whether a test compound reduces one or more symptoms associated with breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. For example, a test compound that reduces tumour burden in a subject having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer will be beneficial for subjects having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

In a particular embodiment, a test compound that reduces the severity of one or more symptoms associated with breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in a human having breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer is selected for further testing or therapeutic use.

Therapeutic and Prophylactic Compositions and their Use

The invention provides methods of treatment (and prophylaxis) comprising administering to a subject an effective amount of a compound of the invention. In a particular aspect, the compound is substantially purified (*e.g.*, substantially free from substances that limit its effect or produce undesired side effects). The subject is, for example, an animal, including but not limited to animals such as cows, pigs, horses, chickens, cats, dogs, etc., and is, for example, a mammal, such as a human. In a specific embodiment, a non-human mammal is the subject.

Formulations and methods of administration that can be employed when the compound comprises a nucleic acid are described above; additional appropriate formulations and routes of administration are described below.

Various delivery systems are known and can be used to administer a compound of the invention, *e.g.*, encapsulation in liposomes, microparticles, microcapsules,

recombinant cells capable of expressing the compound, receptor mediated endocytosis (see, *e.g.*, Wu and Wu, 1987, J. Biol. Chem. 262:4429-4432), construction of a nucleic acid as part of a retroviral or other vector, etc. Methods of introduction can be enteral or parenteral and include but are not limited to intradermal, intramuscular,
5 intraperitoneal, intravenous, subcutaneous, intranasal, epidural, and oral routes. The compounds may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (*e.g.*, oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with other biologically active agents. Administration can be systemic or local. In addition, it may
10 be desirable to introduce the pharmaceutical compositions of the invention into the central nervous system by any suitable route, including intraventricular and intrathecal injection; intraventricular injection may be facilitated by an intraventricular catheter, for example, attached to a reservoir, such as an Ommaya reservoir. Pulmonary administration can also be employed, *e.g.*, by use of an inhaler or nebulizer, and
15 formulation with an aerosolizing agent.

In one aspect of the invention a nucleic acid employed in the invention may be delivered to the dermis, for example employing particle mediated epidermal delivery.

In a specific embodiment, it may be desirable to administer the pharmaceutical compositions of the invention locally to the area in need of treatment; this may be
20 achieved, for example, and not by way of limitation, by local infusion during surgery, topical application, *e.g.*, by injection, by means of a catheter, or by means of an implant, said implant being of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers. In one embodiment, administration can be by direct injection into breast, colorectal, gastric epithelium, liver, lung or
25 pancreatic tissue or at the site (or former site) of a malignant tumour or neoplastic or pre-neoplastic tissue.

In another embodiment, the compound can be delivered in a vesicle, in particular a liposome (*see* Langer, 1990, Science 249:1527-1533; Treat et al., in Liposomes in the Therapy of Infectious Disease and Cancer, Lopez-Berestein and

Fidler (eds.), Liss, New York, pp. 353-365 (1989); Lopez-Berestein, *ibid.*, pp. 317-327; *see* generally *ibid.*)

In yet another embodiment, the compound can be delivered in a controlled release system. In one embodiment, a pump may be used (see Langer, *supra*; Sefton, 1987, CRC Crit. Ref. Biomed. Eng. 14:201; Buchwald et al., 1980, Surgery 88:507; Saudek et al., 1989, N. Engl. J. Med. 321:574). In another embodiment, polymeric materials can be used (see Medical Applications of Controlled Release, Langer and Wise (eds.), CRC Pres., Boca Raton, Florida (1974); Controlled Drug Bioavailability, Drug Product Design and Performance, Smolen and Ball (eds.), Wiley, New York (1984); Ranger and Peppas, J., 1983, Macromol. Sci. Rev. Macromol. Chem. 23:61; see also Levy et al., 1985, Science 228:190; During et al., 1989, Ann. Neurol. 25:351; Howard et al., 1989, J. Neurosurg. 71:105). In yet another embodiment, a controlled release system can be placed in proximity of the therapeutic target, *i.e.*, the breast, colon, stomach, liver, lung or pancreas thus requiring only a fraction of the systemic dose (see, *e.g.*, Goodson, in Medical Applications of Controlled Release, *supra*, vol. 2, pp. 115-138 (1984)).

Other controlled release systems are discussed in the review by Langer (1990, Science 249:1527-1533).

In a specific embodiment where the compound of the invention is a nucleic acid encoding a protein, the nucleic acid can be administered *in vivo* to promote expression of its encoded protein, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, *e.g.*, by use of a retroviral vector (see U.S. Patent No. 4,980,286), or by direct injection, or by use of microparticle bombardment (*e.g.*, a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, or by administering it in linkage to a homeobox-like peptide which is known to enter the nucleus (see *e.g.*, Joliot et al., 1991, Proc. Natl. Acad. Sci. USA 88:1864-1868), etc. Alternatively, a nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination.

The present invention also provides pharmaceutical compositions. Such compositions comprise a therapeutically effective amount of a compound, and a pharmaceutically acceptable carrier. In a specific embodiment, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans. The term "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic is administered. Such pharmaceutical carriers can be sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred carrier when the pharmaceutical composition is administered intravenously. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as liquid carriers, particularly for injectable solutions. Suitable pharmaceutical excipients include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. The composition, if desired, can also contain minor amounts of wetting or emulsifying agents, or pH buffering agents. These compositions can take the form of solutions, suspensions, emulsion, tablets, pills, capsules, powders, sustained-release formulations and the like. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides. Oral formulation can include standard carriers such as pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, etc. Examples of suitable pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E.W. Martin. Such compositions will contain a therapeutically effective amount of the compound, preferably in purified form, together with a suitable amount of carrier so as to provide the form for proper administration to the subject. The formulation should suit the mode of administration.

In one embodiment, for example where one or more antibodies are employed, the composition is formulated in accordance with routine procedures as a

pharmaceutical composition adapted for intravenous administration to human beings. Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Where necessary, the composition may also include a solubilizing agent and a local anaesthetic such as lidocaine to ease pain at the site of the injection. 5 Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachet indicating the quantity of active agent. Where the composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or 10 saline. Where the composition is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients may be mixed prior to administration.

The compounds of the invention can be formulated as neutral or salt forms. Pharmaceutically acceptable salts include those formed with free amino groups such as 15 those derived from hydrochloric, phosphoric, acetic, oxalic, tartaric acids, etc., and those formed with free carboxyl groups such as those derived from sodium, potassium, ammonium, calcium, ferric hydroxides, isopropylamine, triethylamine, 2-ethylamino ethanol, histidine, procaine, etc.

The amount of the compound of the invention which will be effective in the 20 treatment of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer can be determined by standard clinical techniques. In addition, *in vitro* assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness of the disease or disorder, and should be 25 decided according to the judgment of the practitioner and each subject's circumstances. However, suitable dosage ranges for intravenous administration are generally about 20-500 micrograms of active compound per kilogram body weight. Suitable dosage ranges for intranasal administration are generally about 0.01 pg/kg body weight to 1 mg/kg body weight. Effective doses may be extrapolated from dose-response curves

derived from *in vitro* or animal model test systems.

Suppositories generally contain active ingredient in the range of 0.5% to 10% by weight; oral formulations preferably contain 10% to 95% active ingredient.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Optionally associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects (a) approval by the agency of manufacture, use or sale for human administration, (b) directions for use, or both.

Thus, in one aspect the kit comprises antibodies employed in the invention, for example the antibodies may be lyophilized for reconstitution before administration or use. Where the kit is for use in therapy/treatment such as cancer the antibody or antibodies may be reconstituted with an isotonic aqueous solution, which may optionally be provided with the kit. In one aspect the kit may comprise a polypeptide such as an immunogenic polypeptide employed in the invention, which may for example be lyophilized. The latter kit may further comprise an adjuvant for reconstituting the immunogenic polypeptide.

The invention also extends to a composition as described herein for example a pharmaceutical composition and/or vaccine composition for use in inducing an immune response in a subject.

Determining Abundance of Integrin beta 4 by Imaging Technology

An advantage of determining abundance of Integrin beta 4 by imaging technology may be that such a method is non-invasive (save that reagents may need to be administered) and there is no need to extract a sample from the subject.

Suitable imaging technologies include positron emission tomography (PET) and single photon emission computed tomography (SPECT). Visualisation of Integrin beta 4 using such techniques requires incorporation or binding of a suitable label e.g. a

radiotracer such as ^{18}F, ^{11}C or ^{123}I (see e.g. NeuroRx – The Journal of the American Society for Experimental NeuroTherapeutics (2005) 2(2), 348-360 and *idem* pages 361-371 for further details of the techniques). Radiotracers or other labels may be incorporated into Integrin beta 4 by administration to the subject (e.g. by injection) of a suitably labelled specific ligand. Alternatively they may be incorporated into a binding affinity reagent (e.g. an antibody) specific for Integrin beta 4 which may be administered to the subject (e.g. by injection). For discussion of use of Affibodies for imaging see e.g. Orlova A, Magnusson M, Eriksson TL, Nilsson M, Larsson B, Hoiden-Guthenberg I, Widstrom C, Carlsson J, Tolmachev V, Stahl S, Nilsson FY, Tumour imaging using a picomolar affinity HER2 binding affibody molecule, Cancer Res. 2006 Apr 15;66(8):4339-48).

Diagnosis and Treatment of Breast Cancer, Colorectal Cancer, Gastric Cancer, Hepatocellular Carcinoma, Lung Cancer or Pancreatic Cancer using Immunohistochemistry

Immunohistochemistry is an excellent detection technique and may therefore be very useful in the diagnosis and treatment of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer. Immunohistochemistry may be used to detect, diagnose, or monitor breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer through the localization of Integrin beta 4 antigens in tissue sections by the use of labelled antibodies (or other affinity reagents), derivatives and analogues thereof, which specifically bind to Integrin beta 4, as specific reagents through antigen-antibody interactions that are visualized by a marker such as fluorescent dye, enzyme, radioactive element or colloidal gold.

The advancement of monoclonal antibody technology has been of great significance in assuring the place of immunohistochemistry in the modern accurate microscopic diagnosis of human neoplasms. The identification of disseminated neoplastically transformed cells by immunohistochemistry allows for a clearer picture of

cancer invasion and metastasis, as well as the evolution of the tumour cell associated immunophenotype towards increased malignancy. Future antineoplastic therapeutic approaches may include a variety of individualized immunotherapies, specific for the particular immunophenotypical pattern associated with each individual patient's neoplastic disease. For further discussion see e.g. Bodey B, The significance of immunohistochemistry in the diagnosis and therapy of neoplasms, Expert Opin Biol Ther. 2002 Apr;2(4):371-93.

The present invention may also be understood by reference to the following numbered paragraphs:

1. A method of detecting, diagnosing and/or screening for breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer which comprises:
   (a) bringing into contact with a sample to be tested Integrin beta 4, or one or more antigenic or immunogenic fragments thereof; and
   (b) detecting the presence of antibodies (or other affinity reagents) in the subject capable of specific binding to Integrin beta 4, or one or more antigenic or immunogenic fragments thereof.
2. The use of Integrin beta 4 and/or one or more antigenic or immunogenic fragments thereof, in screening for, detecting and/or diagnosing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.
3. A kit for use in the screening for, detection and/or diagnosis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer, which kit comprises Integrin beta 4 and/or one or more antigenic or immunogenic fragments thereof.
4. An antibody or other affinity reagent capable of immunospecific binding to Integrin beta 4.
5. A kit comprising an antibody or other affinity reagent as defined in paragraph 4.

6. A kit comprising a plurality of distinct antibodies or other affinity reagents as defined in paragraph 4.
7. A pharmaceutical composition comprising a therapeutically effective amount of an antibody or other affinity reagent as defined in paragraph 4, or a fragment or derivative thereof which comprises the binding domain of the affinity reagent, and optionally a pharmaceutically acceptable carrier.
8. A method of treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer comprising administering to a subject an antibody or other affinity reagent as defined in paragraph 4, or a fragment or derivative thereof which comprises the binding domain of the affinity reagent.
9. The use of an antibody or other affinity reagent as defined in paragraph 4, a fragment or derivative thereof which comprises the binding domain of the affinity reagent in the manufacture of a medicament for the treatment of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.
10. A method of treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer comprising administering to a subject in need of such treatment or prevention a therapeutically effective amount of nucleic acid encoding Integrin beta 4 or one or more fragments or derivatives thereof.
11. A method of treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer comprising administering to a subject in need of such treatment or prevention a therapeutically effective amount of nucleic acid that inhibits the function or expression of Integrin beta 4.
12. The method of paragraph 11, wherein the nucleic acid is a Integrin beta 4 anti-sense nucleic acid or ribozyme.
13. The use of nucleic acid encoding Integrin beta 4 or one or more fragments or

derivatives thereof, in the manufacture of a medicament for treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

14. The use of nucleic acid that inhibits the function or expression of Integrin beta 4, in the manufacture of a medicament for treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

15. The use of paragraph 14, wherein the nucleic acid is a Integrin beta 4 anti-sense nucleic acid or ribozyme.

16. A method for screening for and/or diagnosis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in a human subject, which comprises the step of identifying the presence or absence of Integrin beta 4, or a fragment thereof, in a biological sample obtained from said human subject.

17. A method for monitoring and/or assessing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer treatment in a human subject, which comprises the step of identifying the presence or absence of Integrin beta 4, or a fragment thereof, in a biological sample obtained from said human subject.

18. A method for identifying the presence or absence of metastatic breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer cells in a biological sample obtained from a human subject, which comprises the step of identifying the presence or absence of Integrin beta 4, or a fragment thereof.

19. A method for monitoring and/or assessing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer treatment in a human subject, which comprises the step of determining whether Integrin beta 4, or a fragment thereof, is increased/decreased in a biological sample obtained from a patient.

20. A method as defined in any one of paragraphs 16 to 19, wherein the method comprises an immunoassay step utilising one or more antibodies or other affinity reagents against Integrin beta 4, or a derivative, homologue or fragment thereof.

5 21. A method as defined in any one of paragraphs 16 to 19, wherein the method comprises the use of nucleic acid probes and/or PCR reactions to amplify nucleic acid coding for Integrin beta 4.

22. A method as defined in any one of paragraphs 16 to 19, wherein a whole body scan of the subject is carried out to determine localisation of breast cancer,
10 colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer cells, particularly metastatic breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer cells.

23. A method as defined in paragraph 22, wherein labelled antibodies or other affinity reagents are employed.

15 24. A diagnostic kit comprising one or more reagents for use in the detection and/or determination of Integrin beta 4 or a fragment thereof.

25. A kit as defined in paragraph 24, which comprises one or more containers with one or more antibodies or other affinity reagents against Integrin beta 4 or a fragment thereof.

20 26. A kit as defined in paragraph 25, which further comprises a labelled binding partner to the or each affinity reagent and/or a solid phase (such as a reagent strip) upon which the or each affinity reagent is/are immobilised.

27. A kit as defined in paragraph 25 which comprises a nucleic acid probe capable of hybridizing to DNA or RNA encoding Integrin beta 4 or a fragment thereof.

25 28. A method for screening, diagnosis or prognosis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in a subject or for monitoring the effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy administered to a subject,

comprising:

(a) analysing a sample from the subject by a protein separation technique, for example one dimensional electrophoresis, to generate a one-dimensional array of features; and

(b) for at least one chosen feature whose relative abundance correlates with the presence or absence of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer, comparing the abundance of each such chosen feature in the sample with the abundance of that chosen feature in a sample from one or more persons free from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer, or with a previously determined reference range,

wherein the relative abundance of the chosen feature or features in the sample indicates the presence or absence of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in the subject.

29. The method of paragraph 28, wherein step (b) comprises quantitatively detecting Integrin beta 4.

30. A method for screening, diagnosis or prognosis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in a subject or for monitoring the effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy administered to a subject, comprising: in a sample from the subject, quantitatively detecting Integrin beta 4.

31. The method of any one of paragraphs 28 to 30, wherein the sample is a sample of breast, colorectal, gastric epithelium, liver, lung or pancreatic tissue.

32. The method according to paragraph 30 or paragraph 31, wherein the step of quantitatively detecting comprises testing the sample, said step of testing

comprising:

(a) contacting the sample with an antibody or other affinity reagent that is immunospecific for Integrin beta 4; and

(b) detecting whether binding has occurred between the affinity reagent and at least one species in the sample.

33. The method according to paragraph 32, wherein the step of quantitatively detecting comprises testing the sample, said step of testing comprising:

(a) contacting the sample with a capture reagent to capture Integrin beta 4; and

(b) detecting the captured Integrin beta 4 using a directly or indirectly labelled detection reagent.

34. The method according to paragraph 33, wherein the capture reagent is an antibody or other affinity reagent.

35. The method according to paragraph 33 or paragraph 34, wherein Integrin beta 4 is in the form of a particular isoform and the capture reagent recognises the component part of that isoform which distinguishes the isoform from other members of the gene family, e.g. lectin for carbohydrate, or phosphotyrosine or phosphoserine/threonine Ab, or methylation or acetylation Ab.

36. The method according to any one of paragraphs 32 to 35, wherein the affinity reagent is a monoclonal antibody.

37. A method of screening for compounds that interact with Integrin beta 4 or biologically active portion thereof, the method comprising:

(a) contacting Integrin beta 4 or biologically active portion thereof with a candidate compound; and

(b) determining the ability of the candidate compound to interact with Integrin beta 4 or biologically active portion thereof.

38. A method of screening for or identifying compounds that modulate the activity of Integrin beta 4 or biologically active portion thereof, the method comprising:

(a) in a first aliquot, contacting a candidate compound with Integrin beta 4

or biologically active portion thereof; and

(b) comparing the activity of Integrin beta 4 or biologically active portion thereof in the first aliquot after addition of the candidate compound with the activity of Integrin beta 4 or biologically active portion thereof in a control aliquot, or with a previously determined reference range.

39. The method of paragraph 37 or 38, wherein Integrin beta 4 or biologically active portion thereof is expressed by a cell.

40. The method of paragraph 37, 38 or 39, wherein Integrin beta 4 or biologically active portion thereof is recombinant.

41. The method of paragraph 40, wherein the polypeptide or biologically active portion thereof is immobilised on a solid phase.

42. A method of screening for compounds that modulate the expression or activity of Integrin beta 4, comprising:

(a) contacting an enzyme which is responsible for the production or degradation of Integrin beta 4 with a candidate compound;

(b) detecting modulation of the activity of said enzyme.

43. A method of screening for compounds that modulate the expression or activity of Integrin beta 4, comprising:

(a) contacting a first group of cells expressing Integrin beta 4 with a candidate compound;

(b) contacting a second group of cells expressing Integrin beta 4 with a control compound; and

(c) comparing the level of Integrin beta 4 or mRNA encoding Integrin beta 4 in the first and second groups of cells, or comparing the level of induction of a cellular second messenger in the first and second groups of cells.

44. A method of screening for or identifying compounds that modulate the expression or activity of Integrin beta 4, the method comprising:

(a) administering a candidate compound to a first group of mammals;

(b) administering a control compound to a second group of mammals; and

(c) comparing the level of expression of Integrin beta 4 or of mRNA encoding Integrin beta 4 in the first and second groups, or comparing the level of induction of a cellular second messenger in the first and second groups.

45. The method of paragraph 44, wherein the mammals are animal models for breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

46. A method for screening, diagnosis or prognosis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in a subject or for monitoring the effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy administered to a subject, comprising:

(a) contacting one or more oligonucleotide probes comprising 10 or more consecutive nucleotides complementary to a nucleotide sequence encoding Integrin beta 4, with an RNA obtained from a biological sample from the subject or with cDNA copied from the RNA, wherein said contacting occurs under conditions that permit hybridization of the probe to the nucleotide sequence if present;

(b) detecting hybridization, if any, between the probe and the nucleotide sequence; and

(c) comparing the hybridization, if any, detected in step (b) with the hybridization detected in a control sample, or with a previously determined reference range.

47. The use of an agent which interacts with, or modulates the activity of Integrin beta 4 in the manufacture of a medicament for the treatment of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

48. A method for the treatment or prophylaxis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in a subject, or of vaccinating a subject against breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer, which comprises the step of administering to the subject an effective amount of Integrin beta 4 and/or one or more antigenic or immunogenic fragments thereof, preferably as a vaccine.
49. The use of Integrin beta 4, one or more fragments or derivatives thereof, or one or more fragments or derivatives thereof, in the manufacture of a medicament for the treatment or prophylaxis of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.
50. A vaccine comprising Integrin beta 4 or derivatives thereof, and/or one or more antigenic or immunogenic fragments thereof.
51. A composition capable of eliciting an immune response in a subject, which composition comprises Integrin beta 4 and/or one or more antigenic or immunogenic fragments thereof, and one or more suitable adjuvants.
52. The use of a composition as defined in paragraph 51 in inducing an immune response in a subject.
53. A method according to any one of the preceding paragraphs wherein the method of determining the abundance of Integrin beta 4, for example a method of quantitatively detecting Integrin beta 4, involves use of an imaging technology.
54. A method according to paragraph 53 wherein the imaging technology involves use of labelled Affibodies.
55. A method according to paragraph 53 wherein the imaging technology involves use of labelled antibodies.

Preferred features of each aspect of the invention are as for each of the other aspects *mutatis mutandis*. The prior art documents mentioned herein are incorporated

to the fullest extent permitted by law.

EXAMPLE 1: IDENTIFICATION OF MEMBRANE PROTEINS EXPRESSED IN BREAST CANCER, COLORECTAL CANCER, GASTRIC CANCER, HEPATOCELLULAR CARCINOMA, LUNG CANCER AND PANCREATIC CANCER BLOOD AND TISSUE SAMPLES

Using the following Reference Protocol, membrane proteins extracted from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer tissue samples were separated by 1D gel and analysed.

1.1 MATERIALS AND METHODS

1.1.1 – Plasma Membrane Fractionation

The cells recovered from the epithelium of a breast adenocarcinoma, colorectal adenocarcinoma, gastric cancer, hepatocellular carcinoma, lung carcinoma and pancreatic adenocarcinoma were lysed and submitted to centrifugation at 1000G. The supernatant was taken, and it was subsequently centrifuged at 3000G. Once again, the supernatant was taken, and it was then centrifuged at 100 000G.

The resulting pellet was recovered and put on 15-60% sucrose gradient.

A Western blot was used to identify sub cellular markers, and the Plasma Membrane fractions were pooled.

The pooled solution was either run directly on 1D gels (see section 1.1.4 below), or further fractionated into heparin binding and nucleotide binding fractions as described below.

1.1.2 – Plasma Membrane Heparin-binding Fraction

The pooled solution from 1.1.1 above was applied to a Heparin column, eluted from column and run on 1D gels (see section 1.1.4 below).

1.1.3 – Plasma Nucleotide-binding Fraction

The pooled solution from 1.1.1 above was applied to a Cibacrom Blue 3GA column, eluted from column and run on 1D gels (see section 1.1.4 below).

1.1.4 – 1D gel technology

5 Protein or membrane pellets were solubilised in 1D sample buffer (1-2 μg/μl). The sample buffer and protein mixture was then heated to 95°C for 3 min.

A 9-16% acrylamide gradient gel was cast with a stacking gel and a stacking comb according to the procedure described in Ausubel F.M. et al., eds., 1989, Current Protocols in Molecular Biology, Vol. II, Green Publishing Associates, Inc., and John
10 Wiley & Sons, Inc., New York, section 10.2, incorporated herein by reference in its entirety.

30-50 micrograms of the protein mixtures obtained from detergent and the molecular weight standards (66, 45 ,31, 21,14 kDa) were added to the stacking gel wells using a 10 microlitre pipette tip and the samples run at 40mA for 5 hours.
15 The plates were then prised open, the gel placed in a tray of fixer (10% acetic acid, 40% ethanol, 50% water) and shaken overnight. Following this, the gel was primed by 30 minutes shaking in a primer solution (7.5% acetic acid (75ml), 0.05% SDS (5ml of 10%)). The gel was then incubated with a fluorescent dye (7.5% acetic acid, 0.06% OGS in-house dye (600μl)) with shaking for 3 hrs. Sypro Red (Molecular
20 Probes, Inc., Eugene, Oregon) is a suitable dye for this purpose. A preferred fluorescent dye is disclosed in U.S. Application No. 09/412,168, filed on October 5, 1999, which is incorporated herein by reference in its entirety.

A computer-readable output was produced by imaging the fluorescently stained gels with an Apollo 3 scanner (Oxford Glycosciences, Oxford, UK). This scanner is
25 developed from the scanner described in WO 96/36882 and in the Ph.D. thesis of David A. Basiji, entitled “Development of a High-throughput Fluorescence Scanner Employing Internal Reflection Optics and Phase-sensitive Detection (Total Internal Reflection, Electrophoresis)”, University of Washington (1997), Volume 58/12-B of

Dissertation Abstracts International, page 6686, the contents of each of which are incorporated herein by reference. The latest embodiment of this instrument includes the following improvements: The gel is transported through the scanner on a precision lead-screw drive system. This is preferable to laying the glass plate on the belt-driven system that is defined in the Basiji thesis as it provides a reproducible means of accurately transporting the gel past the imaging optics.

The gel is secured into the scanner against three alignment stops that rigidly hold the glass plate in a known position. By doing this in conjunction with the above precision transport system and the fact that the gel is bound to the glass plate, the absolute position of the gel can be predicted and recorded. This ensures that accurate co-ordinates of each feature on the gel can be communicated to the cutting robot for excision. This cutting robot has an identical mounting arrangement for the glass plate to preserve the positional accuracy.

The carrier that holds the gel in place has integral fluorescent markers (Designated M1, M2, M3) that are used to correct the image geometry and are a quality control feature to confirm that the scanning has been performed correctly.

The optical components of the system have been inverted. The laser, mirror, waveguide and other optical components are now above the glass plate being scanned. The embodiment of the Basiji thesis has these underneath. The glass plate is therefore mounted onto the scanner gel side down, so that the optical path remains through the glass plate. By doing this, any particles of gel that may break away from the glass plate will fall onto the base of the instrument rather than into the optics.

In scanning the gels, they were removed from the stain, rinsed with water and allowed to air dry briefly and imaged on the Apollo 3. After imaging, the gels were sealed in polyethylene bags containing a small volume of staining solution, and then stored at 4°C.

Apparent molecular weights were calculated by interpolation from a set of known molecular weight markers run alongside the samples.

1.1.5 – Recovery and analysis of selected proteins

Proteins were robotically excised from the gels by the process described in U.S. Patent No. 6,064,754, Sections 5.4 and 5.6, 5.7, 5.8 (incorporated herein by reference), as is applicable to 1D-electrophoresis, with modification to the robotic cutter as follows: the cutter begins at the top of the lane, and cuts a gel disc 1.7mm in diameter from the left edge of the lane. The cutter then moves 2mm to the right, and 0.7mm down and cuts a further disc. This is then repeated. The cutter then moves back to a position directly underneath the first gel cut, but offset by 2.2mm downwards, and the pattern of three diagonal cuts are repeated. This is continued for the whole length of the gel.

NOTE: If the lane is observed to broaden significantly then a correction can be made also sideways i.e. instead of returning to a position directly underneath a previous gel cut, the cut can be offset slightly to the left (on the left of the lane) and/or the right (on the right of the lane). The proteins contained within the gel fragments were processed to generate tryptic peptides; partial amino acid sequences of these peptides were determined by mass spectroscopy as described in WO98/53323 and Application No. 09/094,996, filed June 15, 1998.

Proteins were processed to generate tryptic digest peptides. Tryptic peptides were analyzed by mass spectrometry using a PerSeptive Biosystems Voyager- DETM STR Matrix-Assisted Laser Desorption Ionization Time-of-Flight (MALDI-TOF) mass spectrometer, and selected tryptic peptides were analyzed by tandem mass spectrometry (MS/MS) using a Micromass Quadrupole Time-of-Flight (Q-TOF) mass spectrometer (Micromass, Altrincham, U.K.) equipped with a nanoflowTM electrospray Z-spray source. For partial amino acid sequencing and identification of Integrin beta 4, uninterpreted tandem mass spectra of tryptic peptides were searched using the SEQUEST search program (Eng et al., 1994, J. Am. Soc. Mass Spectrom. 5:976-989), version v.C.1. Criteria for database identification included: the cleavage specificity of trypsin; the detection of a suite of a, b and y ions in peptides returned from the database, and a mass increment for all Cys residues to account for

carbamidomethylation. The database searched was a database constructed of protein entries in the non-redundant database held by the National Centre for Biotechnology Information (NCBI) which is accessible at www.ncbi.nlm.nih.gov. Following identification of proteins through spectral-spectral correlation using the SEQUEST program, masses detected in MALDI-TOF mass spectra were assigned to tryptic digest peptides within the proteins identified. In cases where no amino acid sequences could be identified through searching with uninterpreted MS/MS spectra of tryptic digest peptides using the SEQUEST program, tandem mass spectra of the peptides were interpreted manually, using methods known in the art. (In the case of interpretation of low-energy fragmentation mass spectra of peptide ions see Gaskell et al., 1992, Rapid Commun. Mass Spectrom. 6:658-662).

1.1.6 – Discrimination of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer associated proteins

The process to identify Integrin beta 4 uses the peptide sequences obtained experimentally by mass spectrometry described above of naturally occurring human proteins to identify and organize coding exons in the published human genome sequence.

Recent dramatic advances in defining the chemical sequence of the human genome have led to the near completion of this immense task (Venter, J.C. et al. (2001). The sequence of the human genome. Science 16: 1304-51; International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome Nature 409: 860-921). There is little doubt that this sequence information will have a substantial impact on our understanding of many biological processes, including molecular evolution, comparative genomics, pathogenic mechanisms and molecular medicine. For the full medical value inherent in the sequence of the human genome to be realised, the genome needs to be ‘organised’ and annotated. By this, is meant at least the following three things: (i) The assembly of the sequences of the individual portions of the genome

into a coherent, continuous sequence for each chromosome. (ii) The unambiguous identification of those regions of each chromosome that contain genes. (iii) Determination of the fine structure of the genes and the properties of its mRNA and protein products. While the definition of a 'gene' is an increasingly complex issue (H Pearson: What is a gene? Nature (2006) 24: 399 – 401), what is of immediate interest for drug discovery and development is a catalogue of those genes that encode functional, expressed proteins. A subset of these genes will be involved in the molecular basis of most if not all pathologies. Therefore an important and immediate goal for the pharmaceutical industry is to identify all such genes in the human genome and describe their fine structure.

Processing and integration of peptide masses, peptide signatures, ESTs and Public Domain Genomic Sequence Data to form OGAP® database

Discrete genetic units (exons, transcripts and genes) were identified using the following sequential steps:

1. A 'virtual transcriptome' is generated, containing the tryptic peptides which map to the human genome by combining the gene identifications available from Ensembl and various gene prediction programs. This also incorporates SNP data (from dbSNP) and all alternate splicing of gene identifications. Known contaminants were also added to the virtual transcriptome.
2. All tandem spectra in the OGeS Mass Spectrometry Database are interpreted in order to produce a peptide that can be mapped to one in the virtual transcriptome. A set of automated spectral interpretation algorithms were used to produce the peptide identifications.
3. The set of all mass-matched peptides in the OGeS Mass Spectrometry Database is generated by searching all peptides from transcripts hit by the tandem peptides using a tolerance based on the mass accuracy of the mass spectrometer, typically 20ppm.
4. All tandem and mass-matched peptides are combined in the form of "protein clusters". This is done using a recursive process which groups sequences into clusters based

on common peptide hits. Biological sequences are considered to belong to the same cluster if they share one or more tandem or mass-matched peptide.

5. After initial filtering to screen out incorrectly identified peptides, the resulting clusters are then mapped on the human genome.
6. The protein clusters are then aggregated into regions that define preliminary gene boundaries using their proximity and the co-observation of peptides within protein clusters. Proximity is defined as the peptide being within 80,000 nucleotides on the same strand of the same chromosome. Various elimination rules, based on cluster observation scoring and multiple mapping to the genome are used to refine the output. The resulting 'confirmed genes' are those which best account for the peptides and masses observed by mass spectrometry in each cluster. Nominal co-ordinates for the gene are also an output of this stage.
7. The best set of transcripts for each confirmed gene are created from the protein clusters, peptides, ESTs, candidate exons and molecular weight of the original protein spot.
8. Each identified transcript was linked to the sample providing the observed peptides.
9. Use of an application for viewing and mining the data. The result of steps 1 - 8 was a database containing genes, each of which consisted of a number of exons and one or more transcripts. An application was written to display and search this integrated genome / proteome data. Any features (OMIM disease locus, InterPro etc.) that had been mapped to the same Golden Path co-ordinate system by Ensembl could be cross-referenced to these genes by coincidence of location and fine structure.

Results

The process was used to generate approximately 1 million peptide sequences to identify protein-coding genes and their exons resulted in the identification of protein sequences for 18083 genes across 67 different tissues and 57 diseases including 506 genes in Bladder cancer, 4,713 genes in Breast cancer, 766 genes in Burkitt's lymphoma, 1,371 genes in Cervical cancer, 949 genes in Colorectal cancer, 524 genes in Gastric cancer, 1,782 genes in Hepatocellular carcinoma, 2,424 genes in chronic lymphocytic leukaemia

(CLL), 978 genes in Lung cancer, 1,764 genes in Melanoma, 1,033 genes in Ovarian Cancer, 2,961 genes in Pancreatic cancer and 3,307 genes in Prostate cancer, illustrated here by Integrin beta 4 isolated and identified from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer samples. Following 5 comparison of the experimentally determined sequences with sequences in the OGAP® database, Integrin beta 4 showed a high degree of specificity to breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer indicative of the prognostic and diagnostic nature.

10 1.2 RESULTS

These experiments identified Integrin beta 4, in its five different splice variants, as further described herein. The full-length Integrin beta 4 was detected in the plasma membrane of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer samples and was not detected in the cytosol.

15 Figure 2 shows the Protein Index for Integrin beta 4. For each gene, the protein index uses the mass spectrometry data to assign a score to each disease, relative to the global database. The Protein Index can then be used to identify cancer specific genes with a high score in cancer indications and low/negligible scores in normal and other diseases. The index contains ~ 1 million peptides sequenced via mass spectrometry 20 from 56 diseases. For each gene, this yields a score for each disease and subcellular location. The results are summarized below:

Protein Index Report for Integrin beta 4

25

Indications positive:
Breast cancer
Colorectal cancer
Gastric cancer
30 Hepatocellular carcinoma
Lung cancer
Pancreatic cancer

Disease controls

Acute monocytic leukaemia	Diabetes and Obesity	Migraine, acute
Acute T-cell leukaemia	Diverticulitis	Multiple sclerosis
Alzheimer's Disease	Dyslipidaemia	Neuroblastoma
Arthritis	Emphysema	Normal
Asthma	Focal apocrine metaplasia	Obesity
Atherosclerosis	**Gastric cancer**	Osteoarthritis
B-cell non-Hodgkin's lymphoma	Gaucher disease	Osteosarcoma
Bladder carcinoma	Glioblastoma	Ovarian cancer
Breast cancer	Hepatoblastoma	**Pancreatic cancer**
Breast diseases, benign	**Hepatocellular carcinoma**	Prostate cancer
Burkitt's lymphoma	Hypertension	Prostatic diseases, benign
Bursitis	Kidney cancer	Prostatitis
Cancer, unspecified	Lactational foci	Retinoblastoma
Cervical cancer	Leukaemia, unspecified	Schizophrenia
Chronic lymphocytic leukaemia	Liver cirrhosis	Skin ulcer
Chronic obstructive pulmonary disease	**Lung cancer**	Smoker
Colorectal cancer	Lymphoma, histiocytic	Teratocarcinoma
Dementia, vascular	Melanoma	
Depression	Metabolic syndrome X	

Subcellular location

Birbeck Granules	**Membrane**	Secreted
Cell surface digest	Membrane Glycoprotein Binding Fraction	Soluble Fraction
Chromatin Fraction	Mitochondria	Supernatant
Crude Cell Membrane	**Nucleus**	Whole Cell
Cytosol	Peroxisomes	
Golgi/Mitochondrial	**Plasma Membrane**	

5

Figure 2 shows the Protein Index for Integrin beta 4 is medium in plasma membrane and very low in membrane for breast cancer, very high in colorectal cancer plasma membrane, high in gastric cancer plasma membrane, high in hepatocellular
10 carcinoma plasma membrane, medium in lung cancer plasma membrane and very high in plasma membrane and low in nucleus for pancreatic cancer. It was also detected as very low in membrane and low in plasma membrane in normal samples. Integrin beta 4 was not detected in any other diseases. This indicates that Integrin beta 4 is potentially

a good marker for breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer and pancreatic cancer.

5

All references referred to in this application, including patent and patent applications, are incorporated herein by reference to the fullest extent possible.

10 Throughout the specification and the claims which follow, unless the context requires otherwise, the word 'comprise', and variations such as 'comprises' and 'comprising', will be understood to imply the inclusion of a stated integer, step, group of integers or group of steps but not to the exclusion of any other integer, step, group of integers or group of steps.

15 Embodiments of the invention are described herein, which comprise certain elements. The invention also extends to separate embodiments consisting of or consisting essentially of the same elements, and *vice versa*.

The application of which this description and claims form part may be used as a basis for priority in respect of any subsequent application. The claims of such subsequent 20 application may be directed to any feature or combination of features described herein. They may take the form of product, composition, process, or use claims and may include, by way of example and without limitation, the following claims:

SEQUENCE LISTING

Sequence	Sequence ID

MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAYCTDE MFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQMSPQGLRVR LRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLKKMGQNLARVLSQLT SDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPPFSFKNVISLTEDVDEFRNK LQGERISGNLDAPEGGFDAILQTAVCTRDIGWRPDSTHLLVFSTESAFHYEADGA NVLAGIMSRNDERCHLDTTGTYTQYRTQDYPSVPTLVRLLAKHNIIPIFAVTNYSYS YYEKLHTYFPVSSLGVLQEDSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSK MFQKTRTGSFHIRRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDG LKMDAGIICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGS LSDIQPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGF LCNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGRC HCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEECNFKV KMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLVHKKKDCPP GSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCNRGHMVGFKEDHYM LRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQRPGFATHAASINPTELVP YGLSLRLARLCTENLLKPDTRECAQLRQEVEENLNEVYRQISGVHKLQQTKFRQQ PNAGKKQDHTIVDTVLMAPRSAKPALLKLTEKQVEQRAFHDLKVAPGYYTLTADQ DARGMVEFQEGVELVDVRVPLFIRPEDDDEKQLLVEAIDVPAGTATLGRRLVNITII KEQARDVVSFEQPEFSVSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRD YIPVEGELLFQPGEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLG QPHSTTIIIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWL PPSGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYGAQ GEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEITAYEVCY GLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKARNGAGWGPERE AIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYSDDVLRSPSGSQRPSVSDDT GCGWKFEPLLGEELDLRRVTWRLPPELIPRLSASSGRSSDAEAPHGPPDDGGAG GKGGSLPRSATPGPPGEHLVNGRMDFAFPGSTNSLHRMTTTSAAAYGTHLSPHV PHRVLSTSSTLTRDYNSLTRSEHSHSTTLPRDYSTLTSVSSHDSRLTAGVPDTPT RLVFSALGPTSLRVSWQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVV EDLLPNHSYVFRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPS APGPLVFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDS PESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGSRAGLF QHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVTQEFVSRTLTT SGTLSTHMDQQFFQT	1
MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAYCTDE MFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQMSPQGLRVR LRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLKKMGQNLARVLSQLT SDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPPFSFKNVISLTEDVDEFRNK LQGERISGNLDAPEGGFDAILQTAVCTRDIGWRPDSTHLLVFSTESAFHYEADGA NVLAGIMSRNDERCHLDTTGTYTQYRTQDYPSVPTLVRLLAKHNIIPIFAVTNYSYS YYEKLHTYFPVSSLGVLQEDSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSK MFQKTRTGSFHIRRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDG LKMDAGIICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGS LSDIQPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGF LCNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGRC HCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEECNFKV KMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLVHKKKDCPP GSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCNRGHMVGFKEDHYM LRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQRPGFATHAASINPTELVP YGLSLRLARLCTENLLKPDTRECAQLRQEVEENLNEVYRQISGVHKLQQTKFRQQ PNAGKKQDHTIVDTVLMAPRSAKPALLKLTEKQVEQRAFHDLKVAPGYYTLTADQ DARGMVEFQEGVELVDVRVPLFIRPEDDDEKQLLVEAIDVPAGTATLGRRLVNITII KEQARDVVSFEQPEFSVSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRD	2

YIPVEGELLFQPGEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLG QPHSTTIIIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWL PPSGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYGAQ GEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEITAYEVCY GLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKARNGAGWGPERE AIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYSDDVLRSPSGSQRPSVSDDT EHLVNGRMDFAFPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYN SLTRSEHSHSTTLPRDYSTLTSVSSHDSRLTAGVPDTPTRLVFSALGPTSLRVSW QEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYVFRVRAQ SQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPLVFTALSPDSLQ LSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSPESRLTVPGLSENVPY KFKVQARTTEGFGPEREGIITIESQDGGPFPQLGSRAGLFQHPLQSEYSSITTTHT SATEPFLVDGPTLGAQHLEAGGSLTRHVTQEFVSRTLTTSGTLSTHMDQQFFQT	
MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAYCTDE MFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQMSPQGLRVR LRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLKKMGQNLARVLSQLT SDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPPFSFKNVISLTEDVDEFRNK LQGERISGNLDAPEGGFDAILQTAVCTRDIGWRPDSTHLLVFSTESAFHYEADGA NVLAGIMSRNDERCHLDTTGTYTQYRTQDYPSVPTLVRLLAKHNIIPIFAVTNYSYS YYEKLHTYFPVSSLGVLQEDSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSK MFQKTRTGSFHIRRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDG LKMDAGIICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGS LSDIQPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGF LCNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGRC HCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEECNFKV KMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLVHKKKDCPP GSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCNRGHMVGFKEDHYM LRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQRPGFATHAASINPTELVP YGLSLRLARLCTENLLKPDTRECAQLRQEVEENLNEVYRQISGVHKLQQTKFRQQ PNAGKKQDHTIVDTVLMAPRSAKPALLKLTEKQVEQRAFHDLKVAPGYYTLTADQ DARGMVEFQEGVELVDVRVPLFIRPEDDDEKQLLVEAIDVPAGTATLGRRLVNITII KEQARDVVSFEQPEFSVSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRD YIPVEGELLFQPGEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLG QPHSTTIIIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWL PPSGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYGAQ GEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEITAYEVCY GLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKARNGAGWGPERE AIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYSDDVLRSPSGSQRPSVSDDT EHLVNGRMDFAFPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYN SLTRSEHSHSTTLPRDYSTLTSVSSHGLPPIWEHGRSRLPLSWALGSRSRAQMK GFPPSRGPRDSIILAGRPAAPSWGPDSRLTAGVPDTPTRLVFSALGPTSLRVSWQ EPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYVFRVRAQS QEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPLVFTALSPDSLQL SWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSPESRLTVPGLSENVPY KFKVQARTTEGFGPEREGIITIESQDGGPFPQLGSRAGLFQHPLQSEYSSITTTHT SATEPFLVDGPTLGAQHLEAGGSLTRHVTQEFVSRTLTTSGTLSTHMDQQFFQT	3

MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAYCTDE MFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQMSPQGLRVR LRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLKKMGQNLARVLSQLT SDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPPFSFKNVISLTEDVDEFRNK LQGERISGNLDAPEGGFDAILQTAVCTRDIGWRPDSTHLLVFSTESAFHYEADGA NVLAGIMSRNDERCHLDTTGTYTQYRTQDYPSVPTLVRLLAKHNIIPIFAVTNYSYS YYEKLHTYFPVSSLGVLQEDSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSK MFQKTRTGSFHIRRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDG LKMDAGIICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGS LSDIQPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGF LCNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGRC HCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEECNFKV KMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLVHKKKDCPP GSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCNRGHMVGFKEDHYM LRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQRPGFATHAASINPTELVP YGLSLRLARLCTENLLKPDTRECAQLRQEVEENLNEVYRQISGVHKLQQTKFRQQ PNAGKKQDHTIVDTVLMAPRSAKPALLKLTEKQVEQRAFHDLKVAPGYYTLTADQ DARGMVEFQEGVELVDVRVPLFIRPEDDDEKQLLVEAIDVPAGTATLGRRLVNITII KEQARDVVSFEQPEFSVSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRD YIPVEGELLFQPGEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLG QPHSTTIIIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWL PPSGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYGAQ GEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEITAYEVCY GLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKARNGAGWGPERE AIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYSDDVLRSPSGSQRPSVSDDT EHLVNGRMDFAFPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYN SLTRSEHSHSTTLPRDYSTLTSVSSHDSRLTAGVPDTPTRLVFSALGPTSLRVSW QEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYVFRVRAQ SQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPLVFTALSPDSLQ LSWERPRRPNGDIVGYLVTWPATAFRVDGDSPESRLTVPGLSENVPYKFKVQAR TTEGFGPEREGIITIESQDGGPFPQLGSRAGLFQHPLQSEYSSITTTHTSATEPFLV DGPTLGAQHLEAGGSLTRHVTQEFVSRTLTTSGTLSTHMDQQFFQT	4
MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAYCTDE MFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQMSPQGLRVR LRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLKKMGQNLARVLSQLT SDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPPFSFKNVISLTEDVDEFRNK LQGERISGNLDAPEGGFDAILQTAVCTRDIGWRPDSTHLLVFSTESAFHYEADGA NVLAGIMSRNDERCHLDTTGTYTQYRTQDYPSVPTLVRLLAKHNIIPIFAVTNYSYS YYEKLHTYFPVSSLGVLQEDSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSK MFQKTRTGSFHIRRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDG LKMDAGIICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGS LSDIQPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGF LCNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGRC HCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEECNFKV KMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLVHKKKDCPP GSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCNRGHMVGFKEDHYM LRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQRPGFATHAASINPTELVP YGLSLRLARLCTENLLKPDTRECAQLRQEVEENVRTQELGLAGDVAERGLQADL RCTQAPADQVPAAAQCREKARPHHCGHSADGAPLGQAGPAEAYREAGGTEGLP RPQGGPRLLHPHCRPGRPGHGGVPGGRGAGGRTGAPLYPA	5
ACLALLPCCNR	6
AEEVVVR	7

AFHDLK	8
ALEHVDGTHVCQLPEDQK	9
AQSQEGWGR	10
CERPLQGYSVEYQLLNGGELHR	11
DPDELDR	12
DVVSFEQPEFSVSR	13
DYIPVEGELLFQPGEAWK	14
DYNSLTR	15
DYSTLTSVSSHDSR	16
EAIINLATQPK	17
ECAQLR	18
EDHYMLR	19
EGEDKPCSGR	20
EGIITIESQDGGPFPQLGSR	21
ENLMASDHLDTPMLR	22
FEPLLGEELDLR	23
FEPLLGEELDLRR	24
FHVQLSNPK	25
GEVGIYQVQLR	26
GMVEFQEGVELVDVR	27
GNIHLKPSFSDGLK	28
HNIIPIFAVTNYSYSYYEK	29
HVTQEFVSR	30
IHFNWLPPSGKPMGYR	31
ISGNLDAPEGGFDAILQTAVCTR	32
KIHFNWLPPSGKPMGYR	33
LCTENLLKPDTR	34
LLELQEVDSLLR	35
LNIPNPAQTSVVVEDLLPNHSYVFR	36
LTAGVPDTPTR	37
LTVPGLSENVPYK	38
LVFSALGPTSLR	39
MAGPRPSPWAR	40
MDAGIICDVCTCELQK	41
MDFAFPGSTNSLHR	42
MGQNLAR	43
MLLIENLR	44
MTTTSAAAYGTHLSPHVPHR	45
MVDELK	46
NGAGWGPER	47
NVISLTEDVDEFR	48
PSVSDDTEHLVNGR	49
QDHTIVDTVLMAPR	50
QEVEENLNEVYR	51
QISGVHK	52
QLLVEAIDVPAGTATLGR	53

QQPNAGK	54
RAEEVVVR	55
RCNTQAELLAAGCQR	56
RFHVQLSNPK	57
RGEVGIYQVQLR	58
RPNGDIVGYLVTCEMAQGGGPATAFR	59
RSQMSPQGLR	60
RVTWR	61
SATPGPPGEHLVNGR	62
SCVQCQAWGTGEK	63
SEHSHSTTLPR	64
SFTSQMLSSQPPPHGDLGAPQNPNAK	65
SNLDIR	66
SQMSPQGLR	67
SQVSYR	68
SSDAEAPHGPPDDGGAGGK	69
TGSFHIR	70
TGSFHIRR	71
THQEVPSEPGR	72
TLTTSGTLSTHMDQQFFQT	73
TQDYPSVPTLVR	74
TTEGFGPER	75
VAPGYYTLTADQDAR	76
VCAYGAQGEGPYSSLVSCR	77
VDGDSPESR	78
VLSTSSTLTR	79
VLVDNPK	80
VPLFIRPEDDDEK	81
VPSVELTNLYPYCDYEMK	82
VSVPQTDMRPEK	83
VSWQEPR	84
YWIQGDSESEAHLLDSK	85

CLAIMS:

1. A method for treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer which comprises administering to a subject in need thereof a therapeutically effective amount of a composition comprising an affinity reagent capable of specific binding to Integrin beta 4 or a fragment thereof, and a pharmaceutically acceptable diluent or carrier, wherein Integrin beta 4 is overexpressed in said cancers.
2. An affinity reagent capable of specific binding to Integrin beta 4 or a fragment thereof.
3. An affinity reagent according to claim 2 which contains or is conjugated to a therapeutic moiety.
4. An affinity reagent according to claim 3 wherein the therapeutic moiety is a cytotoxic moiety or a radioactive isotype.
5. An affinity reagent according to claim 2 which contains or is conjugated to a detectable label.
6. An affinity reagent according to any one of claims 2 to 5 which is an antibody.
7. An antibody according to claim 6 which is an isolated monoclonal antibody, or an antigen-binding portion thereof, an antibody fragment, or an antibody mimetic.
8. An isolated monoclonal antibody according to claim 7 wherein said antibody is a full-length antibody of an IgG1, IgG2, IgG3, or IgG4 isotype.
9. An isolated monoclonal antibody according to claim 7 wherein said antibody is selected from the group consisting of: a whole antibody, an antibody fragment, a humanised antibody, a single chain antibody, an immunoconjugate, a defucosylated antibody, and a bispecific antibody.
10. An antibody fragment according to claim 7, wherein the fragment is selected from the group consisting of: a UniBody, a domain antibody and a Nanobody.
11. An antibody mimetic according to claim 7, wherein the mimetic is selected from the group consisting of: an Affibody, a DARPin, an Anticalin, an Avimer, a Versabody, and a Duocalin.

12. A monoclonal antibody according to claim 7, which has cytotoxicity against Integrin beta 4 antigen expressing cells in the presence of a human complement.
13. A monoclonal antibody according to claim 7, which has cytotoxicity against Integrin beta 4 antigen expressing cells in the presence of human immune effector
5 cells.
14. A pharmaceutical composition comprising a therapeutically effective amount of an affinity reagent or a fragment thereof as defined in any one of claims 2 to 13, and a pharmaceutically acceptable diluent or carrier.
15. A pharmaceutical composition according to claim 14 comprising one or more
10 affinity reagents as defined in any one of claims 2 to 13 and a pharmaceutically acceptable excipient.
16. An agent as defined in any one of claims 2 to 13 or a composition as defined in claim 14 or claim 15 for use in treating or preventing disease.
17. An agent according to claim 16 wherein the disease is cancer.
15 18. An agent according to claim 17 wherein the cancer is breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.
19. Integrin beta 4, or a fragment thereof for use in treating or preventing disease.
20. Integrin beta 4, or a fragment thereof according to claim 19 wherein the disease is cancer.
20 21. Integrin beta 4, or a fragment thereof according to claim 20 wherein the cancer is breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.
22. A method for treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer which comprises
25 administering to a subject in need thereof a therapeutically effective amount of an agent as defined in any one of claims 2 to 13 or a composition as defined in claim 14 or claim 15.
23. An isolated nucleic acid molecule encoding the isolated antibody or antigen-binding portion thereof of claim 7.

24. An expression vector comprising the nucleic acid molecule of claim 23.
25. A host cell comprising the expression vector of claim 24.
26. A kit containing one or more affinity reagents according to any one of claims 2 to 13 or a composition as defined in claim 14, wherein said affinity reagent is suitable 5 for use in treatment and/or diagnosis.
27. A kit according to claim 26, which further comprises instructions for use of said affinity reagent as defined in any one of claims 16 to 18.
28. A kit according to claim 26 or claim 27 which further comprises a hybridising agent.
29. A method of screening for compounds that modulate the activity of Integrin beta 4, 10 the method comprising: (a) contacting Integrin beta 4 or a biologically active portion thereof with a candidate compound; and (b) determining whether activity of Integrin beta 4 is thereby modulated.
30. A method according to claim 29 which comprises (a) contacting Integrin beta 4 or a biologically active portion thereof with a candidate compound in a sample; and (b) 15 comparing the activity of Integrin beta 4 or a biologically active portion thereof in said sample after contact with said candidate compound with the activity of Integrin beta 4 or a biologically active portion thereof in said sample before contact with said candidate compound, or with a reference level of activity.
31. A method according to claim 29 or claim 30 which is a method of screening for 20 compounds that inhibit activity of Integrin beta 4.
32. A method according to any one of claims 29 to 31 wherein Integrin beta 4 or a biologically active portion thereof is expressed on or by a cell.
33. A method according to any one of claims 29 to 31 wherein Integrin beta 4 or a biologically active portion thereof is isolated from cells which express it.
34. 25 A method according to claim 33 wherein Integrin beta 4 or a biologically active portion thereof is immobilised onto a solid phase.
35. A method of screening for compounds that modulate the expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4, the method comprising: (a) contacting cells expressing Integrin beta 4 or nucleic acid encoding Integrin beta 4 with a

candidate compound; and (b) determining whether expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4 is thereby modulated.

36. A method according to claim 35 which comprises (a) contacting cells expressing Integrin beta 4 or nucleic acid encoding Integrin beta 4 with a candidate compound in a sample; and (b) comparing the expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4 by cells in said sample after contact with said candidate compound with the expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4 of cells in said sample before contact with said candidate compound, or with a reference level of expression.

37. A method according to claim 35 or claim 36 which is a method of screening for compounds that inhibit expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4.

38. A compound obtainable by a method according to any one of claims 29 to 37.

39. A compound which modulates the activity or expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4.

40. A compound according to claim 39 which inhibits the activity or expression of Integrin beta 4 or nucleic acid encoding Integrin beta 4.

41. A compound according to any one of claims 38 to 40 for use in treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

42. A method for treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer which comprises administering to a subject in need thereof a therapeutically effective amount of a compound according to any one of claims 38 to 40.

43. A hybridizing agent capable of hybridizing to nucleic acid encoding Integrin beta 4 and inhibiting transcription of mRNA.

44. A hybridizing agent according to claim 43 which contains or is conjugated to a detectable label.

45. A pharmaceutical composition comprising one or more hybridizing agents as defined in claim 43 or claim 44 and a pharmaceutically acceptable diluent or carrier.
46. A kit containing one or more hybridizing agents according to any one of claims 43 to 45 wherein said hybridising agent is suitable for use in treatment and/or diagnosis.
47. A kit according to claim 46 further containing reagents capable of detecting and reporting the binding of said hybridizing agents to their binding partners.
48. A hybridizing agent as defined in any one of claim 43 or claim 44 for use in treatment.
49. A hybridizing agent according to claim 48 wherein the treatment is for cancer.
50. A hybridizing agent according to claim 49, wherein the cancer is selected from breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.
51. A method for treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer which comprises administering to a subject in need thereof a therapeutically effective amount of a composition comprising a hybridizing agent capable of hybridizing to nucleic acid encoding Integrin beta 4, and a pharmaceutically acceptable diluent or carrier.
52. An immunogenic composition comprising Integrin beta 4 or an epitope containing fragment thereof, or nucleic acid encoding Integrin beta 4 or a fragment thereof optionally together with an immunostimulant.
53. A vaccine composition comprising Integrin beta 4 or an epitope containing fragment thereof, or nucleic acid encoding Integrin beta 4 or an epitope containing fragment thereof optionally together with an immunostimulant.
54. A method of raising an immune response which comprises administering to a subject a composition according to claim 52.
55. A method for treating or preventing breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer which comprises

administering to a subject in need thereof a therapeutically effective amount of a composition according to claim 52 or claim 53.

56. A composition according to claim 52 or claim 53 for use in preventing or treating breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer.

57. A method of detecting, diagnosing and/or screening for or monitoring the progression of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer or of monitoring the effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy in a subject which comprises detecting the presence or level of Integrin beta 4, or one or more fragments thereof, or the presence or level of nucleic acid encoding Integrin beta 4 or the presence or level of the activity of Integrin beta 4 or which comprises detecting a change in the level thereof in said subject.

58. A method of detecting, diagnosing and/or screening for breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in a candidate subject which comprises detecting the presence of Integrin beta 4, or one or more fragments thereof, or the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of Integrin beta 4 in said candidate subject, in which either (a) the presence of an elevated level of Integrin beta 4 or said one or more fragments thereof or an elevated level of nucleic acid encoding Integrin beta 4 or the presence of an elevated level of Integrin beta 4 activity in the candidate subject as compared with the level in a healthy subject or (b) the presence of a detectable level of Integrin beta 4 or said one or more fragments thereof or a detectable level of nucleic acid encoding Integrin beta 4 or the presence of a detectable level of Integrin beta 4 activity in the candidate subject as compared with a corresponding undetectable level in a healthy subject indicates the presence of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in said subject.

59. A method of monitoring the progression of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in a subject or of monitoring the effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy which comprises detecting the presence of Integrin beta 4, or one or more fragments thereof, or the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of Integrin beta 4 in said candidate subject at a first time point and at a later time point, the presence of an elevated or lowered level of Integrin beta 4 or said one or more fragments thereof or an elevated or lowered level of nucleic acid encoding Integrin beta 4 or the presence of an elevated or lowered level of Integrin beta 4 activity in the subject at the later time point as compared with the level in the subject at said first time point, indicating the progression or regression of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer or indicating the effect or non-effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy in said subject.
60. A method according to any one of claims 57 to 59 wherein the presence of Integrin beta 4, or one or more fragments thereof, or the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of Integrin beta 4 is detected by analysis of a biological sample obtained from said subject.
61. A method according to claim 60 which includes the step of obtaining said sample for analysis from said subject.
62. A method according to claim 60 or claim 61 wherein the sample is a sample of breast, colorectal, gastric epithelium, liver, lung or pancreatic tissue.
63. A method according to any one of claims 57 to 62 wherein the presence of Integrin beta 4, or one or more fragments thereof, or the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of Integrin beta 4 is detected quantitatively.

64. A method according to claim 63 wherein the presence of Integrin beta 4, or one or more fragments thereof, or the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of Integrin beta 4 is detected quantitatively by means involving use of an imaging technology.
65. A method according to any one of claims 57 to 63 involving use of immunohistochemistry on tissue sections in order to determine the presence of Integrin beta 4, or one or more fragments thereof, or the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of Integrin beta 4, and thereby to localise breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer cells.
66. A method according to any one of claims 57 to 59 wherein the presence of Integrin beta 4, or one or more fragments thereof, or the presence of nucleic acid encoding Integrin beta 4 or the presence of the activity of Integrin beta 4 is detected by analysis in situ.
67. A method according to any one of claims 57 to 66 wherein the presence of Integrin beta 4 or one or more epitope-containing fragments thereof is detected.
68. A method according to claim 67 wherein the presence of Integrin beta 4 or one or more fragments thereof is detected using an affinity reagent capable of specific binding to Integrin beta 4 or one or more fragments thereof.
69. A method according to claim 68 wherein the affinity reagent is an antibody.
70. A method according to any one of claims 57 to 66 wherein nucleic acid encoding Integrin beta 4 is detected.
71. A method according to claim 70 wherein nucleic acid encoding Integrin beta 4 is detected using a hybridizing agent capable of hybridizing to nucleic acid encoding Integrin beta 4.
72. A method according to any one of claims 57 to 66 wherein the activity of Integrin beta 4 is detected.
73. A method of detecting, diagnosing and/or screening for or monitoring the progression of breast cancer, colorectal cancer, gastric cancer, hepatocellular

carcinoma, lung cancer or pancreatic cancer or of monitoring the effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy in a subject which comprises detecting the presence or level of antibodies capable of immunospecific binding to Integrin beta 4, or one or more epitope-containing fragments thereof or which comprises detecting a change in the level thereof in said subject.

74. A method of detecting, diagnosing and/or screening for breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in a subject which comprises detecting the presence of antibodies capable of immunospecific binding to Integrin beta 4, or one or more epitope-containing fragments thereof in said subject, in which (a) the presence of an elevated level of antibodies capable of immunospecific binding to Integrin beta 4 or said one or more epitope-containing fragments thereof in said subject as compared with the level in a healthy subject or (b) the presence of a detectable level of antibodies capable of immunospecific binding to Integrin beta 4 or said one or more epitope-containing fragments thereof in said subject as compared with a corresponding undetectable level in a healthy subject indicates the presence of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer in said subject.

75. A method of monitoring the progression of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer or of monitoring the effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy in a subject which comprises detecting the presence of antibodies capable of immunospecific binding to Integrin beta 4, or one or more epitope-containing fragments thereof in said subject at a first time point and at a later time point, the presence of an elevated or lowered level of antibodies capable of immunospecific binding to Integrin beta 4, or one or more epitope-containing fragments thereof in

said subject at the later time point as compared with the level in said subject at said first time point, indicating the progression or regression of breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer or the effect or non-effect of an anti-breast cancer, anti-colorectal cancer, anti-gastric cancer, anti-hepatocellular carcinoma, anti-lung cancer or anti-pancreatic cancer drug or therapy in said subject.

76. A method according to any one of claims 73 to 75 wherein the presence of antibodies capable of immunospecific binding to Integrin beta 4, or one or more epitope-containing fragments thereof is detected by analysis of a biological sample obtained from said subject.

77. A method according to claim 76 which includes the step of obtaining said sample for analysis from said subject.

78. A method according to claim 76 or claim 77 wherein the sample is a sample of breast, colorectal, gastric epithelium, liver, lung or pancreatic tissue.

79. A method according to any one of claims 57 to 78 wherein the level that may be detected in the candidate subject who has breast cancer, colorectal cancer, gastric cancer, hepatocellular carcinoma, lung cancer or pancreatic cancer is 2 or more fold higher than the level in the healthy subject.

Figure 1

Breast cancer

```
Integrin beta-4a      MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 1
Integrin beta-4b      MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY   50
SEQ ID No: 2
Integrin beta-4c      MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 3
Integrin beta-4d      MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 4
Integrin beta-4e      MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 1
Integrin beta-4b      CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ   100
SEQ ID No: 2
Integrin beta-4c      CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 3
Integrin beta-4d      CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 4
Integrin beta-4e      CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 1
Integrin beta-4b      MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK   150
SEQ ID No: 2
Integrin beta-4c      MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 3
Integrin beta-4d      MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 4
Integrin beta-4e      MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 1
Integrin beta-4b      KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP   200
SEQ ID No: 2
Integrin beta-4c      KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 3
Integrin beta-4d      KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 4
Integrin beta-4e      KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 1
Integrin beta-4b      FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG   250
SEQ ID No: 2
```

```
Integrin beta-4c        FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 3
Integrin beta-4d        FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 4
Integrin beta-4e        FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 1
Integrin beta-4b        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR  300
SEQ ID No: 2
Integrin beta-4c        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 3
Integrin beta-4d        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 4
Integrin beta-4e        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 1
Integrin beta-4b        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE  350
SEQ ID No: 2
Integrin beta-4c        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 3
Integrin beta-4d        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 4
Integrin beta-4e        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 1
Integrin beta-4b        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI  400
SEQ ID No: 2
Integrin beta-4c        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 3
Integrin beta-4d        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 4
Integrin beta-4e        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 1
Integrin beta-4b        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI  450
SEQ ID No: 2
Integrin beta-4c        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 3
Integrin beta-4d        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 4
Integrin beta-4e        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
```

```
SEQ ID No: 1
Integrin beta-4b        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI   500
SEQ ID No: 2
Integrin beta-4c        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 3
Integrin beta-4d        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 4
Integrin beta-4e        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 1
Integrin beta-4b        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL   550
SEQ ID No: 2
Integrin beta-4c        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 3
Integrin beta-4d        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 4
Integrin beta-4e        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 1
Integrin beta-4b        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR   600
SEQ ID No: 2
Integrin beta-4c        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 3
Integrin beta-4d        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 4
Integrin beta-4e        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 1
Integrin beta-4b        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE   650
SEQ ID No: 2
Integrin beta-4c        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 3
Integrin beta-4d        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 4
Integrin beta-4e        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 1
Integrin beta-4b        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV   700
SEQ ID No: 2
Integrin beta-4c        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 3
Integrin beta-4d        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 4
Integrin beta-4e        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 5
```

```
                     **************************************************
Integrin beta-4a     HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 1
Integrin beta-4b     HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN  750
SEQ ID No: 2
Integrin beta-4c     HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 3
Integrin beta-4d     HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 4
Integrin beta-4e     HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 5
                     **************************************************
Integrin beta-4a     RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 1
Integrin beta-4b     RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR  800
SEQ ID No: 2
Integrin beta-4c     RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 3
Integrin beta-4d     RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 4
Integrin beta-4e     RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 5
                     **************************************************
Integrin beta-4a     PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 1
Integrin beta-4b     PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN  850
SEQ ID No: 2
Integrin beta-4c     PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 3
Integrin beta-4d     PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 4
Integrin beta-4e     PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 5
                     **************************************************
Integrin beta-4a     LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 1
Integrin beta-4b     LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK  900
SEQ ID No: 2
Integrin beta-4c     LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 3
Integrin beta-4d     LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 4
Integrin beta-4e     VRTQELGLAGDVAERGLQADLRCTQAPADQVPAAAQCREKARPHHCGHSA
SEQ ID No: 5
                     :.     ::*    :  :   :  . .  :: . .      .*. : .
Integrin beta-4a     LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 1
Integrin beta-4b     LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI  950
SEQ ID No: 2
Integrin beta-4c     LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 3
Integrin beta-4d     LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
```

```
SEQ ID No: 4
Integrin beta-4e     DGAPLGQAGPAEAYREAGGTEGLPRPQGGPRLLHPHCRPGRPGHGGVPGG
SEQ ID No: 5
                           :  .  .   : * *    *.   *..   ::. :        .

Integrin beta-4a     RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 1
Integrin beta-4b     RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS  1000
SEQ ID No: 2
Integrin beta-4c     RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 3
Integrin beta-4d     RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 4
Integrin beta-4e     RGAGGRTGAPLYPA------------------------------------
SEQ ID No: 5
                     *   ..      *   *

Integrin beta-4a     VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 1
Integrin beta-4b     VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP  1050
SEQ ID No: 2
Integrin beta-4c     VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 3
Integrin beta-4d     VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 1
Integrin beta-4b     GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI  1100
SEQ ID No: 2
Integrin beta-4c     GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 3
Integrin beta-4d     GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 1
Integrin beta-4b     IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP  1150
SEQ ID No: 2
Integrin beta-4c     IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 3
Integrin beta-4d     IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 1
Integrin beta-4b     SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG  1200
SEQ ID No: 2
```

```
Integrin beta-4c     SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 3
Integrin beta-4d     SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 1
Integrin beta-4b     AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI   1250
SEQ ID No: 2
Integrin beta-4c     AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 3
Integrin beta-4d     AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 1
Integrin beta-4b     TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA   1300
SEQ ID No: 2
Integrin beta-4c     TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 3
Integrin beta-4d     TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 1
Integrin beta-4b     RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS   1350
SEQ ID No: 2
Integrin beta-4c     RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 3
Integrin beta-4d     RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     DDVLRSPSGSQRPSVSDDTGCGWKFEPLLGEELDLRRVTWRLPPELIPRL
SEQ ID No: 1
Integrin beta-4b     DDVLRSPSGSQRPSVSDDT-------------------------------   1369
SEQ ID No: 2
Integrin beta-4c     DDVLRSPSGSQRPSVSDDT-------------------------------
SEQ ID No: 3
Integrin beta-4d     DDVLRSPSGSQRPSVSDDT-------------------------------
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     SASSGRSSDAEAPHGPPDDGGAGGKGGSLPRSATPGPPGEHLVNGRMDFA
SEQ ID No: 1
```

```
Integrin beta-4b    --------------------------------------EHLVNGRMDFA  1380
SEQ ID No: 2
Integrin beta-4c    --------------------------------------EHLVNGRMDFA
SEQ ID No: 3
Integrin beta-4d    --------------------------------------EHLVNGRMDFA
SEQ ID No: 4
Integrin beta-4e    -------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 1
Integrin beta-4b    FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH  1430
SEQ ID No: 2
Integrin beta-4c    FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 3
Integrin beta-4d    FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    SHSTTLPRDYSTLTSVSSH-------------------------------
SEQ ID No: 1
Integrin beta-4b    SHSTTLPRDYSTLTSVSSH-------------------------------  1449
SEQ ID No: 2
Integrin beta-4c    SHSTTLPRDYSTLTSVSSHGLPPIWEHGRSRLPLSWALGSRSRAQMKGFP
SEQ ID No: 3
Integrin beta-4d    SHSTTLPRDYSTLTSVSSH-------------------------------
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 1
Integrin beta-4b    ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS  1477
SEQ ID No: 2
Integrin beta-4c    PSRGPRDSIILAGRPAAPSWGPDSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 3
Integrin beta-4d    ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 1
Integrin beta-4b    WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV  1527
SEQ ID No: 2
Integrin beta-4c    WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 3
Integrin beta-4d    WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5
```

```
Integrin beta-4a       FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 1
Integrin beta-4b       FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL  1577
SEQ ID No: 2
Integrin beta-4c       FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 3
Integrin beta-4d       FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP
SEQ ID No: 1
Integrin beta-4b       VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP  1627
SEQ ID No: 2
Integrin beta-4c       VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP
SEQ ID No: 3
Integrin beta-4d       VFTALSPDSLQLSWERPRRPNGDIVGYLVTW-------PATAFRVDGDSP
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 1
Integrin beta-4b       ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS  1677
SEQ ID No: 2
Integrin beta-4c       ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 3
Integrin beta-4d       ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 1
Integrin beta-4b       RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT  1727
SEQ ID No: 2
Integrin beta-4c       RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 3
Integrin beta-4d       RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 1
Integrin beta-4b       QEFVSRTLTTSGTLSTHMDQQFFQT                           1752
SEQ ID No: 2
Integrin beta-4c       QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 3
Integrin beta-4d       QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 4
```

Integrin beta-4e -------------------------
SEQ ID No: 5

Key:
Recombinant protein
Key Tandem peptides(underline)
Mass Match peptides (bold)
Calx-beta and FNIII domains
Important tyrosine residues(double underline)

Colorectal cancer

Integrin beta-4a SEQ ID No: 1 **MAGPRPSPWAR**LLLAALISVSLSGTLA*NRCKKAPVKSCTECVRVDKDCAY*
Integrin beta-4b SEQ ID No: 2 **MAGPRPSPWAR**LLLAALISVSLSGTLA*NRCKKAPVKSCTECVRVDKDCAY* 50
Integrin beta-4c SEQ ID No: 3 **MAGPRPSPWAR**LLLAALISVSLSGTLA*NRCKKAPVKSCTECVRVDKDCAY*
Integrin beta-4d SEQ ID No: 4 **MAGPRPSPWAR**LLLAALISVSLSGTLA*NRCKKAPVKSCTECVRVDKDCAY*
Integrin beta-4e SEQ ID No: 5 **MAGPRPSPWAR**LLLAALISVSLSGTLA*NRCKKAPVKSCTECVRVDKDCAY*
****

Integrin beta-4a SEQ ID No: 1 *CTDEMFRDR***RCNTQAELLAAGCQR***ESIVVMESSFQITEETQIDTTLRR***SQ**
Integrin beta-4b SEQ ID No: 2 *CTDEMFRDR***RCNTQAELLAAGCQR***ESIVVMESSFQITEETQIDTTLRR***SQ** 100
Integrin beta-4c SEQ ID No: 3 *CTDEMFRDR***RCNTQAELLAAGCQR***ESIVVMESSFQITEETQIDTTLRR***SQ**
Integrin beta-4d SEQ ID No: 4 *CTDEMFRDR***RCNTQAELLAAGCQR***ESIVVMESSFQITEETQIDTTLRR***SQ**
Integrin beta-4e SEQ ID No: 5 *CTDEMFRDR***RCNTQAELLAAGCQR***ESIVVMESSFQITEETQIDTTLRR***SQ**
****

Integrin beta-4a SEQ ID No: 1 **MSPQGLR***VRLRPGEERHFELEVFEPLESPVDLYILMD*FSNSMSDDLDNLK
Integrin beta-4b SEQ ID No: 2 **MSPQGLR***VRLRPGEERHFELEVFEPLESPVDLYILMD*FSNSMSDDLDNLK 150
Integrin beta-4c SEQ ID No: 3 **MSPQGLR***VRLRPGEERHFELEVFEPLESPVDLYILMD*FSNSMSDDLDNLK
Integrin beta-4d SEQ ID No: 4 **MSPQGLR***VRLRPGEERHFELEVFEPLESPVDLYILMD*FSNSMSDDLDNLK
Integrin beta-4e SEQ ID No: 5 **MSPQGLR***VRLRPGEERHFELEVFEPLESPVDLYILMD*FSNSMSDDLDNLK
****

Integrin beta-4a SEQ ID No: 1 K**MGQNLAR**VLSQLTSDYTIGFGKFVDK**VSVPQTDMRPEK**LKEPWPNSDPP
Integrin beta-4b SEQ ID No: 2 K**MGQNLAR**VLSQLTSDYTIGFGKFVDK**VSVPQTDMRPEK**LKEPWPNSDPP 200
Integrin beta-4c SEQ ID No: 3 K**MGQNLAR**VLSQLTSDYTIGFGKFVDK**VSVPQTDMRPEK**LKEPWPNSDPP
Integrin beta-4d SEQ ID No: 4 K**MGQNLAR**VLSQLTSDYTIGFGKFVDK**VSVPQTDMRPEK**LKEPWPNSDPP
Integrin beta-4e SEQ ID No: 5 K**MGQNLAR**VLSQLTSDYTIGFGKFVDK**VSVPQTDMRPEK**LKEPWPNSDPP
****

Integrin beta-4a SEQ ID No: 1 FSFK**NVISLTEDVDEFR**NKLQGER**ISGNLDAPEGGFDAILQTAVCTR**DIG
Integrin beta-4b SEQ ID No: 2 FSFK**NVISLTEDVDEFR**NKLQGER**ISGNLDAPEGGFDAILQTAVCTR**DIG 250

```
Integrin beta-4c       FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 3
Integrin beta-4d       FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 4
Integrin beta-4e       FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 5
                       **************************************************

Integrin beta-4a       WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 1
Integrin beta-4b       WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR  300
SEQ ID No: 2
Integrin beta-4c       WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 3
Integrin beta-4d       WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 4
Integrin beta-4e       WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 5
                       **************************************************

Integrin beta-4a       TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 1
Integrin beta-4b       TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE  350
SEQ ID No: 2
Integrin beta-4c       TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 3
Integrin beta-4d       TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 4
Integrin beta-4e       TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 5
                       **************************************************

Integrin beta-4a       DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 1
Integrin beta-4b       DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI  400
SEQ ID No: 2
Integrin beta-4c       DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 3
Integrin beta-4d       DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 4
Integrin beta-4e       DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 5
                       **************************************************

Integrin beta-4a       RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 1
Integrin beta-4b       RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI  450
SEQ ID No: 2
Integrin beta-4c       RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 3
Integrin beta-4d       RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 4
Integrin beta-4e       RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
```

```
SEQ ID No: 5
                          ************************************************

Integrin beta-4a          ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 1
Integrin beta-4b          ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI   500
SEQ ID No: 2
Integrin beta-4c          ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 3
Integrin beta-4d          ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 4
Integrin beta-4e          ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 5
                          **************************************************

Integrin beta-4a          QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 1
Integrin beta-4b          QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL   550
SEQ ID No: 2
Integrin beta-4c          QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 3
Integrin beta-4d          QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 4
Integrin beta-4e          QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 5
                          **************************************************

Integrin beta-4a          CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 1
Integrin beta-4b          CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR   600
SEQ ID No: 2
Integrin beta-4c          CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 3
Integrin beta-4d          CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 4
Integrin beta-4e          CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 5
                          **************************************************

Integrin beta-4a          CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 1
Integrin beta-4b          CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE   650
SEQ ID No: 2
Integrin beta-4c          CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 3
Integrin beta-4d          CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 4
Integrin beta-4e          CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 5
                          **************************************************

Integrin beta-4a          CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 1
```

```
Integrin beta-4b      CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV  700
SEQ ID No: 2
Integrin beta-4c      CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 3
Integrin beta-4d      CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 4
Integrin beta-4e      CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 1
Integrin beta-4b      HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN  750
SEQ ID No: 2
Integrin beta-4c      HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 3
Integrin beta-4d      HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 4
Integrin beta-4e      HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 1
Integrin beta-4b      RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR  800
SEQ ID No: 2
Integrin beta-4c      RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 3
Integrin beta-4d      RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 4
Integrin beta-4e      RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 1
Integrin beta-4b      PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN  850
SEQ ID No: 2
Integrin beta-4c      PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 3
Integrin beta-4d      PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 4
Integrin beta-4e      PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 1
Integrin beta-4b      LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK  900
SEQ ID No: 2
Integrin beta-4c      LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 3
Integrin beta-4d      LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
```

```
SEQ ID No: 4
Integrin beta-4e       VRTQELGLAGDVAERGLQADLRCTQAPADQVPAAAQCREKARPHHCGHSA
SEQ ID No: 5
                       :.      ::*     :  :   :  . .  :: . .       .*. : .

Integrin beta-4a       LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 1
Integrin beta-4b       LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI   950
SEQ ID No: 2
Integrin beta-4c       LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 3
Integrin beta-4d       LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 4
Integrin beta-4e       DGAPLGQAGPAEAYREAGGTEGLPRPQGGPRLLHPHCRPGRPGHGGVPGG
SEQ ID No: 5
                           :  . .  : * *   *.  *..  ::. :        .

Integrin beta-4a       RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 1
Integrin beta-4b       RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS   1000
SEQ ID No: 2
Integrin beta-4c       RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 3
Integrin beta-4d       RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 4
Integrin beta-4e       RGAGGRTGAPLYPA------------------------------------
SEQ ID No: 5
                       *  ..      *  *

Integrin beta-4a       VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 1
Integrin beta-4b       VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP   1050
SEQ ID No: 2
Integrin beta-4c       VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 3
Integrin beta-4d       VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 1
Integrin beta-4b       GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI   1100
SEQ ID No: 2
Integrin beta-4c       GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 3
Integrin beta-4d       GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5
```

```
Integrin beta-4a    IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 1
Integrin beta-4b    IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP  1150
SEQ ID No: 2
Integrin beta-4c    IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 3
Integrin beta-4d    IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 1
Integrin beta-4b    SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG  1200
SEQ ID No: 2
Integrin beta-4c    SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 3
Integrin beta-4d    SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 1
Integrin beta-4b    AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI  1250
SEQ ID No: 2
Integrin beta-4c    AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 3
Integrin beta-4d    AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 1
Integrin beta-4b    TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA  1300
SEQ ID No: 2
Integrin beta-4c    TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 3
Integrin beta-4d    TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 1
Integrin beta-4b    RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS  1350
SEQ ID No: 2
Integrin beta-4c    RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
```

```
SEQ ID No: 3
Integrin beta-4d     RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     DDVLRSPSGSQRPSVSDDTGCGWKFEPLLGEELDLRRVTWRLPPELIPRL
SEQ ID No: 1
Integrin beta-4b     DDVLRSPSGSQRPSVSDDT-------------------------------  1369
SEQ ID No: 2
Integrin beta-4c     DDVLRSPSGSQRPSVSDDT-------------------------------
SEQ ID No: 3
Integrin beta-4d     DDVLRSPSGSQRPSVSDDT-------------------------------
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     SASSGRSSDAEAPHGPPDDGGAGGKGGSLPRSATPGPPGEHLVNGRMDFA
SEQ ID No: 1
Integrin beta-4b     ---------------------------------------EHLVNGRMDFA  1380
SEQ ID No: 2
Integrin beta-4c     ---------------------------------------EHLVNGRMDFA
SEQ ID No: 3
Integrin beta-4d     ---------------------------------------EHLVNGRMDFA
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 1
Integrin beta-4b     FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH  1430
SEQ ID No: 2
Integrin beta-4c     FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 3
Integrin beta-4d     FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     SHSTTLPRDYSTLTSVSSH-------------------------------
SEQ ID No: 1
Integrin beta-4b     SHSTTLPRDYSTLTSVSSH-------------------------------  1449
SEQ ID No: 2
Integrin beta-4c     SHSTTLPRDYSTLTSVSSHGLPPIWEHGRSRLPLSWALGSRSRAQMKGFP
SEQ ID No: 3
Integrin beta-4d     SHSTTLPRDYSTLTSVSSH-------------------------------
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5
```

```
Integrin beta-4a    ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 1
Integrin beta-4b    ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS   1477
SEQ ID No: 2
Integrin beta-4c    PSRGPRDSIILAGRPAAPSWGPDSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 3
Integrin beta-4d    ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 1
Integrin beta-4b    WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV   1527
SEQ ID No: 2
Integrin beta-4c    WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 3
Integrin beta-4d    WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 1
Integrin beta-4b    FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL   1577
SEQ ID No: 2
Integrin beta-4c    FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 3
Integrin beta-4d    FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP
SEQ ID No: 1
Integrin beta-4b    VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP   1627
SEQ ID No: 2
Integrin beta-4c    VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP
SEQ ID No: 3
Integrin beta-4d    VFTALSPDSLQLSWERPRRPNGDIVGYLVTW-------PATAFRVDGDSP
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 1
Integrin beta-4b    ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS   1677
```

```
SEQ ID No: 2
Integrin beta-4c      ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 3
Integrin beta-4d      ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5


Integrin beta-4a      RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 1
Integrin beta-4b      RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT  1727
SEQ ID No: 2
Integrin beta-4c      RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 3
Integrin beta-4d      RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5


Integrin beta-4a      QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 1
Integrin beta-4b      QEFVSRTLTTSGTLSTHMDQQFFQT                           1752
SEQ ID No: 2
Integrin beta-4c      QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 3
Integrin beta-4d      QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 4
Integrin beta-4e      -------------------------
SEQ ID No: 5
```

Key:
Recombinant protein
Key Tandem peptides(underline)
Mass Match peptides(bold)
Calx-beta and FNIII domains
Important tyrosine residues(double underline)

Gastric cancer

```
Integrin beta-4a     MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 1
Integrin beta-4b     MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY  50
SEQ ID No: 2
Integrin beta-4c     MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 3
Integrin beta-4d     MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 4
Integrin beta-4e     MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 5
                     **************************************************

Integrin beta-4a     CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 1
Integrin beta-4b     CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ  100
SEQ ID No: 2
Integrin beta-4c     CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 3
Integrin beta-4d     CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 4
Integrin beta-4e     CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 5
                     **************************************************

Integrin beta-4a     MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 1
Integrin beta-4b     MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK  150
SEQ ID No: 2
Integrin beta-4c     MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 3
Integrin beta-4d     MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 4
Integrin beta-4e     MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 5
                     **************************************************

Integrin beta-4a     KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 1
Integrin beta-4b     KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP  200
SEQ ID No: 2
Integrin beta-4c     KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 3
Integrin beta-4d     KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 4
Integrin beta-4e     KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 5
                     **************************************************

Integrin beta-4a     FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 1
Integrin beta-4b     FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG  250
SEQ ID No: 2
Integrin beta-4c     FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 3
Integrin beta-4d     FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
```

```
SEQ ID No: 4
Integrin beta-4e        FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 1
Integrin beta-4b        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR  300
SEQ ID No: 2
Integrin beta-4c        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 3
Integrin beta-4d        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 4
Integrin beta-4e        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 1
Integrin beta-4b        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE  350
SEQ ID No: 2
Integrin beta-4c        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 3
Integrin beta-4d        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 4
Integrin beta-4e        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 1
Integrin beta-4b        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI  400
SEQ ID No: 2
Integrin beta-4c        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 3
Integrin beta-4d        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 4
Integrin beta-4e        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 1
Integrin beta-4b        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI  450
SEQ ID No: 2
Integrin beta-4c        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 3
Integrin beta-4d        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 4
Integrin beta-4e        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 1
Integrin beta-4b        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI  500
SEQ ID No: 2
```

```
Integrin beta-4c        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 3
Integrin beta-4d        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 4
Integrin beta-4e        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 1
Integrin beta-4b        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL  550
SEQ ID No: 2
Integrin beta-4c        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 3
Integrin beta-4d        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 4
Integrin beta-4e        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 1
Integrin beta-4b        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR  600
SEQ ID No: 2
Integrin beta-4c        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 3
Integrin beta-4d        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 4
Integrin beta-4e        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 1
Integrin beta-4b        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE  650
SEQ ID No: 2
Integrin beta-4c        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 3
Integrin beta-4d        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 4
Integrin beta-4e        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 1
Integrin beta-4b        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV  700
SEQ ID No: 2
Integrin beta-4c        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 3
Integrin beta-4d        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 4
Integrin beta-4e        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
```

```
SEQ ID No: 1
Integrin beta-4b         HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN  750
SEQ ID No: 2
Integrin beta-4c         HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 3
Integrin beta-4d         HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 4
Integrin beta-4e         HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 5
                         **************************************************

Integrin beta-4a         RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 1
Integrin beta-4b         RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR  800
SEQ ID No: 2
Integrin beta-4c         RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 3
Integrin beta-4d         RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 4
Integrin beta-4e         RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 5
                         **************************************************

Integrin beta-4a         PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 1
Integrin beta-4b         PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN  850
SEQ ID No: 2
Integrin beta-4c         PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 3
Integrin beta-4d         PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 4
Integrin beta-4e         PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 5
                         **************************************************

Integrin beta-4a         LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 1
Integrin beta-4b         LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK  900
SEQ ID No: 2
Integrin beta-4c         LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 3
Integrin beta-4d         LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 4
Integrin beta-4e         VRTQELGLAGDVAERGLQADLRCTQAPADQVPAAAQCREKARPHHCGHSA
SEQ ID No: 5
                         :.     ::*    :  :   : . .  :: . .      .*. : .

Integrin beta-4a         LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 1
Integrin beta-4b         LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI  950
SEQ ID No: 2
Integrin beta-4c         LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 3
Integrin beta-4d         LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 4
Integrin beta-4e         DGAPLGQAGPAEAYREAGGTEGLPRPQGGPRLLHPHCRPGRPGHGGVPGG
SEQ ID No: 5
```

```
                           :  . .  : * *   *.  *..  ::. :      .

Integrin beta-4a       RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 1
Integrin beta-4b       RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS  1000
SEQ ID No: 2
Integrin beta-4c       RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 3
Integrin beta-4d       RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 4
Integrin beta-4e       RGAGGRTGAPLYPA------------------------------------
SEQ ID No: 5
                       *  ..     *  *

Integrin beta-4a       VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 1
Integrin beta-4b       VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP  1050
SEQ ID No: 2
Integrin beta-4c       VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 3
Integrin beta-4d       VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 1
Integrin beta-4b       GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI  1100
SEQ ID No: 2
Integrin beta-4c       GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 3
Integrin beta-4d       GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 1
Integrin beta-4b       IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP  1150
SEQ ID No: 2
Integrin beta-4c       IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 3
Integrin beta-4d       IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 1
Integrin beta-4b       SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG  1200
SEQ ID No: 2
Integrin beta-4c       SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 3
Integrin beta-4d       SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
```

```
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a      AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 1
Integrin beta-4b      AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI  1250
SEQ ID No: 2
Integrin beta-4c      AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 3
Integrin beta-4d      AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a      TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 1
Integrin beta-4b      TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA  1300
SEQ ID No: 2
Integrin beta-4c      TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 3
Integrin beta-4d      TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a      RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 1
Integrin beta-4b      RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS  1350
SEQ ID No: 2
Integrin beta-4c      RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 3
Integrin beta-4d      RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a      DDVLRSPSGSQRPSVSDDTGCGWKFEPLLGEELDLRRVTWRLPPELIPRL
SEQ ID No: 1
Integrin beta-4b      DDVLRSPSGSQRPSVSDDT-------------------------------  1369
SEQ ID No: 2
Integrin beta-4c      DDVLRSPSGSQRPSVSDDT-------------------------------
SEQ ID No: 3
Integrin beta-4d      DDVLRSPSGSQRPSVSDDT-------------------------------
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a      SASSGRSSDAEAPHGPPDDGGAGGKGGSLPRSATPGPPGEHLVNGRMDFA
SEQ ID No: 1
Integrin beta-4b      ---------------------------------------EHLVNGRMDFA  1380
SEQ ID No: 2
```

```
Integrin beta-4c        -------------------------------------EHLVNGRMDFA
SEQ ID No: 3
Integrin beta-4d        -------------------------------------EHLVNGRMDFA
SEQ ID No: 4
Integrin beta-4e        ------------------------------------------------
SEQ ID No: 5

Integrin beta-4a        FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 1
Integrin beta-4b        FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH  1430
SEQ ID No: 2
Integrin beta-4c        FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 3
Integrin beta-4d        FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 4
Integrin beta-4e        ------------------------------------------------
SEQ ID No: 5

Integrin beta-4a        SHSTTLPRDYSTLTSVSSH------------------------------
SEQ ID No: 1
Integrin beta-4b        SHSTTLPRDYSTLTSVSSH------------------------------  1449
SEQ ID No: 2
Integrin beta-4c        SHSTTLPRDYSTLTSVSSHGLPPIWEHGRSRLPLSWALGSRSRAQMKGFP
SEQ ID No: 3
Integrin beta-4d        SHSTTLPRDYSTLTSVSSH------------------------------
SEQ ID No: 4
Integrin beta-4e        ------------------------------------------------
SEQ ID No: 5

Integrin beta-4a        ---------------------DSRLTACVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 1
Integrin beta-4b        ---------------------DSRLTACVPDTPTRLVFSALGPTSLRVS  1477
SEQ ID No: 2
Integrin beta-4c        PSRGPRDSIILAGRPAAPSWGPDSRLTACVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 3
Integrin beta-4d        ---------------------DSRLTACVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 4
Integrin beta-4e        ------------------------------------------------
SEQ ID No: 5

Integrin beta-4a        WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 1
Integrin beta-4b        WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV  1527
SEQ ID No: 2
Integrin beta-4c        WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 3
Integrin beta-4d        WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 4
Integrin beta-4e        ------------------------------------------------
SEQ ID No: 5

Integrin beta-4a        FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
```

```
SEQ ID No: 1
Integrin beta-4b    FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL  1577
SEQ ID No: 2
Integrin beta-4c    FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 3
Integrin beta-4d    FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP
SEQ ID No: 1
Integrin beta-4b    VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP  1627
SEQ ID No: 2
Integrin beta-4c    VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP
SEQ ID No: 3
Integrin beta-4d    VFTALSPDSLQLSWERPRRPNGDIVGYLVTW-------PATAFRVDGDSP
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 1
Integrin beta-4b    ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS  1677
SEQ ID No: 2
Integrin beta-4c    ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 3
Integrin beta-4d    ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 1
Integrin beta-4b    RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT  1727
SEQ ID No: 2
Integrin beta-4c    RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 3
Integrin beta-4d    RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 1
Integrin beta-4b    QEFVSRTLTTSGTLSTHMDQQFFQT                          1752
SEQ ID No: 2
Integrin beta-4c    QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 3
Integrin beta-4d    QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 4
Integrin beta-4e    -------------------------
SEQ ID No: 5
```

Key:
Recombinant protein
Key Tandem peptides (underline)
Mass Match peptides (bold)
Calx-beta and FNIII domains
Important tyrosine residues (double underline)

Hepatocellular carcinoma

```
Integrin beta-4a       MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 1
Integrin beta-4b       MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY  50
SEQ ID No: 2
Integrin beta-4c       MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 3
Integrin beta-4d       MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 4
Integrin beta-4e       MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 5
                       **************************************************

Integrin beta-4a       CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 1
Integrin beta-4b       CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ  100
SEQ ID No: 2
Integrin beta-4c       CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 3
Integrin beta-4d       CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 4
Integrin beta-4e       CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 5
                       **************************************************

Integrin beta-4a       MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 1
Integrin beta-4b       MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK  150
SEQ ID No: 2
Integrin beta-4c       MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 3
Integrin beta-4d       MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 4
Integrin beta-4e       MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 5
                       **************************************************

Integrin beta-4a       KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 1
Integrin beta-4b       KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP  200
SEQ ID No: 2
Integrin beta-4c       KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 3
Integrin beta-4d       KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 4
Integrin beta-4e       KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 5
                       **************************************************

Integrin beta-4a       FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 1
Integrin beta-4b       FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG  250
SEQ ID No: 2
Integrin beta-4c       FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 3
Integrin beta-4d       FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
```

```
SEQ ID No: 4
Integrin beta-4e        FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 1
Integrin beta-4b        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR  300
SEQ ID No: 2
Integrin beta-4c        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 3
Integrin beta-4d        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 4
Integrin beta-4e        WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 1
Integrin beta-4b        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE  350
SEQ ID No: 2
Integrin beta-4c        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 3
Integrin beta-4d        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 4
Integrin beta-4e        TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 1
Integrin beta-4b        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI  400
SEQ ID No: 2
Integrin beta-4c        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 3
Integrin beta-4d        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 4
Integrin beta-4e        DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 1
Integrin beta-4b        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI  450
SEQ ID No: 2
Integrin beta-4c        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 3
Integrin beta-4d        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 4
Integrin beta-4e        RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 1
Integrin beta-4b        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI  500
SEQ ID No: 2
```

```
Integrin beta-4c       ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 3
Integrin beta-4d       ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 4
Integrin beta-4e       ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 5
                       **************************************************

Integrin beta-4a       QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 1
Integrin beta-4b       QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL  550
SEQ ID No: 2
Integrin beta-4c       QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 3
Integrin beta-4d       QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 4
Integrin beta-4e       QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 5
                       **************************************************

Integrin beta-4a       CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 1
Integrin beta-4b       CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR  600
SEQ ID No: 2
Integrin beta-4c       CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 3
Integrin beta-4d       CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 4
Integrin beta-4e       CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 5
                       **************************************************

Integrin beta-4a       CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 1
Integrin beta-4b       CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE  650
SEQ ID No: 2
Integrin beta-4c       CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 3
Integrin beta-4d       CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 4
Integrin beta-4e       CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 5
                       **************************************************

Integrin beta-4a       CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 1
Integrin beta-4b       CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV  700
SEQ ID No: 2
Integrin beta-4c       CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 3
Integrin beta-4d       CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 4
Integrin beta-4e       CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 5
                       **************************************************

Integrin beta-4a       HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
```

SEQ ID No: 1
Integrin beta-4b HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN 750
SEQ ID No: 2
Integrin beta-4c HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 3
Integrin beta-4d HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 4
Integrin beta-4e HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 5
****

Integrin beta-4a RGHMVGFKEDHYMLR**ENLMASDHLDTPMLR**SGNLKGRDVVRWKVTNNMQR
SEQ ID No: 1
Integrin beta-4b RGHMVGFKEDHYMLR**ENLMASDHLDTPMLR**SGNLKGRDVVRWKVTNNMQR 800
SEQ ID No: 2
Integrin beta-4c RGHMVGFKEDHYMLR**ENLMASDHLDTPMLR**SGNLKGRDVVRWKVTNNMQR
SEQ ID No: 3
Integrin beta-4d RGHMVGFKEDHYMLR**ENLMASDHLDTPMLR**SGNLKGRDVVRWKVTNNMQR
SEQ ID No: 4
Integrin beta-4e RGHMVGFKEDHYMLR**ENLMASDHLDTPMLR**SGNLKGRDVVRWKVTNNMQR
SEQ ID No: 5
****

Integrin beta-4a PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLR**QEVEEN**
SEQ ID No: 1
Integrin beta-4b PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLR**QEVEEN** 850
SEQ ID No: 2
Integrin beta-4c PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLR**QEVEEN**
SEQ ID No: 3
Integrin beta-4d PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLR**QEVEEN**
SEQ ID No: 4
Integrin beta-4e PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 5
****

Integrin beta-4a **LNEVYRQISGVHK**LQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 1
Integrin beta-4b **LNEVYRQISGVHK**LQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK 900
SEQ ID No: 2
Integrin beta-4c **LNEVYRQISGVHK**LQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 3
Integrin beta-4d **LNEVYRQISGVHK**LQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 4
Integrin beta-4e VRTQELGLAGDVAERGLQADLRCTQAPADQVPAAAQCREKARPHHCGHSA
SEQ ID No: 5
:.   ::*   : :   : . .   :: . .   .*. : .

Integrin beta-4a LTEKQVEQRAFHDLK**VAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI**
SEQ ID No: 1
Integrin beta-4b LTEKQVEQRAFHDLK**VAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI** 950
SEQ ID No: 2
Integrin beta-4c LTEKQVEQRAFHDLK**VAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI**
SEQ ID No: 3
Integrin beta-4d LTEKQVEQRAFHDLK**VAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI**
SEQ ID No: 4
Integrin beta-4e DGAPLGQAGPAEAYREAGGTEGLPRPQGGPRLLHPHCRPGRPGHGGVPGG
SEQ ID No: 5

```
                              :  . .  : * *   *.  *..  ::. :      .

Integrin beta-4a   RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 1
Integrin beta-4b   RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS  1000
SEQ ID No: 2
Integrin beta-4c   RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 3
Integrin beta-4d   RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 4
Integrin beta-4e   RGAGGRTGAPLYPA------------------------------------
SEQ ID No: 5

                   *  ..      *  *

Integrin beta-4a   VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 1
Integrin beta-4b   VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP  1050
SEQ ID No: 2
Integrin beta-4c   VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 3
Integrin beta-4d   VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 4
Integrin beta-4e   --------------------------------------------------
SEQ ID No: 5


Integrin beta-4a   GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 1
Integrin beta-4b   GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI  1100
SEQ ID No: 2
Integrin beta-4c   GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 3
Integrin beta-4d   GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 4
Integrin beta-4e   --------------------------------------------------
SEQ ID No: 5


Integrin beta-4a   IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 1
Integrin beta-4b   IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP  1150
SEQ ID No: 2
Integrin beta-4c   IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 3
Integrin beta-4d   IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 4
Integrin beta-4e   --------------------------------------------------
SEQ ID No: 5


Integrin beta-4a   SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 1
Integrin beta-4b   SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG  1200
SEQ ID No: 2
Integrin beta-4c   SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 3
Integrin beta-4d   SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
```

```
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a      AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 1
Integrin beta-4b      AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI  1250
SEQ ID No: 2
Integrin beta-4c      AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 3
Integrin beta-4d      AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a      TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 1
Integrin beta-4b      TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA  1300
SEQ ID No: 2
Integrin beta-4c      TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 3
Integrin beta-4d      TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a      RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 1
Integrin beta-4b      RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS  1350
SEQ ID No: 2
Integrin beta-4c      RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 3
Integrin beta-4d      RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a      DDVLRSPSGSQRPSVSDDTGCGWKFEPLLGEELDLRRVTWRLPPELIPRL
SEQ ID No: 1
Integrin beta-4b      DDVLRSPSGSQRPSVSDDT-------------------------------  1369
SEQ ID No: 2
Integrin beta-4c      DDVLRSPSGSQRPSVSDDT-------------------------------
SEQ ID No: 3
Integrin beta-4d      DDVLRSPSGSQRPSVSDDT-------------------------------
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a      SASSGRSSDAEAPHGPPDDGGAGGKGGSLPRSATPGPPGEHLVNGRMDFA
SEQ ID No: 1
Integrin beta-4b      ---------------------------------------EHLVNGRMDFA  1380
SEQ ID No: 2
```

```
Integrin beta-4c    ---------------------------------------EHLVNGRMDFA
SEQ ID No: 3
Integrin beta-4d    ---------------------------------------EHLVNGRMDFA
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 1
Integrin beta-4b    FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH  1430
SEQ ID No: 2
Integrin beta-4c    FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 3
Integrin beta-4d    FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    SHSTTLPRDYSTLTSVSSH-------------------------------
SEQ ID No: 1
Integrin beta-4b    SHSTTLPRDYSTLTSVSSH-------------------------------  1449
SEQ ID No: 2
Integrin beta-4c    SHSTTLPRDYSTLTSVSSHGLPPIWEHGRSRLPLSWALGSRSRAQMKGFP
SEQ ID No: 3
Integrin beta-4d    SHSTTLPRDYSTLTSVSSH-------------------------------
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 1
Integrin beta-4b    ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS  1477
SEQ ID No: 2
Integrin beta-4c    PSRGPRDSIILAGRPAAPSWGPDSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 3
Integrin beta-4d    ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 1
Integrin beta-4b    WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV  1527
SEQ ID No: 2
Integrin beta-4c    WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 3
Integrin beta-4d    WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
```

```
SEQ ID No: 1
Integrin beta-4b    FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL  1577
SEQ ID No: 2
Integrin beta-4c    FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 3
Integrin beta-4d    FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP
SEQ ID No: 1
Integrin beta-4b    VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP  1627
SEQ ID No: 2
Integrin beta-4c    VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP
SEQ ID No: 3
Integrin beta-4d    VFTALSPDSLQLSWERPRRPNGDIVGYLVTW-------PATAFRVDGDSP
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 1
Integrin beta-4b    ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS  1677
SEQ ID No: 2
Integrin beta-4c    ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 3
Integrin beta-4d    ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 1
Integrin beta-4b    RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT  1727
SEQ ID No: 2
Integrin beta-4c    RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 3
Integrin beta-4d    RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 4
Integrin beta-4e    --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a    QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 1
Integrin beta-4b    QEFVSRTLTTSGTLSTHMDQQFFQT                           1752
SEQ ID No: 2
Integrin beta-4c    QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 3
Integrin beta-4d    QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 4
Integrin beta-4e    -------------------------
SEQ ID No: 5
```

Key:
Recombinant protein
Key Tandem peptides (underline)
Mass Match peptides (bold)
Calx-beta and FNIII domains
Important tyrosine residues (double underline)

Lung cancer

```
Integrin beta-4a     MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 1
Integrin beta-4b     MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY  50
SEQ ID No: 2
Integrin beta-4c     MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 3
Integrin beta-4d     MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 4
Integrin beta-4e     MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 5
                     **************************************************

Integrin beta-4a     CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 1
Integrin beta-4b     CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ  100
SEQ ID No: 2
Integrin beta-4c     CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 3
Integrin beta-4d     CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 4
Integrin beta-4e     CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 5
                     **************************************************

Integrin beta-4a     MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 1
Integrin beta-4b     MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK  150
SEQ ID No: 2
Integrin beta-4c     MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 3
Integrin beta-4d     MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 4
Integrin beta-4e     MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 5
                     **************************************************

Integrin beta-4a     KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 1
Integrin beta-4b     KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP  200
SEQ ID No: 2
Integrin beta-4c     KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 3
Integrin beta-4d     KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 4
Integrin beta-4e     KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 5
                     **************************************************

Integrin beta-4a     FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 1
Integrin beta-4b     FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG  250
SEQ ID No: 2
Integrin beta-4c     FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 3
Integrin beta-4d     FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
```

```
SEQ ID No: 4
Integrin beta-4e      FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 1
Integrin beta-4b      WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR  300
SEQ ID No: 2
Integrin beta-4c      WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 3
Integrin beta-4d      WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 4
Integrin beta-4e      WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 1
Integrin beta-4b      TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE  350
SEQ ID No: 2
Integrin beta-4c      TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 3
Integrin beta-4d      TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 4
Integrin beta-4e      TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 1
Integrin beta-4b      DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI  400
SEQ ID No: 2
Integrin beta-4c      DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 3
Integrin beta-4d      DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 4
Integrin beta-4e      DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 1
Integrin beta-4b      RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI  450
SEQ ID No: 2
Integrin beta-4c      RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 3
Integrin beta-4d      RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 4
Integrin beta-4e      RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 1
Integrin beta-4b      ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI  500
```

```
SEQ ID No: 2
Integrin beta-4c        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 3
Integrin beta-4d        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 4
Integrin beta-4e        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 1
Integrin beta-4b        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL  550
SEQ ID No: 2
Integrin beta-4c        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 3
Integrin beta-4d        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 4
Integrin beta-4e        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 1
Integrin beta-4b        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR  600
SEQ ID No: 2
Integrin beta-4c        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 3
Integrin beta-4d        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 4
Integrin beta-4e        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 1
Integrin beta-4b        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE  650
SEQ ID No: 2
Integrin beta-4c        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 3
Integrin beta-4d        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 4
Integrin beta-4e        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 1
Integrin beta-4b        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV  700
SEQ ID No: 2
Integrin beta-4c        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 3
Integrin beta-4d        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 4
Integrin beta-4e        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 5
                        **************************************************
```

```
Integrin beta-4a      HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 1
Integrin beta-4b      HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN  750
SEQ ID No: 2
Integrin beta-4c      HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 3
Integrin beta-4d      HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 4
Integrin beta-4e      HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 1
Integrin beta-4b      RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR  800
SEQ ID No: 2
Integrin beta-4c      RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 3
Integrin beta-4d      RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 4
Integrin beta-4e      RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 1
Integrin beta-4b      PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN  850
SEQ ID No: 2
Integrin beta-4c      PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 3
Integrin beta-4d      PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 4
Integrin beta-4e      PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 1
Integrin beta-4b      LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK  900
SEQ ID No: 2
Integrin beta-4c      LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 3
Integrin beta-4d      LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 4
Integrin beta-4e      VRTQELGLAGDVAERGLQADLRCTQAPADQVPAAAQCREKARPHHCGHSA
SEQ ID No: 5
                      :.     ::*    :  :   : . .  :: . .      .*. : .

Integrin beta-4a      LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 1
Integrin beta-4b      LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI  950
SEQ ID No: 2
Integrin beta-4c      LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 3
Integrin beta-4d      LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 4
```

```
Integrin beta-4e      DGAPLGQAGPAEAYREAGGTEGLPRPQGGPRLLHPHCRPGRPGHGGVPGG
SEQ ID No: 5
                              :  .  .   : * *    *.   *..   ::. :        .

Integrin beta-4a      RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 1
Integrin beta-4b      RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS  1000
SEQ ID No: 2
Integrin beta-4c      RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 3
Integrin beta-4d      RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 4
Integrin beta-4e      RGAGGRTGAPLYPA------------------------------------
SEQ ID No: 5
                      *   ..      *   *

Integrin beta-4a      VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 1
Integrin beta-4b      VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP  1050
SEQ ID No: 2
Integrin beta-4c      VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 3
Integrin beta-4d      VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a      GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 1
Integrin beta-4b      GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI  1100
SEQ ID No: 2
Integrin beta-4c      GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 3
Integrin beta-4d      GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a      IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 1
Integrin beta-4b      IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP  1150
SEQ ID No: 2
Integrin beta-4c      IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 3
Integrin beta-4d      IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 4
Integrin beta-4e      --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a      SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 1
Integrin beta-4b      SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG  1200
SEQ ID No: 2
```

```
Integrin beta-4c       SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 3
Integrin beta-4d       SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 1
Integrin beta-4b       AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI   1250
SEQ ID No: 2
Integrin beta-4c       AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 3
Integrin beta-4d       AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 1
Integrin beta-4b       TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA   1300
SEQ ID No: 2
Integrin beta-4c       TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 3
Integrin beta-4d       TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 1
Integrin beta-4b       RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS   1350
SEQ ID No: 2
Integrin beta-4c       RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 3
Integrin beta-4d       RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       DDVLRSPSGSQRPSVSDDTGCGWKFEPLLGEELDLRRVTWRLPPELIPRL
SEQ ID No: 1
Integrin beta-4b       DDVLRSPSGSQRPSVSDDT-------------------------------   1369
SEQ ID No: 2
Integrin beta-4c       DDVLRSPSGSQRPSVSDDT-------------------------------
SEQ ID No: 3
Integrin beta-4d       DDVLRSPSGSQRPSVSDDT-------------------------------
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       SASSGRSSDAEAPHGPPDDGGAGGKGGSLPRSATPGPPGEHLVNGRMDFA
```

```
SEQ ID No: 1
Integrin beta-4b        ----------------------------------------EHLVNGRMDFA  1380
SEQ ID No: 2
Integrin beta-4c        ----------------------------------------EHLVNGRMDFA
SEQ ID No: 3
Integrin beta-4d        ----------------------------------------EHLVNGRMDFA
SEQ ID No: 4
Integrin beta-4e        ---------------------------------------------------
SEQ ID No: 5

Integrin beta-4a        FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 1
Integrin beta-4b        FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH  1430
SEQ ID No: 2
Integrin beta-4c        FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 3
Integrin beta-4d        FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 4
Integrin beta-4e        --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a        SHSTTLPRDYSTLTSVSSH-------------------------------
SEQ ID No: 1
Integrin beta-4b        SHSTTLPRDYSTLTSVSSH-------------------------------  1449
SEQ ID No: 2
Integrin beta-4c        SHSTTLPRDYSTLTSVSSHGLPPIWEHGRSRLPLSWALGSRSRAQMKGFP
SEQ ID No: 3
Integrin beta-4d        SHSTTLPRDYSTLTSVSSH-------------------------------
SEQ ID No: 4
Integrin beta-4e        --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a        ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 1
Integrin beta-4b        ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS  1477
SEQ ID No: 2
Integrin beta-4c        PSRGPRDSIILAGRPAAPSWGPDSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 3
Integrin beta-4d        ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 4
Integrin beta-4e        --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a        WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 1
Integrin beta-4b        WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV  1527
SEQ ID No: 2
Integrin beta-4c        WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 3
Integrin beta-4d        WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 4
Integrin beta-4e        --------------------------------------------------
SEQ ID No: 5
```

```
Integrin beta-4a     FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 1
Integrin beta-4b     FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL  1577
SEQ ID No: 2
Integrin beta-4c     FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 3
Integrin beta-4d     FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP
SEQ ID No: 1
Integrin beta-4b     VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP  1627
SEQ ID No: 2
Integrin beta-4c     VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP
SEQ ID No: 3
Integrin beta-4d     VFTALSPDSLQLSWERPRRPNGDIVGYLVTW-------PATAFRVDGDSP
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 1
Integrin beta-4b     ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS  1677
SEQ ID No: 2
Integrin beta-4c     ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 3
Integrin beta-4d     ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 1
Integrin beta-4b     RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT  1727
SEQ ID No: 2
Integrin beta-4c     RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 3
Integrin beta-4d     RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 1
Integrin beta-4b     QEFVSRTLTTSGTLSTHMDQQFFQT                           1752
SEQ ID No: 2
Integrin beta-4c     QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 3
Integrin beta-4d     QEFVSRTLTTSGTLSTHMDQQFFQT
```

SEQ ID No: 4
Integrin beta-4e -------------------------
SEQ ID No: 5

Key:
Recombinant protein
Key Tandem peptides (underline)
Mass Match peptides (bold)
Calx-beta and FNIII domains
Important tyrosine residues (double underline)

Pancreatic cancer

```
Integrin beta-4a    MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 1
Integrin beta-4b    MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY   50
SEQ ID No: 2
Integrin beta-4c    MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 3
Integrin beta-4d    MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 4
Integrin beta-4e    MAGPRPSPWARLLLAALISVSLSGTLANRCKKAPVKSCTECVRVDKDCAY
SEQ ID No: 5
                    **************************************************

Integrin beta-4a    CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 1
Integrin beta-4b    CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ   100
SEQ ID No: 2
Integrin beta-4c    CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 3
Integrin beta-4d    CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 4
Integrin beta-4e    CTDEMFRDRRCNTQAELLAAGCQRESIVVMESSFQITEETQIDTTLRRSQ
SEQ ID No: 5
                    **************************************************

Integrin beta-4a    MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 1
Integrin beta-4b    MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK   150
SEQ ID No: 2
Integrin beta-4c    MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 3
Integrin beta-4d    MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 4
Integrin beta-4e    MSPQGLRVRLRPGEERHFELEVFEPLESPVDLYILMDFSNSMSDDLDNLK
SEQ ID No: 5
                    **************************************************

Integrin beta-4a    KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 1
Integrin beta-4b    KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP   200
SEQ ID No: 2
Integrin beta-4c    KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 3
Integrin beta-4d    KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 4
Integrin beta-4e    KMGQNLARVLSQLTSDYTIGFGKFVDKVSVPQTDMRPEKLKEPWPNSDPP
SEQ ID No: 5
                    **************************************************

Integrin beta-4a    FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 1
Integrin beta-4b    FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG   250
SEQ ID No: 2
Integrin beta-4c    FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 3
Integrin beta-4d    FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
```

```
SEQ ID No: 4
Integrin beta-4e      FSFKNVISLTEDVDEFRNKLQGERISGNLDAPEGGFDAILQTAVCTRDIG
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 1
Integrin beta-4b      WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR  300
SEQ ID No: 2
Integrin beta-4c      WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 3
Integrin beta-4d      WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 4
Integrin beta-4e      WRPDSTHLLVFSTESAFHYEADGANVLAGIMSRNDERCHLDTTGTYTQYR
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 1
Integrin beta-4b      TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE  350
SEQ ID No: 2
Integrin beta-4c      TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 3
Integrin beta-4d      TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 4
Integrin beta-4e      TQDYPSVPTLVRLLAKHNIIPIFAVTNYSYSYYEKLHTYFPVSSLGVLQE
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 1
Integrin beta-4b      DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI  400
SEQ ID No: 2
Integrin beta-4c      DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 3
Integrin beta-4d      DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 4
Integrin beta-4e      DSSNIVELLEEAFNRIRSNLDIRALDSPRGLRTEVTSKMFQKTRTGSFHI
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 1
Integrin beta-4b      RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI  450
SEQ ID No: 2
Integrin beta-4c      RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 3
Integrin beta-4d      RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 4
Integrin beta-4e      RRGEVGIYQVQLRALEHVDGTHVCQLPEDQKGNIHLKPSFSDGLKMDAGI
SEQ ID No: 5
                      **************************************************

Integrin beta-4a      ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 1
Integrin beta-4b      ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI  500
SEQ ID No: 2
```

```
Integrin beta-4c        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 3
Integrin beta-4d        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 4
Integrin beta-4e        ICDVCTCELQKEVRSARCSFNGDFVCGQCVCSEGWSGQTCNCSTGSLSDI
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 1
Integrin beta-4b        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL  550
SEQ ID No: 2
Integrin beta-4c        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 3
Integrin beta-4d        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 4
Integrin beta-4e        QPCLREGEDKPCSGRGECQCGHCVCYGEGRYEGQFCEYDNFQCPRTSGFL
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 1
Integrin beta-4b        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR  600
SEQ ID No: 2
Integrin beta-4c        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 3
Integrin beta-4d        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 4
Integrin beta-4e        CNDRGRCSMGQCVCEPGWTGPSCDCPLSNATCIDSNGGICNGRGHCECGR
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 1
Integrin beta-4b        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE  650
SEQ ID No: 2
Integrin beta-4c        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 3
Integrin beta-4d        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 4
Integrin beta-4e        CHCHQQSLYTDTICEINYSAIHPGLCEDLRSCVQCQAWGTGEKKGRTCEE
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 1
Integrin beta-4b        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV  700
SEQ ID No: 2
Integrin beta-4c        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 3
Integrin beta-4d        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 4
Integrin beta-4e        CNFKVKMVDELKRAEEVVVRCSFRDEDDDCTYSYTMEGDGAPGPNSTVLV
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
```

```
SEQ ID No: 1
Integrin beta-4b        HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN  750
SEQ ID No: 2
Integrin beta-4c        HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 3
Integrin beta-4d        HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 4
Integrin beta-4e        HKKKDCPPGSFWWLIPLLLLLLPLLALLLLLCWKYCACCKACLALLPCCN
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 1
Integrin beta-4b        RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR  800
SEQ ID No: 2
Integrin beta-4c        RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 3
Integrin beta-4d        RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 4
Integrin beta-4e        RGHMVGFKEDHYMLRENLMASDHLDTPMLRSGNLKGRDVVRWKVTNNMQR
SEQ ID No: 5
                        **************************************************

Integrin beta-4a        PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 1
Integrin beta-4b        PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN  850
SEQ ID No: 2
Integrin beta-4c        PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 3
Integrin beta-4d        PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 4
Integrin beta-4e        PGFATHAASINPTELVPYGLSLRLARLCTENLLKPDTRECAQLRQEVEEN
SEQ ID No: 5
 *                      *************************************************

Integrin beta-4a        LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 1
Integrin beta-4b        LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK  900
SEQ ID No: 2
Integrin beta-4c        LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 3
Integrin beta-4d        LNEVYRQISGVHKLQQTKFRQQPNAGKKQDHTIVDTVLMAPRSAKPALLK
SEQ ID No: 4
Integrin beta-4e        VRTQELGLAGDVAERGLQADLRCTQAPADQVPAAAQCREKARPHHCGHSA
SEQ ID No: 5
                        :.     ::*    :  :   :  . .  :: . .      .*. : .

Integrin beta-4a        LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 1
Integrin beta-4b        LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI  950
SEQ ID No: 2
Integrin beta-4c        LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 3
Integrin beta-4d        LTEKQVEQRAFHDLKVAPGYYTLTADQDARGMVEFQEGVELVDVRVPLFI
SEQ ID No: 4
Integrin beta-4e        DGAPLGQAGPAEAYREAGGTEGLPRPQGGPRLLHPHCRPGRPGHGGVPGG
SEQ ID No: 5
```

```
                                 :  . .  : * *   *.  *..  ::. :       .

Integrin beta-4a       RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 1
Integrin beta-4b       RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS  1000
SEQ ID No: 2
Integrin beta-4c       RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 3
Integrin beta-4d       RPEDDDEKQLLVEAIDVPAGTATLGRRLVNITIIKEQARDVVSFEQPEFS
SEQ ID No: 4
Integrin beta-4e       RGAGGRTGAPLYPA------------------------------------
SEQ ID No: 5
                       *  ..     *  *

Integrin beta-4a       VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 1
Integrin beta-4b       VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP  1050
SEQ ID No: 2
Integrin beta-4c       VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 3
Integrin beta-4d       VSRGDQVARIPVIRRVLDGGKSQVSYRTQDGTAQGNRDYIPVEGELLFQP
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 1
Integrin beta-4b       GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI  1100
SEQ ID No: 2
Integrin beta-4c       GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 3
Integrin beta-4d       GEAWKELQVKLLELQEVDSLLRGRQVRRFHVQLSNPKFGAHLGQPHSTTI
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 1
Integrin beta-4b       IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP  1150
SEQ ID No: 2
Integrin beta-4c       IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 3
Integrin beta-4d       IIRDPDELDRSFTSQMLSSQPPPHGDLGAPQNPNAKAAGSRKIHFNWLPP
SEQ ID No: 4
Integrin beta-4e       --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a       SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 1
Integrin beta-4b       SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG  1200
SEQ ID No: 2
Integrin beta-4c       SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
SEQ ID No: 3
Integrin beta-4d       SGKPMGYRVKYWIQGDSESEAHLLDSKVPSVELTNLYPYCDYEMKVCAYG
```

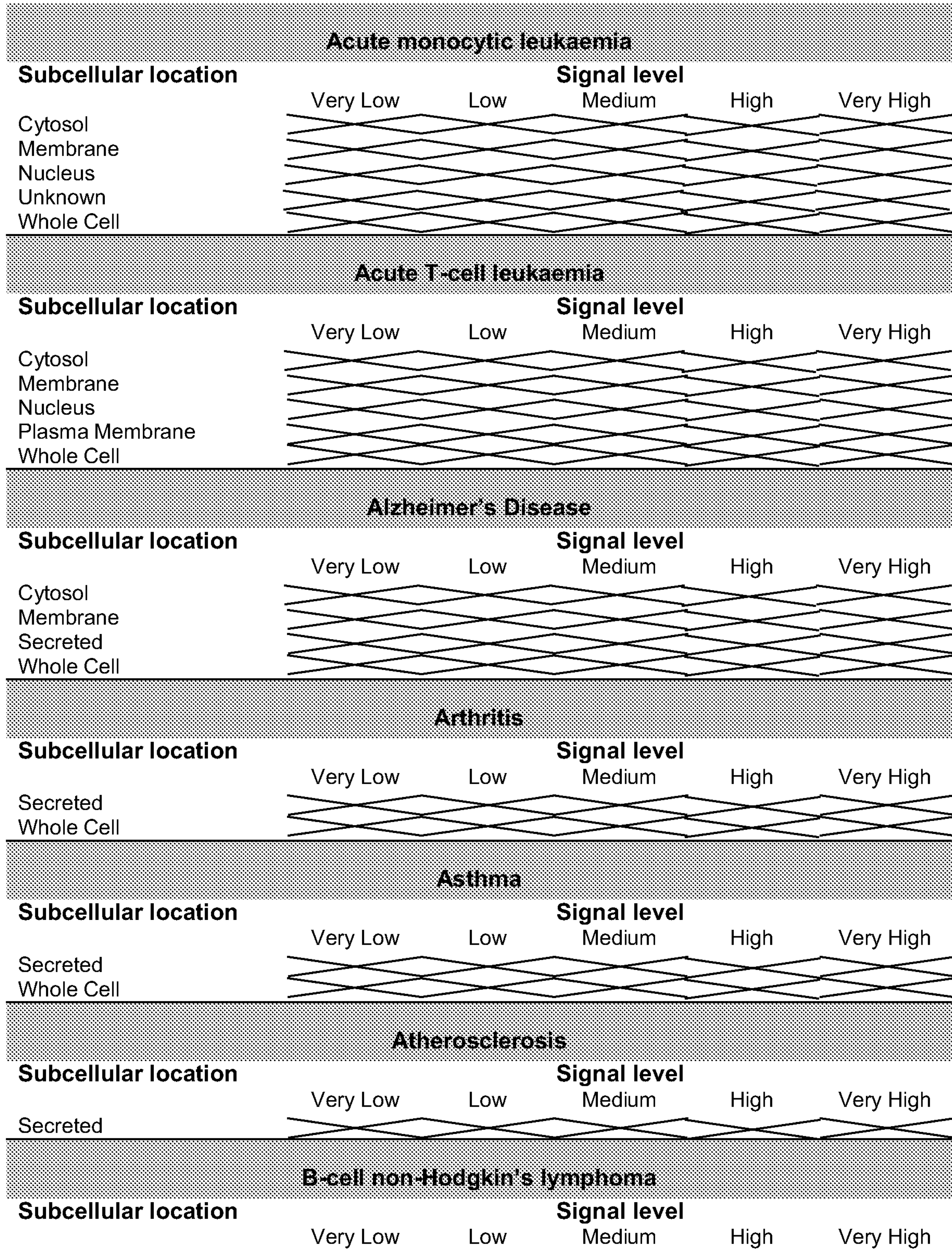
```
SEQ ID No: 4
Integrin beta-4e         -----------------------------------------------------
SEQ ID No: 5

Integrin beta-4a         AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 1
Integrin beta-4b         AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI   1250
SEQ ID No: 2
Integrin beta-4c         AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 3
Integrin beta-4d         AQGEGPYSSLVSCRTHQEVPSEPGRLAFNVVSSTVTQLSWAEPAETNGEI
SEQ ID No: 4
Integrin beta-4e         -----------------------------------------------------
SEQ ID No: 5

Integrin beta-4a         TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 1
Integrin beta-4b         TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA   1300
SEQ ID No: 2
Integrin beta-4c         TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 3
Integrin beta-4d         TAYEVCYGLVNDDNRPIGPMKKVLVDNPKNRMLLIENLRESQPYRYTVKA
SEQ ID No: 4
Integrin beta-4e         -----------------------------------------------------
SEQ ID No: 5

Integrin beta-4a         RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 1
Integrin beta-4b         RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS   1350
SEQ ID No: 2
Integrin beta-4c         RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 3
Integrin beta-4d         RNGAGWGPEREAIINLATQPKRPMSIPIIPDIPIVDAQSGEDYDSFLMYS
SEQ ID No: 4
Integrin beta-4e         -----------------------------------------------------
SEQ ID No: 5

Integrin beta-4a         DDVLRSPSGSQRPSVSDDTGCGWKFEPLLGEELDLRRVTWRLPPELIPRL
SEQ ID No: 1
Integrin beta-4b         DDVLRSPSGSQRPSVSDDT-------------------------------   1369
SEQ ID No: 2
Integrin beta-4c         DDVLRSPSGSQRPSVSDDT-------------------------------
SEQ ID No: 3
Integrin beta-4d         DDVLRSPSGSQRPSVSDDT-------------------------------
SEQ ID No: 4
Integrin beta-4e         -----------------------------------------------------
SEQ ID No: 5

Integrin beta-4a         SASSGRSSDAEAPHGPPDDGGAGGKGGSLPRSATPGPPGEHLVNGRMDFA
SEQ ID No: 1
Integrin beta-4b         ---------------------------------------EHLVNGRMDFA   1380
SEQ ID No: 2
```

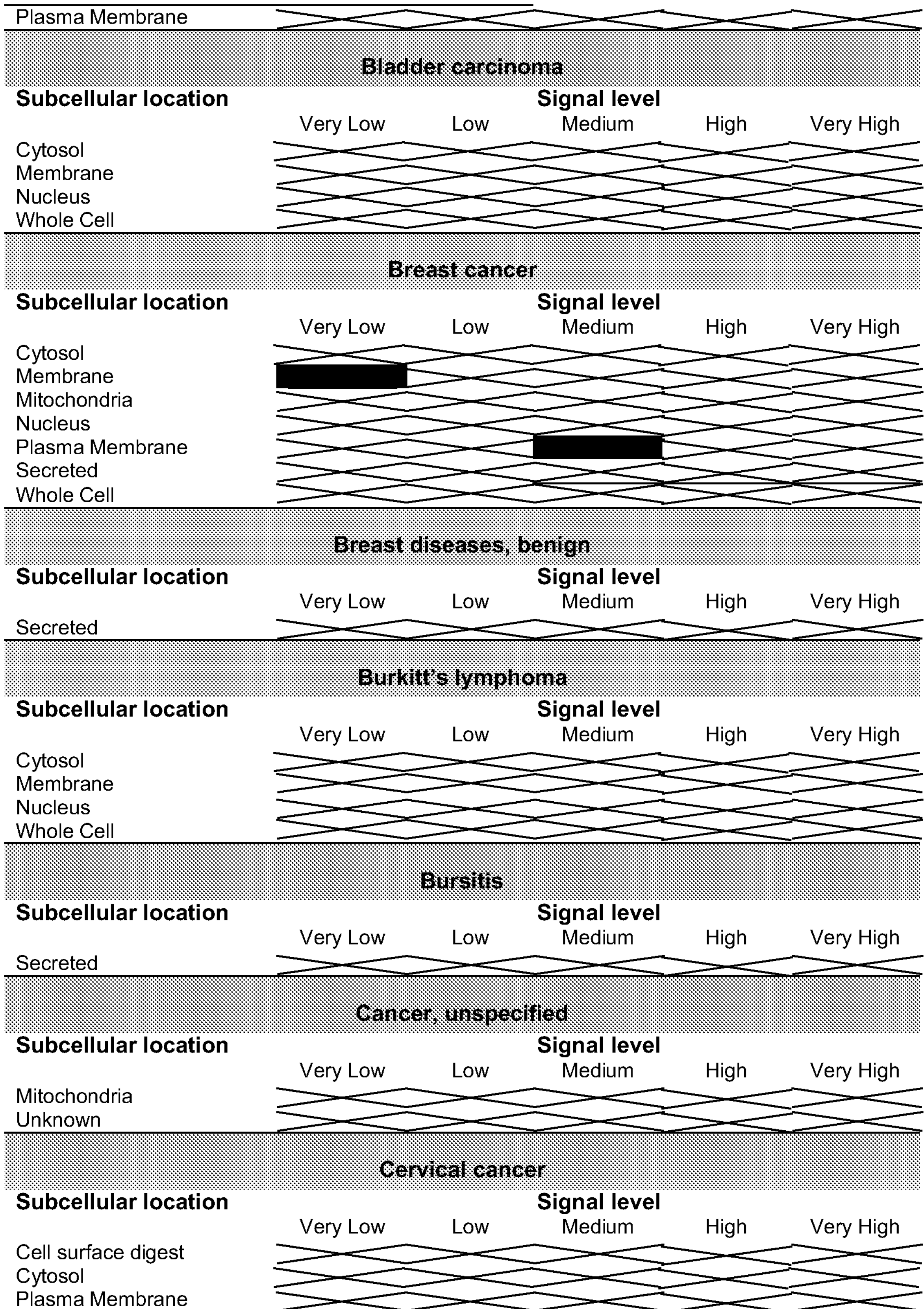
```
Integrin beta-4c     -------------------------------------EHLVNGRMDFA
SEQ ID No: 3
Integrin beta-4d     -------------------------------------EHLVNGRMDFA
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 1
Integrin beta-4b     FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH  1430
SEQ ID No: 2
Integrin beta-4c     FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 3
Integrin beta-4d     FPGSTNSLHRMTTTSAAAYGTHLSPHVPHRVLSTSSTLTRDYNSLTRSEH
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     SHSTTLPRDYSTLTSVSSH-------------------------------
SEQ ID No: 1
Integrin beta-4b     SHSTTLPRDYSTLTSVSSH-------------------------------  1449
SEQ ID No: 2
Integrin beta-4c     SHSTTLPRDYSTLTSVSSHGLPPIWEHGRSRLPLSWALGSRSRAQMKGFP
SEQ ID No: 3
Integrin beta-4d     SHSTTLPRDYSTLTSVSSH-------------------------------
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 1
Integrin beta-4b     ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS  1477
SEQ ID No: 2
Integrin beta-4c     PSRGPRDSIILAGRPAAPSWGPDSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 3
Integrin beta-4d     ----------------------DSRLTAGVPDTPTRLVFSALGPTSLRVS
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 1
Integrin beta-4b     WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV  1527
SEQ ID No: 2
Integrin beta-4c     WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 3
Integrin beta-4d     WQEPRCERPLQGYSVEYQLLNGGELHRLNIPNPAQTSVVVEDLLPNHSYV
SEQ ID No: 4
Integrin beta-4e     --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a     FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
```

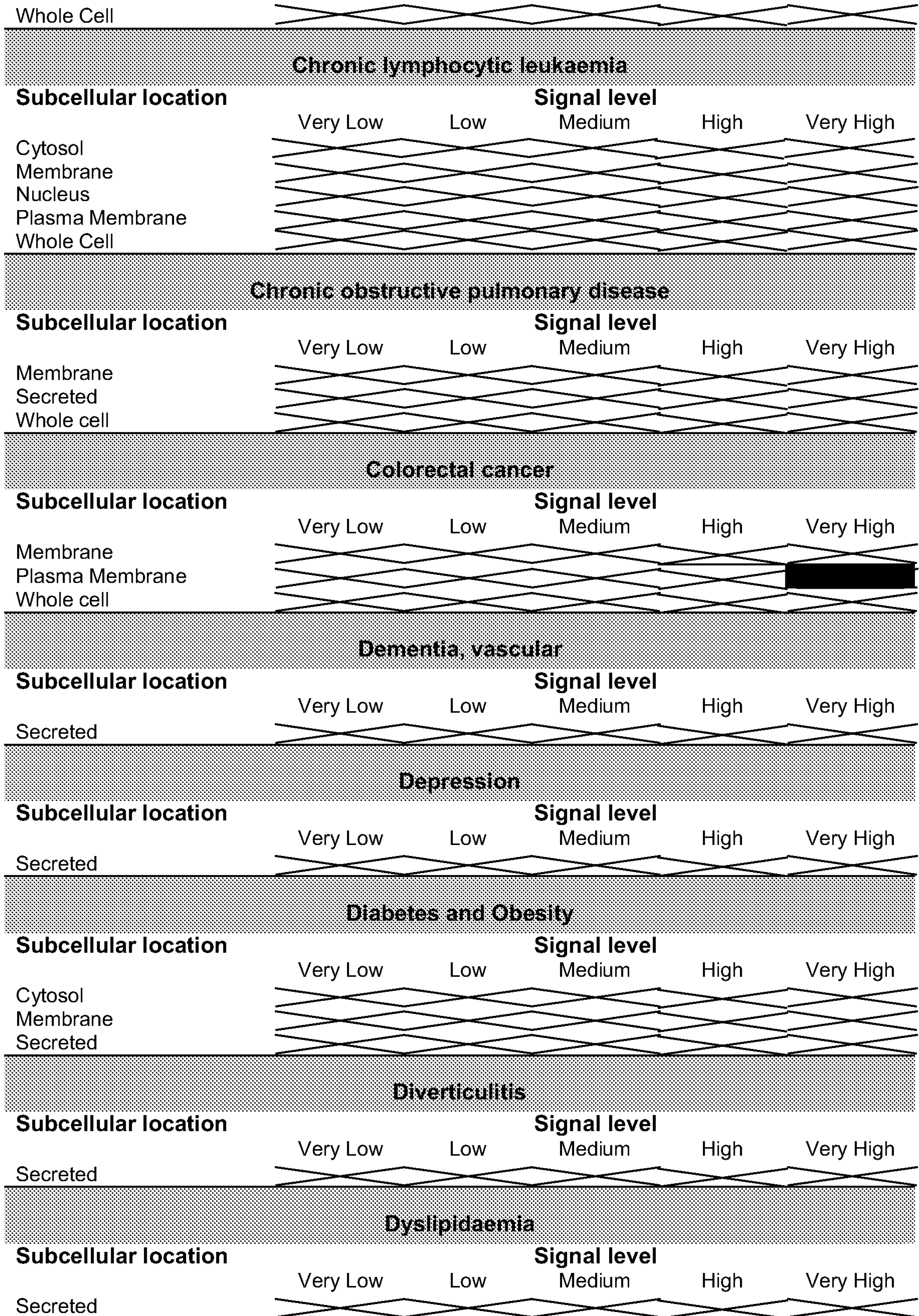
```
SEQ ID No: 1
Integrin beta-4b        FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL  1577
SEQ ID No: 2
Integrin beta-4c        FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 3
Integrin beta-4d        FRVRAQSQEGWGREREGVITIESQVHPQSPLCPLPGSAFTLSTPSAPGPL
SEQ ID No: 4
Integrin beta-4e        --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a        VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP
SEQ ID No: 1
Integrin beta-4b        VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP  1627
SEQ ID No: 2
Integrin beta-4c        VFTALSPDSLQLSWERPRRPNGDIVGYLVTCEMAQGGGPATAFRVDGDSP
SEQ ID No: 3
Integrin beta-4d        VFTALSPDSLQLSWERPRRPNGDIVGYLVTW-------PATAFRVDGDSP
SEQ ID No: 4
Integrin beta-4e        --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a        ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 1
Integrin beta-4b        ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS  1677
SEQ ID No: 2
Integrin beta-4c        ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 3
Integrin beta-4d        ESRLTVPGLSENVPYKFKVQARTTEGFGPEREGIITIESQDGGPFPQLGS
SEQ ID No: 4
Integrin beta-4e        --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a        RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 1
Integrin beta-4b        RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT  1727
SEQ ID No: 2
Integrin beta-4c        RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 3
Integrin beta-4d        RAGLFQHPLQSEYSSITTTHTSATEPFLVDGPTLGAQHLEAGGSLTRHVT
SEQ ID No: 4
Integrin beta-4e        --------------------------------------------------
SEQ ID No: 5

Integrin beta-4a        QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 1
Integrin beta-4b        QEFVSRTLTTSGTLSTHMDQQFFQT                           1752
SEQ ID No: 2
Integrin beta-4c        QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 3
Integrin beta-4d        QEFVSRTLTTSGTLSTHMDQQFFQT
SEQ ID No: 4
Integrin beta-4e        -------------------------
SEQ ID No: 5
```

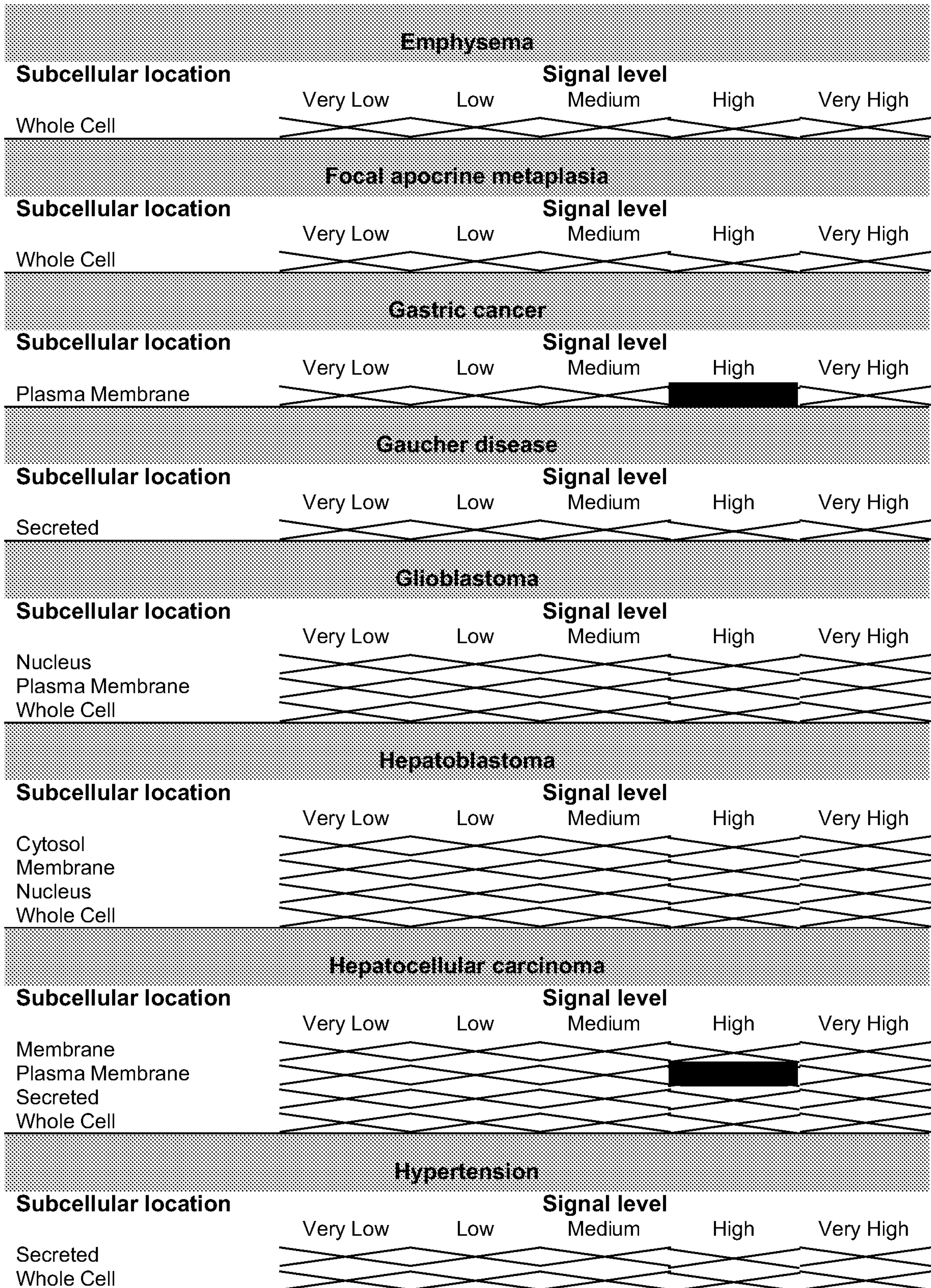
Key:
Recombinant protein
Key Tandem peptide (underline)
Mass Match peptides (bold)
Calx-beta and FNIII domains
Important tyrosine residues (double underline)

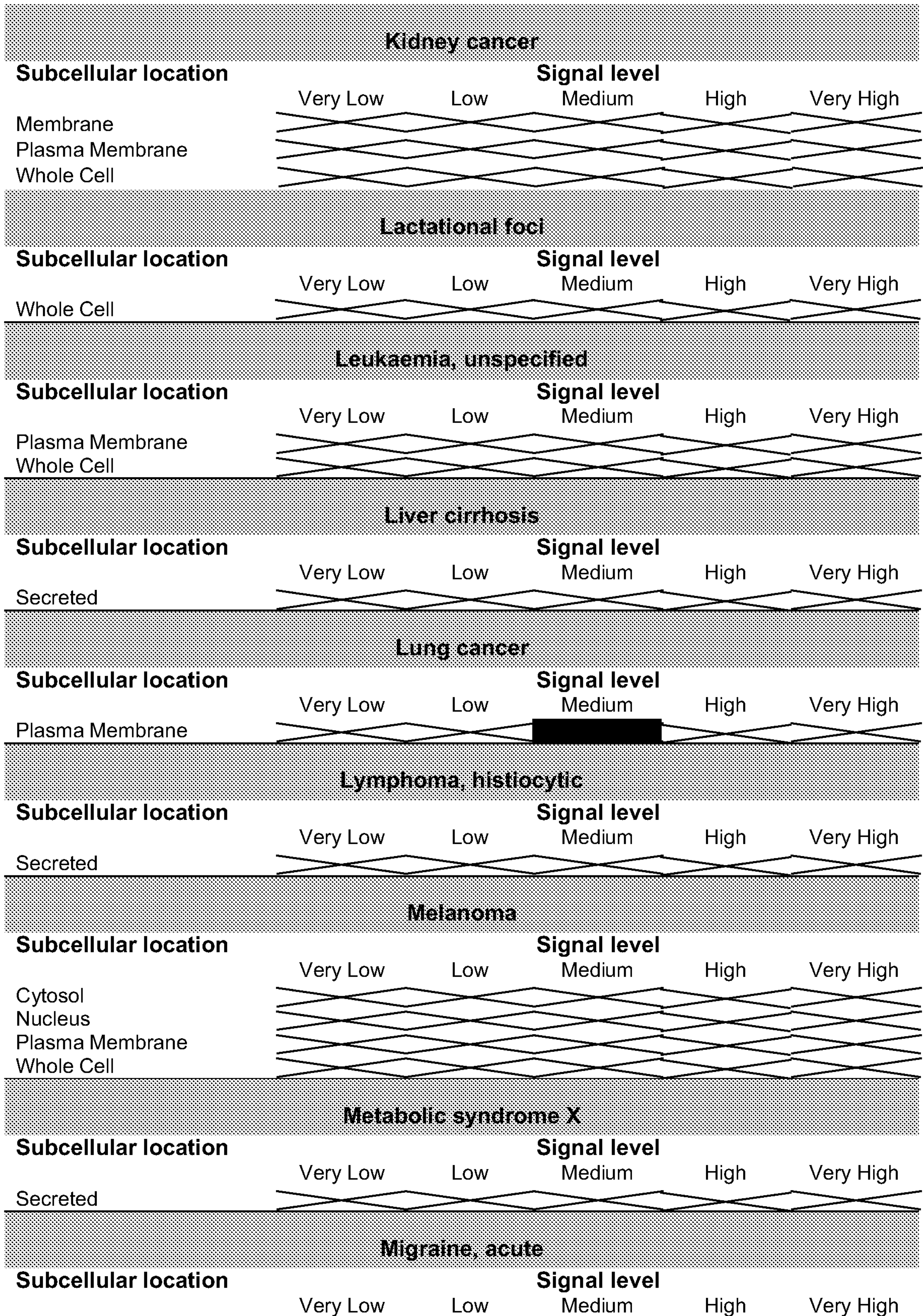
Figure 2

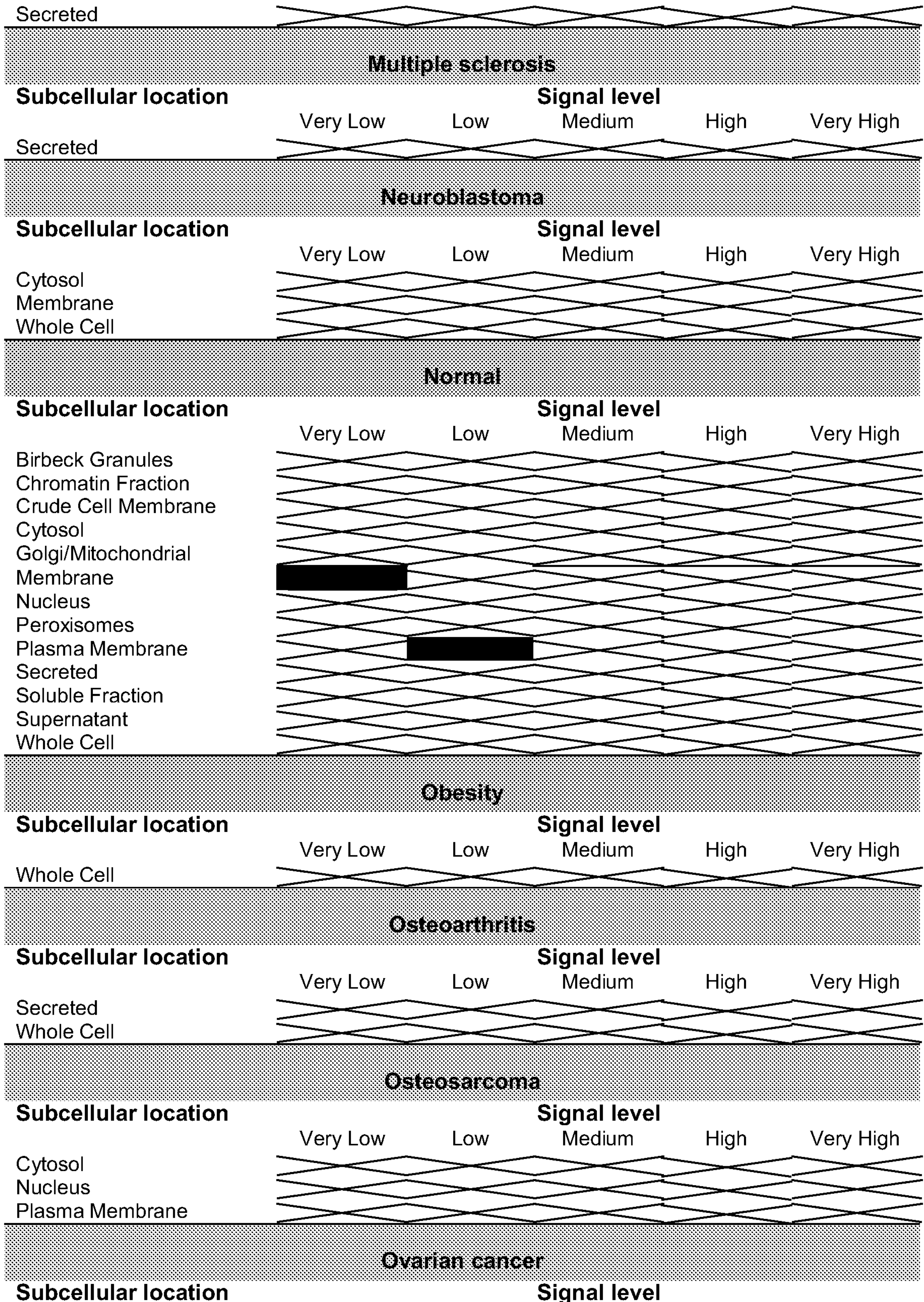


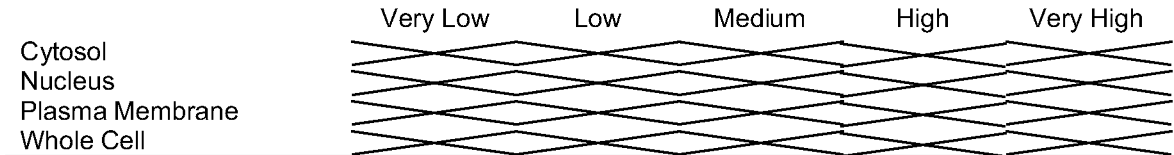

Plasma Membrane
Bladder carcinoma
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Cytosol
Membrane
Nucleus
Whole Cell
Breast cancer
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Cytosol
Membrane
Mitochondria
Nucleus
Plasma Membrane
Secreted
Whole Cell
Breast diseases, benign
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted
Burkitt's lymphoma
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Cytosol
Membrane
Nucleus
Whole Cell
Bursitis
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted
Cancer, unspecified
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Mitochondria
Unknown
Cervical cancer
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Cell surface digest
Cytosol
Plasma Membrane


Whole Cell
Chronic lymphocytic leukaemia
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Cytosol
Membrane
Nucleus
Plasma Membrane
Whole Cell
Chronic obstructive pulmonary disease
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Membrane
Secreted
Whole cell
Colorectal cancer
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Membrane
Plasma Membrane
Whole cell
Dementia, vascular
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted
Depression
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted
Diabetes and Obesity
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Cytosol
Membrane
Secreted
Diverticulitis
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted
Dyslipidaemia
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted

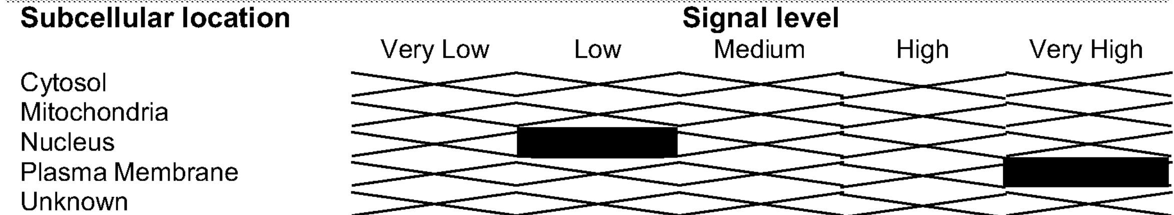

Emphysema
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Whole Cell
Focal apocrine metaplasia
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Whole Cell
Gastric cancer
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Plasma Membrane
Gaucher disease
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted
Glioblastoma
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Nucleus
Plasma Membrane
Whole Cell
Hepatoblastoma
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Cytosol
Membrane
Nucleus
Whole Cell
Hepatocellular carcinoma
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Membrane
Plasma Membrane
Secreted
Whole Cell
Hypertension
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted
Whole Cell


Kidney cancer
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Membrane
Plasma Membrane
Whole Cell
Lactational foci
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Whole Cell
Leukaemia, unspecified
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Plasma Membrane
Whole Cell
Liver cirrhosis
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted
Lung cancer
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Plasma Membrane
Lymphoma, histiocytic
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted
Melanoma
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Cytosol
Nucleus
Plasma Membrane
Whole Cell
Metabolic syndrome X
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted
Migraine, acute
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High

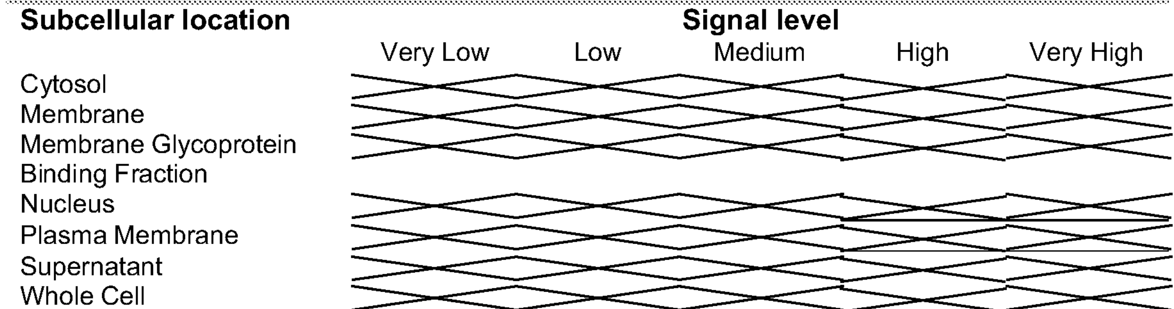

Secreted
Multiple sclerosis
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted
Neuroblastoma
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Cytosol
Membrane
Whole Cell
Normal
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Birbeck Granules
Chromatin Fraction
Crude Cell Membrane
Cytosol
Golgi/Mitochondrial
Membrane
Nucleus
Peroxisomes
Plasma Membrane
Secreted
Soluble Fraction
Supernatant
Whole Cell
Obesity
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Whole Cell
Osteoarthritis
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted
Whole Cell
Osteosarcoma
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Cytosol
Nucleus
Plasma Membrane
Ovarian cancer
Subcellular location
Signal level



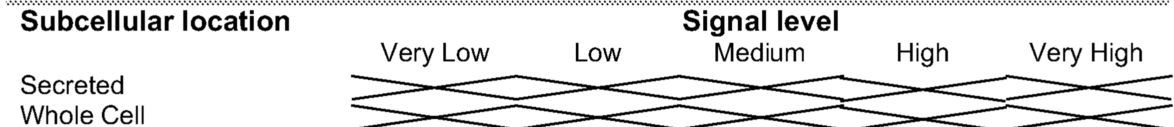
Pancreatic cancer



Prostate cancer

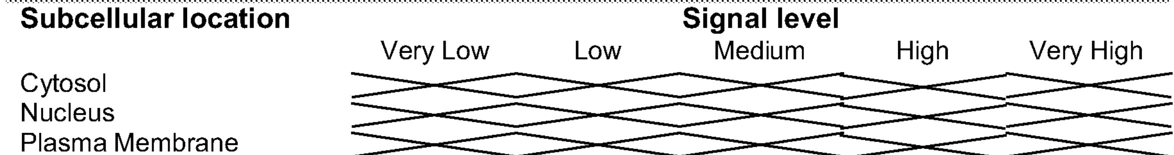


Prostatic diseases, benign

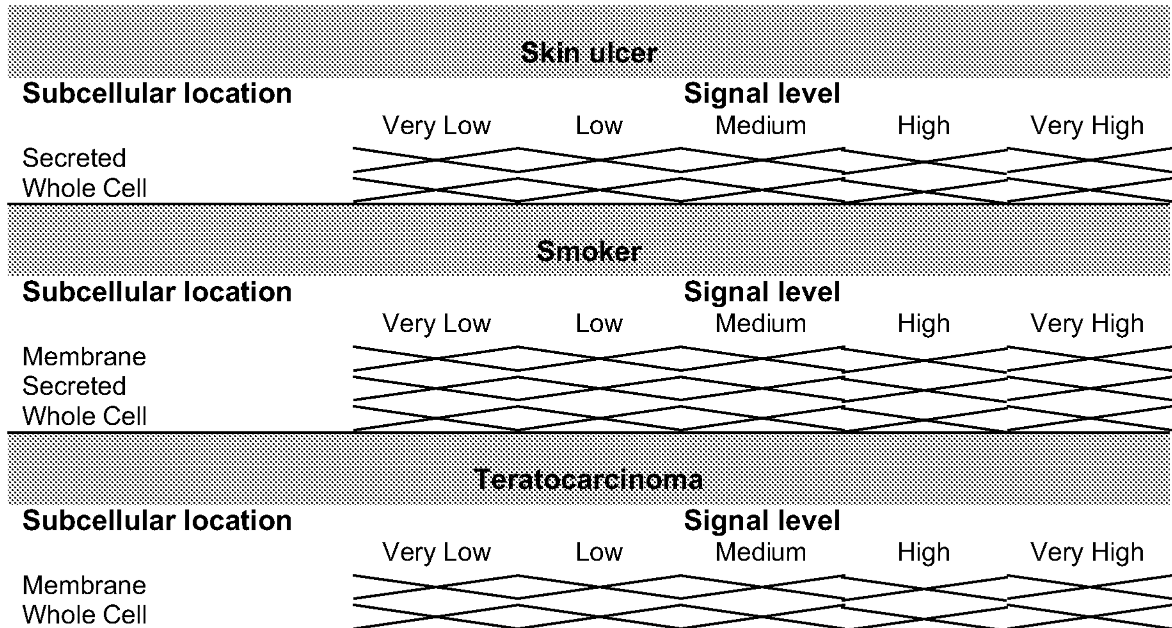


Prostatitis

Retinoblastoma



Schizophrenia


Skin ulcer
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Secreted
Whole Cell
Smoker
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Membrane
Secreted
Whole Cell
Teratocarcinoma
Subcellular location
Signal level
Very Low
Low
Medium
High
Very High
Membrane
Whole Cell

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2008/050902

A. CLASSIFICATION OF SUBJECT MATTER
INV. A61K39/00 C12N15/11

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
A61K C12N

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

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

EPO-Internal, BIOSIS, EMBASE, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2003/224993 A1 (LAND HARTMUT [US] ET AL) 4 December 2003 (2003-12-04) * SEQ ID NO: 17 *claims 2,29 -----	6-13, 19-21, 23-25, 29-37, 43-79
X	WO 02/055659 A (ISIS PHARMACEUTICALS INC [US]; BENNETT FRANK C [US]; FREIER SUSAN M [U) 18 July 2002 (2002-07-18) the whole document -----	6-13, 19-21, 23-25, 29-37, 43-79

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 January 2009	21/01/2009

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Hillenbrand, Guenter

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 320 942 A (QUARANTA VITO [US] ET AL) 14 June 1994 (1994-06-14) * Fig. 9 * the whole document -----	6-13, 19-21, 23-25, 29-37, 43-79
X	WO 01/30854 A (ACTIVE BIOTECH AB [SE]; BRODIN THOMAS N [SE]; KARLSTROEM PIA J [SE]; O) 3 May 2001 (2001-05-03) * Fig 5B * the whole document -----	6-13, 19-21, 23-25, 29-37, 43-79
X	DAEMI NOUCHA ET AL: "Anti-beta4 integrin antibodies enhance migratory and invasive abilities of human colon adenocarcinoma cells and their MMP-2 expression" INTERNATIONAL JOURNAL OF CANCER, vol. 85, no. 6, 15 March 2000 (2000-03-15), pages 850-856, XP002507477 ISSN: 0020-7136 the whole document -----	6-13, 19-21, 23-25, 29-37, 43-79
A	GIANCOTTI ET AL: "Targeting integrin beta 4 for cancer and anti-angiogenic therapy" TRENDS IN PHARMACOLOGICAL SCIENCES, ELSEVIER, HAYWARTH, GB, vol. 28, no. 10, 2 October 2007 (2007-10-02), pages 506-511, XP022282053 ISSN: 0165-6147 the whole document -----	
A	HOGERVORST F ET AL: "CLONING AND SEQUENCE ANALYSIS OF BETA-4 COMPLEMENTARY DNA AN INTEGRIN SUBUNIT THAT CONTAINS A UNIQUE 118-KD CYTOPLASMIC DOMAIN" EMBO (EUROPEAN MOLECULAR BIOLOGY ORGANIZATION) JOURNAL, vol. 9, no. 3, 1990, pages 765-770, XP002507478 ISSN: 0261-4189 -----	

4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB2008/050902

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

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

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

2. ☒ Claims Nos.: 1-5,14-18,22, 26-28,38-42
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

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

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

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

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

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

Remark on Protest

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

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

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box II.2

Claims Nos.: 1-5,14-18,22, 26-28,38-42

The present claims 1-5, 14-18, 22,26-28 and 38-42 encompass affinity reagents/kits/compounds defined only by their desired function, contrary to the requirements of clarity of Article 6 PCT, because the result-to-be-achieved type of definition does not allow the scope of the claim to be ascertained. The fact that any affinity reagents/kits/compounds could be screened does not overcome this objection, as the skilled person would not have knowledge beforehand as to whether it would fall within the scope claimed. Undue experimentation would be required to screen compounds randomly. This non-compliance with the substantive provisions is to such an extent, that a meaningful search for these compounds and the use of these completely undefined affinity reagents/kits/compounds in a method for treating or preventing breast cancer could not be carried out. The search was therefore limited to claims directed to antibodies as affinity reagent/ agents/kits.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guideline C-VI, 8.2), should the problems which led to the Article 17(2)PCT declaration be overcome.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2008/050902

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2003224993 A1	04-12-2003	NONE	
WO 02055659 A	18-07-2002	AU 2002245190 A1	24-07-2002
		US 2006172957 A1	03-08-2006
		US 6355482 B1	12-03-2002
US 5320942 A	14-06-1994	US 5344919 A	06-09-1994
WO 0130854 A	03-05-2001	AU 773568 B2	27-05-2004
		AU 1424801 A	08-05-2001
		BR 0015103 A	16-07-2002
		CA 2386926 A1	03-05-2001
		CN 1384840 A	11-12-2002
		EE 200200224 A	16-06-2003
		EP 1232186 A2	21-08-2002
		HU 0203297 A2	28-01-2003
		JP 2003513024 T	08-04-2003
		NO 20021436 A	19-06-2002
		NZ 518300 A	26-03-2004
		PL 354508 A1	26-01-2004
		US 7189816 B1	13-03-2007
		US 2007112179 A1	17-05-2007
		ZA 200202221 A	19-03-2003