DISCOVERY OF EARLY SERUM BIOMARKERS OF GASTRIC CANCER IN A RAT MODEL

LAM WING KEI

M.Phil

The Hong Kong

Polytechnic University

2009

The Hong Kong Polytechnic University

Department of Applied Biology and Chemical Technology

Discovery of Early Serum Biomarkers of Gastric Cancer in a Rat Model

LAM WING KEI

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Philosophy

December 2008

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____ (Signed)

_____ (Name of student)

Abstract

Gastric cancer has significant morbidity and mortality worldwide. According to the World Health Organization, stomach cancer, which accounts for 866,000 deaths, is the second most common cause of cancer-related death world-wide in 2007. Early diagnosis has a proven impact on improving disease outcome compared with generally poor prognosis after disease progression. Thus, early detection of dysplasia is crucial. The most widely used screening procedures for gastric cancer include endoscopy and biopsy. However, as these invasive screening methods are physically discomfort, they are not suitable for large-scale screening purposes. Therefore, non-invasive serum- or plasma-based diagnostic screening for early gastric cancer detection are likely to be far more acceptable than the current screening options and would likely greatly increase the percentage of the population screened.

The most frequently used gastric tumor markers are carcinoembryonic antigen (CEA), carbohydrate antigen 19-9 (CA 19-9), alpha-fetoprotein antigen (AFP) and human chorionic gonadotropin (hCG). However, they are not sensitive biomarkers as only a modest proportion of patients with gastric cancer have elevated levels of these proteins. Therefore, there is an urgent need for specific biomarkers indicative of early gastric cancer for diagnostic purposes. New emerging proteomic technology using high resolution two-dimensional gel electrophoresis (2-DE) and mass spectrometry (MS) provides a promising option. One of the major hindrances of biomarker discovery is the presence of high abundance proteins. Some abundant proteins such as albumin, immunoglobulins, haptoglobin, antitrypsin and transferrin constitute around 90% of all protein masses present in serum samples. These proteins mask over the relatively low abundance proteins, resulting in difficulties in discovering candidate biomarkers. Another hurdle for early biomarker discovery will be the availability of early stage gastric cancer samples. It should be stressed that when gastric cancer diagnosed by classical endoscopy or biopsy method, it

is already in the late stage. Therefore, in this study, rat gastric cancer was induced experimentally. Serially serum samples were taken and these rats were examined periodically by high resolution mammography to check for gastric lesion. It was intended to correlate possible gastric lesion seen with differentially expressed proteins found in the serum samples. The removal of high abundance proteins by a customized immunoaffinity column was also attempted. Comparative proteomic analysis was used to search for characteristic alternations in the sera of rats with dysplasia and early adenocarcinoma. The various stages of stomach carcinogenesis were confirmed by histological examination. Finally, a panel of serum candidate biomarkers was identified by MALDI-TOF-TOF mass spectrometry. Some of these biomarkers had not been reported to be related to gastric cancer.

List of selected publication and poster presentations:

Publication

LAM, K.W.K. and LO, S.C.L. Discovery of diagnostic serum biomarkers of gastric cancer using proteomics. Proteomics Clinical Applications. vol. 2, pp. 219-228 (2008)

Conference presentation

LAM, K.W.K. and LO, S.C.L. Protein biomarkers associated with experimentally induce gastric cancer in rats. 7th Annual World Congress, Human Proteome Organization, Netherlands, 16-20 Aug 2008

LAM, K.W.K., LAI, Y.M., CHAN, P.S.T., HUI, P.K. and LO, S.C.L. Putative protein biomarkers found in experimentally induced gastric cancer in rats. The Annual Symposium of Hong Kong Proteomics Society, Hong Kong, 17 Nov 2008

ACKNOWLEGEMENTS

I would like to take this opportunity to express of heartfelt gratitude to all the people who have supported me in my MPhil study. Without their help, I would never be able to achieve my study.

First of all, I would like to offer my sincere thanks to my chief supervisor, Professor Chun-lap Lo, for his incredible support, generous encouragement, invaluable advice and patient guidance in my study.

I would also like to express my greatest appreciation to Dr Yau-ming Lai, Dr Suk-tak Chan, Ms Ho-yan Lin and Dr Pak-kwan Hui, for their kindness and generosity in providing technical assistance and advice in the X-ray and histology studies.

I am eternally grateful for the friendship and sharing from my labmates, Yoki, Fred, Kin and Daniel. They have given me tremendous support, generous help and precious advice throughout the project.

I owe a deep thanks to Shing for his endless care and infinite patience. His love and understanding have helped me to go through many difficult times during my research.

Last but not least, to my dearest Dad, Mom and sister who have given me the utmost love and understanding throughout the different stages of my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abbreviations

1-DE	One Dimension electrophoresis
2-DE	Two Dimension electrophoresis
2D-DIGE	Two Dimensional difference gel electrophoresis
3D	Three-dimensional
AFP	Alpha fetoprotein antigen
AHSG	α-2-HS glycoprotein
AML	Acute myeloid leukaemia
AMP	Antrum mucosal protein
BP	Band pass
BSA	Bovine serum albumin
CEA	Carcinoembryonic antigen
CFI	Complement factor I
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanes ulfonate hydrate
CNBr	Cyanogen bromide
Da	Dalton
ddH_2O	Deionized, distilled water
DIA	Differential In-gel Analysis
DMF	N-dimethylformamide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid, disodium salt
ELISA	Enzyme-Linked Immuno Sorbent Assay
g	Gram
GMP	Guanosine monophosphate
HAPs	High abundance proteins
hCG	Human chorionic gonadotropin
HCl	Hydrochloric acid
HDL	High density lipoprotein
HMG-1	High-mobility group-1 protein
HPLC	High Pressure Liquid Chromatography

HRG	Horseradish peroxidase
HSP	Heat shock protein
IAA	Indole-3-acetic acid
ICAT	Isotope-coded affinity tag
IEF	Isoelectric focusing
IgA	Immunoglobulin A
IgD	Immunoglobulin D
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IgY	Immunoglobulin Y
IPG	Immobilized pH gradient
KNG	T-kininogen
kVh	Kilovolt.hour
L	Litre
M	Molar
mA	Milliampere
mAs	Milliampere seconds
MALDI-TOF	Matric-assisted laser desorption/ionization time-of-flight mass spectrometer
MNG	N-methyl-N'-nitroguanidine
MNNG	N-methyl-N'-Nitro-N-Nitrosoguanidine
Mr	Molecular weight
mRNA	messenger Ribonucleic acid
MARS	Multiple affinity removal system
MS	Mass spectrometry
NaCl	Sodium chloride
NADPH	β-Nicotinamide adenine dinucleotide phosphate, reduced tetra(cyclohexylammonium) salt
NAF	Nipple aspirate fluid
NC	Nitrocellulose
NCBInr database	NCBI non-redundant database
O.D.	Optical density
PBS	Phosphate-buffered saline

pI	Isoelectric point
PMF	Peptide mass fingerprint
ppm	Parts per million
PSA	Prostate cancer antigen
PVDF	Polyvinylidene difluoride
RT-PCR	Reverse transcription polymerase chain reaction
S.D	Standard deviation
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SELDI-TOF	Surface-enhanced laser desorption/ionization time-of-flight
SOD	Superoxide dismutase
TBS	Tris-buffered saline
TBST	Tris-buffered saline-Tween
TFA	Trifluoroacetic acid
TTR	Transthyretin
U	Enzyme unit
UV	Ultra-violet
V	Volt
VLDL	Very low-density lipoprotein
v/v	Volume by volume
w/v	Weight by volume

m	milli-	(10^{-3})
μ	micro-	(10^{-6})
n	nano-	(10^{-9})
ρ	pico-	(10^{-12})

Chapter 1: Literature review

1.1 Introduction

This chapter begins with the background information on gastric cancer. Studies employing proteomic technologies to search for cancer biomarkers will be addressed. Discussions on the difficulties in the discovery process will also be elaborated. The purpose of this review is to provide an update on results obtained from previous studies, as well as to provide a rationale for the choice of animal cancer model used.

1.2. Gastric cancer

1.2.1. Prevalence

According to the World Health Organization, stomach cancer is the second most common cause of cancer-related death worldwide in 2007, which accounts for 866,000 deaths. In Hong Kong, stomach cancer ranks as the fifth most common cancer (Figure 1.1 and 1.2). According to the official statistics from the Hong Kong Cancer Registry, the incidence of gastric cancer in 2005 was 1028 (male, 643 and female, 385). The number of gastric cancer deaths in 2005 was 635 (male, 396 and female, 239). The mortality to incidence ratios of stomach cancer of men and of women were both 0.62. This ratio is higher than the overall mortality to incidence ratio of all cancers, which were 0.61 and 0.46 for men and women respectively. These figures show that patients with stomach cancer has a poor prognosis and thus a higher chance to die of it as compared to other kinds of cancers in the context of Hong Kong.

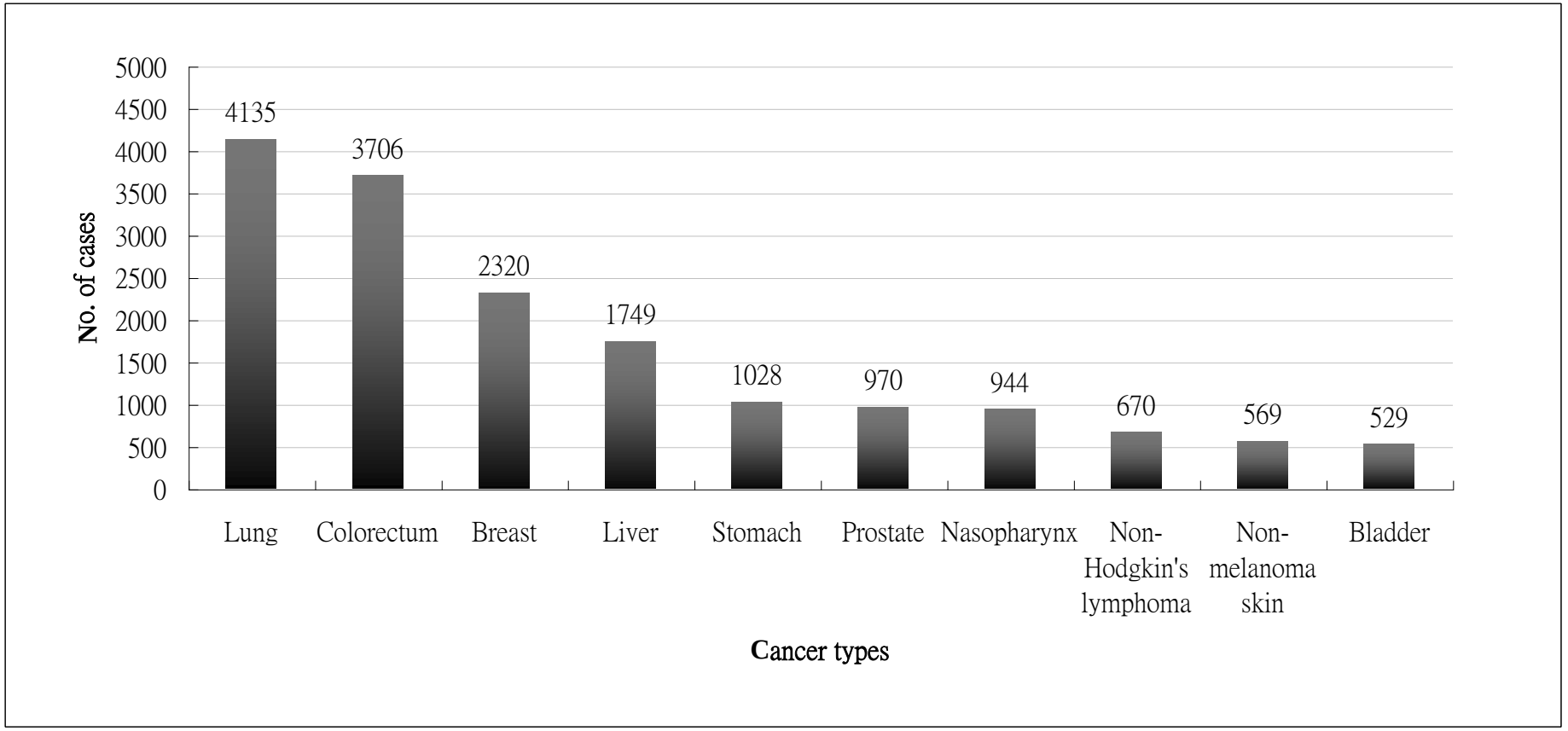


Figure 1.1: Ten leading cancers in Hong Kong 2005.

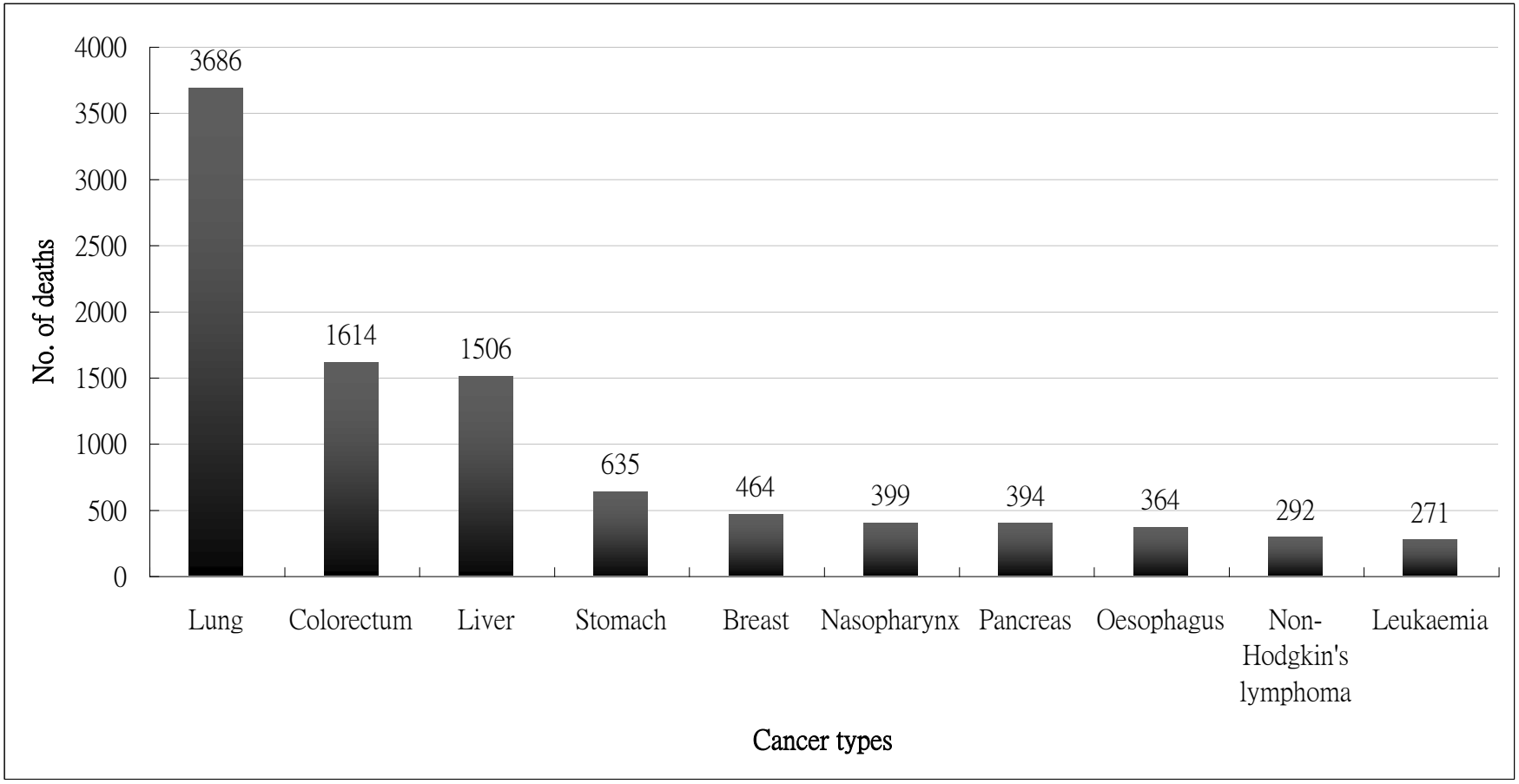


Figure 1.2: The ten major causes of cancer deaths in Hong Kong 2005.

According to the Cancer Facts and Figures in the United States, there were approximately 22,280 new cases of stomach cancer and 11,430 stomach cancer related deaths in 2006. It ranked 14th in incidence among the major types of cancer malignancies. According to the American Cancer Society, the overall survival rate of these patients in 5 years ranges from almost no survival for patients with disseminated disease to an almost 50% survival for patients with localized distal gastric cancers confined to respectable regional diseases. However, even with apparent localized disease, the 5-year survival rate of patients with proximal gastric cancer is only 10-15%. Treatment of disseminated gastric cancer may result in palliation of symptoms and some prolongation of survival, nevertheless, long remissions are unfortunately uncommon.

1.2.2. Types

Gastric carcinogenesis is a progressive sequence of gastric lesions from non-atrophic gastritis to glandular atrophy, metaplasia, dysplasia, and finally to adenocarcinoma (Correa, 1988). According to the morphological features of tumor by Lauren (Lauren, 1965), there are two major histological types of gastric cancer: the intestinal type and the diffuse type. For the intestinal type adenocarcinoma, tumor cells are seen as having irregular tubular structures, harbouring pluristratification, multiple lumens, and reduced stroma. Association with intestinal metaplasia in the neighbouring mucosal is commonly seen. For the diffuse type adenocarcinoma, tumor cells are found to be discohesive and mucus-secreting. When delivered in the interstitium, the tumor cells produced large pools of mucus/colloid.

1.2.3. Dysplasia

Dysplasia in the stomach is commonly used as a marker for the early detection of gastric cancer due the several reasons. Firstly, it is the final step of the progressive genetic and phenotypic changes that modify the original cellular morphology, eventually generating a biologically new cell type that is characterized by uncontrolled growth with the potential to migrate and

implant in locations beyond its original fixed site (Genta and Rugge, 2006). Secondly, it is the earliest carcinogenic lesions in the gastric mucosal recognizable by light microscope (Kunze *et al.*, 1979). Thirdly, the majority (50-80%) of patients have dysplasia progress to invasive carcinoma (Tan and Fielding, 2006). Cellular atypia, abnormal differentiation, and disorganised mucosal architecture are the major histological and cytological features of epithelial dysplasia (You *et al.*, 1993). These histological changes can take place in the intestinal metaplasia and gastric epithelium, both of which may be the source of carcinoma (Morson *et al.*, 1980).

1.2.4. Risk factors

There are a number of acknowledged factors that increase risk of having gastric cancer: *helicobacter pylori* gastric infection (Scheiman and Cutler, 1999); advanced age; the male gender (You *et al.*, 1993); diet low in fruits and vegetables; diet high in salted, smoked, or preserved foods; cigarette smoking; chronic atrophic gastritis; intestinal metaplasia; pernicious anaemia; family history of gastric cancer (Barber *et al.*, 2006); familial adenomatous polyposis and Menetrier's disease (giant hypertrophic gastritis) (Tan and Fielding, 2006). Elaboration of the (postulated) mechanisms of these risk factors to the occurrence of gastric cancer is outside the scope of this thesis.

1.2.5. Symptoms

Gastric cancer is often asymptomatic or causes only non-specific symptoms in its early stages. By the time when more severe symptoms occur, the cancer has usually metastasized to other parts of the body, which accounts for its poor prognosis. Some of the signs may include: indigestion and heartburn; loss of appetite, especially for meat; abdominal pain or vague abdominal discomfort; nausea and vomiting; diarrhoea or constipation; stomach bloating or sense of fullness after eating; weight loss; weakness and fatigue; bleeding in vomit or stool; stool that has turned black or tarry; and fluid swelling in abdomen.

1.2.6. Diagnosis

Double contrast X-ray examination is the initial test for gastric cancer and suspicious lesions will be further investigated by endoscopy (Tan and Fielding, 2006). Endoscopic evaluation is widely used and is the golden standard for the diagnosis of gastrointestinal neoplasm. However, the cost effectiveness of this aggressive approach remains questionable. This is because endoscopy is an invasive procedure with its own risks of morbidity and mortality. Furthermore, this invasive method is time-consuming, relatively more expensive and not suitable for large-scale screening purposes (Feldman *et al.*, 2004).

1.3. Urge for biomarker discovery

As an alternative to endoscopy, detection of serum biomarkers that are indicative of the occurrence of gastric cancer is preferred. Nowadays, the most frequently used gastric tumor markers are carcinoembryonic antigen (CEA), carbohydrate antigen 19-9 (CA 19-9), alpha fetoprotein antigen (AFP) and human chorionic gonadotropin (hCG). However, all of them have inadequate sensitivity and specificity with low predictive value for population screening. Although most gastric oncologists would welcome a blood test for the specific detection of the occurrence of gastric cancer and in spite of the large number of groups that are searching for reliable biomarkers, the progress in this front is very slow. Partly as a reflection of how difficult the task is, the rate of introduction of new protein biomarker tests approved by the United States Food and Drug Administration (FDA) has paradoxically declined over the last decade to less than one new protein diagnostic biomarker per year (Anderson and Anderson, 2002). Although the figure is inclusive of all cancers, implication on the difficulty in finding a reliable biomarker is apparent! The various kinds of difficulties encountered will be discussed later (see Section 1.5.4). One of the biggest hindrances in this type of study is the availability of the right samples. In order to solve this problem, a rat cancer model was used in this study (see Section 1.6 for details).

1.4. Proteomic technology

Studies using genomics approach to screen for cancer-related biomarkers have been launched for decades. Nevertheless, there are still questions about the correlation between the expression levels of mRNA and the corresponding changes in the expression levels of the proteins expressed. It is reasoned that once the cascade of a particular biochemical signal starts off, it may not require the continuous presence of mRNA information. Furthermore, the advance in genomic DNA and mRNA researches has improved our understanding of cell regulation. Yet, these studies cannot illustrate precisely the physiological state of the living cell and its modification at the functional level. This is why only a handful of genes have been approved by FDA for cancer diagnosis purposes. While protein is physiologically closely related to cellular functions, it is more likely to be involved in the key pathways in the biological system and regulations of the pathogenic processes. Thus proteomic investigation on disease states will be the most appropriate and informative approach in the search of reliable biomarkers for cancer.

1.4.1. Two-dimensional electrophoresis (2-DE) with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS)

Two-dimensional electrophoresis (2-DE) was firstly described by Klose (Klose, 1975) and O'Farrell (O'Farrell, 1975) in 1975. It consists of two steps: isoelectric focusing (IEF) and sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE). Proteins are firstly separated according to their pI points in the first IEF step (Figure 1.3), and subsequently separated by their molecular weights during SDS-PAGE (Figure 1.4). At present, the use of immobilized pH gradients strip (IPGs) (Gorg *et al.*, 2000) becomes the choice material for running IEF. It can overcome limitations of the chromatographic type of carrier ampholytes used in the first generation 2-DE process. The IPG based 2-DE, which has ampholytes immobilised on a plastic support and hence will not migrate during the electrophoretic step,

provides better reproducibility, handling, resolution, and separation of very acidic and/or basic proteins.

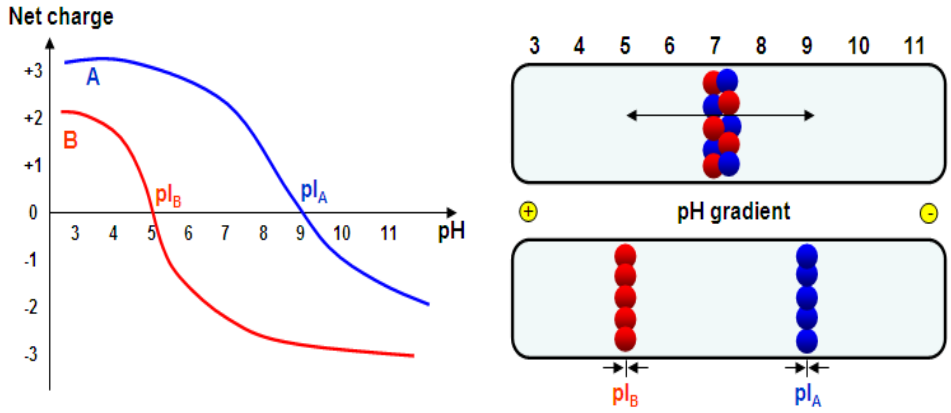


Figure 1.3: Principle of isoelectric focusing (IEF). In the left hand side panel, the 2 proteins (A and B) having 2 different isoelectric points (pI_A and pI_B), they will be resolved from each other after the IEF step (right hand side panel).

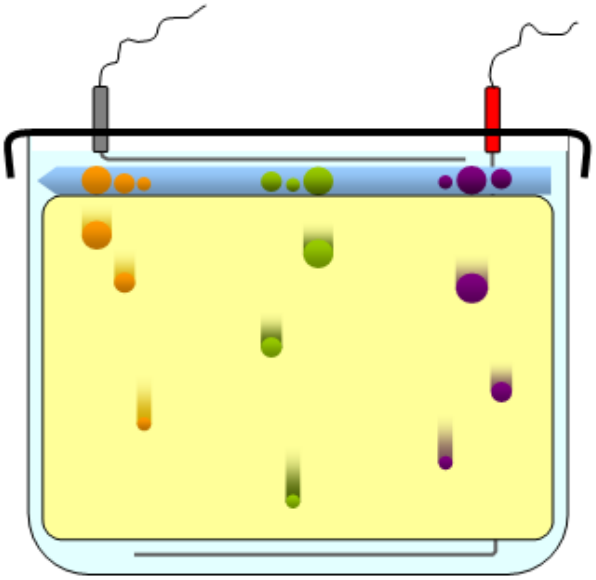


Figure 1.4: Principle of gel electrophoresis. The pores of the gel sieve proteins according to their size. The smaller the size, the faster a protein will go, and vice versa.

After the electrophoretic separation of proteins, the gel is stained for visualizing of the proteins before visual comparison. Coomassie blue is one of the most commonly used stains. However, with the limitation of insufficient sensitivity and a detection range of 200 - 500ng protein per spot, this stain does not permit the detection of low abundance proteins. Thus, silver staining and fluorescence detection have become popular substitutes. These methods are far more sensitive with detection limits as low as 0.1ng protein per spot or less.

After image analysis of gels containing samples with different physiological stages (e.g. normal against cancer), gel plugs containing proteins of interest (e.g. those that are differentially expressed) will be excised before being digested by specific protease with high sequence specificity, such as trypsin. The eluted analyte (i.e. tryptic digest of the protein of interest) is co-crystallised with an UV absorbing matrix on a target plate before being analyzed by a matrix assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer.

The principle of mass spectrometry is to measure the mass (m) to charge (z) ratio of ions for an analyte in the gas phase. Detailed elaboration of the function and setup of a MALDI-TOF mass spectrometer is out of the scope of this thesis. Briefly, a MALDI-TOF mass spectrometer consists of three essential components: an ion source, a mass analyzer and a detector. The analyte-matrix crystal increases the effectiveness in volatizing and ionizing the proteins when it is irradiated by nitrogen laser pulse. The ionized analyte is then directed and separated by the time-of-flight analyzer, and the detector records the flight time of all ions before generating a mass spectrum. The MALDI-TOF MS allows high accuracy in mass to charge ratio and provides spectra of very high resolution (Figure 1.5).

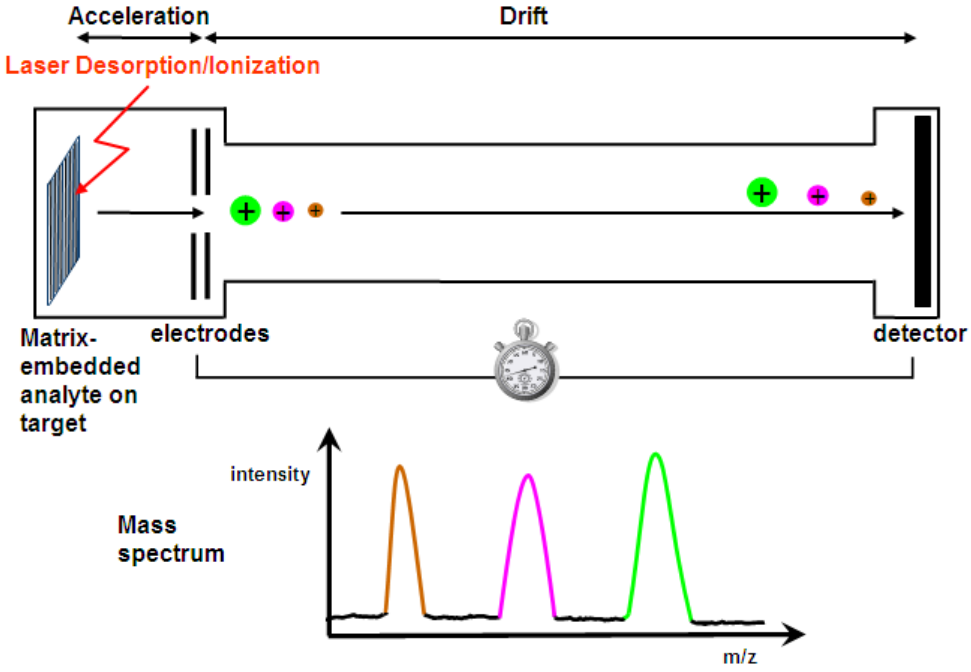


Figure 1.5: Schematic representation of the setup of a MALDI-TOF MS.

The spectrum generated for a protein is the peptide mass fingerprint (PMF) of that particular protein. Except for proteins that share with highly similar sequence homologies, the PMF of every single protein is unique. This experimental profile is then compared to the theoretical PMF masses derived from *in silico* digestion at the same enzyme cleavage sites of the proteins in the database for a match. A significant match will be either based on calculated scoring scheme or based on a random probability algorithm. Finally, the identity of the protein can be elucidated. Although highly sensitive, it should be stressed that this method is both labour intensive and low in throughput (Seibert *et al.*, 2005).

1.4.2. Surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF MS)

This technique aims to reduce sample complexity to a level at which whole protein mass spectrometric profiles can yield a series of supposedly fairly reproducible features. It is a commonly used non-gel based method, which is based on pretreatment of samples with various types of solid matrix called protein chips. On their surfaces, protein extraction based on hydrophobic, ion-exchange, metal binding, or other interactions can be performed. The bound proteins are then eluted before being subject to MS analysis (Diamandis, 2004).

It should be stressed that not all proteins can be well visualized by SELDI-TOF MS, especially those with a molecular weight larger than 20kDa. The sensitivity of these high molecular weight proteins is low, resulting in a smaller number of signals seen (Seibert *et al.*, 2005). Detection of low abundance proteins is another problem. When SELDI-TOF analysis is performed with unfractionated serum, due to the presence of a tremendous array of extremely high abundance to low abundance proteins, these proteins will compete for binding to the same matrix. The low abundance proteins, being the most informative, are highly unlikely to be able to be immobilized onto the chips and detected. For example, it is inexplicable that prostate cancer antigen (PSA), a validated biomarker for prostate cancer, cannot be identified in any of the published studies using SELDI-TOF-MS. Besides that, there are also discrepancies among these studies. Take studies on prostate cancer as an example, five studies had employed SELDI-TOF MS technology for the discovery of biomarkers indicative of the occurrence of prostate cancer. Different sets of "distinguishing peaks" between cancer and non-cancer patients were found in these 5 studies. None of the distinguishing peaks among the five studies for prostate cancer agreed with one another (Diamandis, 2004). This highlights the weaknesses of the SELDI-TOF MS technology in attempts to find reliable disease biomarkers.

1.5. Perspectives of proteomic approach in the discovery of serum biomarkers of gastric cancer

1.5.1. Serum sample analysis

Serum, being the best protein sample representative of the whole body condition, is a favorite sample for investigation. It is easily accessible with a low degree of invasiveness. The possible presence of serum biomarkers rest on the hypothesis that as blood/plasma/serum contains tens of thousands of different proteins, along with their cleaved or modified forms, they can reflect the ongoing physiological and pathological events. Specifically, when blood and/or lymph perfuse the carcinogenic tissues, they will carry "neoplasm-related" proteins and protein fragments passively or actively into the systemic circulation. The complex chemistry of the tumor-host microenvironment and/or reactions of the body to the neoplasm should generate unique signatures in the blood microenvironment. Therefore, these specific patterns-specific serum biomarkers can be used to screen patients with early gastric cancers.

Some researchers compared serum proteomes from patients having gastric cancers with those from healthy controls. Any differentially expressed proteins would be taken as possible candidates of gastric cancer serum biomarkers. One of the major advantages of this strategy is that no pre-selection of biomarker candidates is required as many molecules are screened in a single experiment. Furthermore, no biological knowledge about the patho-physiology of the disease is required (Villar-Garea *et al.*, 2007). In a study done by Chan et al. (Chan *et al.*, 2007), serum samples from gastric cancer patients and healthy controls were compared by traditional 2-DE before MALDI-TOF MS analysis. A markedly increased serum amyloid A (SAA) was identified by this comparative proteomic approach in the cancer patients. ELISA assay was also carried out to verify the above findings. Results showed that the mean SAA concentration was higher in the gastric cancer patients (88.54 $\pm$ 50.44 mg/L) when compared to that of the healthy subjects (3.36 $\pm$ 2.29 mg/L) ($P < 0.05$). Further

validation of the up-regulation of SAA was performed using immunohistochemical staining which showed visible immunoreactivity in resected tumor specimens from the patients with gastric cancer. No immunoreactivity could be seen in samples from the control healthy subjects. These results suggested that SAA was useful in predicting survival of patients with gastric cancer, and was a valuable tool for postoperative follow-up.

Serum samples from 14 gastric cancer patients and 14 healthy individuals were analyzed by Ebert et al. (Ebert *et al.*, 2006) with prefractionation using magnetic bead-based C8-hydrophobic chromatography resins. The eluates were analyzed subsequently by MALDI-TOF-MS. Fibrinopeptide A was found to have increased in the gastric cancer patients. This result was further confirmed by performing ELISA test on serum fibrinopeptide A levels. However, increased fibrinopeptide A serum levels were not specific for gastric cancer; the same results were also found in hepatocellular, ovarian and urothelial cancers. One of the reasons for this non-specificity is the fact that during tissue injury, fibinogen is cleaved to fibrin by thrombin with the release of fibrinopeptides A and B. Therefore, any injury, whether it is from gastric cancer or not, is likely to increase the level of fibrinopeptide A level in blood.

Back to the criticism on earlier proteomic studies that only abundant proteins in the plasma proteome were found, Liu et al. (Liu *et al.*, 2007b) employed an immunoaffinity removal method to clear most of the highly abundant proteins from serum samples. Serum samples from 20 gastric cancer patients and 10 healthy controls were collected and processed with ProteoExtract Albumin/IgG Removal Kit to remove albumin and IgG. After comparative analysis using 2-DE before MALDI-TOF MS identification, three proteins including complement C4-B precursor, complement factor I (CFI) precursor and haptoglobin precursor were found to be differentially expressed in sera from the gastric cancer patients. In validation experiments using Western blotting and in comparison with sera from the healthy

controls, it was found that C4-B precursor was over-expressed in sera from the patients with gastric cancer, while both CFI precursor and haptoglobin precursor were under-expressed in the gastric cancer sera. According to the authors, further investigation on the possible relationship between these proteins and disease is required before these biomarkers could be put into common use.

One of the main reasons why many researchers are cautious in saying that their candidate biomarkers are ready for common use is because there are too many variations in their samples. Limitations over the availability and suitability of human samples used in any of these investigations, such as the date and time of collection, drugs administered, heterogeneity due to differences in diet, race and age etc. are difficult to control and/or optimized. Therefore, some researchers resorted to use animal models. Juan et al. (Juan *et al.*, 2004) developed a strategy that combines a tumor xenotransplantaion model in nude mice with comparative proteomic technology. Human gastric cancer cell line SC-M1 was used to inoculate into nude mice. One month after the inoculation, plasma was collected for proteomic analysis using 2-DE and subsequent MALDI-TOF MS. Plasma proteomes from these inoculated mice were compared with those obtained using plasma from untreated mice. Various acute phase proteins such as haptoglobin and SAA were found to be up-regulated in the tumor-bearing mice. SAA and haptoglobin were confirmed to originate from the mouse hosts and not from the human cancer cell line donors by immunohistochemistry. Up-regulation of SAA was further validated with the serum samples obtained from patients with gastric cancer. One of the concerns of this study is the fact that injection of phosphate-buffered-saline (PBS) into these nude mice was taken as controls to those nude mice receiving the human cancer cells. Instead of using PBS, injection of some non-cancerous cell lines should have been used as control. SAA is a known acute phase protein that is related to inflammation. Whether SAA production in the SC-M1-injected nude mice is due to rudiment activities of their much impaired immune system is currently unknown. Further, there was no absorption procedure to remove abundant serum proteins such as

albumin. Therefore, major abundant proteins including SAA were seen over-expressed in the 2-DE gels of these patients. This topic will be elaborated more in Section 1.6 below.

1.5.2. Tissue sample analysis

Human gastric tissues are also widely used in the search for differential protein expression. Some researchers compared proteomes from gastric tumor tissues with those from adjacent matching non-cancerous samples taken from the same patients. Quantities of differentially expressed proteins identified in the gastric tumor samples will then be measured in the patients' sera, with the expectation that changes in protein amount in the tissue samples and serum samples will agree with each other. Generally, this approach is based on the hypothesis that the proteins found are either secreted proteins from the tumor so that they will go into blood, or specific surface proteins in the tumor which will leak out into blood. These proteins and/or their fragments can then be used as serum biomarkers for gastric cancer.

Ebert et al. (Ebert *et al.*, 2005) screened gastrectomy specimens obtained from ten patients with gastric cancer. After enriching an epithelial cell fraction of these samples, cellular homogenates were resolved by 2-DE. They compared the cancer-proteome to that obtained from matched normal non-malignant mucosal preparation. One hundred ninety-one differentially expressed protein spots were found by 2-DE and identified by MALDI-TOF MS. Among the 191 differentially expressed protein spots that were identified, cathepsin B was found to be over-expressed in 60% of the cancer patients. (The other proteins that were listed in the publication are shown in Table 1.1) Western blotting confirmed that the active form of cathepsin B was over-expressed, while immunohistochemistry showed strong cytoplasmic staining in cancer tissues of 98% of the patients. Further, the serum levels of cathepsin B were significantly higher in the gastric cancer patients (mean concentration 129.41 $\pm$ 79.55 pmol/L) compared with the control patients without evidence of a malignant disease (mean 56.92 $\pm$ 67.4

pmol/L; p = 0.0026). Besides, there was also a significant association between the size of the primary tumor and the presence of distant metastases. Furthermore, high cathepsin B serum levels were also associated with a poor survival in patients undergoing cancer surgery (Petricoin *et al.*, 2002). However, as this study is retrospective in nature, up-regulated proteins such as cathepsin B are not early diagnostic biomarkers. They are rather prognosis biomarkers. In addition, cathepsin B is also over-expressed in other cancer patients, such as patients with prostate cancer (Sinha *et al.*, 2007), colon cancer (Alencar *et al.*, 2007) and breast cancer (Turashvili *et al.*, 2005). Its potential as a specific biomarker for gastric cancer is yet to be confirmed.

Ryu et al. also compared the proteome of human gastric cancer tissue samples with those from non-cancerous region (Ryu *et al.*, 2003). Among the 140 spots identified, seven up-regulated proteins in cancer patients were found. They were NPS3, transgelin, prohibitin, heat shock protein (hsp) 27 and variant, protein disulfide isomerase A3, an unnamed protein product, and glucose-regulated protein. In addition, another seven proteins that were under-expressed were also observed. They were apolipoprotein A-I, p20, nucleoside diphophate isomerase A, α-1-antitrypsin, desmin, serum albumin, and serotransferrin. The authors neither discussed the possible reasons why specific proteins were up- or down-regulated nor suggested the use of any of the differentially expressed proteins as specific biomarkers. However, it should be stressed that among the differentially expressed proteins, increased expression of HSP27 was also found by other researchers (see below).

Undertaking a similar strategy, another group, He et al. (He *et al.*, 2004) also employed 2-DE to analyze primary gastric cancer tissues. In the gastric cancer tissue proteome, they noticed multiple protein alterations when compared to the proteomes of non-tumor tissues. These alternations included variations in the expression of cytoskeleton proteins, such as an increase in cytokeratin 8 and a tropomyosin isoform as well as a decrease in

cytokeratin 20. Co-up-regulations of HSPs and glycolytic enzymes were also observed in these tumor tissues. According to the authors, these up-regulations indicate the self-protective efforts of cells and the growing energy requirement during malignant transformation. Down-regulations were also seen in proteins involved in cell proliferation and differentiation. These proteins included GMP reductase 2, creatine kinase B, and proteins bearing potential tumor suppressor activities, including prohibitin and selenium binding protein 1. Lastly, a human stomach-specific protein, 18 kDa antrum mucosal protein (AMP-18 with NCBI Accession number = 26000208, AAN75447 and experimental Mr/pI = 20kDa / 6.0), was found to be dramatically down-regulated in cancer tissues. According to the authors, the decreased AMP-18 production implicates a possible special pathological role for this protein in gastric carcinogenesis. In the same year, Oien et al. (Oien *et al.*, 2004) reported that this AMP-18 was also down-regulated in their cohort of gastric cancer patients. According to these authors, the protein was renamed as gastrokine 1 by the Human Gene Nomenclature Committee in 2003. These authors found that gastrokine 1 was highly expressed in normal stomach, where it was located in the superficial/foveolar gastric epithelium. However, gastrokine 1 was absent from gastric carcinomas. Furthermore, gastrokine 1 was found only in epithelia showing gastric metaplasia, such as Barrett's oesophagus, the ulcer-associated cell lineage and ovarian mucinous neoplasms. Although there is no solid proof, these authors speculated that gastrokine 1 (also called CA11, AMP-18 or foveolin) has a role in mucosal protection. Further research is required to see if gastrokine-1 could be used as a specific cancer biomarker.

Yoshihara and coworkers also tried to minimize individual variations by studying protein alternations between tumor and non-tumor tissues from the same gastric cancer patients (Yoshihara *et al.*, 2006). It was found that a non-histone chromosomal protein called high-mobility group-1 protein (HMG-1) was elevated in tumor tissue. HMG-1 is a DNA-binding protein that regulates the transcription of various genes and the structural stabilization of chromosomes. It was reported to be associated with

carcinogenesis and metastasis in colorectal and breast cancer. On the other hand, these authors reported a remarkable decrease in the level of foveolin precursor (FOV, Accession number = 38488935; AAR21211 with Mr/pI = 20kDa/5.7). This protein was called gastrokine-1 in the report by Oien et al. (Oien *et al.*, 2004). RT-PCR was performed and the results revealed significant down-regulation of FOV mRNA expression in tumor tissues (Yoshihara *et al.*, 2006). Hence, results of Yoshihara et al. reinforced the observations made by He et al. (He *et al.*, 2004) as well as Oien et al. (Oien *et al.*, 2004). Therefore, a decrease in gastrokine-1/FOV may be seen as a specific phenomenon for the occurrence of gastric cancer.

Ten pairs of samples consisting of gastric cancer tissues and their corresponding adjacent sections were analyzed by Zhang (Zhang *et al.*, 2007). After resolving by 2-DE, the gastric cancer associated proteins were identified by MALDI-TOF-TOF MS; they included twelve up-regulated and thirteen down-regulated proteins. Extra attention was paid to one of the down-regulated protein, MAWBP. Quantitatively, the expression of MAWBP was measured with real-time PCR and semi-quantitatively with Western blot. Immunohistochemical staining also confirmed the existence of MAWBP in the gastric cancer tissues.

After proteomic analysis of 14 paired samples of gastric carcinomas and their corresponding non-cancerous gastric mucosa, Nishigaki et al. found nine proteins with an increased expression and 13 with a decreased expression in gastric cancer (Nishigaki *et al.*, 2005). The two most notable groups included proteins involved in mitotic checkpoint (MAD1L1 and EB1) and mitochondrial functions (CLPP, COX5A, and ECH1). These results suggested that there were links between dysfunctions in these processes and gastric carcinogenesis. More importantly, immunohistochemical analyses confirmed that the levels of expression of MAD1L1 and CYR61 were decreased in gastric carcinoma tissues while HSP27 was increased in gastric carcinoma tissues.

A different approach was adopted by Chen et al. (Chen *et al.*, 2004), who used a well described N-methyl-N'-Nitro-N-Nitrosoguanidine (MNNG) intoxicated Wistar rat gastric cancer model to investigate the protein profiling of carcinogenesis and metastasis in the induced gastric cancer tissues. After 2-DE and MALDI-TOF-MS analysis of the cancer tissues and matching normal tissues, they found that 11 proteins were up-regulated and two proteins were down-regulated (Table 1.1). In particular, the over-expression of HSP27 in gastric cancer was confirmed by immunohistochemical analysis of human gastric cancer specimens. HSP27 is a known molecular chaperone with an ability to interact with a large number of proteins (Kato *et al.*, 1992). It was said that its production is accelerated in non-physiological conditions and that it will aid cell survival. Thus, it was speculated that overproduction of HSPs could protect malignantly transformed cells from apoptotic cell death and foster resistance to chemotherapeutic agents and irradiation (Witkin, 2001; Concannon *et al.*, 2003). Its expression is associated with poor prognosis (Kapranos *et al.*, 2002). However, higher expression of HSP27 protein had been previously reported in liver, breast, colon and gastric cancers (Lemieux *et al.*, 1997; Garrido *et al.*, 1998; Yu *et al.*, 2000; Kapranos *et al.*, 2002; Ryu *et al.*, 2003). Therefore, HSP27 may not be a specific biomarker for the occurrence of gastric cancer.

Given the retrospective nature of most gastric cancer biomarkers studies discussed above, researchers had tried to the improve quality of data with samples from gastric cancer cell lines. Takikawa et al. (Takikawa *et al.*, 2006) used 2D-DIGE coupled with MALDI-TOF MS to identify specific proteins differentially expressed between a highly metastatic stomach cancer cell line MKN-45-P and its parental cell line MKN-45. Their results revealed up-regulation of eight proteins (IFN-induced Mx protein, Gly-, Tyr-, Trp-tRNA synthetase, the flavoprotein subunit of complex II, the ts11 cell cycle protein, keratin 5 and adenylate kinase) and down-regulation of five proteins (pyruvate kinase, cytokeratin 8, didydrodiol dehydrogenase, annexin I, and carbonic anhydrase II). It was found that these 13 proteins were mainly involved in protein synthesis, metabolism, receptor and signal

transduction, the cytoskeleton and cell cycling. It was hoped that these results could be informative and applied in large-scale validation studies.

On the other hand, Lee et al. (Lee *et al.*, 2005) tried to identify the target antigen of a reported stomach cancer specific antibody, MG7 mAb, using KATO III and MKN-45 gastric carcinoma cell lines. MG7 antigen was found to have decreased in gastric cancer patients who underwent gastrectomy. This allowed MG7 detection to be used in diagnosis of the presence of gastrointestinal cancer as well as to evaluate the effectiveness of and treatment outcome after cancer therapy. With 2-DE and Western blotting, two proteins of 35 kDa were consistently detected by the MG7 antibody. Followed by MALDI-TOF MS, these MG7 immunoreactive proteins were identified as the herterogeneous nuclear ribonucleoprotein A2/B1 (hnRNP A2/B1). Four of the five gastrectomy samples from patients with gastric cancer showed positive results in Western blotting with commercially available hnRNAP A2/B1 antibodies. However, hnRNAP upregulation is also found in other types of cancer (Turck *et al.*, 2004; Zech *et al.*, 2006).

1.5.3. Gastric juice analysis

Besides gastric tissues as well as blood related body fluid, other types of samples were also investigated. Gastric juice is a unique body fluid containing most of the chemical components secreted from the stomach. Gastric juice has become a choice of study because it is speculated that its composition and relative concentration will reflect the functional state of the stomach.

Lee et al. (Lee *et al.*, 2004) utilized gastric juice in 2-DE analysis aiming to obtain disease specific protein expression patterns. In healthy subjects, pepsin A, pepsin C and gastric lipase were the major proteins, but they were not detected in 60% of gastric cancer cases. Conversely, an

extraordinary amount of α-1 antitrypsin was observed in gastric cancer patients, while it was only detected in 5% of the healthy group.

Another group, Hsu et al. (Hsu *et al.*, 2007), investigated the protein components in gastric juice by 2-DE and MALDI-TOF MS. They found that frequency of occurrence of the specific α-1-antitrypsin precursor was higher in gastric cancer patients (93%) compared to that of healthy subjects (6%).

Research group	Proteins	Expression
Chan et al. (Chan *et al.*, 2007)	Serum amyloid A	↑
Chen et al. (Chen *et al.*, 2004)	Catecholamine-O-methyltransferase	↑
	Protein disulfide isomerase	↑
	Alpha-1-antitrypsin precursor	↑
	Heat shock 27 kDa protein	↑
	Haptoglobin precursor	↑
	Desmin	↑
	Albumin	↑
	3-mercaptopyruvate sulfurtransferase	↓
	Calreticulin	↓
	ATP synthase beta-subunit	↓
Ebert et al. (Ebert *et al.*, 2005)	Myosin light chain alkali, non-muscle isoform/myosin light chain alkali, smooth-muscle	↑
	Cathepsin B	↑
	Proteasome activator hPA28 subunit beta	↑
	Enhancer of rudimentary homolog	↑
	PDZ and LIM domain protein 1	↑
	Transitional ER ATPase	↑
	Peroxiredoxin 2	↓
	CA11 protein	↓
	Cathepsin E precursor	↓
	Gastricsin precursor	↓
	Complement component1, Q subcomponent binding protein	↓
	BC007716) unknown / (AF058954) GTPspecific succinyl-CoA synthetase beta subunit	↓
	ATP synthase D chain	↓
	Dihydrolipoamide dehydrogenase, mitochondrial precursor	↓
	Electron transfer flavoprotein-ubiquinone oxidoreductase	↓
	Fumarate hydratase, mitochondrial precursor	↓
	NADH-ubiquinone oxidoreductase 42 kDa subunit	↓
	Succinate dehydrogenase [ubiquinone] flavoprotein	↓
He et al. (He *et al.*, 2004)	150 kDa oxygenregulated protein precursor	↓
	Cytokeratin 8	↑
	Prohibitin	↑
	Tropomyosin isoform	↑
	Enolase 1	↑
	Trisoephosphate isomerase	↑
	Phosphoglycerate mutase 1	↑

Research group	Proteins	Expression
He et al. (He *et al.*, 2004)	Pyruvate kinase	↑
	Chaperonin containing TCP1	↑
	Heat shock protein 60	↑
	Heat shock cognate 71 protein	↑
	Protein disulfide isomerase	↑
	Translation elongation factor EF-Tu	↑
	Alpha-1-antitrypsin	↓
	Apolipoprotein A-1	↓
	GMP reductase 2	↓
	Creatine kinase B	↓
	Selenium binding protein 1	↓
	Cytokeratin 20	↓
	Carbonic anhydrase I	↓
	Carbonic anhydrase II	↓
	18kDa Antrum mucosa protein	↓
Hus et al. (Hsu *et al.*, 2007)	Alpha-1-antitrypsin	↑
Juan et al. (Juan *et al.*, 2004)	Haptoglobin	↑
	Serum amyloid A	↑
Lee et al. (Lee *et al.*, 2004)	Alpha-1-antitrypsin	↑
Liu et al. (Liu *et al.*, 2007b)	Complement C4-B precursor	↑
	Complement factor I	↓
	Haptoglobin	↓
Nishigaki et al. (Nishigaki *et al.*, 2005)	Annexin 5	↑
	Bisphosphate 3-nucleotidase	↑
	Microtubule-associated protein, RP/EB family, member 1	↑
	GST	↑
	Heat shock 27 kDa protein 1	↑
	MADS box transcription enhancer factor 2, polypeptide C	↑
	Nicotinamide *N*-methyltransferase	↑
	UDP-glucuronosyltransferase	↑
	ApoA-I binding protein	↓
	Serum amyloid P component precursor	↓
	CA11 protein	↓
	ATP-dependent proteolytic subunit *E. coli* homolog	↓
	Cytochrome c oxidase subunit Va	↓

Research group	Proteins	Expression
Nishigaki et al. (Nishigaki *et al.*, 2005)	Tumor RMS cell line RD specific product	↓
	Enoyl Coenzyme A hydratase 1, mitochondrial and peroxisomal	↓
	Serine proteinase inhibitor, clade B (ovalbumin), member 1	↓
	Ferritin, heavy chain	↓
	Cytosolic inorganic pyrophosphatase	↓
	Mitotic checkpoint protein isoform MAD1a	↓
	Mannose-6-phosphate isomerase	↓
Oien et al. (Oien *et al.*, 2004)	Gastrokine 1	↓
Ryu et al. (Ryu *et al.*, 2003)	NPS3	↑
	Transgelin	↑
	Prohibitin	↑
	Heat shock protein (hsp) 27 and variant	↑
	Protein disulfide isomerase A3	↑
	Unnamed protein product	↑
	Glucose-regulated protein	↑
	Apolipoprotein A-1	↓
	P20	↓
	Nucleoside diphophate isomerase A	↓
	α1 antitrypsin	↓
	Desmin	↓
	Serum albumin	↓
	Serotransferrin	↓
Takikawa et al. (Takikawa *et al.*, 2006)	Interferon-induced Mx protein	↑
	Glycyl-tRNA synthetase	↑
	Flavoprotein unbunit of complex II	↑
	Ts11 cell cycle protein	↑
	Adenylate kinase 2	↑
	Tyrosyl-tRNAsynthetase	↑
	Keratin 5	↑
	Transfer RNA-Trp synthetase	↑
	Cytokeratin 8	↓
	Dihydrodiol dehydrogenase isoform DD1	↓
	Annexin I	↓
	Annexin A2, isoform 2	↓
	Pyruvate kinase	↓
	Carbonic anhydrase II	↓

Research group	Proteins	Expression
Yoshihara et al. (Yoshihara *et al.*, 2006)	Manganese superoxide dismutase	↑
	Nonhistone chromosomal protein HMG-1	↑
	Carbonic anhydrase I	↓
	Carbonic anhydrase II	↓
	Aspartate aminotransferase 2 precursor	↓
	Glutathione S-transferase	↓
Zhang et al. (Zhang *et al.*, 2007)	SM22	↑
	α-tropomysin	↑
	cathepsin D	↑
	ATP synthase	↑
	HSP27	↑
	HnRNP A2	↑
	actin alpha 2	↑
	SEPT2 protein	↑
	calponin 1	↑
	ACTG protein	↑
	aldehyde dehydrogenase	↑
	NADH dehydrogenase Fe-S protein 8	↑
	acyl-coA dehydrogenase	↓
	NDUFV2 protein	↓
	MAWBP	↓
	hypothetical protein	↓
	isocitrate dehydrogenase	↓
	aldehyde dehydrogenase 6A1	↓
	annexin I	↓
	pyruvate kinase 3 isoform 1	↓
	aldo-keto reductase familty 1 B10	↓
	dihydrodiol dehydrogenase isoform DD1	↓
	carbonyl reductase 1	↓
	HnRNP-E1	↓
	HSDL2 protein	↓

Table 1.1: Differential proteins identified by different proteomics research groups regarding gastric cancer. (↑: up-regulated in cancer, ↓: down-regulated in cancer)

In the studies discussed above, several proteins were repeatedly found to be differentially expressed in serum or tissue samples from patients with gastric cancer when compared to non-neoplastic serum/tissue samples. They were apolipoprotein A-I, α-1-antitrypsin and HSP27. However, we know that apolipoprotein A-I and α-1-antitrypsin are abundant proteins present in human blood. Therefore, their roles as specific biomarkers for gastric cancer have to be verified. Furthermore, minor proteins such as HSP27 and hnRNP A2/B1 are not specifically related to gastric cancer. Taken overall, we are yet to have a specific (set of) early diagnostic protein biomarker for a large scale screening of gastric cancer.

1.5.4. Difficulties encountered

Although researchers are working hard on biomarker discovery, identifying sensitive and specific serum biomarkers for early diagnosis of gastric cancer still remains a daunting challenge. The biggest problem is the presence of the high abundant proteins which interfere with the exhibition of some low abundant proteins that may be very important. All of the current, clinically utilized cancer serum biomarkers are low abundance proteins that usually constitute less than one billionth the concentration of albumin. Based on these precedents, specific cancer biomarkers will almost certainly be low abundance proteins (Hoffman *et al.*, 2007), as shown in Table 1.2. A few proteins such as albumin, immunoglobulins, haptoglobin, antitrypsin and transferrin constitute around 90% of all protein masses present in serum samples (Yu *et al.*, 2005). Serum samples should therefore be treated to concentrate these relatively rare biomarkers for the discovery process. Multiple approaches for simplifying the serum proteome have been done and compared. These serum fractionation schemes include N-linked glycopeptide enrichment, cysteinyl-peptide enrichment, magnetic bead separation (C3, C8 and WCX), size fractionation, protein A/G depletion, and immunoaffinity column depletion of abundant serum proteins. After comparing with the results obtained from the unfractionated human serum, it was found that the immunoaffinity subtraction is the most effective means

for simplifying the serum proteome while maintaining a reasonable sample throughput (Whiteaker *et al.*, 2007).

Cancer biomarker	Reference range (ng/mL)	Difference from albumin
AFP	0 - 15	10^6
PSA	0 - 4	10^7
CEA	0 - 3	10^7
hCG	0 - 0.1	10^8

Table 1.2: Concentration of some classical cancer biomarkers in serum and their comparison to albumin (Hoffman *et al.*, 2007).

Another difficulty for biomarker discovery is the intrinsic variation of serum samples within individuals (Juan *et al.*, 2004). The proteome pattern of serum depends on a number of factors, including the genetic background, sex, age, nutritional status, lifestyle, medical treatment and bed rest. The methods of samples collection, protocols for preparation and storage of samples are also extremely sensitive to the proteome patterns. Lastly, given the significant research efforts worldwide that are in place to discover useful cancer biomarkers, the process is still slow and difficult most of the time. This is because the validation of a biomarker is more difficult and time-consuming than its discovery. Unfortunately, validation has to be performed and successfully validated before the candidate biomarkers can be translated from the discovery phase to become viable clinical screening tools. Therefore, the challenge is still on.

1.6. Gastric cancer model in rat

Given that human variations are huge and clinical relevant samples are difficult to obtain, animal cancer models provide a reasonable alternative. Since the 1930s, attempts to experimentally induce gastric cancers in rat had been performed by researchers using several different carcinogens such as benzo[a]pyrene, 3-methylcholanthrene, and 2-acethylaminofluorene (Tsukamoto *et al.*, 2007). However, the incidences of experimentally induced gastric cancer were low. It was not until 1967 when Sugimura and Fujimura were able to report good yields of adenocarcinomas in the glandular stomachs of rats fed with a chemical carcinogen called *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG) (Sugimura and Fujimura, 1967). Rat gastric cancers have histological features very similar to that of human cancers, and they can closely reflect the effects of various tumor promoters, such as salt and sex hormones, and those of preventive agents such as green tea components (Niwa *et al.*, 2005). Thereafter, rats have been used most extensively in gastric carcinogenesis studies with the cancer induced with MNNG ingestion (Kobori *et al.*, 1976; Tahara *et al.*, 1976; Kobori *et al.*, 1977; Matsukura *et al.*, 1978; Kobori *et al.*, 1984; Ohgaki *et al.*, 1984; Ohgaki *et al.*, 1991; Yuasa *et al.*, 1994; Schwab *et al.*, 1997). Besides using rats with MNNG-induced gastric cancers to investigate the gastric carcinogenesis process, this model is also used in chemopreventive agents intervention studies (Ganapathy *et al.*, 2008; Manikandan *et al.*, 2008).

1.6.1. Wistar rats

It has been ascertained that the susceptibility of the gastric epithelium to MNNG varies among several strains of rats under similar experimental conditions. Wistar rats are the most susceptible one, while Buffalo strain is the most resistant (Kobori *et al.*, 1976). Besides, male rats are generally more susceptible to gastric carcinogenesis induced by MNNG than females (Ohgaki *et al.*, 1983).

The sequential morphological changes of the glandular stomach of rats intoxicated with MNNG in drinking water had been described (Kobori *et al.*, 1976). In the first 2 weeks of MNNG administration, erosions with severe inflammatory reaction are rapidly induced. These erosions will gradually subside in weeks 6-8. In the following 6 weeks, regenerative changes are mainly observed. Subsequently, atypical changes are then observed in weeks 15-32. Finally, adenocarcinomas of the glandular stomach will be induced in weeks 35-72 (Yamashita *et al.*, 2002).

1.6.2. Gastric cancer induction mechanism by MNNG

MNNG is used for gastric cancer induction because it is a direct acting mutagenic and carcinogenic compound. It induces the methylation of nucleic acids, forming mainly 7-methylguanine with a tiny amount of 3-methyladenine, 1-methyladenine, 3-methylcytosine, and O^6-methyguanine (Craddock, 1968; McCalla, 1968; Singer *et al.*, 1968; Lawley and Thatcher, 1970; Lawley and Shah, 1972). MNNG methylates proteins as well. It is found that O^6-methyguanine is one of the most crucial adducts for mutation and can be repaired by O^6-methyguanine DNA methyltransferase. When the pyloric mucosa of the rat glandular stomach is exposed to MNNG, erosive lesions occur with disordering of glandular structures and proliferation of pyloric mucosa. Atypical glands and stomach cancer cells are detected subsequently. In the end, both differentiated and undifferentiated stomach carcinomas mimicking the histological types of human stomach cancer are induced (Sugimura and Fujimura, 1967).

MNNG is target tissue specific when given in drinking water, it selectively causes induction of tumors in the glandular stomach but not in the oesophagus and fore-stomach of rats. As MNNG is related to methylation, this phenomenon shows that gastric cancer induction in MNNG-ingested rats has a great relationship with the extent of DNA methylation. It has been reported that after a single oral administration in rats, the concentration of methylpurines in the glandular stomach was 9 times higher than that in the fore-stomach, 20 times higher than that in the

oesophagus, and 1.3 times higher than that in duodenum (Wiestler *et al.*, 1983). The regional differences in DNA methylation are believed to be related to the concentrations of cellular thiols which boost the decomposition of MNNG and the extent of its macromolecular binding. The concentration of thiols in the oesophagus and fore-stomach are 25% and 10% of that in glandular stomach, respectively.

Tumor incidence in the small intestine is found to be comparatively lower as MNNG is rapidly converted to non-carcinogenic N-methyl-N'-nitroguanidine (MNG) owing to the acidic conditions of gastric juice (Mc and Wright, 1948).

1.6.3. Enhancement effect of sodium chloride

Both the incidence and the size of gastric tumors in rats increased when 10% of sodium chloride was added in the diet during the treatment with MNNG (Takahashi *et al.*, 1983). It signifies that NaCl acts as a co-carcinogen in the commencement of gastric carcinogenesis induced by MNNG in rats. Sodium chloride also acts as a gastric carcinogenesis promoter as the incidences of gastric carcinomas were found to be higher in rats ingesting NaCl after termination of MNNG treatment than those given MNNG alone (Takahashi *et al.*, 1984). Sodium chloride has been proven to induce ornithine decarboxylase activity and semiconservative DNA replication in the gastric mucosal of rats, suggesting promotion activity of NaCl in the stomach (Ohgaki *et al.*, 1984).

Given the important accessory role of NaCl in inducing gastric cancer in rats with MNNG ingestion, it is adopted in our study. Details of the MNNG-ingestion protocol in the presence of NaCl will be discussed in Chapter 2.

1.7. Aims and objectives

Since there are urging needs in the discovery of early biomarkers of gastric cancer, the present study aims to discover new biomarker candidates with high sensitivity and specificity with proteomic technologies. The study is mainly divided into three parts.

The first part of the study was to establish a gastric cancer model in rat with the induction of a chemical carcinogen.

For the second part, chicken antibody was raised and employed in an immunodepletion column. This immunodepletion column will be used to deplete most of the high abundance proteins presence in serum for further analysis.

Comparative proteomics were done in the last part of the study for the identification of differentially expressed proteins. Results obtained from conventional 2-DE and 2D-DIGE were compared.

Chapter 2: Materials and methods

2.1. Introduction

This chapter describes the research methodology, methods and materials used in this study. The establishment of the animal model, sampling methods, and cancer detection methods were described. The use of antibody production in hens approach and a comparison of different antibody purification methods were included. The preparation of the affinity column and the proteins depletion process were stated. Various protein detection and analysis methods were also described.

2.2. Materials

Unless stated otherwise, all chemicals used are from Sigma (USA) with at least Analytical Grade (AR). Further, organic solvents including acetonitrile, methanol and trifluoroacetic acid are at least of HPLC grade. Plastic wares are of the highest quality to avoid leakage of the polymers. Eppendorf vials are of the highest analytical grade from Eppendorf (USA).

2.3. Animal model

2.3.1. Wistar rats

Twenty-four male Wistar rats of 6 weeks old were purchased from the Laboratory Animal Services Centre (LASEC) of The Chinese University of Hong Kong. Animals were housed in polycarbonate cage under standard laboratory condition, with room temperature 23±2℃, relative humidity 60± 5%, 12h light/dark cycle. The animals were given food and water *ad libitum*. All the procedures were performed according to protocols approved by the Animal Subjects Ethics Subcommittee of The Hong Kong Polytechnic University. (Ethics Approval no.: ASECS No.06/25).

2.3.2. Gastric cancer induction by N-methyl-N'-nitro-N-nitrosoguandine (MNNG)

After acclimatization to the laboratory conditions for 2 weeks, gastric cancer induction by MNNG started when the rats were 8-week old. 19 rats in the experimental group were given drinking water containing 0.01% MNNG freshly prepared everyday. 1mL 10% NaCl was given weekly by oral gavage in the initial 6 weeks to initiate gastric carcinogenesis induced by MNNG (Chen *et al.*, 2004). The solution was stored in lightproof bottles to avoid the degradation of MNNG by light. 5 control rats were given normal tap water. Body weight of each rat was recorded every 2 weeks.

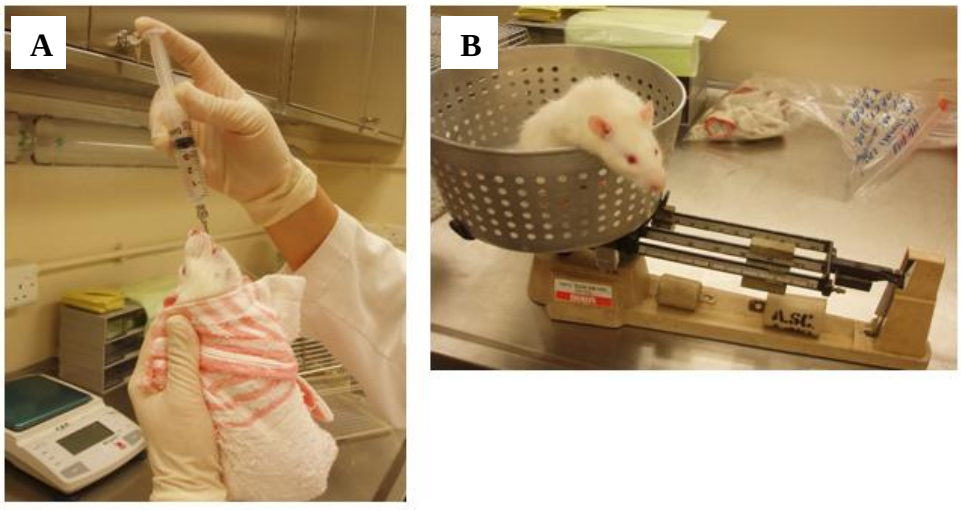


Figure 2.1: (A) Rat fed with 1mL 10% NaCl weekly and (B) weighed bi-weekly.

2.3.3. Blood sample collection

Starting from week 8 after gastric cancer induction by MNNG, blood sample was collected from each rat biweekly. Under slight ether anesthesia, about 1.5mL blood was taken from the vein underneath the eye with a sharp glass capillary tube. The blood was allowed to clot by standing at room temperature for 1 hour. It was then centrifuged at 2,000g at 4℃, for 20mins. The serum obtained was aliquot and was stored at -80℃ until use. None of the samples were stored more than 12 months.

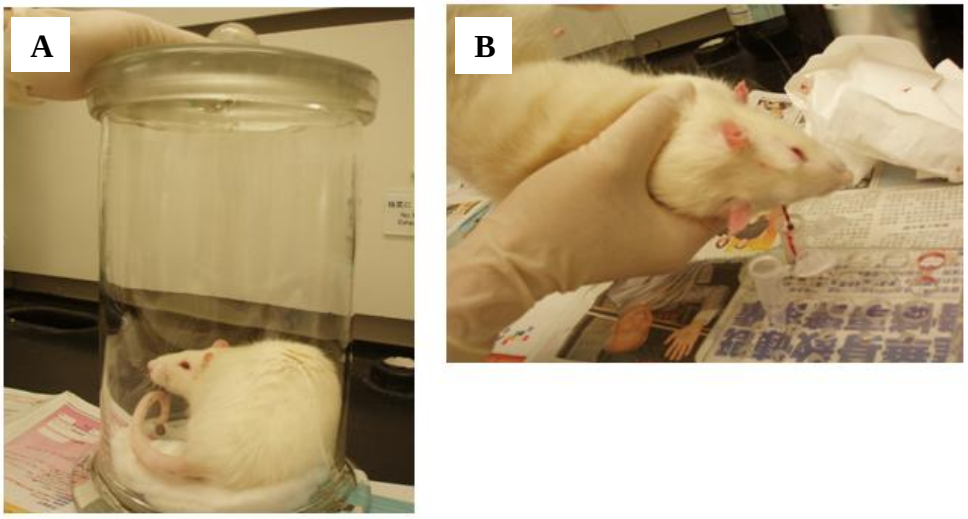


Figure 2.2: (A) Rat was lightly anesthesized with ether before (B) collecting blood from the vein underneath the eye.

2.3.4. High resolution X-ray analysis of stomachs

Starting from week 19, the experimental rats were examined by high resolution X-ray monthly. The mammography machine, MAMMOMAT 3000 (Siemens, Germany), was provided by the Radiographic Clinic in The Hong Kong Polytechnic University. Under anesthesia with ketamine (90mg/kg) and zylaxine (10mg/kg), 1mL 10% gastrografin was given by oral gavage to coat the stomach wall. Subsequenhtly, about 50mL of air was administrated orally to inflate the stomach for better observation. The rat was put in prone position for X-ray radiation (Figure 2.3). The X-ray condition was set with voltage ranging from 23kV to 35kV, and milli-Ampere seconds varied from 2mAs to 5mAs.

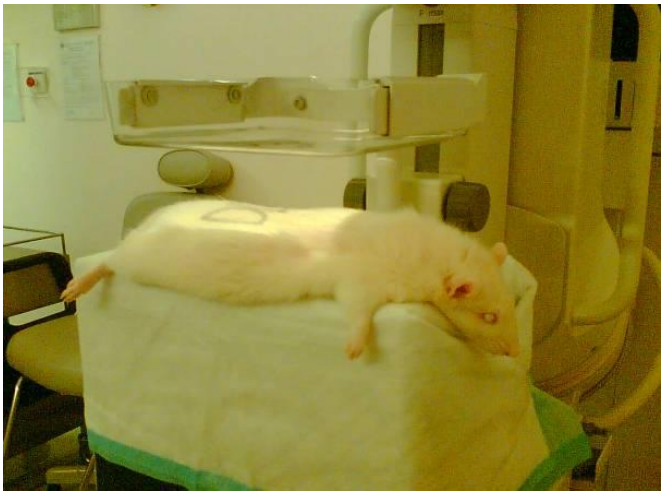


Figure 2.3: Rat in prone position for high-resolution X-ray examination.

2.3.5. Histological examination of gastric tissues

In cases where gastric tissues of rats were required, the stomachs were removed from the animals before they were fixed in 10% buffered formalin to stabilize the tissues to prevent decay. Each stomach was cut into 4 strips, including 1 strip in proximal the stomach, 2 strips in the body of stomach and 1 strip in the distal stomach (as shown in Figure 2.4). They were immersed into formalin at 53.7℃ for 6 hrs. The samples were then immersed in multiple baths of ascending concentration of ethanol to dehydrate the tissues, followed by xylene, a clearing agent before being finally put in hot molten paraffin wax. The sampe was then embedded in a mould and allowed to cool and harden. The tissue was sectioned into 4μm using microtome. The slices were placed on a glass slide for staining with hematoxylin and eosin. Hematoxylin stains the nuclei blue, while eosin stains the cytoplasm pink. The specimens were analyzed under a light microscope. Occurrence of dysplasia and/or adenocarcinoma was examined before being confirmed by Dr. Pak-kwan Hui, Consultant Pathologist and Chief-of-Service of the Pathology Department of The Kwong Wah Hospital.

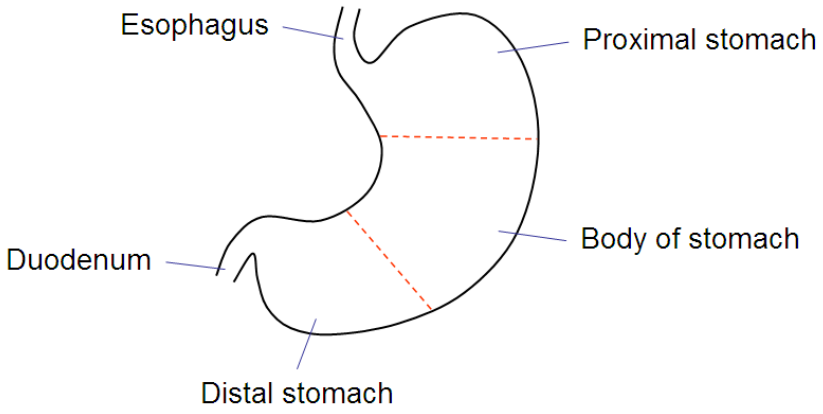


Figure 2.4: Stomach resected from the rat is cut along the red dot line for histological study.

2.4. Antibody production in hens and purification of IgY from egg yolk

2.4.1. Antibody production

200μg of rat serum from normal Wistar rats diluted with PBS (0.5 ml final volume) was mixed 1:1 (v/v) with either Complete Freund's Adjuvant (in the initial injection) or with Incomplete Freund's Adjuvant in subsequent boosters (Ruan *et al.*, 2005). 1mL of the antigen emulsion was injected into the pectoral muscle of hens at a 3-week interval for 4 times, and the last two times at a 6-week interval. Eggs were collected daily, marked and stored at 4 ℃.

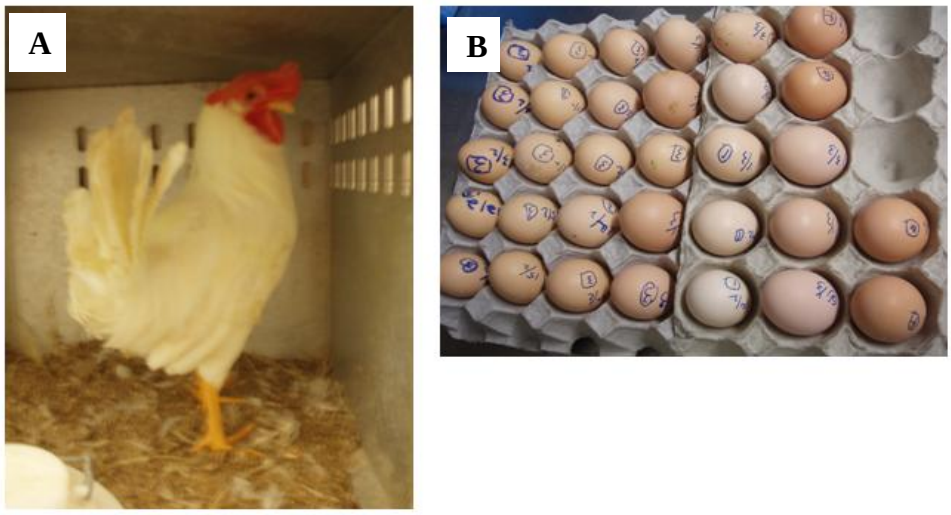


Figure 2.5: (A) The immunized hen with (B) some of the eggs collected.

2.4.2. IgY purification

In order to obtain the highest amount of IgY (against the target antigens) of eggs, 3 different IgY purification methods were performed.

2.4.2.1. EGGstract Kit

EGGstract IgY purification (Promega, USA) was performed according to the instructions from the manufacturer. Briefly, the yolk from each egg was

obtained and weighed before 3 volumes of solution A were slowly added into it at room temperature. The mixture was mixed for 5min. Lipid portion of the yolk was precipitated after the act and the yolk mixture was then centrifuged for 15min at 10000g and 4^{o}C. 1/3 volume of solution B was added before it was slowly stirred for 5min. The suspension was centrifuged again for 15min at 10000g and 4^{o}C. Subsequently, the pellet was resuspended again in 1/3 volume of solution B and stirred slowly for 5min and centrifuged for the last around in the condition same as above. Finally, the pellet was resuspended in the original volume of the egg yolk of PBS.

2.4.2.2. Ammonium sulphate precipitation

The egg yolk was diluted with water in a 1:2 (v/v) ratio and then frozen at -20℃ overnight. The mixture was then thawed and centrifuged at 10000g for 10min at 4℃. Ammonium sulphate was added to the supernatant to a final concentration of 50% (w/v). The solution was mixed for 1 hr at 4℃ before centrifugation at 5000g for 20min. After centrifugation, the pellet was resuspended in 0.01M Tris-HCl to a volume equal to half of that of the supernatant. The sample was then precipitated by adding ammonium sulphate to a final concentration of 33%. The solution was again stirred for 1h at 4℃ before centrifugation for 20min at 5000g. Finally, pellet was resuspended in PBS after centrifugation (Ruan *et al.*, 2005).

2.4.2.3. Sodium sulphate precipitation

The egg yolk was diluted with water in a 1:9 dilution and allowed to incubate at 4℃ for 6 hrs. The mixture was then centrifuged at 10000g for 25min and 4 ℃. Subsequently, the supernatant was collected and precipitated by adding 19% (w/v) sodium sulphate. The suspension was stirred for 2hrs at room temperature. Pellet was collected after another round of centrifugation at 10000g for 25min at 4℃ and dissolved in PBS (Akita and Nakai, 1993).

2.5. Affinity column

2.5.1. Preparation of the IgY affinity column

Cyanogen bromide activated sepharose was prepared according to the manufacturer's instruction. Briefly, the CNBr-activated sepharose was washed and swelled in cold 1mM HCl for 30min in several aliquots. Subsequently, the CNBr-activated sepharose was washed with 5-10 column volumes of distilled water before being washed with coupling buffer. The suspension was then filtered quickly through a sintered glass filter. Immediately afterwards, the IgY in the coupling buffer was transferred to the resin. The IgY was mixed with the gel and allowed to incubate overnight at 4℃. The suspension was again filtered by a sintered glass filter. The unreacted ligand was washed away from the sepharose with the coupling buffer. Subsequently, blocking reagent was added to block the unreacted sites for 16 hrs at 4℃. Finally, the gel was washed extensively first with the coupling buffer and then the acetate buffer repeatedly for five times. The IgY-affinity-gel was then ready for use.

2.5.2. Depletion of high abundance proteins (HAPs) in the serum samples using the IgY affinity column

The IgY affinity column (10mL in 4.91cm^2 x 5cm column) was firstly equilibrated with at least 10 column volumes of equilibrating buffer (10mM Tris pH 7.4). The serum samples were loaded onto the top of the column. Usually 4mL fractions were collected. The "flow through" fractions were collected as the "HAPs depleted" serum. Subsequently, acidic elution was carried out with elution buffer (0.1M Glycine, pH 2.5). Proteins eluted from the column were immediately neutralized with neutralization buffer (1M Tris, pH 8). Finally, the column pH was restored by extensive washing with 20mM Tris buffer pH 7.4. The "flow-through/bound" samples collected were then concentrated either by Amicon Ultra 20 (Millipore, USA) or Amicon Ultra 4 (Millipore, USA) with cut-off 10kDa.

2.6. Protein concentration determination

Protein concentration was determined by Bradford assay (Bio-Rad, USA) according to the manufacturer's instructions. The sample absorbance values were read at 595nm. The protein concentration of each sample was determined from a standard curve generated by a range of bovine serum albumin (BSA) concentrations from 0-10mg/ml.

2.7. One-dimensional gel electrophoresis (1-DE)

Protein samples were prepared by mixing with an equal volume of reducing sample buffer (0.5M Tris-HCl pH 6.8, 0.025% SDS (w/v), 0.25% glycerol (v/v), trace amount of 0.1% bromophenol blue and 0.125% β-mercaptoethanol (v/v)) or non-reducing sample buffer (same as reducing sample buffer, except without the addition of β-mercaptoethanol), then heated at 95℃ for 15min. The sample mixtures were resolved using 10% acrylamide slab gel with 5% stacking gel. Power was applied with a constant voltage of 100V. Electrophoresis was performed until the dye front reached the bottom of the separating gel.

2.8. Two-dimensional gel electrophoresis (2-DE)

2.8.1. First dimension - Isoelectric focusing (IEF)

An immobilized pH gradient (IPG) strip of linear pH 3-10 or 4-7 (18cm, Bio-Rad, Hercules, CA, USA) was rehydrated passively with protein in a modified rehydration buffer containing 7M urea, 2M thiourea, 4% (w/v) CHAPS, 10% (w/v) isopropanol, 5% (v/v) glycerol, 64mM DTT and 1% (v/v) IPG buffer in the rehydration cassette of IEF at 20℃ for 12hr. The strip was then transferred to the IEF focusing tray with the gel facing the electrodes. The IPG strip was overlaid with mineral oil to avoid dehydration. IEF was performed at 20℃ with a PROTEAN IEF CELL (Bio-Rad, USA). Voltage was applied accordingly as follows: 1 hr at 500V, 1 hr at 1000V, 2

hrs at 4000V, 4 hrs at 8000V, 80kVh at 8000V and a final step and hold at 50V. A total of about 100kV was applied to the strip.

2.8.2. Equilibration of IPG strip

After IEF, the strip was submerged in the equilibration buffer (50mM Tris-HCl pH 8.8, 6M urea, 39% (v/v) glycerol, 2% (w/v) SDS and 0.006% (w/v) bromophenol blue) with 1% (w/v) DTT for reduction for 15min with continuous shaking. Subsequently, the strip was put in another lot of equilibration buffer with 2.5% (w/v) iodoacetamide for alkylation for another 15min with continuous shaking.

2.8.3. Second dimention – SDS-PAGE

The equilibrated strip was rinsed with distilled water and then placed on top of the 10% acrylamide slab gel. 0.5% agarose solution was added on top of the strip in order to seal the strip with connection to the gel. Electrophoresis was performed at 15mA/gel for the first 15min, and then increased to 35mA/gel, until the dye front reached the bottom of the separating gel.

2.8.4. Silver staining

After electrophoresis, the gel was stained with a mass spectrometry compatible silver staining protocol. Briefly, the gel was fixed in fixation solution containing 40% (v/v) ethanol and 10% (v/v) acetic acid for 1 hr. Then, it was sensitized in 30% (v/v) ethanol, 0.2% (w/v) sodium thiosulphate for 30min. The gel was washed with distilled water 3 times for 5min each. Subsequently, the gel was stained with 0.25% (w/v) silver nitrate for 20min. The gel was then washed again twice with distilled water for 1min per wash. Finally, the gel was developed by shaking the gel in 2.5% (w/v) sodium carbonate containing 0.04% (v/v) formaldehyde. When the protein spots reached a desirable level of darkness, the development process was stopped by adding 1.5% (w/v) EDTA.

2.8.5. Image analysis

The stained gels were scanned by a conventional scanner (Plustek, USA). The images were then analyzed by 2-DE gel analysis software Melanie version 4.0 (Gene-bio, Sweden).

2.9. Two dimensional Difference Gel Electrophoresis (2D-DIGE)

This method labels the protein with cyanine (Cy) dyes prior to IEF and allows quantitative comparisons between 2 samples (control and disease) on the same 2-D gel when different fluorescent labels are used for each sample. This reduces spot pattern variability and the number of gels in an experiment making spot match much more simple and accurate. The minimal labelling ensures that only a single lysine per protein molecule and not more than 3% of the protein molecules in extracts are labelled. The single positive charge of the CyDye replaces the single positive charge present in the lysine, thus keeping the isoelectric point of the protein unchanged, but causes a size increase of approximately 500Da.

The individual protein data from the 2 samples (Cy3 and Cy5) are normalized against the Cy2 labelled sample, Cy5:Cy2 and Cy3:Cy2. These logarithm abundance ratios between the control and disease samples from all the gels are then compared using statistical analysis.

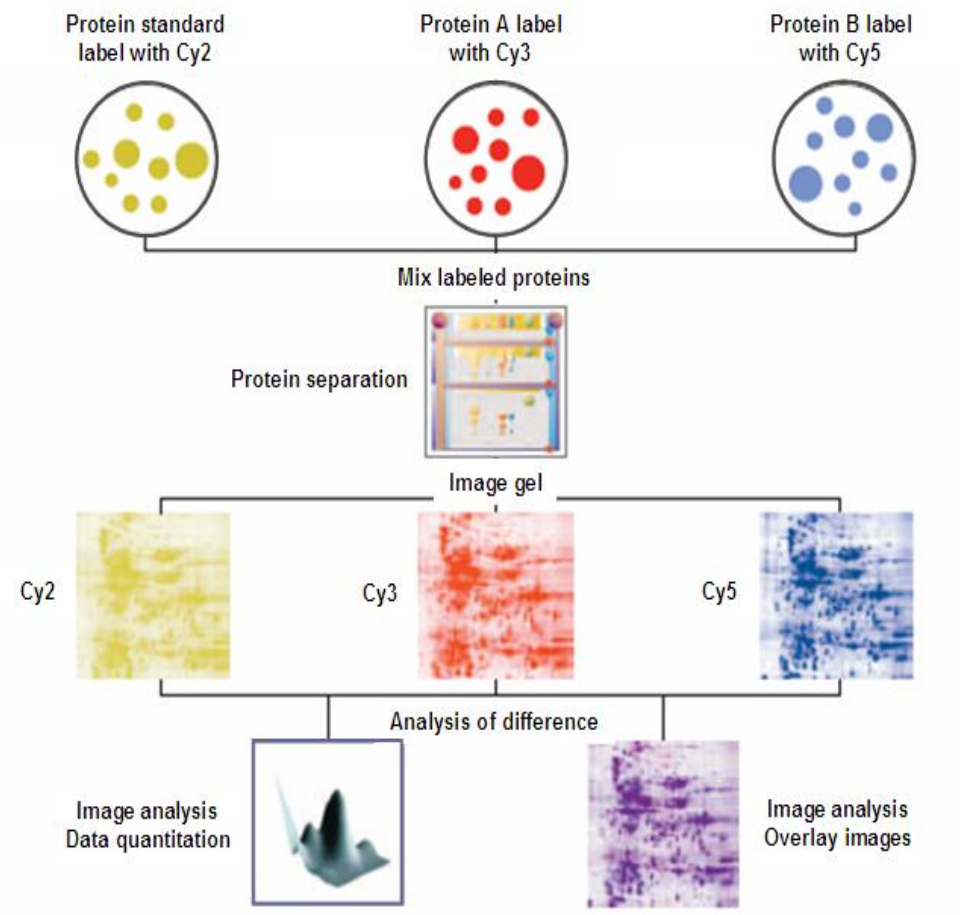


Figure 2.6: Diagram illustrating the workflow of typical 2D-DIGE experiment.

2.9.1. Sample labelling

The fluorescence labelling was performed according to the manufacturer's instruction. Briefly, CyDyes were reconstituted in high-grade N, N-dimethylformamide (DMF) to a stock solution with a concentration of 1mM. The stock solution, stored at -20^{o}C, was supposed to be stable for a few months. In our case, the stock solution was used within 2 months. The 2D-DIGE principle is diagrammatically represented in Figure 2.6 above. Briefly, the CyDye stock solution was diluted with DMF to a final concentration of 400pmol/µl just prior to labelling. The pH of the protein sample had to be adjusted to 8.5 and at the concentration between 5 and 10 µg/µl. The protein samples were labelled with either the 3 Cy-dyes: Cy2, Cy3 and Cy5. The Cy2-labelled sample served as an internal control on the

gel for the DeCyder analysis. The samples were vortexed before being incubated on ice and in dark for 30min. Then, 10mM lysine was added to the samples in order to stop the labelling reaction. The samples were then mixed and incubated on ice and in dark for another 10min. The labelled samples were either directly subjected to 2-DE or stored at -80℃ until use.

When performing 2-DE, the samples labelled with different CyDyes were pooled and then passively rehydrated into the strip with the modified rehydration buffer. The 2D-DIGE was performed with the same procedures as a regular 2-DE as described above. Except for the special low fluorescent Pyrex glass plates were used to minimize the background fluorescent and maximize the signal to noise ratio during image scan, the rest of the equipment setup was the same as regular 2-DE. Moreover, once the samples were labelled with CyDyes, the subsequent steps had to be performed in the dark, as the CyDyes had to be protected from light in order to minimize photo-bleaching.

2.9.2. Protein detection and image analysis

After the second dimension, the gels were directly scanned while they were still in the low-fluorescence glass cassettes. The exterior of the glass plate was cleaned with ddH_2O and dried with KimWipes before the gel cassette was positioned on the Typhoon 9400 Variable Mode Imager (GE Healthcare, USA). Cy2 images were scanned at an excitation wavelength of 488ηm and a 520ηm band pass (BP) 40 emission filter. Cy3 images were scanned at an excitation wavelength of 532ηm and a 580ηm BP 40 emission filter. Cy5 images were scanned at an excitation wavelength of 633ηm laser and a 670ηm BP 30 emission filter. The gels were scanned with power adjustment with 1000μm (pixel size) resolution to allow full utilization of the image dynamic range before saturation. Subsequently, the gel images were converted to 16-bit TIF files before being transferred to the ImageQuant V5.0 (GE Healthcare, USA) for gel alignment and cutting. Afterwards, spots detection was processed using the DeCyder software V6.0 (GE Healthcare,

USA), by which the optimal settings for spot detection and exclusion were determined. The estimated number of spot detection was set to 2500 spots in most cases. Higher values usually resulted in the inclusion of spots of non-protein origin. Statistics and quantitation of protein expression were carried out in DeCyder Differential In-gel Analysis (DIA) module. Only those protein spots with over 3-fold changes in volume after normalization were considered as differentially expressed.

2.10. Matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS)

2.10.1. In-gel digestion

In-gel protein trypsin-digestion was performed according to the manual "Proteomic protocol for mass spectrometry" from the manufacturer of Bruker Daltonick, Germany. Briefly, the protein spots of interest on the silver-stained gel were excised and destained by washing with milli-Q water for 20min. This was followed by incubation in the destaining solution (30mM potassium ferricyanide and 64mM sodium thiosulfate) for 15min before being washed repeatedly with 25mM ammonium bicarbonate and until the gel pieces became colourless. Acetonitrile was then added to dehydrate the gel pieces. The surplus acetonitrile was removed when the gel pieces shrunk. Subsequently, the gel pieces were allowed to dry under vacuum. After drying, the gel pieces were incubated with reducing solution (25mM ammonium bicarbonate with 10mM dithiothreitol) at 56℃ for 45min. Afterwards, the gel pieces of interest were incubated with an alkylating solution (25mM ammonium bicarbonate with 55mM iodoacetamide) for 30min at room temperature and in the dark. Moreover, the gel pieces were washed with 50% (v/v) acetonitrile containing 50mM ammonium bicarbonate for 15min, and 100% acetonitrile was added to shrink them again. Freshly prepared 3μl of 10μg/ml trypsin (Promega) in 25mM ammonium bicarbonate was added to the gel pieces and the gel-solution mixture was allowed to incubate at 56℃ for 90min. Tryptic

fragments in the gel plugs were eluted the addition of 50% (v/v) acetonitrile and 1% (v/v) trifluoroacetic acid into the gel plugs. Diffusion of the peptide fragments into the solution was facilitated with ultra-sonication.

2.10.2. Peptide mass fingerprint (PMF) analysis

PMF fingerprints of the proteins to be identified were analysed using the Bruker Autoflex III MALDI-TOF-TOF mass spectrometer equipped with a reflector (Bruker Daltonic, Germany). In the MALDI-TOF MS reflector mode, ions generated by a pulsed UV laser beam were accelerated to a kinetic energy of 23.5kV. In the MALDI-TOF-TOF MS mode, precursor ions were accelerated to 8kV and selected in a time–ion-gate. The fragments were further accelerated by 19kV in the LIFT cell and their masses were analyzed after the ion reflector passage. Matrix and peptide samples were prepared according to the manual called "AnchorChip^{TM} Technology Revision 2.0" from Bruker Daltonics, Germany. Briefly, 1μl of matrix solution α-cyano-4-hydroxycinnamic acid, saturated in acetonitrile:0.1% TFA (1:1), was spotted on the anchor-chip target plate. After the matrix droplet was air-dried, sample solution of the same volume was applied, and allowed to dry again. The sample spot was washed with 1μl of 1% TFA for 30sec. The droplet was gently blown off the sample spot with compressed air. The resulting diffused sample spot was recrystallized using 1μl of ethanol:acetone:0.1%TFA (6:3:1). Finally, the target plate was placed into the Autoflex III machine for PMF analysis. Spectra were calibrated using Mass Calibration Kit (Bruker Daltonics, Germany) with an external calibration standard.

2.10.3. Database searching

The generated PMF spectrum was submitted to an in-house NCBI database before being searched against the Rodentia taxonomy using the Mascot search engine. Generally, fixed carbamidomethyl modification, variable oxidation modification, one mis-cleavage limitation and 200ppm of molecular weight tolerance were set as the searching criteria.

2.11. Western blot

After SDS-PAGE separation, the resolved protein samples were transferred onto a 0.45μm nitrocellulose (NC) membrane (Amersham Biosciences, US) at 100V for 1 hr at 4℃. Subsequently, the blotted membrane was blocked with 5% (w/v) fat-free milk powder in Tris Bufferd Saline (TBS) buffer at 4 ℃ overnight. The membrane was then washed with 20mM Tris-Tween Buffered Saline (TBST) for 30min in a 5-min cycle to remove any excess BSA. Anti-rat-whole-serum IgY (1:10000) in TBS buffer was added to the membrane and incubated for 1 hr at room temperature. The membrane was then washed with 20mM TBST for 30min in 5-min cycle. Afterwards, the membrane was probed with a 1:10000 dilution of horseradish peroxidase (HRG)-conjugated rabbit anti-IgY heavy and light chain antibody in TBS buffer. The membrane-buffer setup was allowed to incubate for 1 hr at room temperature. The membrane was again washed with 20mM TBST for 30min in a 5-min cycle. Finally, the blots were developed by using SuperSignal chemiluminescent substrate detection kit (Pierce, USA).

Chapter 3: Establishment of gastric cancer model in Wistar rat

3.1. Introduction

Gastric cancer in human is usually diagnosed at a late stage, and in such cases, the collection of serum samples at the early stage is difficult. Moreover, as elaborated in the earlier sections, there are tremendous variations in the human tissue/body-fluid samples collected for biomarkers discovery. These variations are very difficult to control and/or optimize. Furthermore, results of genomic studies only tell the chance of developing cancer to those who possess oncogene(s), but do not indicate the time of occurrence. Many changes in gene expression from benign and malignant tumors are due to post-translational modifications, which cannot be detected by DNA or RNA analysis. Therefore, proteomic studies using rat model provide a promising alternative. As described in the literature and reviewed in Chapter 1, gastric cancer induced in Wistar rat by administration of MNNG has proven to be a good model for human differentiated-type gastric cancer. This is because rat gastric cancer has histological features very similar to those of human cancer. In addition, rats with MNNG-induced cancers showed very similar responses to that of human cancer when treated with various tumor promoters and preventive agents (Manikandan *et al.*, 2008). On the basis of oligonucleotide microarray analysis, in rat gastric cancer, the genes that are involved in differentiated phenotypes of the stomach are down-regulated while those involved in extracellular matrix remodeling and immune responses are up-regulated. These expression profiles are similar to that of human gastric cancer (Abe *et al.*, 2003). Therefore, as elaborated earlier in Chapter 1, MNNG-intoxicated Wistar rats were used to establish a gastric cancer model for the discovery of relevant biomarkers in serum. All the relevant procedures were described in Chapter 2.

3.2. Results

3.2.1. Changes in body weight

The body weights of the experimental rats as well as the controls were recorded every 2 weeks and all exhibited a net gain in weight (Figure 3.1).

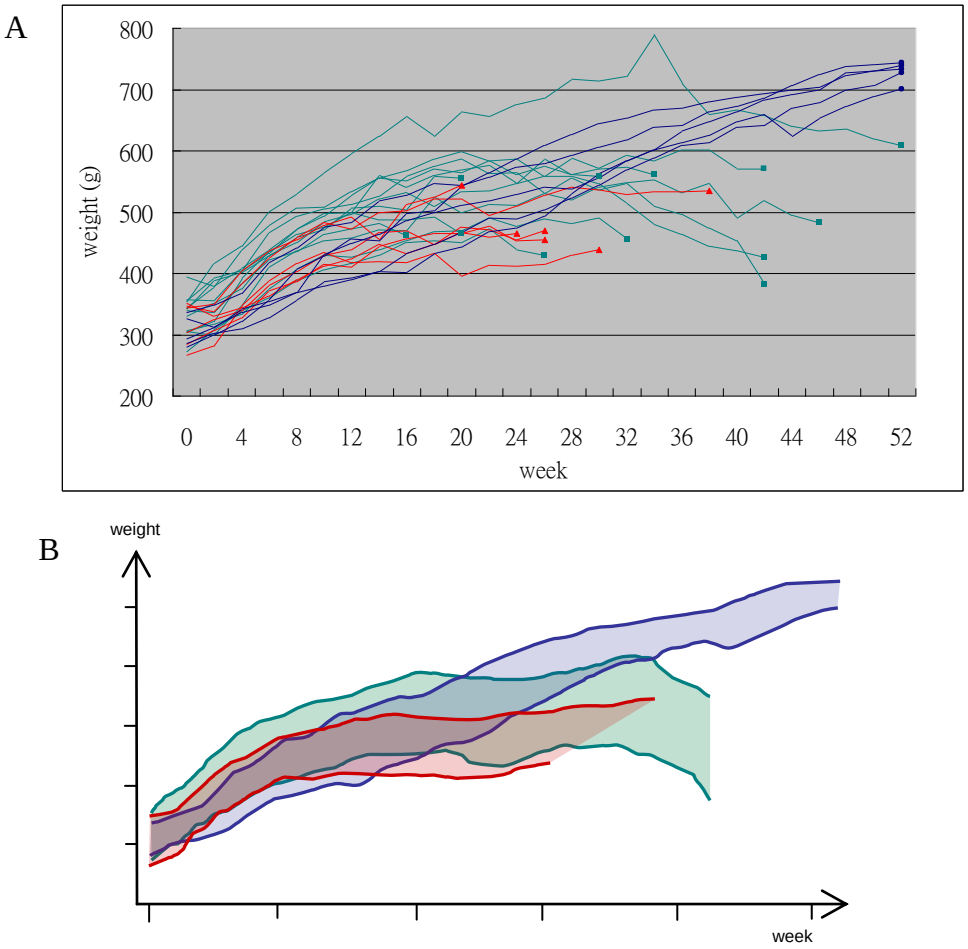


Figure 3.1: (A) Individual body weights of rats in the experiment group during gastric induction and in the control group. (B) A simplified graph showing the trends of rat weight in the various groups during gastric induction. (■: dysplasia rat; ▲: normal rat; ●: control rat)

However, a distinct loss of weight was observed in rats in the experimental group before the detection of dysplasia. This weight gain pattern agrees with the previous studies (Nagai *et al.*, 1984; Schwab *et al.*, 1997).

3.2.2. High resolution X-ray examination of the gastric region

In order to gauge on the induction of gastric cancer in the experimental rats (i.e. rats taking MNNG and NaCl), a high-resolution bio-imaging method was developed aiming to observe any possible macro-lesion in the stomachs of these rats. In order to obtain a sharp X-ray image of the rat stomach, different methods to inflate the rat stomach, such as air, different percentage of a fluorescence agent called gastrografin, and different parameters setting in the X-ray energy were tried. This part of the thesis was performed under the expert help from Dr. Patrick Lai and Dr. Phoebe Chan of the Department of Health Technology and Informatics, The Hong Kong Polytechnic University. Representative images are shown below (Figure 3.2).

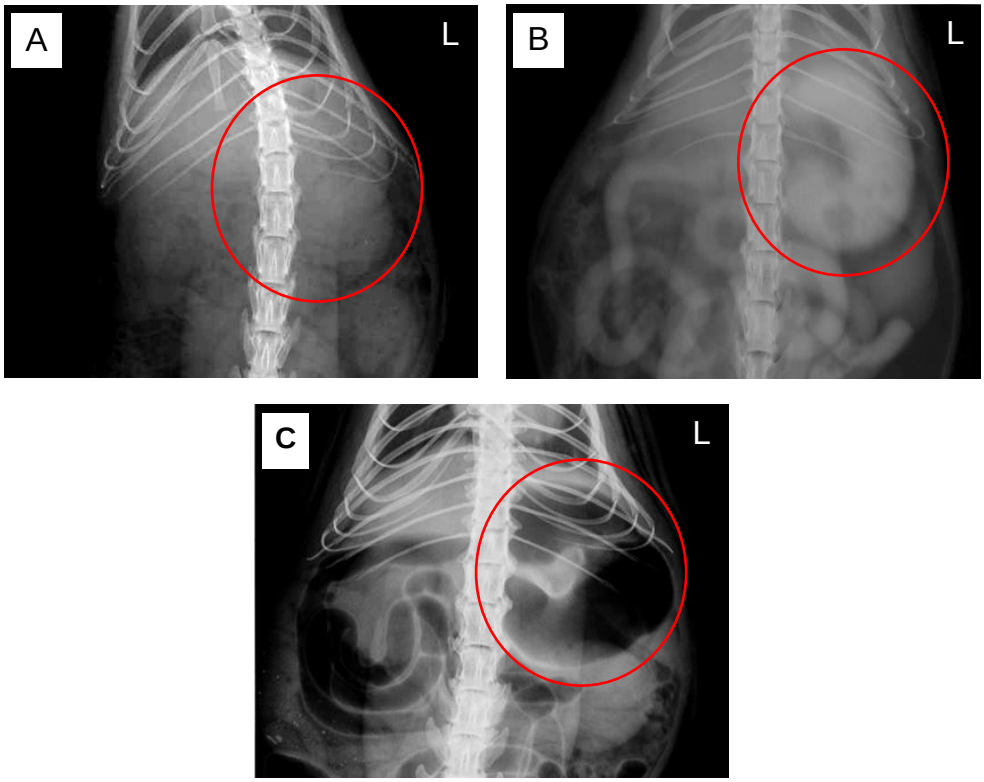


Figure 3.2: Representative high resolution X-ray images of rats in the stomach region (circled in red). (A) Rat stomach filled with 50 ml of air. (B) Rat stomach filled with 1% gastrografin. (C) Double contrast-enhanced rat stomach.

When air was alone used to inflate the stomach, its image was hardly seen in the X-ray film. For stomach filled with gastrografin only, the whole stomach image gave light and could been seen in the film. However, since the aim of taking the x-ray image was to observe if there was any irregularity taking place in the stomach wall, using air or gastrografin alone was not satisfactory. Therefore, double contrast with gastrografin and air together was employed. This gave a sharp out-lined stomach image, and the smoothness on the stomach wall could be easily seen.

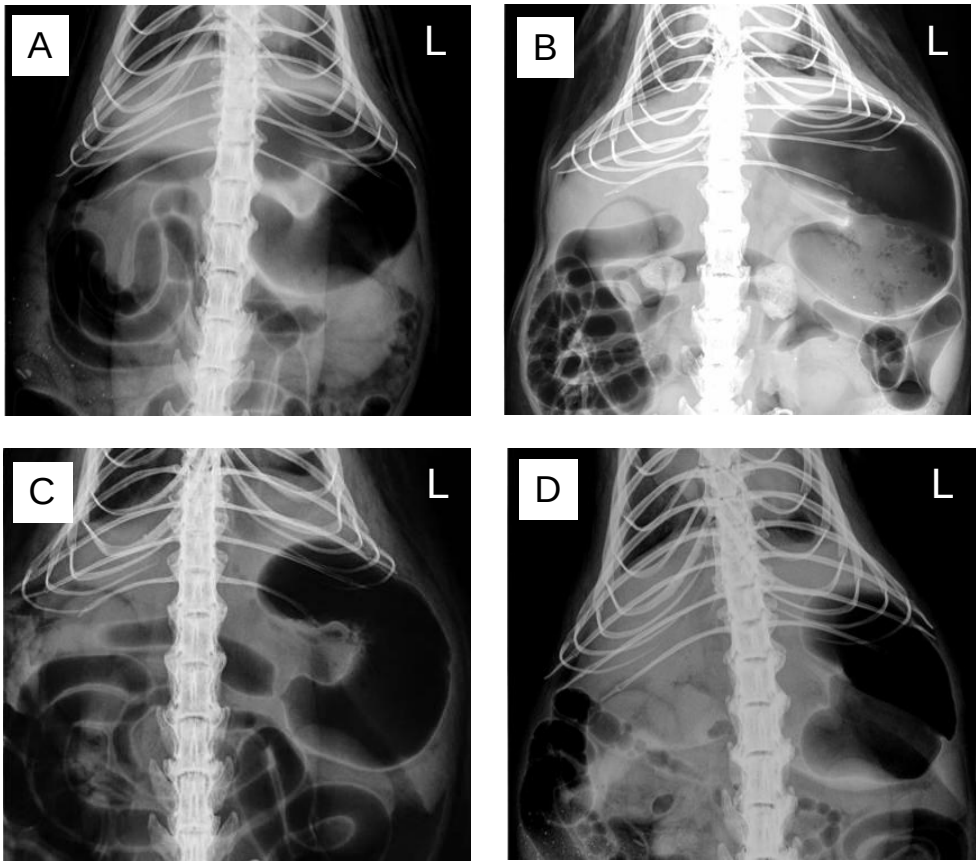


Figure 3.3: X-ray films of rats. (A) normal; (B) dysplasia; (C) dysplasia with early adenocarcinoma; (D) adenocarcinoma. The kidney shaped area on the left is the inflated stomach.

X-ray photos of rats were taken monthly. It was expected that taking X-ray photos for these rats would mimic provision of the endoscopic investigations. It was hoped that earlier stages of gastric carcinogenesis could be observed in the X-ray examination.

However, the x-ray films showed no differences among the stomachs obtained from rats with normal, dysplasia, early adenocarcinoma and adenocarcinoma. Therefore, X-ray study is not a good method to investigate the early stages of gastric carcinogenesis in rats.

3.2.3. Histological study

The presence of dysplasia was also investigated microscopically with tissue sections of stomachs. The main histological and cytological features of dysplasia include cellular atypia (nuclear pleomorphism, hyperchromasia, nuclear stratification, increase nuclear-cytoplasmis ratio, increased cytoplasmic basophilia, and loss of nuclear polarity); abnormal differentiation (lack/reduced numbers of goblet cells and Paneth cells in the metaplastic intestinal epithelium, and disappearance/reduction of secretory products from the gastric epithelium); and disorganized mucosal architecture (irregularity of crypt structure, back-to-back gland formation, budding and branching of crypts, and intraluminal and surface papillary growths).

After obtaining the stomachs from rats, they were processed as described in Chapter 2. After fixing and various staining procedures, histopathological study was undertaken and the stages of carcinogenesis were determined. Part of this work was kindly performed by Dr. Pak-kwan Hui, Consultant Pathologist and Chief of Service of Pathology Department, Kwong Wah Hospital.

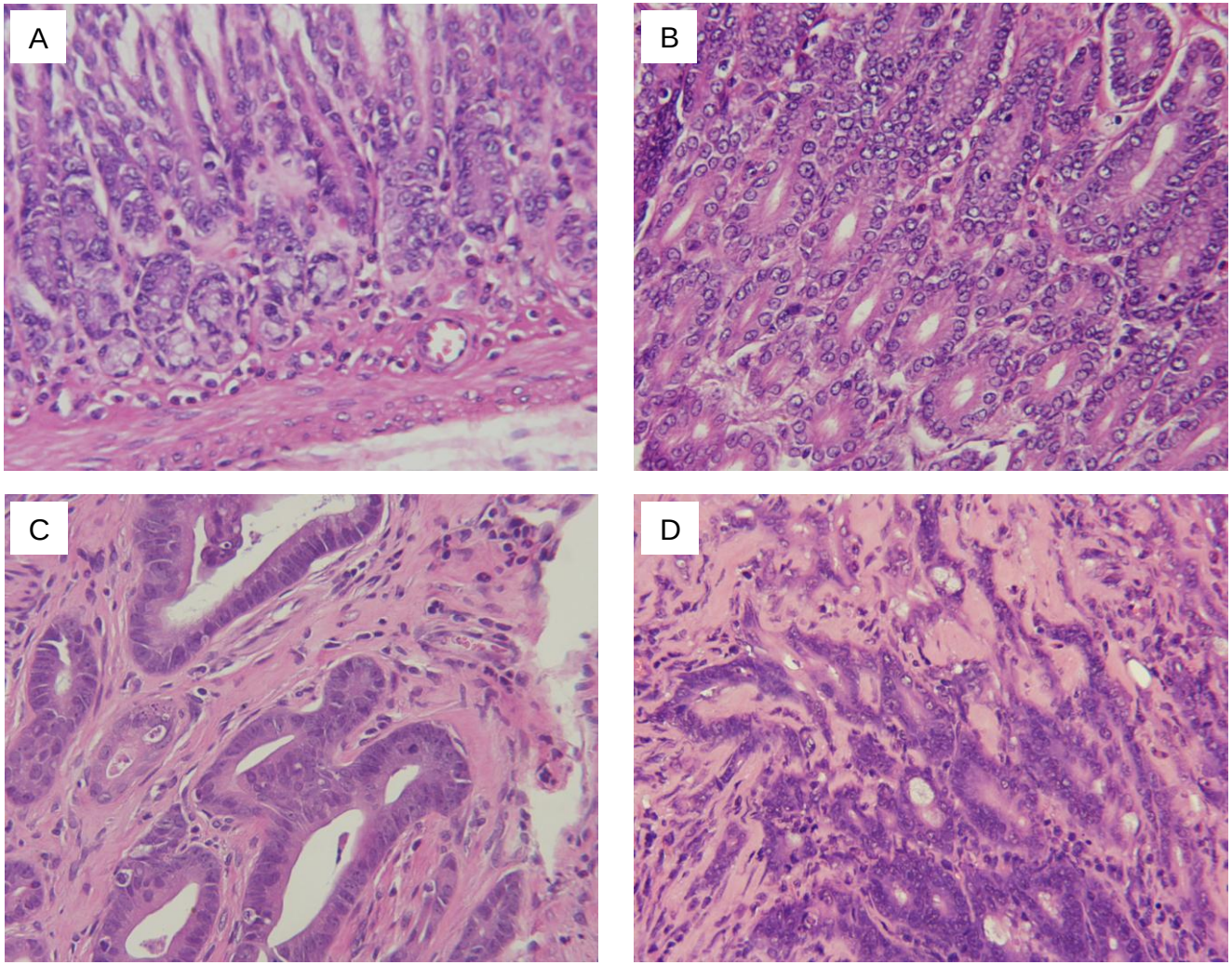


Figure 3.4: Microscopic photos showing stomachs of rats, H&E stain (250x).

Stomach A showed normal histology. For stomach B, there was a loss of gland differentiation. Foveolar glands showed increased nuclear size, vesicular nuclei and increased mitotic figures. Dysplasia occurred. Severe dysplasia occurred in stomach C, early adenocarcinoma showed dysplastic glands infiltrating the lamina propria and muscularis. In stomach D, there was a tumor mass composed of severely dysplastic mucin-secreting glands in keeping with well differentiated adenocarcinoma.

According to the universal classification of gastric carcinoma (Japanese Gastric Cancer, 1998), the health status of rats in the experimental group can be classified into four stages. The first stage was that the rats remained normal. At the second stage, dysplasia, the foveolar glands had increased nuclear size, vesicular nuclei, increased mitotic rate, and there was a loss of gland differentiation. At the third stage, besides severe dysplasia occurred, early adenocarcinoma also showed dysplastic glands infiltrating the lamina propria and muscularis. At the fourth stage, a tumor mass was present. The tumor mass was composed of severely dysplastic mucin-secreting glands in keeping with well differentiated adenocarcinoma.

From the histology results obtained in the rats, it was also found that dysplasia and adenocarcinoma mainly happened in the antrum/pylorus of the stomach, and sometimes also in the body of the stomach. In human, dysplasia is frequently found in the antrum/pylorus, and is also found in the fundus to some extent. Whereas in rat, dysplasia is confined to antrum/pylorus, and not found in fundus (Matsukura *et al.*, 1978).

3.2.4. Health status of rats after gastric cancer induction by MNNG

Two rats in the experimental group were excluded from the study, since they died soon after gastric cancer induction due to complications from anesthesia. Neither dysplasia nor adenocarcinoma was observed in the control rats throughout the experiment. For the experimental group, after feeding with 0.01% MNNG in drinking water, the health status of the 19 Wistar rats was classified into the 4 status from the histological findings of their resected stomachs.

Health status	Number of rats
Remain Normal	7
Dysplasia	8
Dysplasia + Early adenocarcinoma	3
Adenocarcinoma	1

Table 3.1: Classification of rats health status after gastric cancer induction with MNNG.

Starting from the 17th week after ingesting MNNG, 12 (63%) out of 19 rats were histologically confirmed to have developed dysplasia in the stomach, in which 3 of them had also developed to the stage of early adenocarcinoma, and 1 adenocarcinoma with a tumor; while 6 other rats remained normal. When compared with that described in the literature, our results are similar to the incidence rate of gastric cancer induction by MNNG with Wistar rats (55%-73%) (Bralow *et al.*, 1973; Martin *et al.*, 1974) and the time required to induce cancer. Some of the variations may due to the use of highly inbred strains of rats (Kobori *et al.*, 1976).

As the main aim of my project is to find candidate biomarkers for early gastric cancer, serum samples from the group with D&A will be used for later part of the study.

3.3. Discussion

Gastric cancer model in rats using MNNG is well established and provides reproducible data (Schwab *et al.*, 1997). Although there are fundamental differences between human and rat in clinical symptoms, for example, the vast majority of human gastric dysplasia cases are combined with chronic gastritis; and the dysplasia induced by MNNG is irreversible in contrast to the dysplasia in human stomach (Kunze *et al.*, 1979); MNNG-induced gastric tumors in Wistar rats show many of the common features observed in human stomach cancer. Firstly, at the gene level, over-expression of Bcl-2 and mutations in the tumor suppressor gene p53 have been documented in humans, as well as in MNNG-induced gastric cancer (Kaneko *et al.*, 2002). Secondly, the major risk factors associated with human stomach cancer, such as high salt and *Helicobacter pylori* infection, were again found to promote MNNG-induced gastric carcinogenesis (Newberne *et al.*, 1987). Besides, like human stomach tumors, MNNG-induced rat stomach carcinomas also exhibit infiltrative capacity and loss of differentiation (Abe *et al.*, 2003). These similarities between human stomach cancer and MNNG-induced gastric tumors provide the basis for analyzing the proteome changes during gastric cancer development.

According to the literatures examining Wistar rat with MNNG ingestion, dysplasia is first observed at around the 20^{th} week, and progresses to early adenocarcinoma between the 25^{th} and the 35^{th} week. After the 35^{th} week, the carcinoma begins to infiltrate deep into the muscularis propria and subserosa showing a definite morphological malignancy (Saito *et al.*, 1970; Kobori *et al.*, 1977).

The solid dysplasia-carcinoma sequence is not only present in the case of rats, dysplasia has also been accepted as a marker for increased cancer risk in human (Grundmann, 1975; Ming, 1979). In human, besides having dysplasia, there are a number of precancerous conditions and lesion of stomach, such as, atrophic gastritis, gastric ulcer, pernicious anaemia, gastric stumps, gastric polyps, and Menetrier's disease.

It should be noted that dysplasia, a controversial term in pathology, is used loosely and commonly. Taken overall, it is generally considered to have a positive predictive value for malignancy. Gastric epithelial dysplasia can be divided into three grades: mild, moderate, and severe. Mild dysplasia demonstrates minimally to slightly atypically arranged glands with slight dyscaryosis and slightly increased mitotic activity. Moderate dysplasia is distinguished from the mild one on the basis of an increased glandular distortion and cellular atypia. Severe dysplasia is characterized by highly distorted, atypically branched glands or tubules in back-to-back arrangement.

From my results, it could be seen that gastric dysplasia with or without adenocarcinoma was successfully established in 12 rats. Serum samples from these rats can then be used for proteomic investigations aiming to find candidate biomarkers.

3.4. Conclusion

An animal model with dysplasia was successfully established with 0.01% MNNG in drinking water with the administration of 10% NaCl in diet, which has been proved to increase the incidence of gastric cancer in rats (Ohgaki *et al.*, 1991). With the fact that there is a dysplasia –carcinoma sequence, the rats with dysplasia are chosen as the candidate for the serum biomarker discovery in the next stage of the project.

Chapter 4: Removal of high abundance proteins (HAPs) in serum samples

4.1. Introduction

The presence of a large number of proteins in serum makes it an excellent material for uncovering disease biomarkers. However, it also represents a tremendous analytical challenge because serum proteome contains a large number of proteins spanning a wide dynamic range of concentration. Because of the constraints of sample size for most advanced analytical techniques, the presence of a few very abundant proteins make the detection of low abundance proteins difficult. Therefore, removing the high abundance proteins (HAPs) is crucial in biomarker discovery.

Among the several commonly used techniques, the use of antibodies for the removal of albumin, immunoglobulins and other high abundance proteins seems to be the most efficient and specific method. However, as mammals are commonly used to raise antibodies, such as rats, rabbits and dogs, a lot of cross-reactivity induced interference occurred. Fortunately, a commercial immunodepletion column made up of avian polyclonal IgY antibodies has become available recently. The avian polyclonal antibodies have quickly gained popularity as the amount of antibodies produced is many times higher than that of mammals. This is because the immune response in an antibody-producing animal tends to increase as its phylogenetic difference with the animal used as the antigen source increases. Furthermore, chicken antibodies recognize more epitopes on mammalian protein than the corresponding rabbit antibody does (Schade *et al.*, 1996). This makes the production of antibodies against highly conserved mammalian proteins in chicken more successful than in other mammals (Tini *et al.*, 2002). In addition, a comparatively much lower amount of antigen is required to obtain high and long-lasting IgY titers in the yolk than in rabbits (Song *et al.*, 1985; Hatta *et al.*, 1993). Besides being found in avian serum, avian antibodies are in high abundance in egg yolk. Therefore,

collection of the avian antibodies is easily done by collecting eggs (Fischer *et al.*, 1996). This act causes no bleeding to the host. The immunization of chicken with complete Freund's adjuvant is well tolerated and produces no local inflammatory reaction (Ruan *et al.*, 2005). IgY production and purification are simple and efficient. There are various methods and technologies that can be employed to meet different requirements and needs, such as purity, yield, activity and the saving of time. Moreover, yolk antibodies are stable in cold, heat and acid for long periods (Camenisch *et al.*, 1999).

	Mammalian IgG	Avian IgY
Antibody sampling	Invasive	Non-invasive
Antibody yield	5mg IgG per mL (40mL blood)	50-100mg IgY per egg (5-7 eggs)
Antibody yield per month	200mg	~1500mg
Specific antibody yield	~5%	2-10%
Protein A/protein G binding	Yes	No
Interference with mammalian IgG	Yes	No
Interference with rheumatoid factor	Yes	No
Activation of mammalian complement	Yes	No

Table 4.1: Comparison of the characteristics of mammalian IgG and avian IgY (Schade *et al.*, 2005).

4.2. Results

4.2.1. IgY production

Demonstration that maternal antibodies are transferable from the immunized hen to its eggs was performed by Klemperer more than 100 years ago (Klemperer, 1893). The use of chicken egg yolk for polyclonal antibody production is well recognized and the IgY concentration in yolk is higher than in serum of the hen (Rose *et al.*, 1974; Larsson *et al.*, 1993; Woolley and Landon, 1995).

In our experiment, among the five immunized hens, two hens did not lay eggs at around the 10^{th} week after the first immunization with the immunogens (rat serum). Hence, these hens were not used in the experiment. The other three hens were kept for 39 weeks and a total of 308 immunized eggs were collected. This work was performed in the Laboratory Animal Services Centre (LASEC) of The Chinese University of Hong Kong.

4.2.2. IgY purification

As mentioned earlier, there are several commonly used protocols for purification of IgY. In order the select the most suitable method for IgY purification, three egg yolks were mixed, separated into 3 identical sets and purified using three of the most commonly used methods. The three methods to be evaluated were: straight ammonium sulphate precipitation, straight sodium sulphate precipitation, and a commercially available method called the EGGstract kit. After performing the recommended procedures, the efficiencies of the 3 different methods were evaluated using non-reducing SDS-PAGE. Non-reducing gel was used because the major contaminants of IgY purification are approximately of 70kDa and 30kDa, which are very close to the heavy and light chains of IgY, which are 67kDa and 23kDa, respectively. Proteins resolved were visualized by silver stain (Figure 4.1).

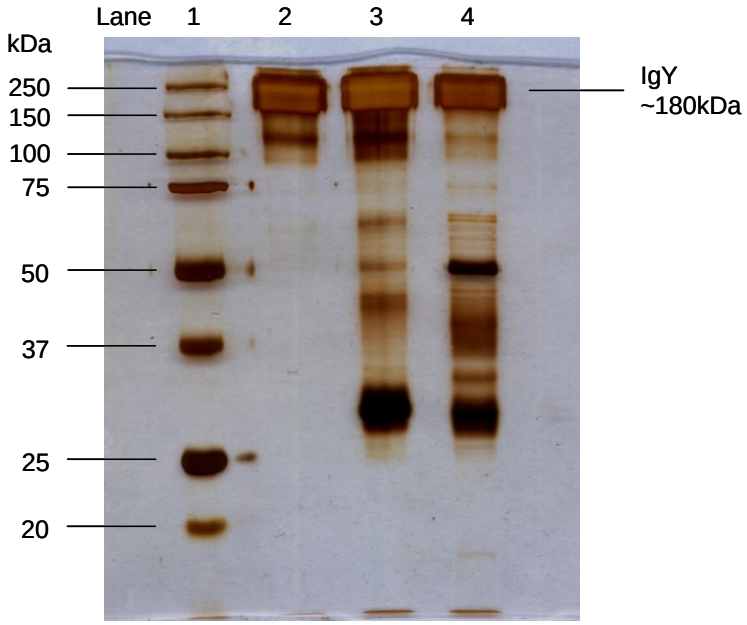


Figure 4.1: Non-reducing SDS-PAGE showing the purity of IgY (MW ~180kDa), obtained with different purification methods. 1μg of protein is loaded into each lane. Lane 1 is molecular weight markers: Lanes 2-4 represented proteins purified using the EGGstract kit, ammonium sulphate precipitation and sodium sulphate precipitation respectively.

As shown in Figure 4.1, proteins presented in Lane 2 had much less contaminating proteins, especially in the 30-50kDa range. Using anti-IgY antibodies, the protein band at around 180kDa is confirmed to be IgY. Therefore, it seems that the efficiency of purification using EGGstract kit is much more superior when compared to the other two methods.

4.2.3. IgY recovered in eggs collected during the experimental period

It is known that specific serum antibodies are transported to the yolk with a delay of about three to six days. Yolk IgY concentration also varies significantly among individuals and genetic lines or breeds (Schade *et al.*, 2005). Further analysis of the IgY isolated from our experiments found that IgY concentration varied in the range of 1.5-4mg/ml throughout the whole experimental period. Figure 4.2 showed the average concentrations of IgY found. Similar levels have been reported by others (Bizhanov and Vyshniauskis, 2000).

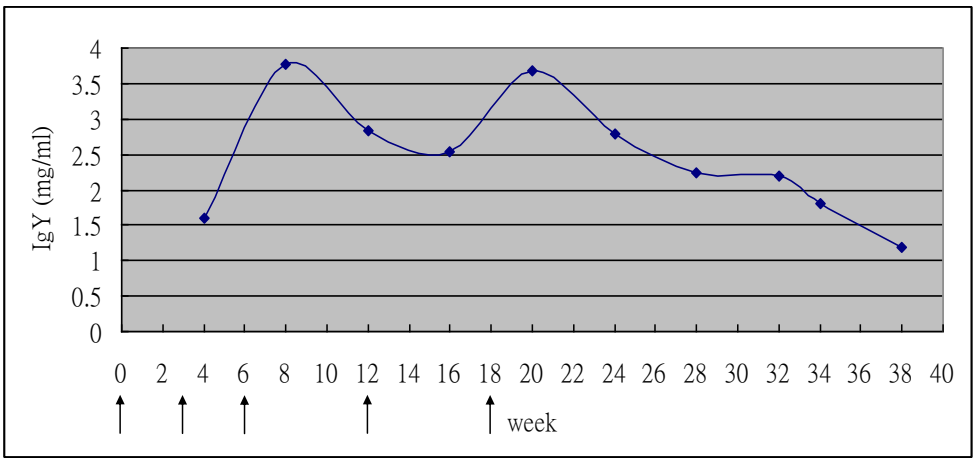


Figure 4.2: IgY concentration obtained from eggs in different weeks. (↑antigen injection)

4.2.4. Specificity of the IgY collected

The IgY raised from the hens with rat serum (as antigen) injection was investigated using Western blot against rat serum. For comparison, the normal IgY from hens without antigen injection was also probed as control.

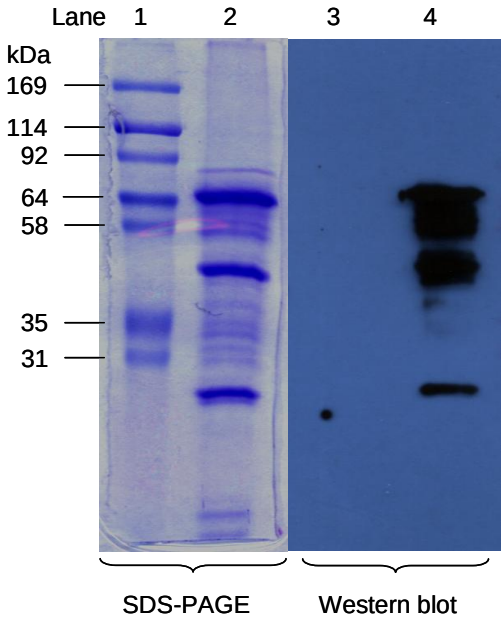


Figure 4.3: 10% SDS-PAGE and Western blot using rat serum as sample. Lane 1 is molecular weight markers. Lane 2 is 5μg of rat serum stained with Coomassie Blue. Lane 3 is a corresponding Western blot probed with IgY collected from a control egg. Lane 4 is a corresponding Western blot probed with IgY raised in immunized hens and at the dilution of 1:10000.

From the Western blot, it can be seen that Lane 3 (which was probed with control IgY) has no band while Lane 4 (which was probed with IgY isolated from immunized eggs) has many bands indicating positivity. These results showed that the IgY raised in immunized hens was specific to the rat serum. Furthermore, some of the abundant proteins seen in the SDS-PAGE with Coomassie blue staining are also present in high abundance in the Western blot. These results also indicate that proteins that were detected by the IgY are proteins of high abundance in the normal rat serum. The identities of proteins that the hens raised against will be indicated in the later part of this chapter.

4.2.5. Binding of the target antigens to the IgY coupled column with the aim to immunodeplete the major proteins from rat serum samples

Active IgY purified from the eggs immunized against rat whole serum were pooled together and coupled to cyanogen bromide-activated sepharose using procedures recommended by the manufacturer. The amount of IgY before being loaded onto the sepharose and the amount of IgY remained in the supernatant after coupling were measured by Bradford test. The amount of IgY bound to the sepharose can then be estimated.

Volume of CNBr-activated sepharose used (mL)	Active IgY mixed with the CNBr-activated sepharose (mg)	Unbound IgY found (mg)	Estimated amount of active IgY that binds to the CNBr-activated sepharose (mg/mL)
30	352	8	11.47

Table 4.2: Amount of active IgY coupled to the CNBR-activated sepharose.

Crude rat serum (3mg) was loaded into the IgY affinity column (fabricated as described above) with a flow rate of 1mL/min. Four mL fractions were collected. The HAPs should have been bounded by the IgY when they came through the column. The less abundant proteins will come through the column un-absorbed. They are the “flow through” fractions. On the other hand, the bound HAPs could be eluted by acid elution and they are the “bound” fractions. A representative chromatography using the active IgY affinity column to remove the HAPs is shown in Figure 4.4.

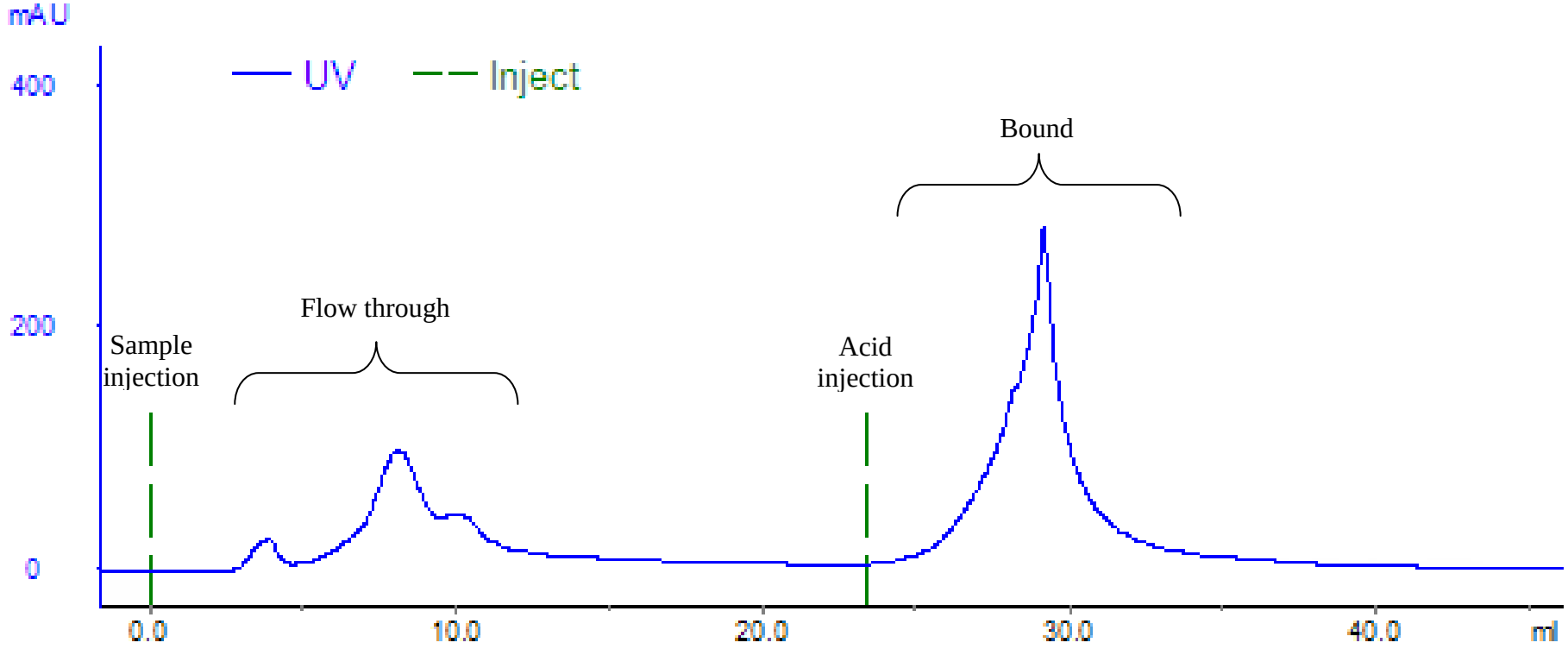


Figure 4.4: Representative chromatographic profile of immunodepletion of rat serum using the customized IgY column designed to absorb the high abundance proteins (HAPs).

4.2.6. Reproducibility of the IgY affinity column

As differentially expressed proteins were found by matching images of gels obtained from different samples with different runs, it is crucial that the IgY affinity absorption column has high reproducibility. Reproducibility measures the ability to repeatedly achievable by same performance in the column chromatographic process, and is also an indication of the endurance of the column with the ability to be generated without loss of either capacity or specificity. In order to test the reproducibility, twenty identical aliquots of rat serum samples were loaded to the IgY column sequentially. Both the "flow-through" and the "bound" fractions were collected with a standardized condition for binding, elution, and analyzed by SDS-PAGE. The results of some of the representative fractions are shown in Figure 4.5. they indicate that the IgY bound sepharose column gave highly reproducible results and the column was reusable.

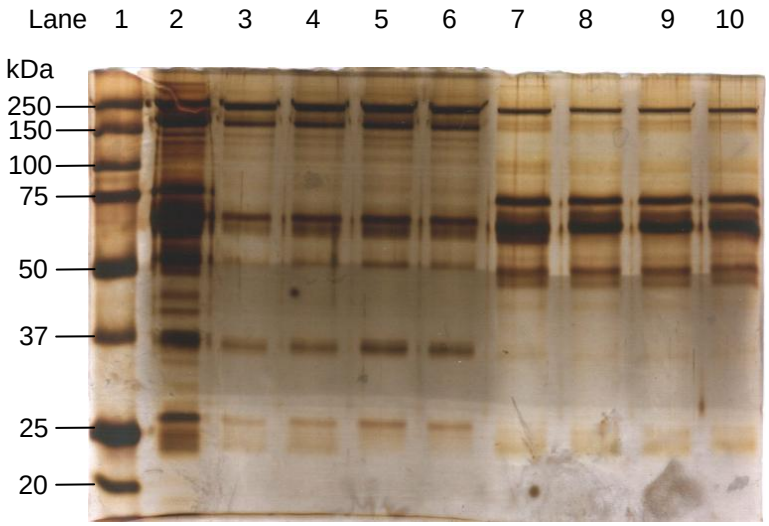


Figure 4.5: SDS-PAGE of selected samples (5^{th}, 10^{th}, 15^{th} 20^{th}) in the "flow-through" and "bound" fractions (out of the 20 runs in successions) are shown. Lane 1: is the Molecular weight markers, Lane 2: Crude serum; Lane 3: "Flow through" fraction in the 5^{th} cycle; Lane 4: "Flow through" fraction in the 10^{th} cycle; Lane 5: "Flow through" fraction in the 15^{th} cycle; Lane 6: "Flow through" fraction in the 20^{th} cycle; Lane 7: "Bound" fraction in the 5^{th} cycle; Lane 8: "Bound" fraction in the 10^{th} cycle; Lane 9: "Bound" fraction in the 15^{th} cycle; Lane 10: "Bound" fraction in the 20^{th} cycle.

4.2.7. Depletion efficiency of the column

This immunodepletion column was designed to remove most of the high abundance proteins in serum. The depletion efficiency of this column could be easily assessed by the total amount of proteins left behind after the depletion process. The crude serum was either passed through the column once (single depletion) or twice (double depletion as the flow through from the first run was passed onto the column a second time). The total protein amounts after both single depletion and double depletion were measured. The results are shown below:

	Crude serum (μg)	Depleted serum (μg)	Depletion efficiency
Single depletion	3000	300	90%
Double depletion	3000	30	99%

Table 4.3: Protein amount regarding the depletion process.

The results show that the depletion efficiency of the column is up to 99% (of the original protein concentration) after double depletion, while it is about 90% for single depletion. Among all the high abundance proteins, albumin is the most abundant. Therefore, the removal of albumin was also investigated to provide an estimation of the HAPs removal process. The results of Western blotting against albumin are shown in Figure 4.6.

After single run immunodepletion by the IgY affinity column, the amount of albumin removed was around 50x less than that in the crude serum. For doubly immunodepleted serum, albumin seems to be totally absent in the sample (position highlighted by the white circle). Using the 500x diluted crude serum sample results (Lane 7), it was estimated that the albumin present (if there is) must be less than 1/500, less than 0.2% of that of the original.

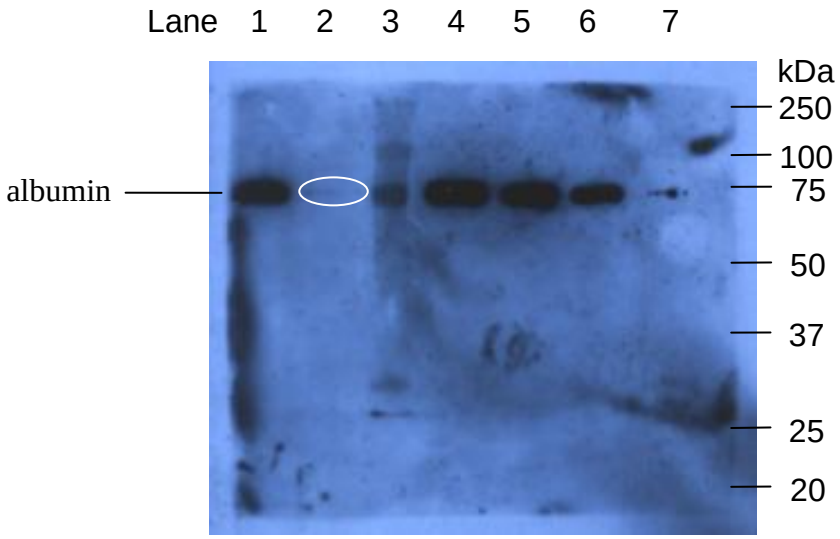


Figure 4.6: A representative western blotting result of immunodepletion of rat serum using anti-albumin antibody as the primary antibodies at the dilution of 1:2000. Lane 1: Depleted serum (1µl) from single depletion; Lane 2: Depleted serum (1µl) from double depletion; Lane 3: Molecular weight markers; Lane 4: Serial dilution (50x) of crude serum (without the HAPs depletion process); Lane 5: Serial dilution (100x) of crude serum (without the HAPs depletion process); Lane 6: Serial dilution (200x) of crude serum (without the HAPs depletion process); Lane 7: Serial dilution (500x) of crude serum (without the HAPs depletion process).

4.2.8. Identities of the high abundance proteins (HAPs) removed by the column

It is of interest to find the identities of HAPs removed by the IgY-affinity column. Therefore, bound proteins were eluted from the IgY column after the first run was resolved by 2-DE. A representative 2-DE gel of one of the bound-and-subsequently-eluted protein fractions was shown in Figure 4.7. The proteins of interest (denoted as 1-16) were identified using MALDI-TOF mass spectrometry.

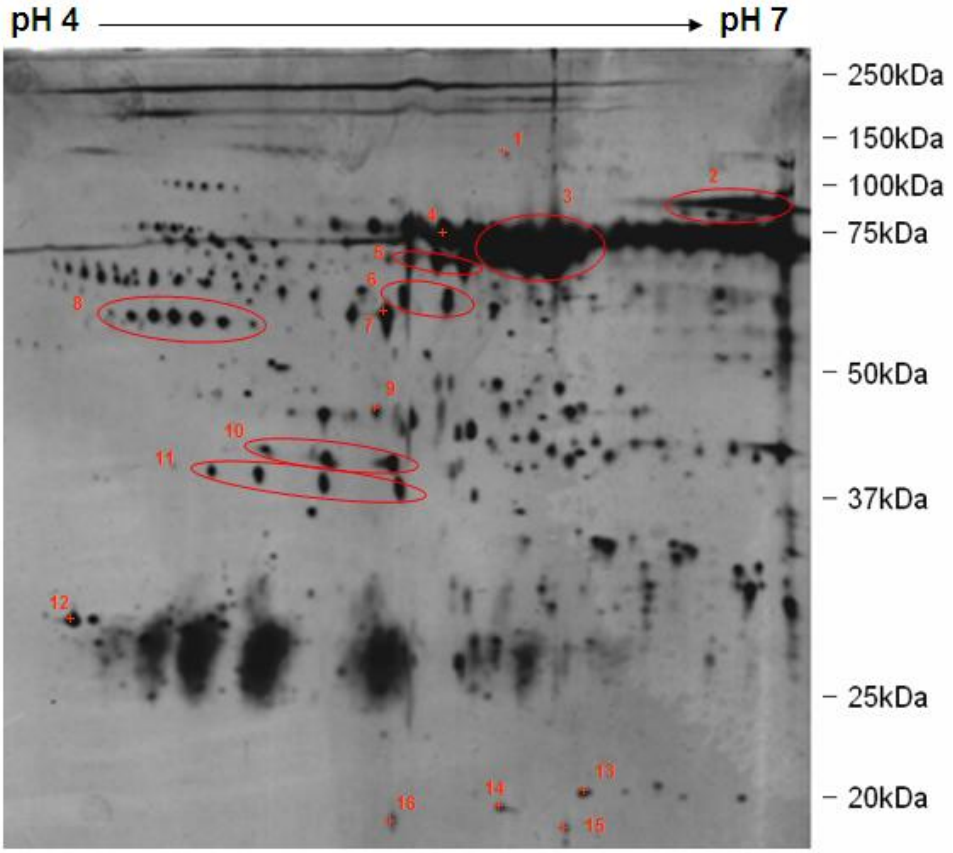


Figure 4.7: Representative 2-DE gel with 40μg protein from the bound-and-subsequently-eluted fraction.

Proteins identities are shown in Table 4.4, and details of the sequence coverage and peptide mass fingerprints are attached in Appendix I. All of these proteins are known to be existed in high abundance in normal serum in both rats and humans.

Spot no.	Name of protein	Accession number	Mr (kDa)/pI	Mascot score*	Sequence coverage (%)*
1	Coagulation factor 2	gi|12621076	71.792/6.28	75	17
2	Transferrin	gi|61556986	78.512/7.14	207	35
3	Albumin	gi|19705431	70.67/6.09	326/94	49/2
4	Hemopexin precursor	gi|122065203	52.06/7.58	151/53	28/2
5	Apolipoprotein H	gi|57528174	39.743/8.58	105/33	35/4
6	Fetuin beta	gi|17865327	42.361/6.71	114	17
7	Vitamin-D binding protein	gi|203941	55.079/5.76	91/30	24/1
8	α-1-antitrypsin	gi|203063	45.978/5.7	98/62	24/2
9	Complement component 3	gi|554423	32.414/5.73	139/75	39/3
10	α-1-macroglobulin	gi|81872093	168.388/6.46	107	8
11	Haptoglobulin	gi|60097941	39.052/6.1	122/70	27/3
12	C-reactive protein	gi|8393197	25.737/4.89	97	41
13	Plasma glutathione peroxidase	gi|6723180	25.653/8.26	88	29
14	Plasma retinol-binding protein	gi|132407	23.547/5.69	46/92	4/8
15	Apolipoprotein M	gi|9506391	21.841/5.73	145	36

Spot no.	Name of protein	Accession number	Mr (kDa)/pI	Mascot score	Sequence coverage (%)
16	Apolipoprotein E	gi|1703338	35.788/5.23	131	29

*Data in BLACK are MS results; data in RED are MS/MS results

Table 4.4: Identities of proteins that were depleted by the immunodepletion column (denoted as 1-16 in Figure 4.7).

4.2.9. Proteome maps of rat serum before and after immunodepletion

Comparison of rat serum proteomes before and after immunodepletion was carried out according to the procedures described in Chapter 2 with passive rehydration with 18cm, pH 4-7 strip.

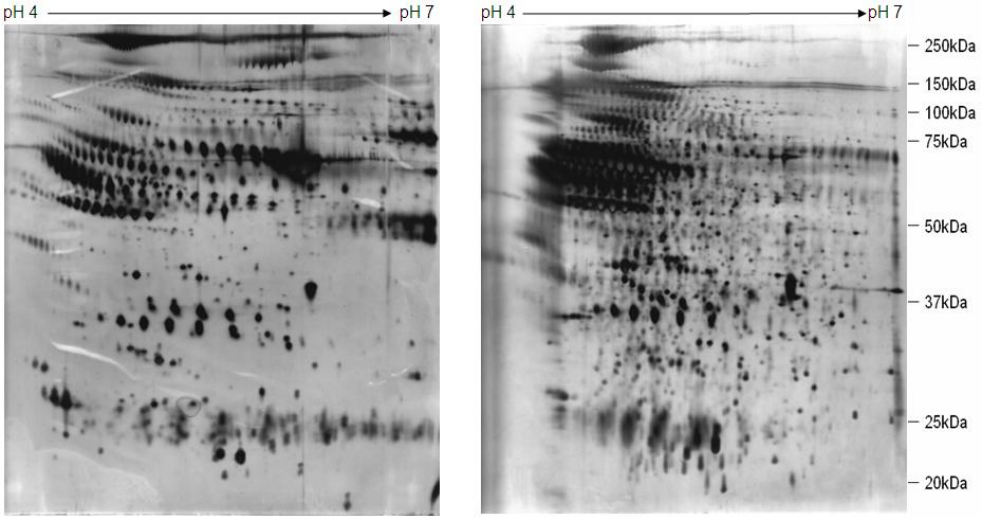


Figure 4.8: Representative 10% 2-DE gels of 40µg of crude rat serum (left) and double depleted rat serum (flow through) (right) with silver stain.

As shown in Figure 4.8 above, the proteomes of serum samples before and after double depletion were very different, showing that most of the high abundance proteins that masked the less abundant proteins in the 2DE gel had been removed. Approximately 700 and 1000 spots were detected in the gels of crude rat serum and double depleted rat serum respectively. This enables "easier" detection of the low abundance candidate cancer biomarker proteins during comparative proteomic studies.

4.3. Discussion

Disease biomarkers typically appear at low concentrations, while serum proteome has a large dynamic range of individual protein concentrations (ten orders of magnitude). Therefore, identification of low-copy-number proteins of interest is difficult due to the confounding presence of higher-abundance proteins. Therefore, to enhance detection and quantification of proteome components, it is necessary to prepare protein samples via specific prefractionation or depletion methods. These procedures aim to separate highly abundant members from those in low abundance before detailed comparative proteomic studies.

A classical depletion method is the use of the hydrophobic dye Cibacron blue, a chlorotriazine dye which has high affinity for albumin. However, Cibacron blue columns show low specificity with removal of significant amounts of non-targeted proteins (Leatherbarrow and Dean, 1980; Gianazza and Arnaud, 1982). This strategy of removing albumin is still being used because of its relatively low cost (Ahmed *et al.*, 2003; Li and Lee, 2004). The second most abundant proteins in the serum are the immunoglobulins. Immunoaffinity separation of proteins has proven to be one of the most developed and effective approaches for this purpose. Protein A/G columns are typically used to deplete them. Assuming most of the albumin and immunoglobulins have been removed, it is still necessary to deplete more other protein in order to enhance the detection of very low abundance proteins. More recently, the human Multiple Affinity Removal System (MARS) from Agilent Technologies is available commercially. Six high abundant proteins: albumin, transferrin, IgG, IgA, antitrypsin, and haptoglobin can be depleted. A subsequent commercial immunodepletion column is MIXED 12 from GenWay Biotech (USA). It can remove the 12 most abundant proteins including albumin, transferrin, IgG, IgA, anti-trypsin, and haptoglobin. Up to now, the most powerful depletion kit available commercially is the ProteoPrep 20 Plasma Immunodepletion Kit from Sigma-aldrich. It can deplete 20 abundant proteins: albumin, transferrin, fibrinogen, IgG, IgA, IgD, IgM, α-1-antitrypsin, α-2-macroglobulin,

complemtent C3, complemtent C4, complemtent C1q, haptoglobin, apolipoprotein A1, apolipoprotein A2, apolipoprotein B, acid-1-glycoprotein, ceruloplasmin, prealbumin, and plasminogen. These commercially available immunodepleted columns utilized IgY antibodies that were raised in egg-laying hens.

Avian polyclonal IgY antibodies as a choice to make up immunodepletion column has been widely used recently (Huang *et al.*, 2005). The immune response in an antibody-producing animal tends to increase as its phylogenetic difference with the animal used as the antigen source increases. Thus, chicken antibodies recognize more epitopes on a mammalian protein than the corresponding rabbit antibody does, making it more advantageous to use IgY in immunological assays of mammalian proteins. This is especially true when the antigen is a highly conserved protein, like hormone. Polyclonal antibodies are preferred over monoclonal antibodies as they are much more likely to recognize and deplete most forms of the targeted abundant proteins, including molecules with different post-translational modifications and proteolytic fragments (Hoffman *et al.*, 2007). Therefore, avian polyclonal IgY antibodies was employed for making the immunodepletion column. Given the nature of the composition of serum, which the high abundance proteins make up 90% of the serum, serum from rat was directly used as an antigen to inject into the hens. Due to the different antigenicity of the serum proteins, the antibodies raised in hens were mainly against the high abundance proteins as shown above. The eggs that contained antibodies with high titre were sorted out, purified and utilized to make the immunodepletion column. High abundance proteins in the serum sample bound to their corresponding antibodies when loaded into the column, leaving the relatively low abundance proteins retained in the "flow through" fraction. The HAPs originally bound to the column were eluted out as the "bound' fraction when there was change of pH in the environment (buffer). With the IgY affinity column, every serum sample was applied to the column and the fractions were collected with identical procedure.

Regarding the reproducibility of the immunodepletion column, 1D gel was run on consecutive depleted samples. The protein patterns of both the "flow through" and "bound" fractions of sample in the 5^{th}, 10^{th}, 15^{th} and 20^{th} cycles showed an incredibly reproducing pattern. This indicated that the immuodepletion column chromatograph was highly reproducible and it was suitable to apply it for comparative quantitative biomarker discovery studies. The depletion efficiency is another crucial factor affecting the biomarker discovery process. It is common to perform a second time depletion procedure when using commercial depletion column. A double depletion is therefore also performed in this project. Regarding the total protein amount, the double depletion allows up to 99% depletion efficiency. Albumin, as the most abundant member in the serum, makes up to around 80% of the serum, was chosen as another criterion to determine the column depletion efficiency. The amount of albumin in the double depleted sample is more than 500 times less than that in the crude serum, which corresponds to 99.8% depletion efficiency. A totally different proteome of the double depleted serum in comparison with untreated sample, with enhanced visualization of the relatively low abundance proteins indicated the success of the immunodepletion column. Most of the abundance proteins in the serum have been effectively removed by the immunodepletion column. These include albumin, α-1-macroglobulin, α-1-antitrypsin, apolipoprotein E, apolipoprotein H, apolipoprotein M, transferrin, complement component 3, haptoglobulin, hemopexin precursor, vitamin-D binding protein, plasma retinol-binding protein, C-reactive protein, coagulation factor 2, plasma glutathione peroxidase and fetuin beta.

4.4. Conclusion

Regardless of the success of singly and/or doubly protein depletion schemes, biomarker detection is usually facilitated by the removal of as many high abundance proteins as possible. An antibody-based (IgY) affinity column has been successfully developed to deplete the high abundance proteins from serum.

Chapter 5: Comparison of different sample preparation methods and identification of biomarker candidates

5.1. Introduction

Besides the challenge of detection limit and the dynamic range issues of using 2-DE with different types of stains, including silver and fluorescent stains, in the biomarker discovery process, the normalization of samples after the depletion process remains controversial. Conventionally, researchers have been labelling proteins after the samples had been depleted. However, there are concerns about variations between the different depletion runs of the same and/or different samples. Therefore, comparisons of 2-DE chromatograms visualized by silver staining and CyDyes fluorescent tags with sample normalization before and after depletion were carried. The results are presented in this chapter. Identification of differentially expressed proteins in the "best" samples preparatory steps which represent possible biomarker candidates of gastric dysplasia will also be performed. The results will also be presented in this chapter.

5.2. Results

5.2.1. Optimization of 2-DE gel

In order to perform comparative proteomics, proteins in the serum proteome need to be well resolved. In the preliminary experiments, proteins in the serum proteome was first resolved in an 18cm IPG strip with pH 3-10 as shown in Figure 5.1. It was found that most of the proteins were concentrated at the acidic region. Therefore, 2-DE was performed again using an 18cm IPG strip with pH 4-7. The results are shown in Figure 5.2. Proteins on the gel appeared to be well resolved and properly focused. Therefore, IPG strips of 18cm and pH 4-7 were adopted in all subsequent experiments.

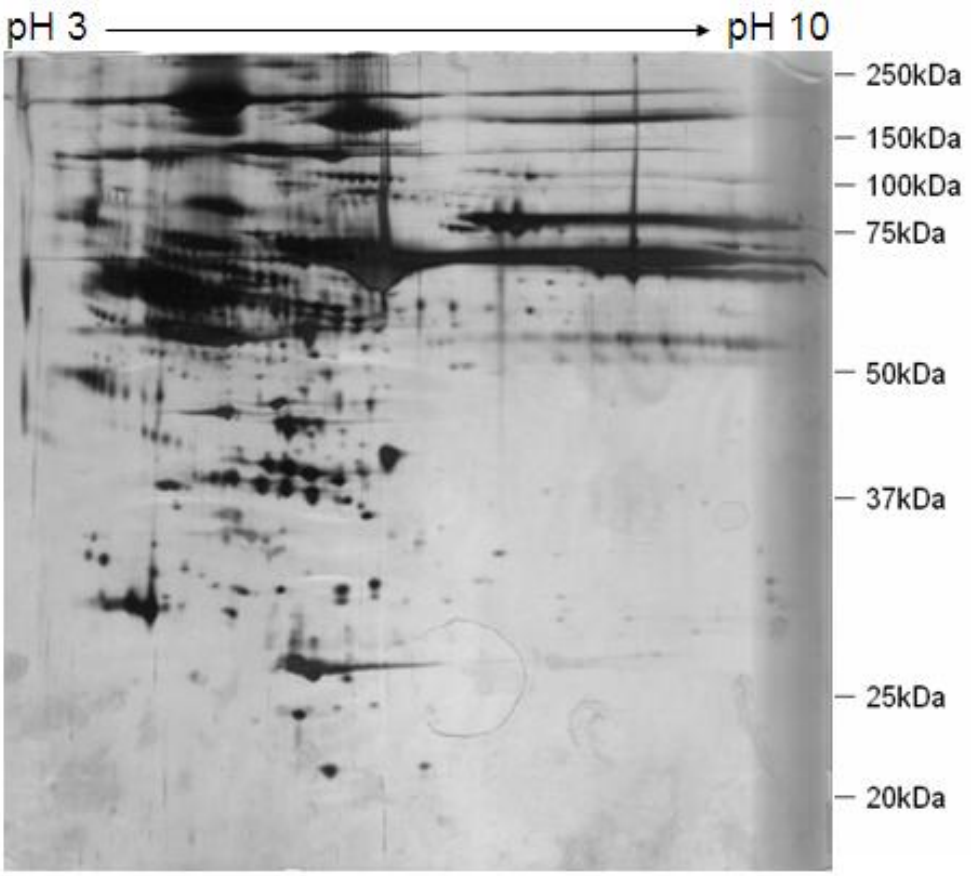


Figure 5.1: Representative 2-DE of the proteome of 40µg rat serum sample resolved in an 18cm IPG strip with pH 3-10, 10% SDS-PAGE was used as the second dimension. (IEF protocol was described in Section 2.8.1.)

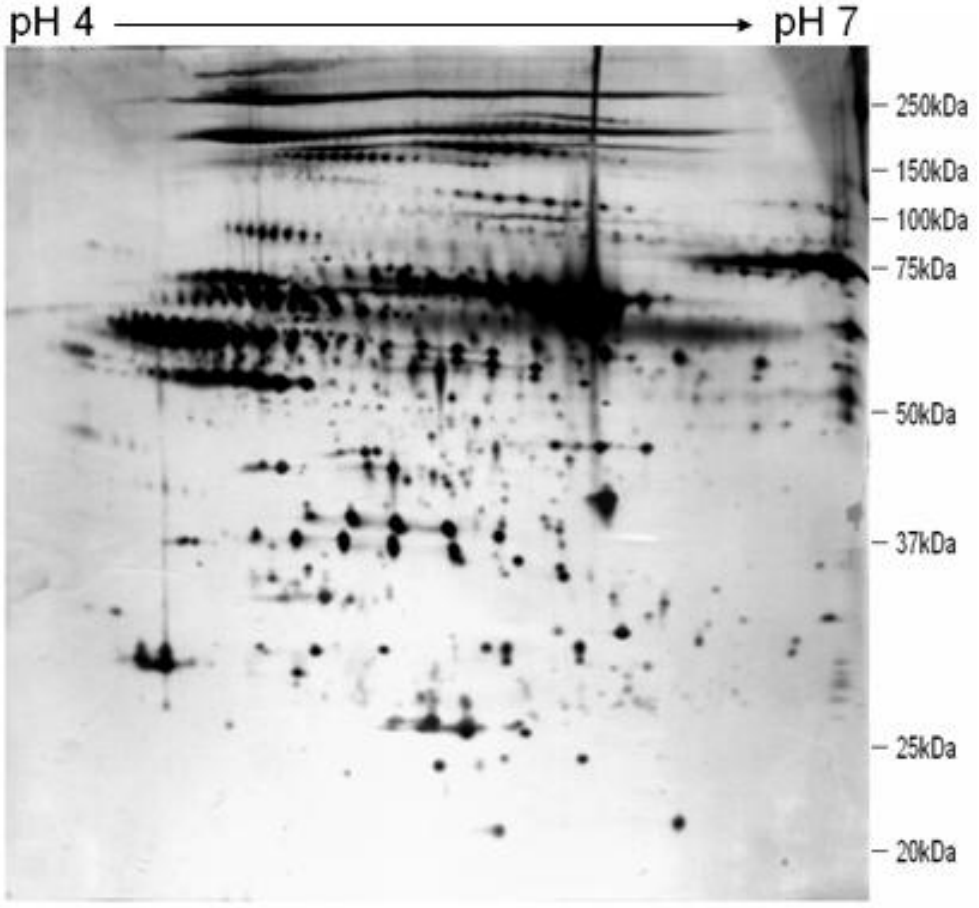


Figure 5.2: Representative 2-DE of the proteome of 40µg rat serum sample resolved in an 18cm IPG strip with pH 4-7, 10% SDS-PAGE as the second dimension. (IEF protocol was described in Section 2.8.1.)

5.2.2. 2-DE with silver staining

The silver staining visualization method with its detection limit of around 1-10ng is commonly used in 2-DE. In my project, serum samples underwent doubly depletion and 40μg protein were resolved on a linear pH 4-7 IPG strip (18cm) in the first dimension, followed by a 10% SDS-PAGE as the second dimension. With silver staining method, approximately 1000 spots were detected by the Melanie software.

In comparing serum proteomes of normal control and dysplasia with adenocarcinoma (D&A), all the experiments were performed in six replicates. After silver staining, differential protein spots detected by the Melanie software with more than 3 folds difference were noted and counted. A total of 15 spots were found to be differentially expressed by at least 3 folds. Among the 15 spots, 11 were successfully identified and they belonged to 6 protein species. Identities of the 4 spots remain unknown. The protein spots were numbered and are shown in Figure 5.3. Details of the sequence coverage and peptide mass fingerprints are attached in Appendix II.

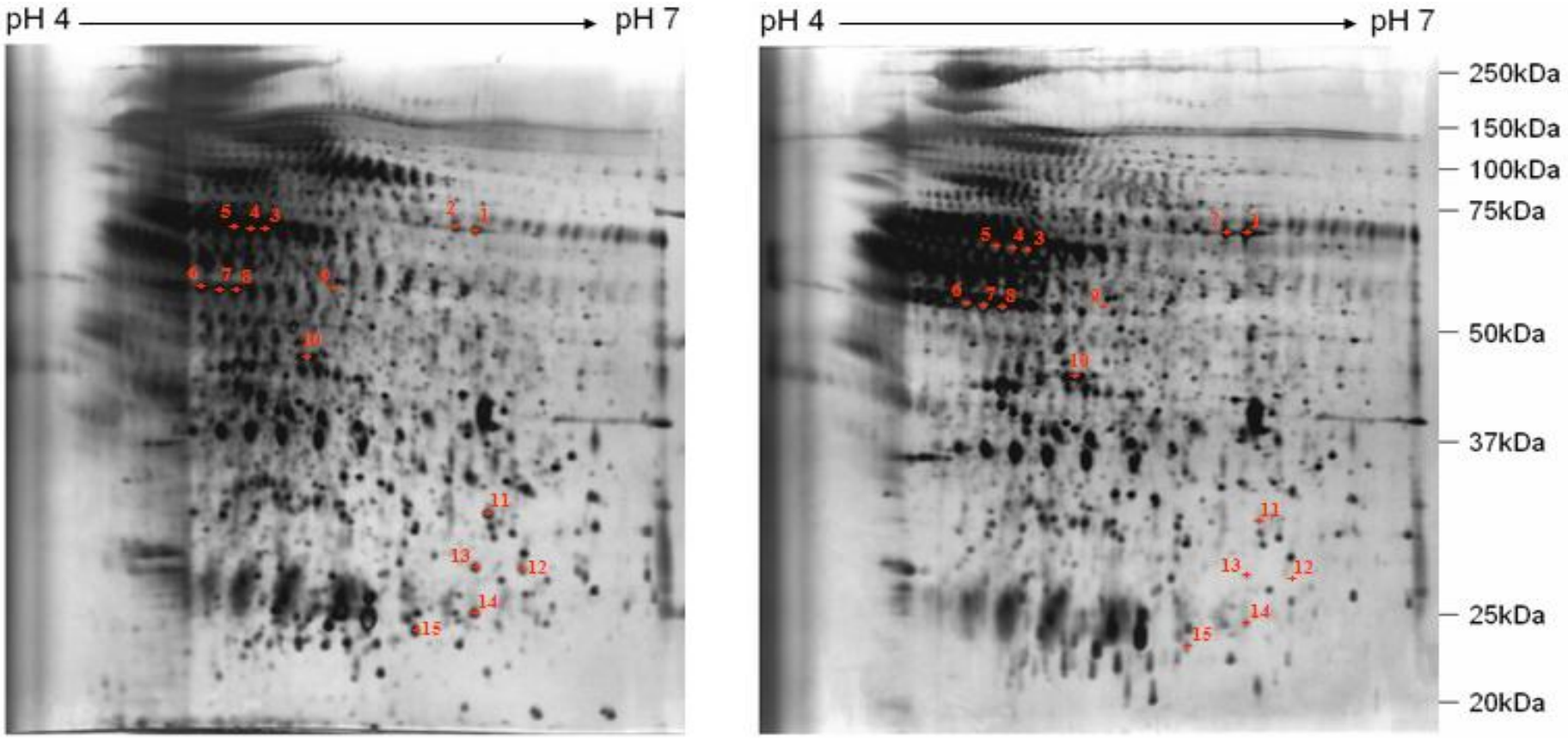


Figure 5.3: Representative 2-DE gels of rat serum patterns after removal of HAPs as visualized with silver stain. Normal (left) and dysplasia with adenocarcinoma (right) serum samples (40µg).

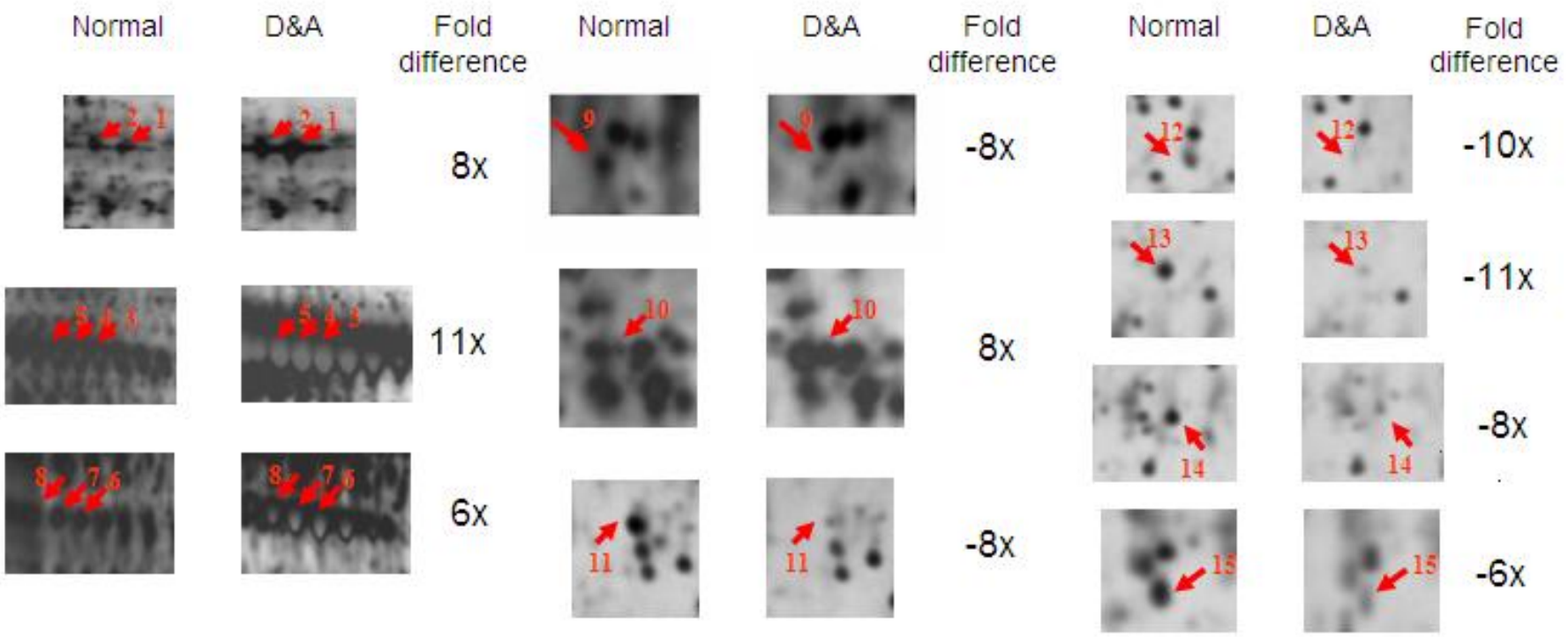


Figure 5.4: Enlarged regions of 2-DE gels showing the 15 proteins that were differentially expressed consistently for more than 3 folds shown in Figure 5.3.

Spot No.	Protein identity	Accession number	Mr (kDa)/pI	Mascot score*	Sequence coverage (%)*	Expression
1,2	Albumin	gi|19705431	70.67/6.09	326/94	49/2	↑
3,4,5	T-kininogen I	gi|60392582	48.828/6.08	93/94	19/3	↑
6,7,8	α-1-antitrypsin	gi|203063	45.978/5.7	98/62	24/2	↑
9	Stress 70 protein	gi|62664205	73.984/5.87	61/NA	13/NA	↓
10	γ-actin	gi|109507063	42.109/5.31	143/107	42/4	↑
11	Inter-α-inhibitor H4 heavy chain	gi|9506819	103.885/6.08	61/NA	5/NA	↓
12	Unknown	NA	NA	NA	NA	↓
13	Unknown	NA	NA	NA	NA	↓
14	Unknown	NA	NA	NA	NA	↓
15	Unknown	NA	NA	NA	NA	↓

Table 5.1: Identities of differential proteins (Figure 5.4) found in rat D&A serum sample of gastric cancer. (*the first number indicates the value by MS analysis, the latter represents the MS/MS analysis) (↑: up-regulated in cancer, ↓: down-regulated in cancer) (NA: data not available)

5.2.3. 2D-DIGE

One of the major disadvantages of silver staining is the inaccurate quantitative analysis of differential expression levels due to its narrow linear dynamic range. Fluorescence dyes, which provide more breath to the dynamic range of detection, was discovered in 2001 and marketed commercially. The most commonly used method that uses fluorescence dyes is DIGE labelling (Tonge *et al.*, 2001). Moreover, taking advantage of the sample labelling of DIGE, comparison of differential proteins generated by column depletion prior to and after labelling could be performed. In additional, reciprocal labelling experiments were conducted to account for any preferential proteins labelling by CyDyes, to further validate the application of DIGE technique to the samples.

5.2.3.1.Sample normalization (DIGE labelling) after immunodepletion

After immunodepletion, 50μg of the D&A sample were labelled with Cy3, 50μg of the normal sample were labelled with Cy5, and an equal mixture of the protein of the normal and the D&A sample was labelled with Cy2 as the internal control. As a result, each gel consisted of two experimental samples plus their mixture as the internal standard. After the procedures of conventional 2-DE, the Cy2, Cy3 and Cy5 channels were imaged individually using mutually exclusive excitation and emission wavelengths. The images were analyzed using the differential in-gel analysis (DIA) module of the DeCyder software, the expression ratio of each protein spot can be calculated against the pool internal standard.

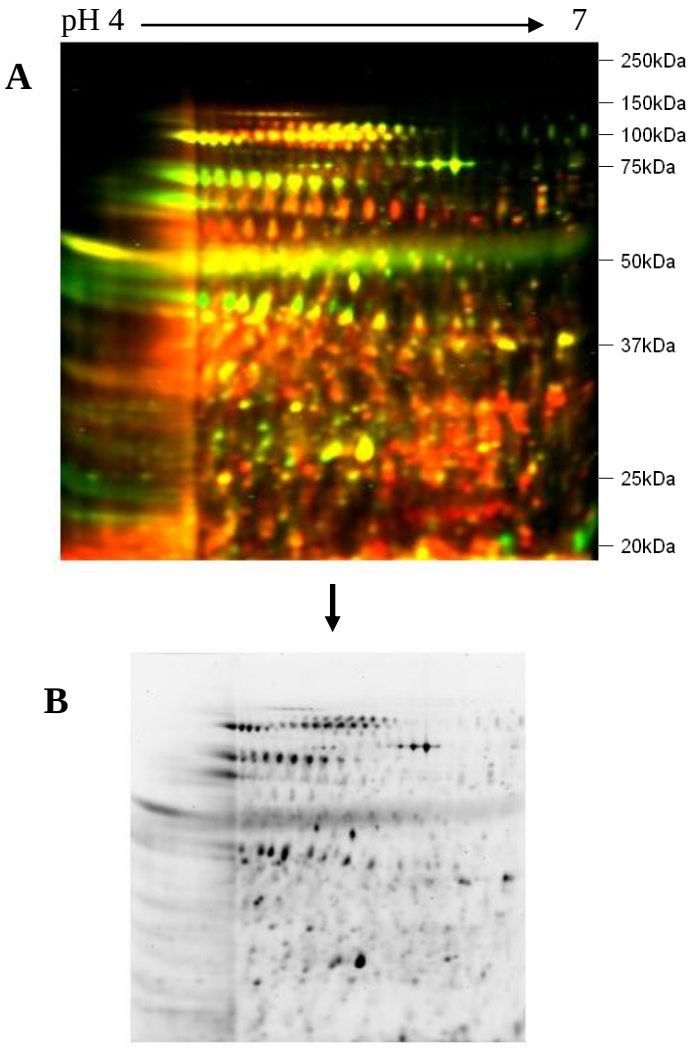
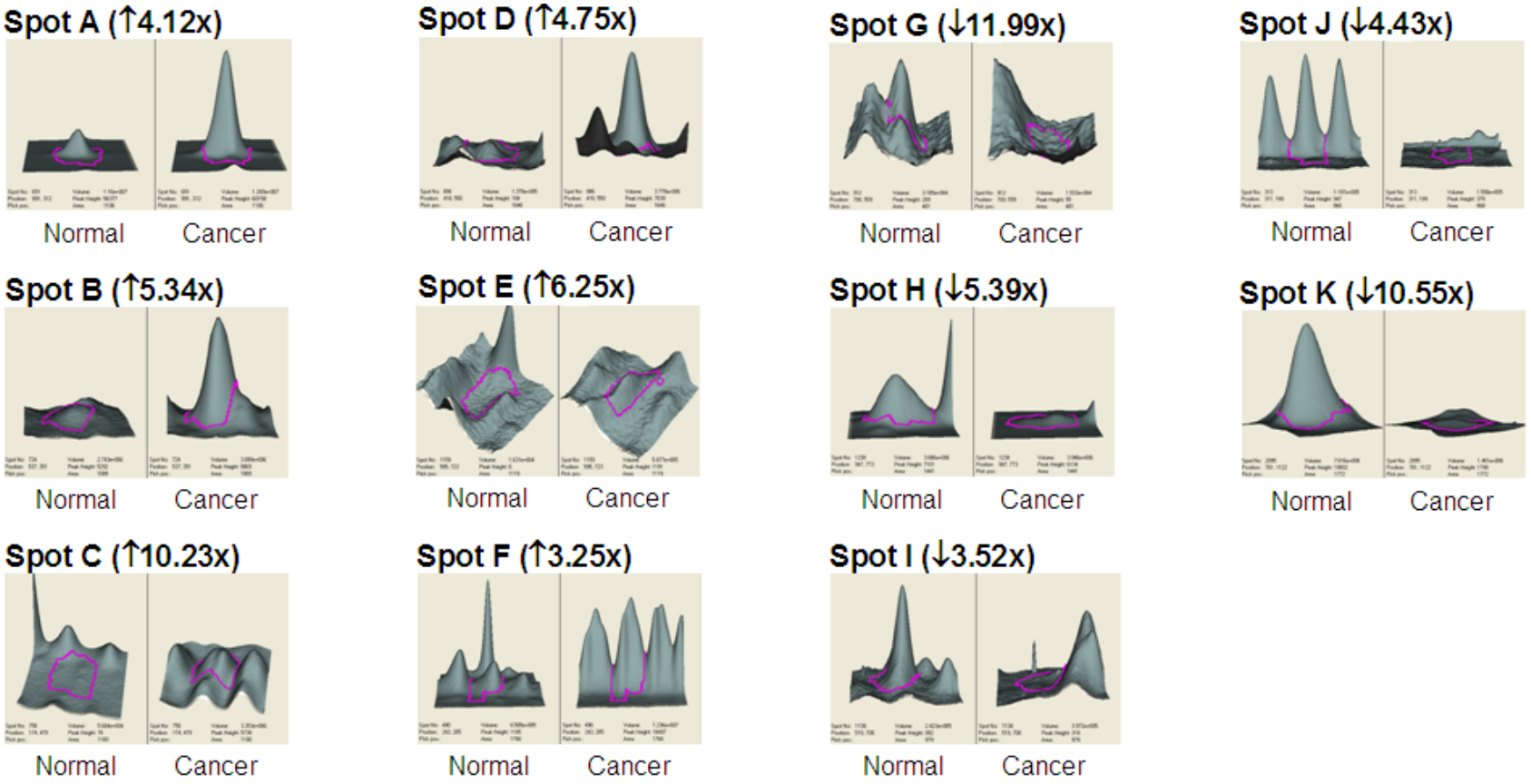


Figure 5.5: Representative 2D-DIGE gel (pH 4-7) of total 150µg serum samples labelled after immunodepletion. (A) Overlay image of Cy 3 (serum sample of D&A) and Cy 5 (control serum). (B) Separate CyDye images: Cy 2 (internal standard), Cy 3 and Cy 5.

For the normal and D&A that were depleted separately before DIGE labelling for comparison, 1597 spots were detected by the software. 35 spots were found to be differentially expressed for at least 3 folds, of which 23 were up-regulated and 12 down-regulated. The intensities analyzed by the DeCyder software of the successfully identified 11 proteins were shown in 3D images in Figure 5.6. Details of the sequence coverage and peptide mass fingerprints are attached in Appendix II.



*Fold difference is shown in bracket. ↑: up-regulation, ↓: down-regulation.

Figure 5.6: 3D images of 11 spots analyzed by DeCyder that were differentially expressed consistently for more than 3 folds.

Spot No.	Protein identity	Accession number	Mr (kDa)/pI	Mascot score*	Sequence coverage (%)*	Expression
A	Albumin	gi|19705431	70.67/6.09	326/94	49/2	↑
B	T-kininogen I	gi|60392582	48.828/6.08	93/94	19/3	↑
C	α-2-HS glycoprotein	gi|231468	38.757/6.05	73/NA	26/NA	↑
D	α-1-antitrypsin	gi|203063	45.978/5.7	98/62	24/2	↑
E	γ-actin	gi|109507063	42.109/5.31	143/107	42/4	↑
F	Afamin	gi|27229290	71.172/5.87	213/144	30/2	↑
G	Stress 70 protein	gi|62664205	73.984/5.87	61/NA	13/NA	↓
H	Transthyretin	gi|3212532	13.122/6.04	92/37	73/19	↓
I	Apolipoprotein A-IV	gi|114008	44.429/5.12	71/62	24/3	↓
J	Murinoglobulin	gi|12831225	166.590/5.68	96/NA	9/NA	↓
K	Apolipoprotein A-I	gi|6978515	30.100/5.52	185/123	50/5	↓

Table 5.2: Identities of differential proteins (Figure 5.6) found in rat D&A serum sample of gastric cancer. (*the first number indicates the value by MS analysis, the latter represents the MS/MS analysis) (↑: up-regulated in cancer, ↓: down-regulated in cancer) (NA: data not available)

5.2.3.2.Sample normalization (DIGE labelling) before immunodepletion

In order to minimize the chance of bias to any set of samples that go through 'different identical' immunodepletion columns, DIGE labelling of samples was performed before loading into the column. This allowed the same initial amount of normal and D&A samples to be loaded to and removed by the immunodepletion column together. 1mg of protein in each sample was labelled with Cy 3 and Cy 5, and the mixture normal and D&A was labelled with Cy 2. A total of 3mg of proteins were loaded into the column and immunodepleted together. After the samples were immunodepleted by the IgY-affinity column fabricated earlier, the gel was then analyzed by DeCyder. The result showed that 60 spots were differentially expressed (for at least 3 folds), of which 40 were up-regulated and 20 down-regulated. The intensities analyzed by the DeCyder software of the successfully identified 11 proteins were shown in 3D images in Figure 5.6. Details of the sequence coverage and peptide mass fingerprints are attached in Appendix II.

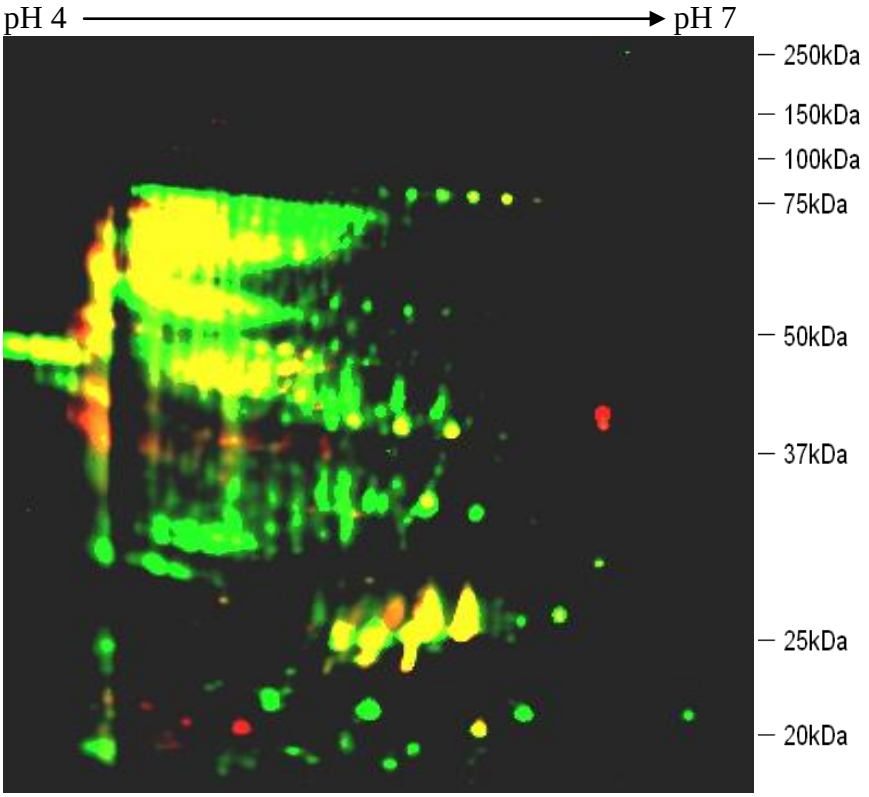
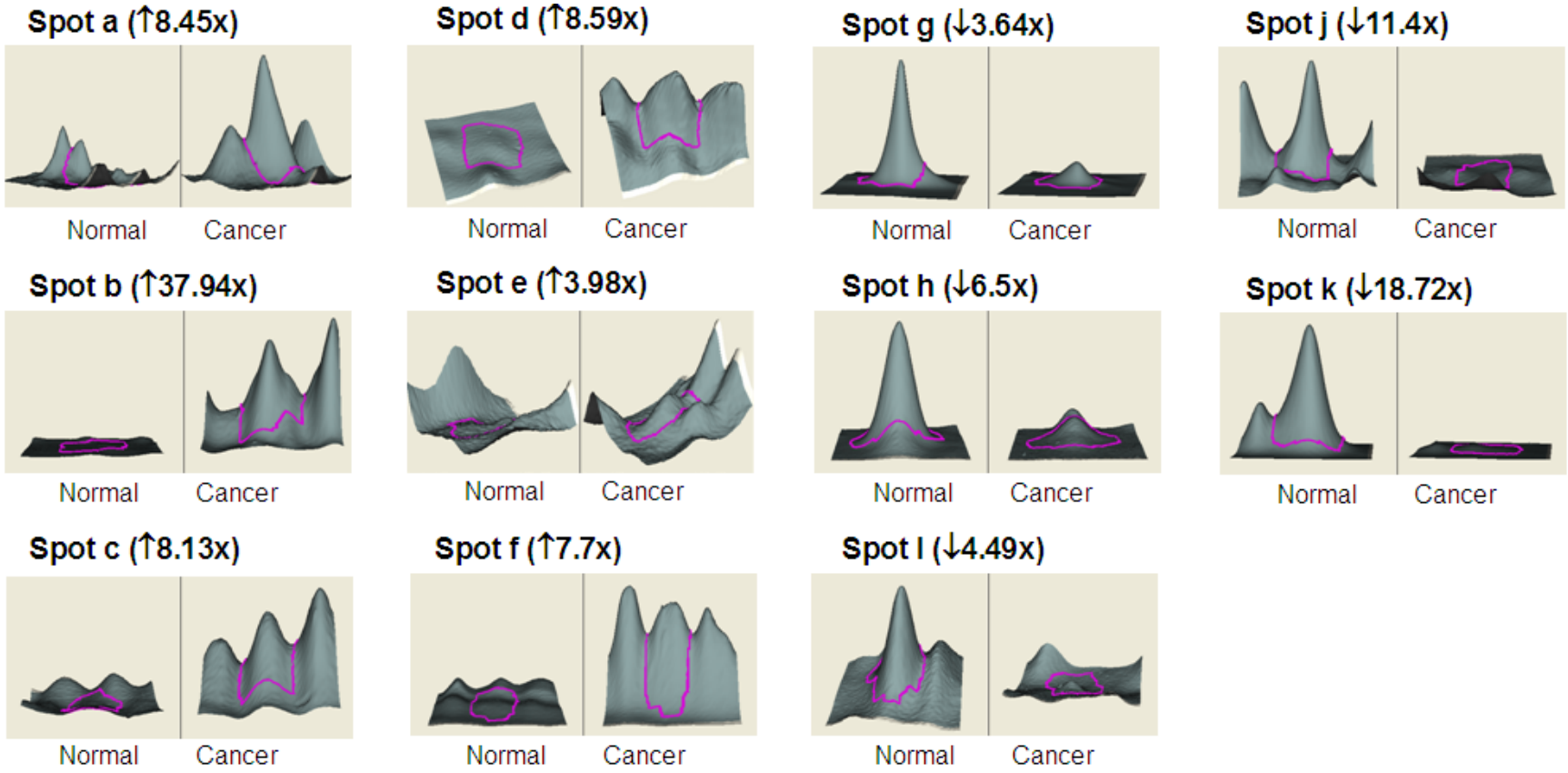


Figure 5.7; Representative overlay image 2D-DIGE (pH 4-7) with initially 3mg protein labelled before immunodepletion.



*Fold difference is shown in bracket. ↑: up-regulation, ↓: down-regulation.

Figure 5.8: 3D images of 11 spots analyzed by DeCyder that were differentially expressed consistently for more than 3 folds.

Spot No.	Protein identity	Accession number	Mr (kDa)/pI	Mascot score*	Sequence coverage (%)*	Expression
a	Albumin	gi|19705431	70.67/6.09	326/94	49/2	↑
b	T-kininogen I	gi|60392582	48.828/6.08	93/94	19/3	↑
c	α-2-HS glycoprotein	gi|231468	38.757/6.05	73/NA	26/NA	↑
d	α-1-antitrypsin	gi|203063	45.978/5.7	98/62	24/2	↑
e	γ-actin	gi|109507063	42.109/5.31	143/107	42/4	↑
f	Afamin	gi|27229290	71.172/5.87	213/144	30/2	↑
g	Stress 70 protein	gi|62664205	73.984/5.87	61/NA	13/NA	↓
h	Transthyretin	gi|3212532	13.122/6.04	92/37	73/19	↓
i	Apolipoprotein A-IV	gi|114008	44.429/5.12	71/62	24/3	↓
j	Murinoglobulin	gi|12831225	166.590/5.68	96/NA	9/NA	↓
k	Apolipoprotein A-I	gi|6978515	30.100/5.52	185/123	50/5	↓

Table 5.3: Identities of differential proteins (Figure 5.8) found in rat D&A serum sample of gastric cancer. (*the first number indicates the value by MS analysis, the latter represents the MS/MS analysis) (↑: up-regulated in cancer, ↓: down-regulated in cancer) (NA: data not available)

5.2.4. Comparison of differentially expressed proteins with different sample preparation procedures

A comparison was made between sample normalization before and after immunodepletion. The differences in their procedures are shown in Table 5.4. The number of spots detected was summarized in Table 5.5.

Procedures	Sample normalization (DIGE labelling) after immunodepletion	Sample normalization (DIGE labelling) before immunodepletion
Step 1	Immunodepletion of normal and D&A samples separately	*Labelling of a total of **3mg** proteins (normal and D&A samples)*
Step 2	Concentrating samples	*Immunodepletion of the labelled samples*
Step 3	*Labelling of a total of **150µg** of depleted samples (normal and D&A samples)*	*Concentrating samples*
Step 4	*Rehydration*	*Rehydration*
Step 5	*IEF*	*IEF*
Step 6	*SDS-PAGE*	*SDS-PAGE*

Steps performed in light Steps performed in dark

Table 5.4: Comparison of different procedures in two 2D-DIGE methods: labelling after and before immunodepletion.

	2-DE visualized with silver staining	Sample normalization (DIGE labelling) after immunodepletion	Sample normalization (DIGE labelling) before immunodepletion
Total no. of spots	1002	1597	1632
No. of up-regulated spots	9	23	40
No. of down-regulated spots	6	12	20

Table 5.5: Comparison on the number of spots being detected by traditional 2-DE with silver staining and the two 2D-DIGE methods: labelling after and before immunodepletion.

DIGE labelling before immunodepletion yielded a higher number of differentially expressed proteins. It is also worthwhile to compare the list of proteins that was found to be differentially expressed by the three different sample preparation methods. Some of the proteins appeared to be commonly detected in all/some of the 3 methods, while others were only detected with the most sensitive method. In order to compare the agreement of differentially expressed proteins found by these 3 methods, a Venn diagram was constructed. It should be noted that only differential proteins with known identities were employed for the construction of this Venn diagram.

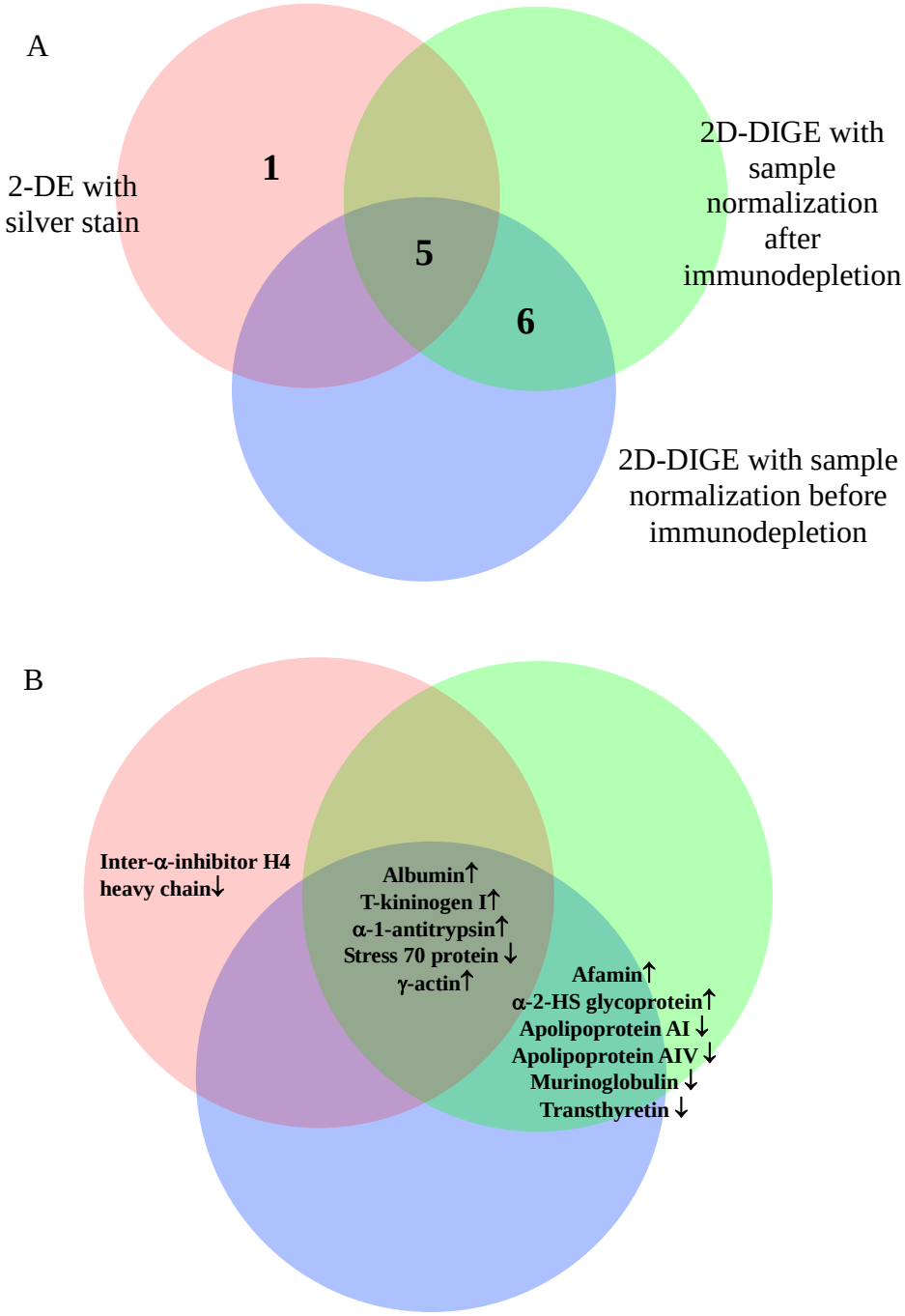


Figure 5.9: Venn diagram showing (A) the number of differential proteins identified using 3 different sample preparation methods and (B) their identities.

In total, 12 proteins with differential expression were identified by these 3 different methods. Among the 6 proteins detected by 2-DE with silver stain, 5 of them were also found in the other two 2D-DIGE methods; and both 2D-DIGE with sample normalization after before immunodepletion methods identified 6 more proteins. It is not surprising that the 2D-DIGE method allows more differential proteins to be detected, as this method has a higher sensitivity and a wider dynamic range than the 2-DE with silver stain. However, some of the protein spots pinpointed by DIGE could not be identified by MALDI-TOF-TOF MS as the protein amounts available for MALDI-TOF-TOF MS identification were not adequate. The proteins detected by the DeCyder cannot be observed on the gel after silver staining, making identification not practical for most laboratories with standard proteomic setup.

5.2.5. Identification of differentially expressed proteins by MALDI-TOF-TOF MS

All the differential proteins (at least 3 folds) from the above three sets of experiments were subjected to MALDI-TOF-TOF MS analysis. The identified proteins are summarized in Table 5.6 below. Details of the sequence coverage and peptide mass fingerprints are attached in Appendix II.

Protein identity	Accession number	Mr (kDa)/pI	Function	Expression
Albumin	gi|19705431	70.67/6.09	Acute phase response, Transport	↑
T-kininogen I	gi|60392582	48.828/6.08	Acute phase response	↑
α-2-HS glycoprotein	gi|231468	38.757/6.05	Acute phase response	↑
α-1-antitrypsin	gi|203063	45.978/5.7	Acute phase response	↑
Afamin	gi|27229290	71.172/5.87	Transport	↑
γ-actin	gi|109507063	42.109/5.31	Cell motility	↑
Stress 70 protein	gi|62664205	73.984/5.87	Protein folding	↓
Apolipoprotein A-I	gi|6978515	30.100/5.52	Transport	↓
Apolipoprotein A-IV	gi|114008	44.429/5.12	Transport	↓
Transthyretin	gi|3212532	13.122/6.04	Transport	↓
Murinoglobulin	gi|12831225	166.590/5.68	Acute phase response	↓
Inter-α-inhibitor H4 heavy chain	gi|9506819	103.885/6.08	Acute phase response	↓

Table 5.6: Identities of differential proteins found in rat D&A serum sample of gastric cancer. (↑: up-regulated in cancer, ↓: down-regulated in cancer)

Among these 12 differential proteins, 6 were up-regulated and the other 6 down-regulated. They were grouped according to their functionality as shown in Figure 5.10, and most of them are involved in transportation and acute phase response. Further discussion of individual proteins of interest will be in the next section.

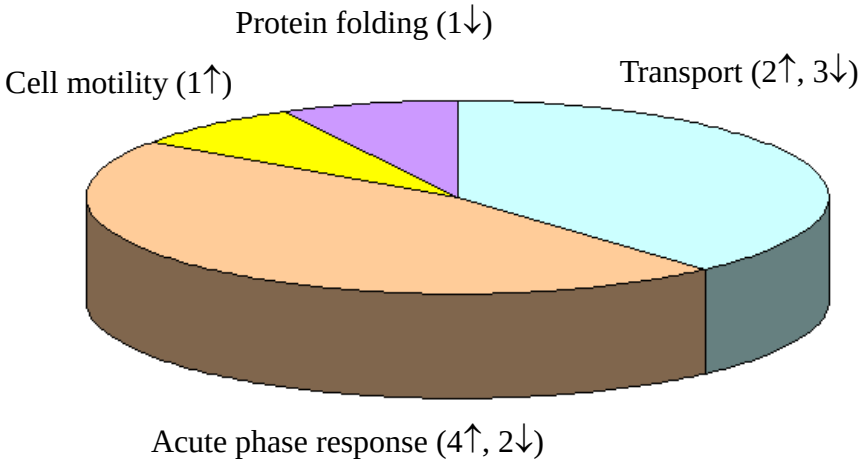


Figure 5.10: The 12 differential proteins were grouped according to their functionality.

5.3. Discussion

Sensitivity, detection dynamic range of different staining methods and accurate normalization of samples are very crucial in pinpointing differential proteins. Differential proteins found by 2-DE with silver stain is often challenged for its limited dynamic range. Its dynamic range is only one order of magnitude, which is unsuitable for quantitative analysis. Another concern is the irreproducibility of 2-DE gels. Since no two gels run identically corresponding spots between two gels have to be matched prior to quantification. This brings another difficulty to normalization, especially in the case of silver staining where staining is protein-dependent. These aforementioned factors all add variability to the system that makes it unsuitable for accurate quantitation.

Difference gel electrophoresis (DIGE) circumvents many of the issues associated with traditional 2-DE with silver stain, such as limited dynamic range and gel-to-gel variation, and allows a more accurate and sensitive quantitative analysis. This technique involves the labelling of up to 3 different samples with fluorescent cyanine dyes, which have nearly identical masses and charge. The protein samples are pre-labelled with fluorescent dyes and then resolved on a single gel. The gel is scanned by laser scanners and the corresponding protein spot patterns are visualized by multichannel scanning of the same gel. Since experimental and control samples are separated on a single gel, the gel-to-gel variation that represents the major limitation of conventional 2-DE gels is eliminated. In addition, running repetitive gels that swap the dyes used to label sample will control for any dye-specific effects that might result from preferential labelling. Furthermore, the Cy2, Cy3 and Cy5 dyes also have a wide dynamic range and are more sensitive than the silver stain.

2D-DIGE, however, still has its limitations. They are mainly associated with the labelling chemistry required for attaching the dye to the proteins. With the minimal dyes labelling strategy, proteins with a high percentage of lysine residues could be labelled more efficiently compared

with proteins containing few or no lysines. Therefore, a high abundance protein spot in a conventional gel system could be a medium or low abundance protein spot in the DIGE system because of its low lysine content. In addition, the technique is not applicable to those proteins without lysine.

Inaccurate sample normalization often occurs when there is extensive sample handling during the depletion process. This increases the chance of sample loss, protein degradation, and modification artifacts, resulting in substantial sample-to-sample variation. Moreover, when samples are diluted during the depletion process, making them less suitable for many downstream analyses, steps must be taken to concentrate the proteins, which again can result in the loss of certain proteins and introduce analysis deviation.

In order to solve this problem, pre-labelling of samples individually prior to immunodepletion was conducted. The normal and dysplasia with early adenocarcinoma serum samples were pre-labelled individually, mixed well, and applied to the depletion column. The HAPs of both the normal and the dysplasia with early adenocarcinoma serum samples will bind to the column randomly and therefore, being depleted. The sample in the “flow through” fraction can directly be concentrated and applied to the IEF strip, without any extra normalization step. This strategy eliminates the defects of column irreproducibility, and thus inaccurate sample normalization.

When comparing the results of 2-DE with silver stain and those of 2D-DIGE, as expected, the latter showed much more differential proteins, yet only 11 of them could be successfully identified. The major difficulty was the incising of gel plug of differential protein. After analysis with the DeCyder software, the protein spots on the 2D-DIGE gel had to be visualized by silver stain method. Since the sensitivity of DIGE was much higher than that of silver stain, most of the differential proteins that could be observed using the DeCyder were seen after silver staining, thus incising their gel plugs for identification was not feasible. Even if the incising is a

possibility, the amount of protein present is not enough for identification with most MALDI-TOF MS.

For sample normalization before CyDyes labelling, the gel image was smudge and the spots were not sharp, as the protein spots appeared to be clumped together. This phenomenon is difficult to explain, since there is no publication on DIGE labelling before any depletion step has been reported. One possible guess is the additional manipulation on the CyDyes. After the labelling of proteins, the labelled sample still has to going through the depletion column and the Amicon for concentrating the sample. This probably affects the performance of the CyDyes. Another suggestion is the difficulty in controlling the amount of protein to be labelled initially, in order to have sufficient amount when the depleted serum is coming out from the column and carry on to the conventional 2-DE for detection. Although the image is not ideal, it happens that the differential spots identified with this method are the same as those identified with the sample normalization after immunodepletion method.

All the 12 identified proteins have been reported to be differentially expressed in different kinds of cancers. However, regarding gastric cancer, only albumin, α-1-antitrypsin, γ-actin and apolipoprotein A-I have been reported.

Albumin is a soluble, monomeric protein which comprises about one-half of the serum protein. It functions primarily as a carrier protein for steroids, fatty acids, and thyroid hormones and plays a role in regulation of the colloidal osmotic pressure of blood by stabilizing the extracellular fluid volume. It is synthesized in the liver as preproalbumin which has an N-terminal peptide that is removed before the nascent protein is released from the rough endoplasmic reticulum. Proalbumin, is in turn cleaved in the Golgi vesicles to produce the secreted albumin. An up-regulation of albumin has also been detected by Chen (Chen *et al.*, 2004) in gastric cancer. However, it should be stressed that albumin, being a member of the HAP,

have to be studied further before it can be confirmed as a biomarker candidate of gastric cancer.

T-kininogen (KNG) acts as a precursor of active peptide bradykinin, and is involved in smooth muscle contraction, induction of hypotension and vascular permeability. It is an inhibitor of thiol protesase and plays a role in blood coagulation. A reduced expression of cleaved high-molecular-weight kininogen (light chain) has been detected in two types of cerviacal cancer, squmaous cell cervical carcinoma and cervical adenocarcinoma (Abdul-Rahman *et al.*, 2007). In addition, recent studies have demonstrated that kininogen was down-regulated in breast cancer patients (Gabrijelcic *et al.*, 1992). In constrast, research conducted on Nagase analbuminemic rats bearing intestinal tumors induced by azoxymethane have displayed an significant higher of a T-kininogen isotype level in serum than normal controls. Elevation of the serum level of that T-kininogen isotype has also been observed in Sprague-Dawley rats bearing intestinal tumors, skin tumors, subcutaneous sarcomas, or mammary tumors and in ACI rats bearing urinary bladder tumors (Kanda *et al.*, 1990).

α-2-HS glycoprotein (AHSG), a glycoprotein present in the serum, is synthesized by hepatocytes in liver and osteoblasts, and is concentrated into the mineralized tissue. It has also been grouped with the fetuins, a family of proteins that occur in fetal plasma in high concentration owing to its extensive sequence identity. The AHSG molecule consists of two polypeptide chains, which are both cleaved from a proprotein encoded from a single mRNA. It is involved in several functions, such as acute phase response, regulation of inflammatory response, negative regulation of bone mineralization and positive regulation of phagocytosis. It also plays a vital role in blocking transforming growth factor-beta 1 signal transduction, which is associated with tumor development. The protein is commonly present in the cortical plate of the immature cerebral cortex and bone marrow hemopoietic matrix, therefore it has been postulated that it participates in the development of the tissues. However, its exact

significance is still obscure. AHSG has been reported as a biomarker in several cancers. Pawlik et al. utilized isotope-coded affinity tag (ICAT) tandem mass spectrometry (MS) to identity and quantify differences in specific protein expression between nipple aspirate fluid (NAF) samples from tumor-bearing and from disease-free breasts. It was found that in women with early-stage breast cancer, the amount of AHSG in the nipple aspirate fluid has been observed to be reduced in terms of ICAT ratio (Pawlik *et al.*, 2006). Besides breast cancer, in lung squamous cell carcinoma a decreased level of AHSG compared with normal sera (Dowling *et al.*, 2007) was also detected. In addition, AHSG has also been observed to be depleted in acute myeloid leukaemia (AML), compared with normal tissue. It has been postulated that the down-regulation of AGSH might be due to the abnormal development of bone marrow in AML patient (Kwak *et al.*, 2004).

α-1-antitrypsin is an inhibitor of serine protease. Its primary target is elastase, but it also has a moderate affinity for plasmin and thrombin. It inhibits trypsin, chymotrypsin and plasminogen activator. The aberrant form inhibits insulin-induced NO synthesis in platelets, decreases coagulation time and has proteolytic activity against insulin and plasmin. This protein is frequently detected to be up-regulated in various kinds of cancers. In the studies of the serum samples from patients with gastric cancer(Chen *et al.*, 2004; Lee *et al.*, 2004; Hsu *et al.*, 2007), lung cancer (Patz *et al.*, 2007), prostate cancer and breast cancer (El-Akawi *et al.*, 2008), up-regulation of α-1-antitrypsin were observed. Over-expression of α-1-antitrypsin was also detected in tissue examinations on primary acinar cell carcinoma of liver (Hervieu *et al.*, 2008) and pancreatic ductal adenocarcinoma (Chung *et al.*, 2008).

Afamin, which belongs to the albumin family, comprises of albumin, α-fetoprotein and vitamin D binding protein. It has a predicted mass of ~65 kDa and apparent molecular mass of 87kDa due to N-glycosylation. Afamin is expressed in the liver and secreted into the bloodstream. Afamin, which

has been recently identified as a potential biomarker for ovarian cancer due to its complementarity with CA-125, is currently the most widely used marker, in longitudinal monitoring of ovarian cancer patients. Studies confirmed by Western blotting and ELISA have shown that the total afamin concentration in serum samples of patients with ovarian cancer sharply decreased when compared with the normal controls and with patients with benign disease (Jackson *et al.*, 2007).

Actin is the most abundant and one of the most highly conserved proteins in all eukaryotic cells. It is a globular, roughly 42kDa protein. Actin microfilaments are involved in many important cellular functions, such as cell locomotion, signal transduction, contractile ring formation during cytokinesis, organelle movement, secretion, mRNA processing and activity of membrane channels. Of its six isoforms, γ-actin mainly exists in stress fibres, and play an important role in the maintenance of cell shape, differentiation, and mechanical resistance. The expression of the actin isoforms can be up- or down-regulated in neoplastic transformation and other cell pathologies. In a recent study by Zhang, γ-actin was found to be over-expressed. It has been illustrated that there is an increased level of γ-actin in tumor cells measured on western blots in the cytosolic fractions of human melanoma cell line (Radwanska *et al.*, 2008). The concentration of γ-actin was found to be lower in human salivary gland adenocarcinoma cell clones, which acquired metastatic capacity, in comparison with its original clone HSGc lacking metastatic ability. It has been suggested that such a difference is one of main factors that contribute to a decrease in cell adhesiveness and an increase in cell motility, which in turn is probably a major cause for acquisition of metastatic potential (Suzuki *et al.*, 1998). It has also been reported that cytoskeleton protein γ-actin is over-expressed in hepatocellular carcinoma cell lines with high metastatic potentials (Zhang *et al.*, 2006).

Stress 70 protein belongs to the heat shock protein 70 family. It is implicated in the control of cell proliferation and cellular aging. It may also

act as a chaperone. Stress 70 protein is over-expressed in malignant melanoma (Kalogeraki *et al.*, 2006) and under-expressed in renal cell cancer (Ramp *et al.*, 2007).

Apolipoprotein A-I is the major protein component of high density lipoprotein (HDL) in plasma. It participates in the reverse transport of cholesterol from tissues to the liver for excretion by promoting cholesterol efflux from tissues and by acting as a cofactor for the lecithin cholesterol acyltransferase. Altered plasma/serum levels of it have been reported in ovarian and prostate cancer. Apolipoprotein-AI levels were found to be significantly lower in gastric cancer (Ryu *et al.*, 2003; He *et al.*, 2004) and pancreatic cancer. Furthermore, a down-regulated expression of apolipoprotein AI is commonly detected in patients of hepatocellular cancer, where abnormal liver function damages lipoprotein synthesis and metabolism (Yan *et al.*, 2008).

Apolipoprotein A-IV, a major component of HDL and chylomicrons, is synthesized by the human intestine and secreted in the plasma. It is responsible for chylomicrons and VLDL secretion and catabolism, and lipid transport. In human serum, apolipoprotein A-IV was found to be of higher abundance in squamous cell carcinoma patients compared to healthy individuals (Dowling *et al.*, 2007), whereas it has a reduced expression in hepatocellular carcinoma patients after radiofrequency ablation (Kawakami *et al.*, 2005). It has also been observed that there is an elevated expression of apolipoprotein A-IV in the advanced stage neuroblastoma samples as compared to controls (Sandoval *et al.*, 2007).

Transthyretin (TTR), a thyroid hormone-binding protein, transports thyroxine from the bloodstream to the brain and is involved in retinol metabolism. It is synthesized in the liver, choroid plexus and retinal pigment epithelium and islet A and B cells. It has been reported that the average total TTR level for lung cancer serum samples is significantly lower than that in normal individuals, whereas it is higher than in the benign lung diseases samples. This level difference demonstrates that TTR could also

discriminate lung cancer patients from benign lung diseases patients and normal individuals. Changes in the levels of TTR in serum samples are mainly due to liver activity, which could be converted to the synthesis of acute-phase response proteins and result in a significant decrease in TTR (Liu *et al.*, 2007a). In addition, recent studies have shown that the level of TTR is reduced in the sera of patients with ovarian cancer and advanced cervical, endometrial carcinomas, pancreatic cancer (Yan *et al.*, 2008) as well as head and neck squamous cell carcinoma (Dowling *et al.*, 2008).

Murinoglobulin belongs to the protease inhibitor I39 family. It activates the inhibitor by specific proteolysis in the bait region, which, by an unknown mechanism leads to reaction at the cysteinyl-glutamyl internal thiol ester site and to a conformational change, whereby the proteinase is trapped and/or covalently bound to the inhibitor. While in the tetrameric preoteinase inhibitors steric inhibition is sufficiently strong, monomeric forms need a covalent linkage between the activated glutamyl residue of the original thiol ester and and a terminal amino group of a lysine or another nucleophilic group on the proteinase for the inhibition to become effective. Its expression level in the liver as well as the protein level in serum are down-regulated by up to 70% during acute inflammation or tumor development (Sugiyama *et al.*, 1989; Regler *et al.*, 1991).

Inter-alpha-inhibitor H4 heavy chain is involved in acute phase reactions. It was reported that this protein is up-regulated in serum of lung cancer patients (Heo *et al.*, 2007), while it is down-regulated in the study of plasma sample of melanoma patients with SELDI-TOF MS analysis (Caputo *et al.*, 2005).

Most of the 12 differential proteins are involved in acute phase response. Acute phase proteins are molecules in blood that either increase or decrease in response to chronic inflammation or cancer. It is envisioned that when the body is injured, immune cells will flood to the injured area to help healing and/or fighting off the harmful substances, and causing inflammation in the body. The acute phase proteins are responsible for

helping the body to respond to injury by destroying the harmful substances or helping the blood clot and preventing blood loss. Links between cancer and inflammation were first discovered in the nineteenth century on the basis of observations that tumors often appeared at sites of chronic inflammation and that inflammatory cells were present in biopsied samples from tumors (Balkwill and Mantovani, 2001). Several epidemiologic (Caruso *et al.*, 2004) and experimental studies (Dannenberg *et al.*, 2005) implicated that inflammation is an important factor in neoplastic progression via production of oxygen and nitrogen radical oxidants, production of growth-promoting cytokines, tumor suppressor inhibition and stimulation of signal transduction pathways. There is evidence suggesting that several cancers are linked to inflammatory origin, such as prostate (Nelson *et al.*, 2004), colon (Dannenberg *et al.*, 2005), breast (Chang *et al.*, 2005) and lung (Mao *et al.*, 2005) cancers. The pathways that connect inflammation and cancer are shown below. Briefly, cancer and inflammation are connected by two pathways, the intrinsic pathway and extrinsic pathway. The intrinsic pathway is activated by genetic events that cause neoplasia. These events include the activation of various types of oncogene by mutation, chromosomal rearrangement or amplification, and the inactivation of tumor-suppressor genes. Inflammatory mediators are produced under this situation, thereby generating an inflammatory microenvironment in tumors. In contrast, inflammatory or infectious conditions augment the risk of developing cancer at certain anatomical sites in the extrinsic pathway. When the two pathways converge, they activate the transcription factors to coordinate the production of inflammatory mediators, such as cytokines. Then the cytokines further activate the same key transcription factors in inflammatory cells, stromal cells and tumor cells, resulting in more inflammatory mediator being produced and a cancer-related inflammatory microenvironment being generated (Mantovani *et al.*, 2008). The connection is diagrammatically shown in Figure 5.11.

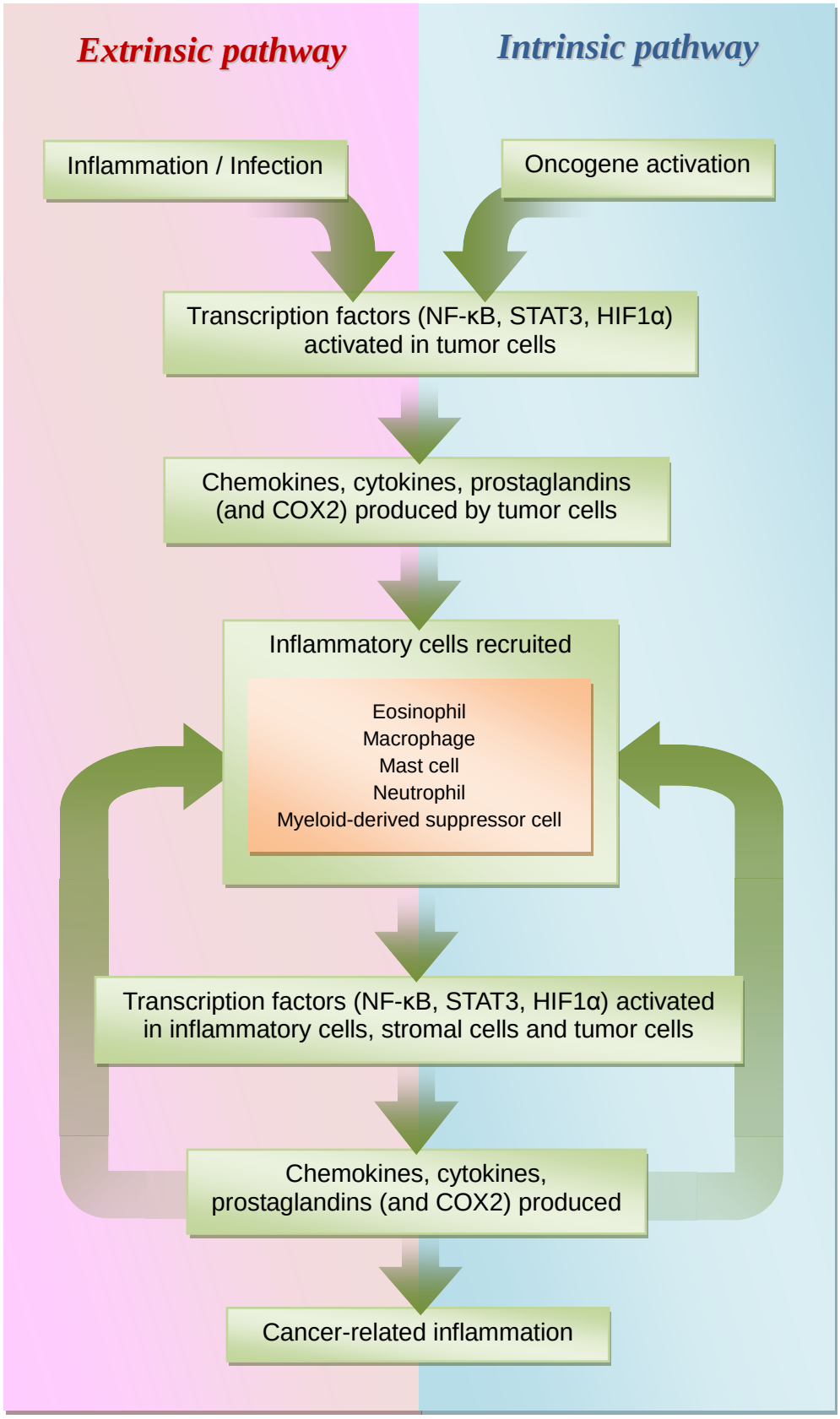


Figure 5.11: Pathways that connect inflammation and cancer (Mantovani *et al.*, 2008).

5.4. Conclusion

In the first part of this chapter, it was seen that IPG strip of pH 4-7 was chosen for the analysis of rat serum proteome. Differential proteins generated by 3 different sample separation methods were compared. Obstacles appeared when performing with the third method, sample normalization with DIGE labelling before immunodepletion. This method has not been reported previously mainly due to the high cost of CyDyes. Several experiments had been tried in altering the sample size, but an ideal image could hardly be obtained and analysis with the DeCyder was difficult. Therefore, it can be concluded that additional manipulation of sample after DIGE labelling should be avoided. Labelling process should be executed after the depletion processes.

After analyzing the data, a total of 12 proteins with at least 3 folds differential expression were successfully identified by MALDI-TOF-TOF MS. Among them, 6 proteins were up-regulated (albumin, T-kininogen, α-2-HS glycoprotein, α-1-antitrypsin, afamin and γ-actin), while the other 6 proteins were down-regulated (stress 70 protein, apolipoprotein A-I, apolipoprotein A-IV, transthyretin, murinoglobulin and inter-α-inhibitor H4 heavy chain). Classification of the proteins according to their functions shows that most of them are involved in the acute phase response. This phenomenon can be explained by the connection among acute phase response, inflammation and cancer.

Chapter 6: Concluding remarks and the way forward

With the gastric cancer model in rat, 12 proteins were found to be differentially expressed in serum samples from rats with histologically confirmed dysplasia and adenocarcinoma. This stage is one of the early stages of gastric cancer. Most of the 12 proteins were involved in the acute phase response. Therefore, my results supposed the notion that early stages of cancer can be related to inflammation.

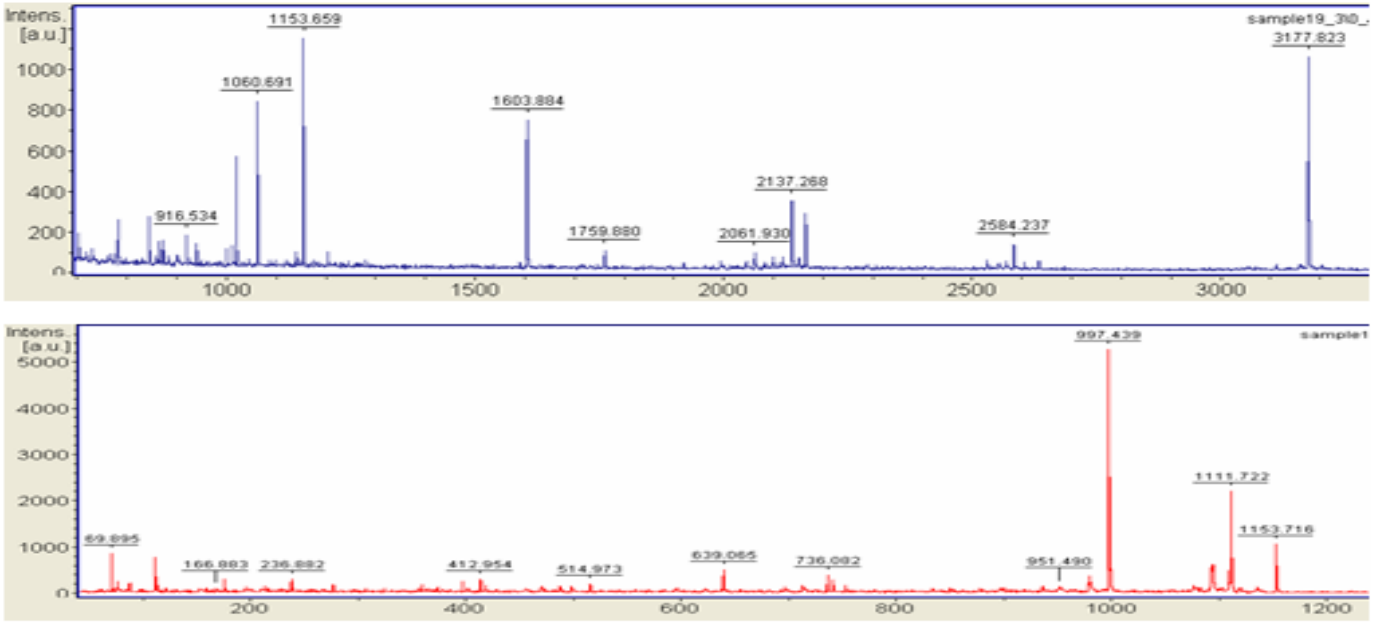
Results of this project also demonstrated that the feasibility of applying 2D-DIGE technique to enhance the discovery process by overcoming some technical bottlenecks associated with classical 2-DE. In my hands, DIGE labelling of samples before depletion was not practical. One of the main concerns is the cost of CyDyes. As shown in Table 5.1 in the previous chapter, the amount of CyDyes needed to label the samples prior to immunodepletion were 20 times more compared to the alternate method. Moreover, all the procedures have to be performed in the dark to prevent photo-bleaching of CyDyes, which increases the complexity of the experiment. However, an increment of spots was detected with this method, indicating that it is possible to find more differential proteins as biomarkers. Therefore, this method may be attainable for other research groups with more resources.

The establishment of the rat gastric cancer model provides us with a great advantage as serial serum samples, which have been collected bi-weekly during the gastric cancer induction process, are available for further investigation. This availability of serum samples in different stages of gastric cancer allows a detailed analysis on the expression levels of the candidate biomarkers throughout the whole process of carcinogenesis. The inter-relationship between these proteins and the carcinogenic process can be studied.

Although a full picture of the protein interactions in the gastric cancer remains unclear, this panel of biomarkers may be employed in the early diagnosis and screening of high risk individuals of gastric cancer. Small (n=100-200) and large (n>1000) scale validation of these protein biomarkers in patients with gastric cancer, as well as other types of cancers including that of colorectal, lung, liver and leukaemia, can be carried out. Further elucidation of their exact roles in the carcinogenesis process may hold vital keys for development of treatment protocols in the future.

Appendix I - Details of the sequence coverage and peptide mass finger prints of proteins identified in Chapter 4

α-1-antitrypsin



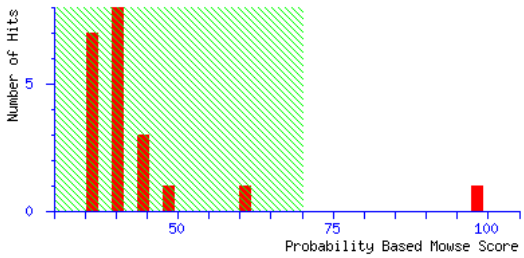
PMF of α-1-antitrypsin (top). MS/MS spectrum of its 1153.659 peptide (bottom).

Mascot Search Results

User : Katie
Email :
Search title : database search
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Mammalia (mammals) (525125 sequences)
Timestamp : 29 Oct 2008 at 02:57:36 GMT
Top Score : 98 for gi|203063, alpha-1-antitrypsin precursor

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 70 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help
Significance threshold p< 0.05 | Max. number of hits 5

Re-Search All | Search Unmatched

1. gi|203063 **Mass:** 45978 **Score:** 98 **Expect:** 8.1e-005 **Queries matched:** 8
alpha-1-antitrypsin precursor
gi|51036655 **Mass:** 46264 **Score:** 98 **Expect:** 8.7e-005 **Queries matched:** 8
serine protease inhibitor alpha 1 [Rattus norvegicus]
gi|112889 **Mass:** 46278 **Score:** 98 **Expect:** 8.7e-005 **Queries matched:** 8
Alpha-1-antiproteinase precursor (Alpha-1-antitrypsin) (Alpha-1-proteinase inhibito

Protein View

Match to: **gi|203063** Score: **98** Expect: **8.1e-005**
alpha-1-antitrypsin precursor

Nominal mass (M_r): **45978**; Calculated pI value: **5.70**
NCBI BLAST search of gi|203063 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **13**
Number of mass values matched: **8**
Sequence Coverage: **24%**

Matched peptides shown in **Bold Red**

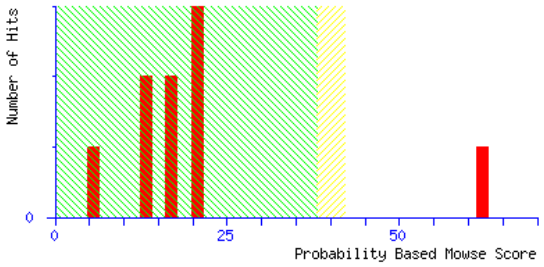
```
   1 SISRGLLLLA ALCCLAPSFL AEDAQETDTS QQDQSPTYRK ISSNLADFAF
  51 SLYRELVHQS NTSNIFFSPM SITTAFAMLS LGSKGDTRKQ ILEGLEFNLT
 101 QIPEADIHKA FHHLLQTLNR PDSELQLNTG NGLFVNKNLK LVEKFLEEVK
 151 NNYHSEAFSV NFADSEEAKK VINDYVEKGT QGKIVDLMKQ LDEDTVFALV
 201 NYIFFKGKWK RPFNPEHTRD ADFHVDKSTT VKVPMMNRLG MFDMHYCSTL
 251 SSWVLMMDYL GNATAIFLLP DDGKMQHLEQ TLTKDLISRF LLNRQTRSAI
 301 LYFPKLSISG TYNLKTLLSS LGITRVFNND ADLSGITEDA PLKLSQAVHK
 351 AVLTLDERGT EAAGATVVEA VPMSLPPQVK FDHPFIFMIV ESETQSPLFV
 401 GKVIDPTR
```

{MATRIX SCIENCE} Mascot Search Results

```
User            : Katie
Email           :
Search title    : database search
MS data file    : DATA.TXT
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Mammalia (mammals) (525125 sequences)
Timestamp       : 29 Oct 2008 at 03:01:10 GMT
Protein hits    : gi|112889     Alpha-1-antiproteinase precursor (Alpha-1-antitrypsin)
                  gi|40254285   interleukin 28 receptor alpha [Mus musculus]
                  gi|119901152  PREDICTED: similar to OTTHUMP00000017175 [Bos taurus]
                  gi|109505146  PREDICTED: hypothetical protein [Rattus norvegicus]
                  gi|94407699   PREDICTED: hypothetical protein [Mus musculus]
```

Probability Based Mowse Score

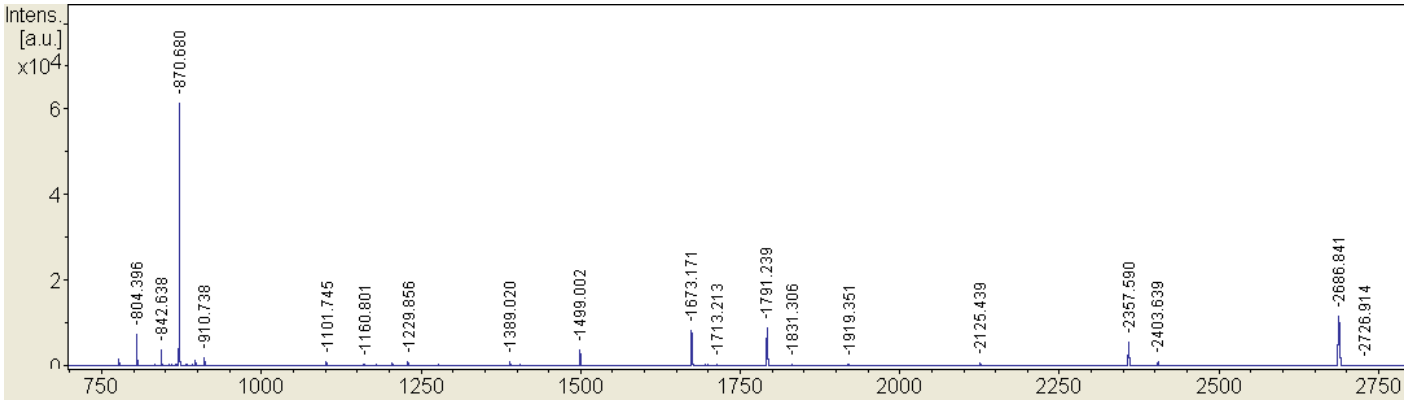
Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 38 indicate peptides with significant homology.
Individual ions scores > 42 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

```
Match to: gi|203063 Score: 62
alpha-1-antitrypsin precursor
Found in search of DATA.TXT

Nominal mass (Mr): 45978; Calculated pI value: 5.70
NCBI BLAST search of gi|203063 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: 2%

Matched peptides shown in Bold Red

     1 SISRGLLLLA ALCCLAPSFL AEDAQETDTS QQDQSPTYRK ISSNLADFAF
    51 SLYRELVHQS NTSNIFFSPM SITTAFAMLS LGSKGDTRKQ ILEGLEFNLT
   101 QIPEADIHKA FHHLLQTLNR PDSELQLNTG NGLFVNKNLK LVEKFLEEVK
   151 NNYHSEAFSV NFADSEEAKK VINDYVEKGT QGKIVDLMKQ LDEDTVFALV
   201 NYIFFKGKWK RPFNPEHTRD ADFHVDKSTT VKVPMMNRLG MFDMHYCSTL
   251 SSWVLMMDYL GNATAIFLLP DDGKMQHLEQ TLTKDLISRF LLNRQTRSAI
   301 LYFPKLSISG TYNLKTLLSS LGITRVFNND ADLSGITEDA PLKLSQAVHK
   351 AVLTLDERGT EAAGATVVEA VPMSLPPQVK FDHPFIFMIV ESETQSPLFV
   401 GKVIDPTR
```

α-1-macroglobulin



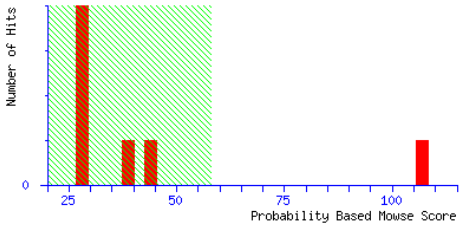
PMF of α-1-macroglobulin

Mascot Search Results

User : katie
Email :
Search title : database search
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Rattus (39483 sequences)
Timestamp : 6 Feb 2009 at 06:18:13 GMT
Top Score : 107 for gi|81872093, Alpha-1-macroglobulin precursor (Alpha-1-M) [Contains: Alpha-1-macroglobulin 165 kDa subunit; Alpha-1-macroglobulin 45 kDa subunit]

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help
Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|81872093 **Mass:** 168388 **Score:** 107 **Expect:** 7.9e-007 **Queries matched:** 11
Alpha-1-macroglobulin precursor (Alpha-1-M) [Contains: Alpha-1-macroglobulin 165 kDa subunit; Alpha-1-macroglobulin 45 kDa subunit]

Protein View

Match to: **gi|81872093** Score: **107** Expect: **7.9e-007**
Alpha-1-macroglobulin precursor (Alpha-1-M) [Contains: Alpha-1-macroglobulin 165 kDa subunit;

Nominal mass (M_r): **168388**; Calculated pI value: **6.46**
NCBI BLAST search of gi|81872093 against nr
Unformatted sequence string for pasting into other applications

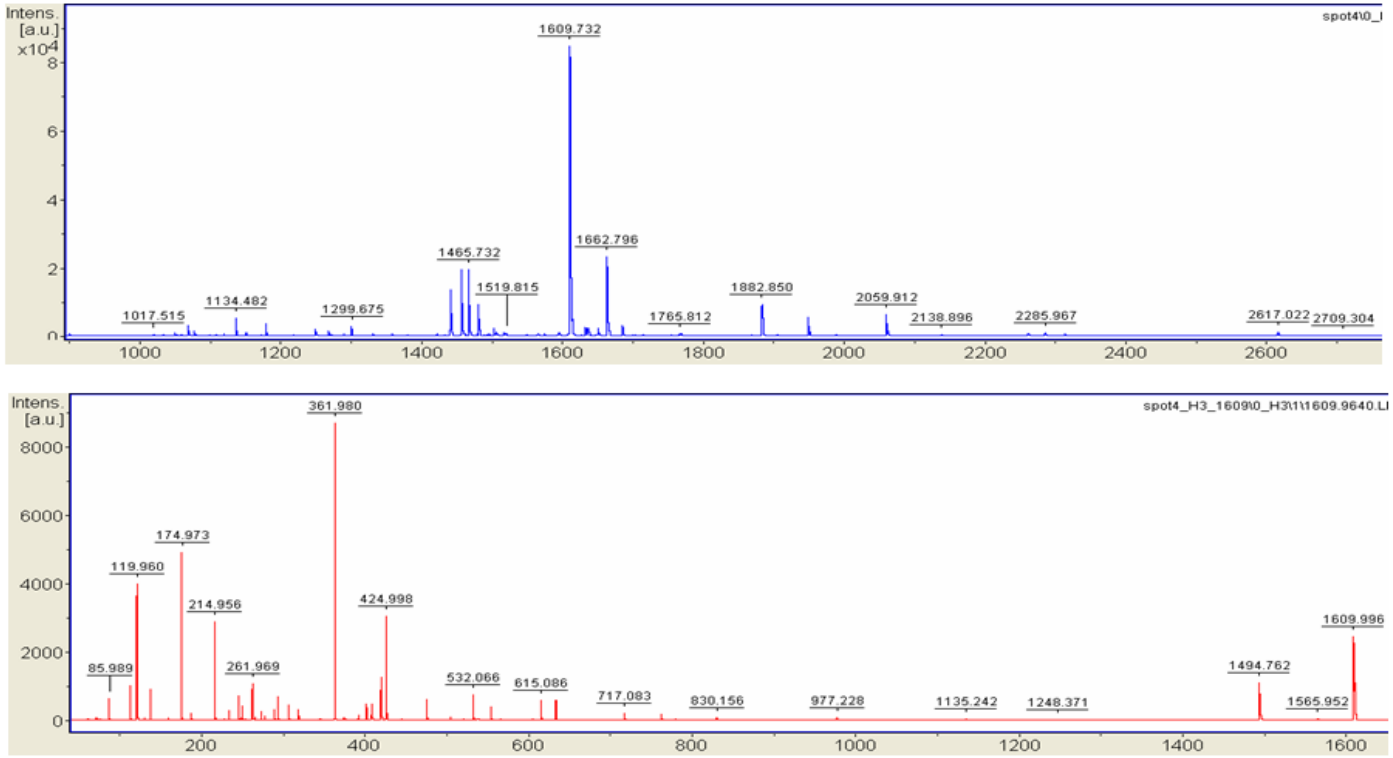
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|202857 from Rattus norvegicus
gi|739596 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **12**
Number of mass values matched: **11**
Sequence Coverage: **8%**

Matched peptides shown in **Bold Red**

```
   1 MRRNQLPIPV FLLLLLLLPR DATAATGKPR YVVLVPSELY AGVPEKVCVH
  51 LNHLNETVTL NVTLEYGVQY SNLLIDQAVD KDSSYCSSFT ISRPLSPSAL
 101 IAVEIKGPTH HFIKKKSMWI TKAESPVFVQ TDKPIYKPGQ TVKFRVVSVD
 151 ISFRPVNETF PVVYIENPKR NRIFQWQNVD LPGGLHQLSF PLSVEPALGI
 201 YKVVVQKDSG KKIEHSFEVK EYVLPKFEVQ VKMPKTMAFL EEELVVTACG
 251 LYTYGKPVPG LVTMKVCRKY TQSYSNCHGQ HSKSICEEFS KQADEKGCFR
 301 QVVKTKVFQP RQKGYDMKIE VEAKIKEDGT GIELTGTGSC EIANTLSKLK
 351 FTKANTFYRP GLPFFGQVLL VDEKGQPIPN KNLTVQVNSV RSQFTFTTDE
 401 HGLANILIDT TNFTFSFMGI RVIYKQNNIC FDNWWVDEYH TQADHSAARI
 451 FSPSRSYIQL ELVLGTLACG QTQEIRIHFL LNEDALKDAK DLTFYYLIKA
 501 RGSIFNSGSH VLPLEQGKVK GVVSFPIRVE PGMAPVAKLI VYTILPNEEL
 551 IADVQKFDIE KCFANTVNLS FPSAQSLPAS DTHLTVKATP LSLCALTAVD
 601 QSVLLLKPEA KLSPQSIYNL LPQKAEQGAY LGPLPYKGGE NCIKAEDITH
 651 NGIVYTPKQD LNDNDAYSVF QSIGLKIFTN TRVHKPRYCP MYQAYPPLPY
 701 VGEPQALAMS AIPGAGYRSS NIRTSSMMMM GASEVAQEVE VRETVRKYFP
 751 ETWIWDMVPL DLSGDGELPV KVPDTITEWK ASAFCLSGTT GLGLSSTISH
 801 KVFQPFFLEL TLPYSVVRGE AFILKATVLN YMPHCIRIHV SLEMSPDFLA
 851 VPVGSHEDSH CICGNERKTV SWAVTPKSLG EVNFTATAEA LQSPELCGNK
 901 VAEVPALVQK DTVVKPVIVE PEGIEKEQTY NTLLCPQDAE LQENWTLDLP
 951 ANVVEGSARA TQSVLGDILG SAMQNLQNLL QMPYGCGEQN MVLFVPNIYV
1001 LEYLNETQQL TEAIKSKAIS YLISGYQRQL NYQHSDGSYS TFGDRGMRHS
1051 QGNTWLTAFV LKAFAQAQSY IYIEKTHITN AFNWLSMKQR ENGCFQQSGS
1101 LLNNAMKGGV DDEVTLSAYI TIALLEMPLP VTHSVVRNAL FCLETAWASI
1151 SNSQESHVYT KALLAYAFAL AGNRAKRSEV LESLNKDAVN EEESVHWQRP
1201 KNVEENVREM RSFSYKPRAP SAEVEMTAYV LLAYLTSASS RPTRDLSSSD
1251 LTTASKIVKW ISKQQNSHGG FSSTQDTVVA LQALSKYGAA TFTKSNKEVS
1301 VTIESSGTVS GTLHVNNGNR LLLQEVRLAD LPGNYITKVS GSGCVYLQTS
1351 LKYNILPEAE GEAPFTLKVN TLPLNFDKAE HHRKFQIHIN VSYIGERPNS
1401 NMVIVDVKMV SGFIPVKPSV KKLQDQSNIQ RTEVNTNHVL IYIEKLTNQT
1451 MGFSFAVEQD IPVKNLKPAP VKVYDYYETD EFAIEEYSAP FSSDSEQGNA
1501
```

Albumin



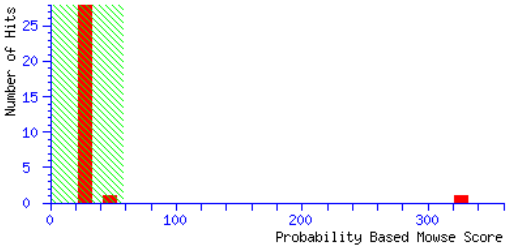
PMF of albumin (top). MS/MS spectrum of its 1609.732 peptide (bottom)

Mascot Search Results

```
User            : Katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 29 Oct 2008 at 03:54:36 GMT
Top Score       : 326 for gi|19705431, albumin [Rattus norvegicus]
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|19705431 **Mass:** 70670 **Score:** 326 **Expect:** 9.9e-029 **Queries matched:** 28
albumin [Rattus norvegicus]
gi|124028612 **Mass:** 70682 **Score:** 326 **Expect:** 9.9e-029 **Queries matched:** 28
Serum albumin precursor
gi|55391508 **Mass:** 70710 **Score:** 308 **Expect:** 6.3e-027 **Queries matched:** 27
Albumin [Rattus norvegicus]

Protein View

Match to: **gi|19705431** Score: **326** Expect: **9.9e-029**
albumin [Rattus norvegicus]

Nominal mass (M_r): **70670**; Calculated pI value: **6.09**
NCBI BLAST search of gi|19705431 against nr
Unformatted sequence string for pasting into other applications
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|55628 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **41**
Number of mass values matched: **28**
Sequence Coverage: **49%**

Matched peptides shown in **Bold Red**

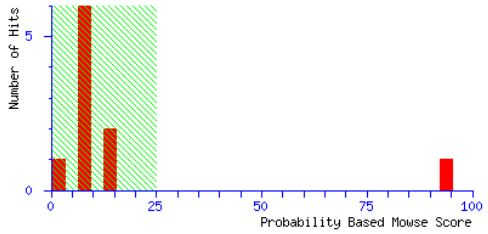
```
     1 MKWVTFLLLL FISGSAFSRG VFRREAHKSE IAHRFKDLGE QHFKGLVLIA
    51 FSQYLQKCPY EEHIKLVQEV TDFAKTCVAD ENAENCDKSI HTLFGDKLCA
   101 IPKLRDNYGE LADCCAKQEP ERNECFLQHK DDNPNLPPFQ RPEAEAMCTS
   151 FQENPTSFLG HYLHEVARRH PYFYAPELLY YAEKYNEVLT QCCTESDKAA
   201 CLTPKLDAVK EKALVAAVRQ RMKCSSMQRF GERAFKAWAV ARMSQRFPNA
   251 EFAEITKLAT DVTKINKECC HGDLLECADD RAELAKYMCE NQATISSKLQ
   301 ACCDKPVLQK SQCLAETEHD NIPADLPSIA ADFVEDKEVC KNYAEAKDVF
   351 LGTFLYEYSR RHPDYSVSLL LRLAKKYEAT LEKCCAEGDP PACYGTVLAE
   401 FQPLVEEPKN LVKTNCELYE KLGEYGFQNA VLVRYTQKAP QVSTPTLVEA
   451 ARNLGRVGTK CCTLPEAQRL PCVEDYLSAI LNRLCVLHEK TPVSEKVTKC
   501 CSGSLVERRP CFSALTVDET YVPKEFKAET FTFHSDICTL PDKEKQIKKQ
   551 TALAELVKHK PKATEDQLKT VMGDFAQFVD KCCKAADKDN CFATEGPNLV
   601 ARSKEALA
```

Mascot Search Results

User	: **Katie**
Email	:
Search title	: **database search**
MS data file	: **DATA.TXT**
Database	: **NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy	: **Rattus (39483 sequences)**
Timestamp	: **29 Oct 2008 at 03:58:02 GMT**
Protein hits	: gi|19705431 albumin [Rattus norvegicus]
	gi|109506635 PREDICTED: similar to MEGF10 protein [Rattus norvegicus]
	gi|9845251 rapamycin and FKBP12 target-1 protein [Rattus norvegicus]
	gi|115485 Calcitonin gene-related peptide 1 precursor (Calcitonin gene-related
	gi|109474014 PREDICTED: similar to cell adhesion molecule with homology to L1CAM
	gi|16128844 conserved protein with nucleoside triphosphate hydrolase domain [Esc
	gi|34860298 PREDICTED: similar to matrix metalloproteinase 1a (interstitial coll
	gi|34878065 PREDICTED: similar to Putative pre-mRNA-splicing factor ATP-dependen
	gi|16128551 bacteriophage N4 receptor, outer membrane subunit [Escherichia coli
	gi|1710906 coatomer beta subunit [Rattus norvegicus]

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 25 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|19705431** Score: **94**
albumin [Rattus norvegicus]
Found in search of DATA.TXT

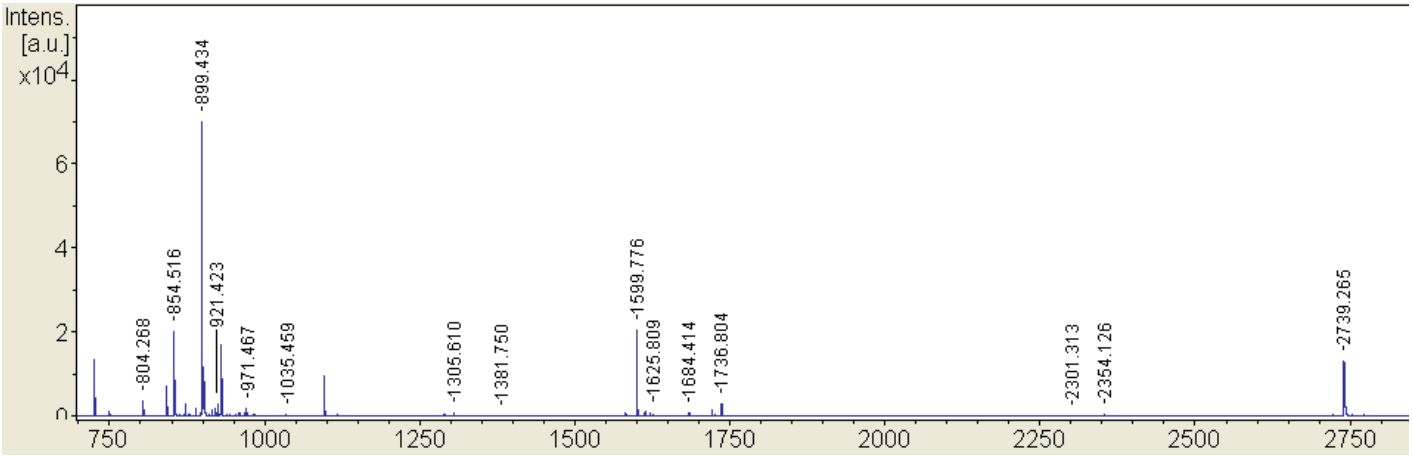
Nominal mass (M_r): **70670**; Calculated pI value: **6.09**
NCBI BLAST search of gi|19705431 against nr
Unformatted sequence string for pasting into other applications
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|55628 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **2%**

Matched peptides shown in **Bold Red**

```
   1 MKWVTFLLLL FISGSAFSRG VFRREAHKSE IAHRFKDLGE QHFKGLVLIA
  51 FSQYLQKCPY EEHIKLVQEV TDFAKTCVAD ESAENCDKSI HTLFGDKLCA
 101 IPKLRDNYGE LADCCAKQEP ERNECFLQHK DDNPNLPPFQ RPEAEAMCTS
 151 FQENPTSFLG HYLHEVARRH PYFYAPELLY YAEKYNEVLT QCCTESDKAA
 201 CLTPKLDAVK EKALVAAVRQ RMKCSSMQRF GERAFKAWAV ARMSQRFPNA
 251 EFAEITKLAT DVTKINKECC HGDLLECADD RAELAKYMCE NQATISSKLQ
 301 ACCDKPVLQK SQCLAETEHD NIPADLPSIA ADFVEDKEVC KNYAEAKDVF
 351 LGTFLYEYSR RHPDYSVSLL LRLAKKYEAT LEKCCAEGDP PACYGTVLAE
 401 FQPLVEEPKN LVKTNCELYE KLGEYGFQNA VLVRYTQKAP QVSTPTLVEA
 451 ARNLGRVGTK CCTLPEAQRL PCVEDYLSAI LNRLCVLHEK TPVSEKVTKC
 501 CSGSLVERRP CFSALTVDET YVPKEFKAET FTFHSDICTL PDKEKQIKKQ
 551 TALAELVKHK PKATEDQLKT VMGDFAQFVD KCCKAADKDN CFATEGPNLV
 601 ARSKEALA
```

Apolipoprotein E



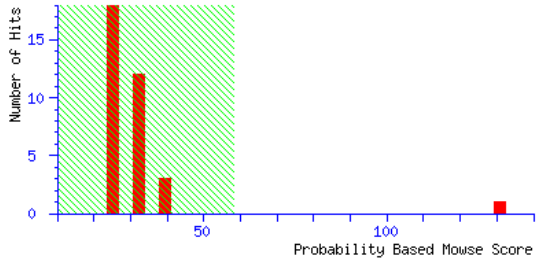
PMF of apolipoprotein E

Mascot Search Results

User : **katie**
Email :
Search title : **database search**
Database : **NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy : **Rattus (39483 sequences)**
Timestamp : **26 Dec 2008 at 05:29:06 GMT**
Top Score : **131 for gi|1703338, Apolipoprotein E precursor (Apo-E)**

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help
Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|1703338 **Mass:** 35788 **Score:** 131 **Expect:** 3.1e-009 **Queries matched:** 15
Apolipoprotein E precursor (Apo-E)
gi|37805241 **Mass:** 35798 **Score:** 131 **Expect:** 3.1e-009 **Queries matched:** 15
Apolipoprotein E [Rattus norvegicus]

Protein View

Match to: **gi|1703338** Score: **131** Expect: **3.1e-009**
Apolipoprotein E precursor (Apo-E)

Nominal mass (M_r): **35788**; Calculated pI value: **5.23**
NCBI BLAST search of gi|1703338 against nr
Unformatted sequence string for pasting into other applications

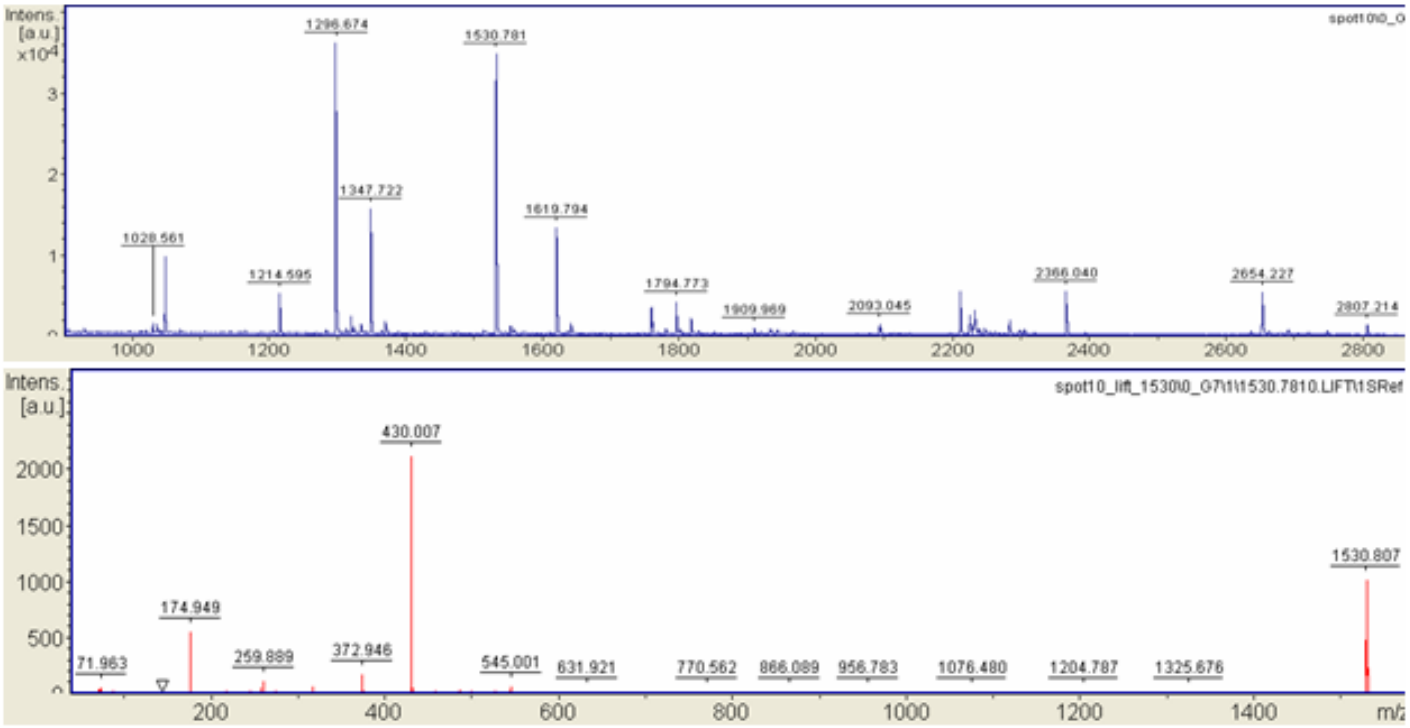
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|202958 from Rattus norvegicus
gi|55824759 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **31**
Number of mass values matched: **15**
Sequence Coverage: **29%**

Matched peptides shown in **Bold Red**

```
     1 MKALWALLLV PLLTGCLAEG ELEVTDQLPG QSDQPWEQAL NRFWDYLRWV
    51 QTLSDQVQEE LQSSQVTQEL TVLMEDTMTE VKAYKKELEE QLGPVAEETR
   101 ARLAKEVQAA QARLGADMED LRNRLGQYRN EVNTMLGQST EELRSRLSTH
   151 LRKMRKRLMR DADDLQKRLA VYKAGAQEGA ERGVSAIRER LGPLVEQGRQ
   201 RTANLGAGAA QPLRDRAQAL SDRIRGRLEE VGNQARDRLE EVREQMEEVR
   251 SKMEEQTQQI RLQAEIFQAR IKGWFEPLVE DMQRQWANLM EKIQASVATN
   301 SIASTTVPLE NQ
```

Apolipoprotein H



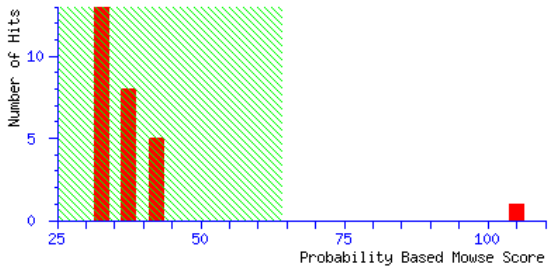
PMF of apolipoprotein H (top). MS/MS spectrum of its 1530.807 peptide (bottom).

Mascot Search Results

User : Katie
Email :
Search title : database search
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Rodentia (Rodents) (149892 sequences)
Timestamp : 29 Oct 2008 at 04:08:03 GMT
Top Score : 105 for gi|57528174, apolipoprotein H [Rattus norvegicus]

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 64 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help
Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|57528174 **Mass:** 39743 **Score:** 105 **Expect:** 4.7e-006 **Queries matched:** 8
apolipoprotein H [Rattus norvegicus]
gi|74381906 **Mass:** 30401 **Score:** 98 **Expect:** 2.2e-005 **Queries matched:** 7
apolipoprotein H [Rattus norvegicus]

Protein View

Match to: **gi|57528174** Score: **105** Expect: **4.7e-006**
apolipoprotein H [Rattus norvegicus]

Nominal mass (M_r): **39743**; Calculated pI value: **8.58**
NCBI BLAST search of gi|57528174 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|56971279 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **36**
Number of mass values matched: **8**
Sequence Coverage: **35%**

Matched peptides shown in **Bold Red**

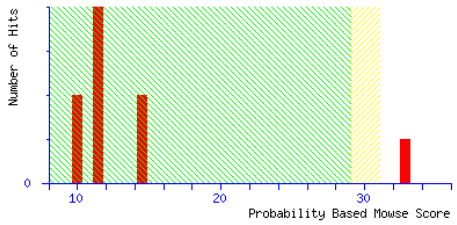
```
     1 MISPALIFFS AFLCHVAIAG RTCPKPDELP FAVVVPLKTF YDPGEQIVYS
    51 CKPGYVSRGG MRRFTCPLTG MWPINTLKCI PRVCPFAGIL ENGVVRYTTF
   101 EYPNTIGFAC NPGYYLNGTS SSKCTEEGKW SPELPVCARI TCPPPPIPKF
   151 AALKEYKTSV GNSSFYQDTV VFKCLPHFAM FGNDTVTCTA HGNWTQLPEC
   201 REVKCPFPSR PDNGFVNYPA KPVLSYKDKA VFGCHETYKL DGPEEVECTK
   251 TGNWSALPSC KASCKLSVKK ATVLYQGQRV KIQDQFKNGM MHGDKVHFYC
   301 KNKEKKCSYT EEAQCIDGTI EIPKCFKEHS SLAFWKTDAS DVTPC
```

{MATRIX SCIENCE} Mascot Search Results

User	:	**Katie**
Email	:	
Search title	:	**database search**
MS data file	:	**DATA.TXT**
Database	:	**NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy	:	**Rodentia (Rodents) (149892 sequences)**
Timestamp	:	**29 Oct 2008 at 04:11:00 GMT**
Protein hits	:	gi|1351954 Beta-2-glycoprotein 1 precursor (Beta-2-glycoprotein I) (Apolipoprotein H)
		gi|119361073 Abhydrolase domain-containing protein 9 precursor
		gi|71043724 proteasome (prosome, macropain) subunit, beta type 10 [Rattus norvegicus]
		gi|350611 transaminase,Glu oxaloacetic
		gi|930347 dishevelled-1 protein [Mus musculus]
		gi|15803709 ribosome-binding factor A [Escherichia coli O157:H7 EDL933]
		gi|26341928 unnamed protein product [Mus musculus]
		gi|26326891 unnamed protein product [Mus musculus]
		gi|16130780 predicted oxidoreductase, Fe-S subunit [Escherichia coli K12]

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 29 indicate peptides with significant homology.
Individual ions scores > 31 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|1351954** Score: **33**
Beta-2-glycoprotein 1 precursor (Beta-2-glycoprotein I) (Apolipoprotein H) (Apo-H) (B2GPI)
Found in search of DATA.TXT

Nominal mass (M_r): **34316**; Calculated pI value: **8.59**
NCBI BLAST search of gi|1351954 against nr
Unformatted sequence string for pasting into other applications

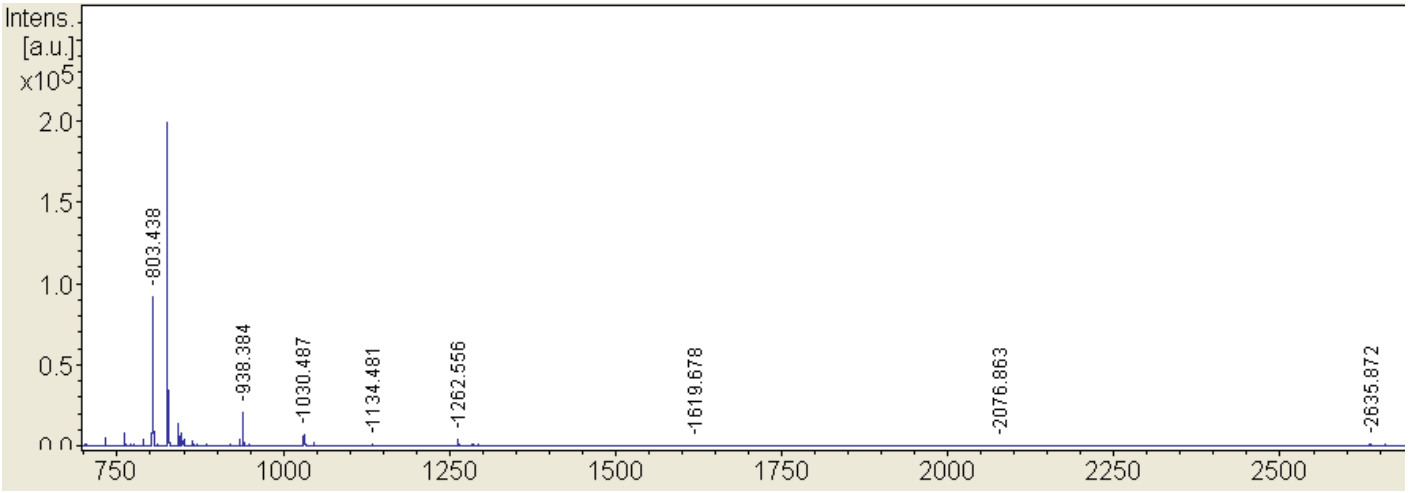
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|57525 from Rattus rattus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **4%**

Matched peptides shown in **Bold Red**

```
     1 MISPALIFFS AFLCHVAIAG RRMWPINTLK CTPRVCPFAG ILENGVVRYT
    51 TFEYPNTIGF ACNPGYYLNG TSSSKCTEEG KWSPELPVCA RITCPPPPIP
   101 KFAALKEYKT SVGNSSFYQD TVVFKCLPHF AMFGNDTVTC TAHGNWTQLP
   151 ECREVKCPFP SRPDNGFVNY PAKPVLSYKD KAVFGCHETY KLDGPEEVEC
   201 TKTGNWSALP SCKASCKLSV KKATVLYQGQ RVKIQDQFKN GMMHGDKVHF
   251 YCKNKEKKCS YTEEAQCIDG TIEIPKCFKE HSSLAFWKTD ASDVTPC
```

Apolipoprotein M



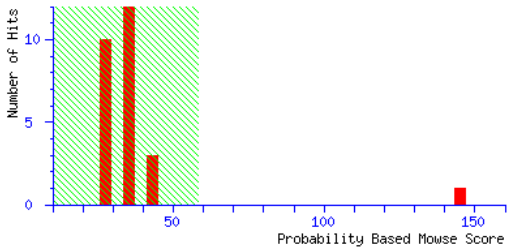
PMF of apolipoprotein M.

Mascot Search Results

```
User            : katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 26 Dec 2008 at 05:39:02 GMT
Top Score       : 145 for gi|9506391, apolipoprotein M [Rattus norvegicus]
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|9506391 **Mass:** 21841 **Score:** 145 **Expect:** 1.2e-010 **Queries matched:** 10
apolipoprotein M [Rattus norvegicus]

Protein View

Match to: **gi|9506391** Score: **145** Expect: **1.2e-010**
apolipoprotein M [Rattus norvegicus]

Nominal mass (M_r): **21841**; Calculated pI value: **5.73**
NCBI BLAST search of gi|9506391 against nr
Unformatted sequence string for pasting into other applications

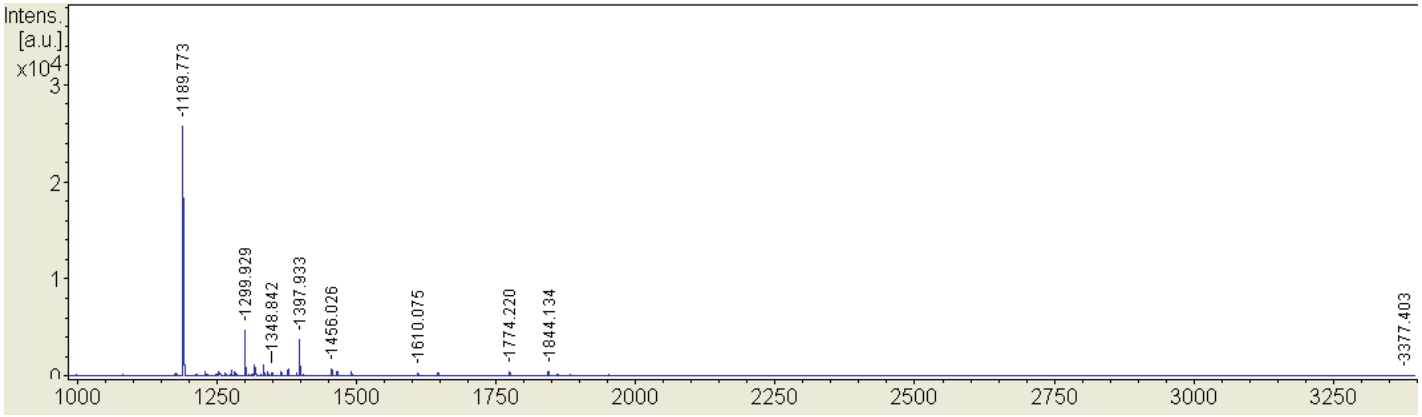
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|19856159 from Rattus norvegicus
gi|6683949 from Rattus norvegicus
gi|46237618 from Rattus norvegicus
gi|59808177 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **14**
Number of mass values matched: **10**
Sequence Coverage: **36%**

Matched peptides shown in **Bold Red**

```
     1 MFHQVWAALL YLYGLLFNSM NQCPEHSQLM TLGMDDKETP EPHLGLWYFI
    51 AGAAPTMEEL ATFDQVDNIV FNMAAGSAPR QLQLRATIRT KNGVCVPRKW
   101 TYHLTEGKGN TELRTEGRPD MKTDLFSISC PGGIMLKETG QGYQRFLLYN
   151 RSPHPPEECV EEFQSLTSCL DFKAFLVTPR NQEACPLSSK
```

Coagulation factor 2



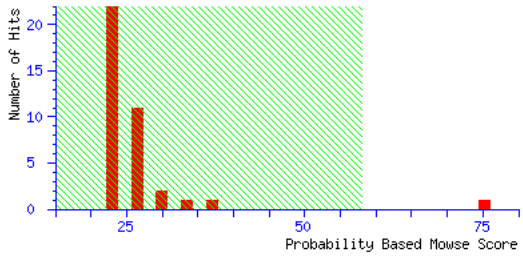
PMF of coagulation factor 2

Mascot Search Results

```
User            : katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 26 Dec 2008 at 04:50:37 GMT
Top Score       : 75 for gi|12621076, coagulation factor 2 [Rattus norvegicus]
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 20

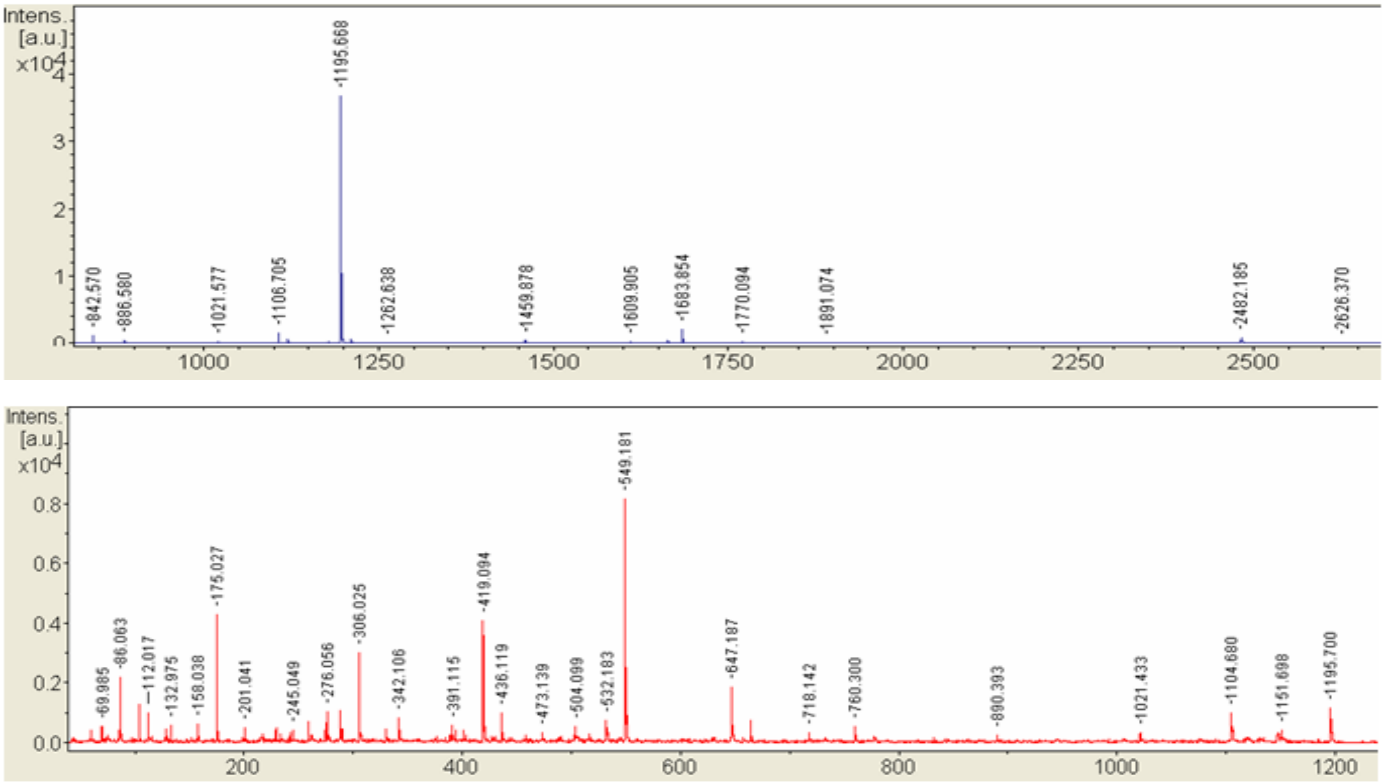
Re-Search All | Search Unmatched

1. gi|12621076 **Mass:** 71792 **Score:** 75 **Expect:** 0.0012 **Queries matched:** 11
coagulation factor 2 [Rattus norvegicus]

Protein View

```
Match to: gi|12621076 Score: 75 Expect: 0.0012
coagulation factor 2 [Rattus norvegicus]

Nominal mass (Mr): 71792; Calculated pI value: 6.28
NCBI BLAST search of gi|12621076 against nr
Unformatted sequence string for pasting into other applications
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|135809 from Rattus norvegicus
gi|56970 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: 22
Number of mass values matched: 11
Sequence Coverage: 17%

Matched peptides shown in Bold Red

     1 MLHVRGLGLP GCLALAALAS LVHSQHVFLA PQQALSLLQR VRRANSGFLE
    51 ELRKGNLERE CVEEQCSYEE AFEALESPQD TDVFWAKYTV CDSVRKPRET
   101 FMDCLEGRCA MDLGLNYHGN VSVTHTGIEC QLWRSRYPHR PDINSTTHPG
   151 ADLKENFCRN PDSSTSGPWC YTTDPTVRRE ECSIPVCGQE GRTTVKMTPR
   201 SRGSKENLSP PLGECLLERG RLYQGNLAVT TLGSPCLAWD SLPTKTLSKY
   251 QNFDPEVKLV QNFCRNPDRD EEGAWCFVAQ QPGFEYCSLN YCDEAVGEEN
   301 HDGDESIAGR TTDAEFHTFF DERTFGLGEA DCGLRPLFEK KSLTDKTEKE
   351 LLDSYIDGRI VEGWDAEKGI APWQVMLFRK SPQELLCGAS LISDRWVLTA
   401 AHCILYPPWD KNFTENDLLV RIGKHSRTRY ERNVEKISML EKIYIHPRYN
   451 WRENLDRDIA LLKLKKPVPF SDYIHPVCLP DKQTVTSLLQ AGYKGRVTGW
   501 GNLRETWTTN INEIQPSVLQ VVNLPIVERP VCKASTRIRI TDNMFCAGFK
   551 VNDTKRGDAC EGDSGGPFVM KSPYNHRWYQ MGIVSWGEGC DRNGKYGFYT
   601 HVFRLKRWMQ KVIDQHR
```

Complement component 3



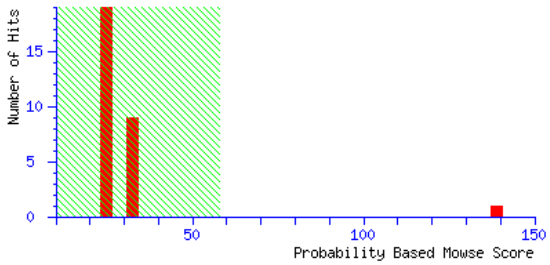
PMF of complement component 3 (top). MS/MS spectrum of its 1195.668 (bottom).

Mascot Search Results

```
User            : katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 17 Dec 2008 at 03:45:38 GMT
Top Score       : 139 for gi|554423, complement component C3
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|554423 **Mass:** 32414 **Score:** 139 **Expect:** 5e-010 **Queries matched:** 11
complement component C3
gi|8393024 **Mass:** 187825 **Score:** 69 **Expect:** 0.0053 **Queries matched:** 11
complement component 3 [Rattus norvegicus]

Protein View

Match to: **gi|554423** Score: **139** Expect: **5e-010**
complement component C3

Nominal mass (M_r): **32414**; Calculated pI value: **5.73**
NCBI BLAST search of gi|554423 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **19**
Number of mass values matched: **11**
Sequence Coverage: **39%**

Matched peptides shown in **Bold Red**

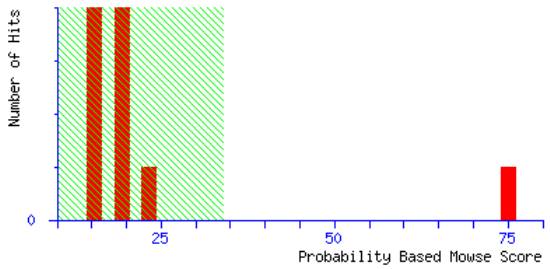
```
     1 PGSLLRSEET KQNEGFSLTA KGKGQGTLSV VTVYHAKVKG KTTCKKFDLR
    51 VTIKPAPETA KKPQDAKSSM ILDICTRYLG DVDATMSILD ISMMTGFIPD
   101 TNDLELLSSG VDRYISKYEM DKAFSNKNTL IIYLEKISHS EEDCLSFKVH
   151 QFFNVGLIQP GSVKVYSYYN LEESCTRFYH PEKDDGMLSK LCHNEMCRCA
   201 EENCFMHQSQ DQVSLNERLD KACEPGVDYV YKTKLTTIEL SDDFDEYIMT
   251 IEQVIKSGSD EVQAGQERRF ISHVKCRNAL K
```

Mascot Search Results

User	:	**katie**
Email	:	
Search title	:	**database search**
MS data file	:	**DATA.TXT**
Database	:	**NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy	:	**Rattus (39483 sequences)**
Timestamp	:	**17 Dec 2008 at 03:53:00 GMT**
Protein hits	:	gi|554423 complement component C3
		gi|16131124 serine endoprotease, periplasmic [Escherichia coli K12]
		gi|15803093 ribonuclease III [Escherichia coli O157:H7 EDL933]
		gi|16129553 predicted DNA-binding transcriptional regulator [Escher
		gi|109480979 PREDICTED: similar to slit homolog 1 [Rattus norvegicus]
		gi|109475443 PREDICTED: similar to microfilament and actin filament
		gi|109464341 PREDICTED: similar to GTP-binding protein ARD-1 (ADP-ri
		gi|109508663 PREDICTED: hypothetical protein isoform 1 [Rattus norve
		gi|109487537 PREDICTED: similar to ankyrin repeat and MYND domain co
		gi|109503443 PREDICTED: similar to deleted in liver cancer 1 isoform

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 34 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|554423** Score: **75**
complement component C3
Found in search of DATA.TXT

Nominal mass (M_r): **32414**; Calculated pI value: **5.73**
NCBI BLAST search of gi|554423 against nr
Unformatted sequence string for pasting into other applications

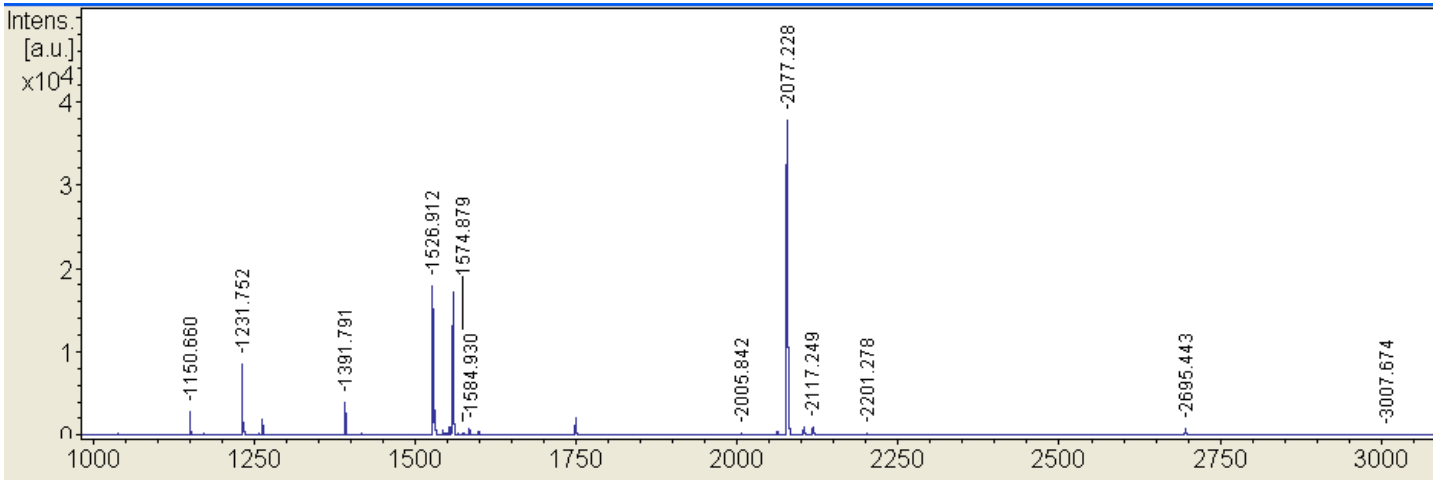
Taxonomy: Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: 3%

Matched peptides shown in **Bold Red**

```
     1 PGSLLRSEET KQNEGFSLTA KGKGQGTLSV VTVYHAKVKG KTTCKKFDLR
    51 VTIKPAPETA KKPQDAKSSM ILDICTRYLG DVDATMSILD ISMMTGFIPD
   101 TNDLELLSSG VDRYISKYEM DKAFSNKNTL IIYLEKISHS EEDCLSFKVH
   151 QFFNVGLIQP GSVKVYSYYN LEESCTRFYH PEKDDGMLSK LCHNEMCRCA
   201 EENCFMHQSQ DQVSLNERLD KACEPGVDYV YKTKLTTIEL SDDFDEYIMT
   251 IEQVIKSGSD EVQAGQERRF ISHVKCRNAL K
```

C-reactive protein



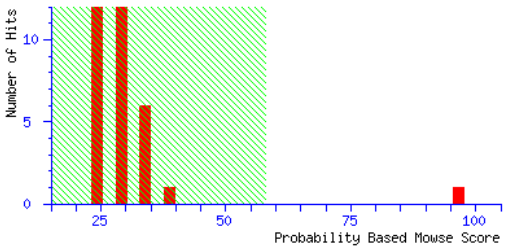
PMF of C-reactive protein

Mascot Search Results

```
User            : katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 26 Dec 2008 at 05:21:48 GMT
Top Score       : 97 for gi|8393197, C-reactive protein, petaxin related [Rattus norvegicus]
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary | Help
Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|8393197 **Mass:** 25737 **Score:** 97 **Expect:** 8.4e-006 **Queries matched:** 7
C-reactive protein, petaxin related [Rattus norvegicus]

Protein View

Match to: **gi|8393197** Score: **97** Expect: **8.4e-006**
C-reactive protein, petaxin related [Rattus norvegicus]

Nominal mass (M_r): **25737**; Calculated pI value: **4.89**
NCBI BLAST search of gi|8393197 against nr
Unformatted sequence string for pasting into other applications

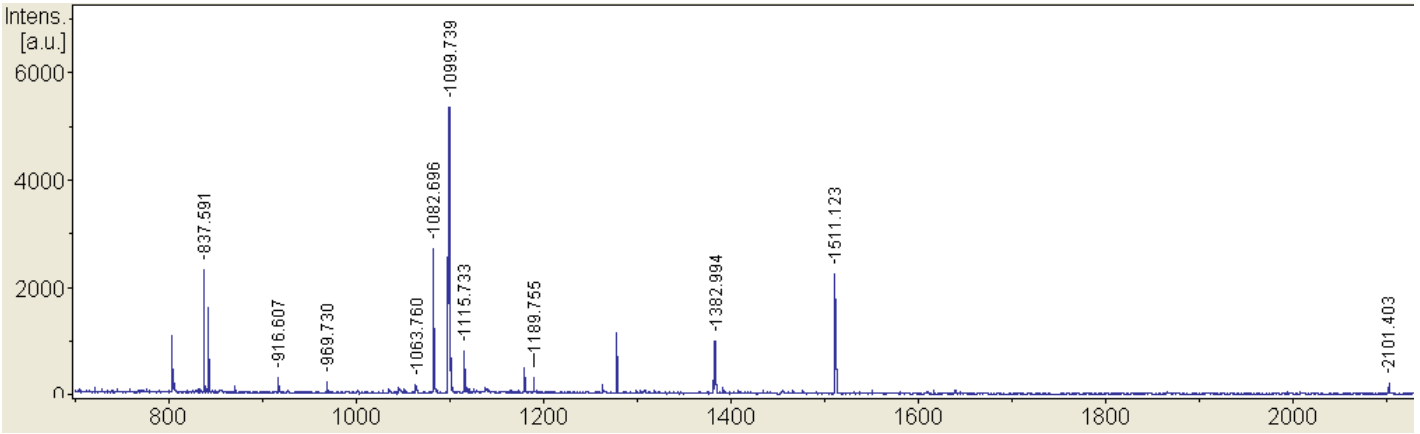
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|1345834 from Rattus norvegicus
gi|203592 from Rattus norvegicus
gi|60552095 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **19**
Number of mass values matched: **7**
Sequence Coverage: **41%**

Matched peptides shown in **Bold Red**

```
     1 MEKLLWCLLI TISFSQAFGH EDMSKQAFVF PGVSATAYVS LEAESKKPLE
    51 AFTVCLYAHA DVSRSFSIFS YATKTSFNEI LLFWTRGQGF SIAVGGPEIL
   101 FSASEIPEVP THICATWESA TGIVELWLDG KPRVRKSLQK GYIVGTNASI
   151 ILGQEQDSYG GFDANQSLV GDIGDVNMWD FVLSPEQINA VYVGRVFSPN
   201 VLNWRALKYE THGDVFIKPQ LWPLTDCCES
```

Fetuin beta



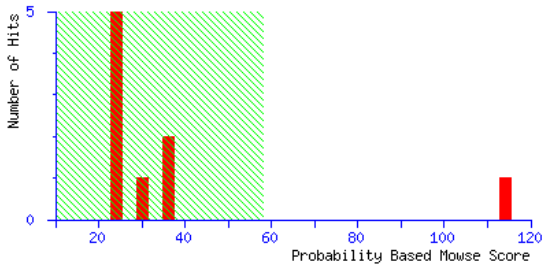
PMF of fetuin beta

{MATRIX SCIENCE} Mascot Search Results

```
User            : katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 26 Dec 2008 at 04:55:41 GMT
Top Score       : 114 for gi|17865327, fetuin beta [Rattus norvegicus]
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|17865327 **Mass:** 42361 **Score:** 114 **Expect:** 1.6e-007 **Queries matched:** 9
fetuin beta [Rattus norvegicus]
gi|47682636 **Mass:** 44054 **Score:** 113 **Expect:** 2e-007 **Queries matched:** 9
Fetub protein [Rattus norvegicus]

Protein View

Match to: **gi|17865327** Score: **114** Expect: **1.6e-007**
fetuin beta [Rattus norvegicus]

Nominal mass (M_r): **42361**; Calculated pI value: **6.71**
NCBI BLAST search of gi|17865327 against nr
Unformatted sequence string for pasting into other applications

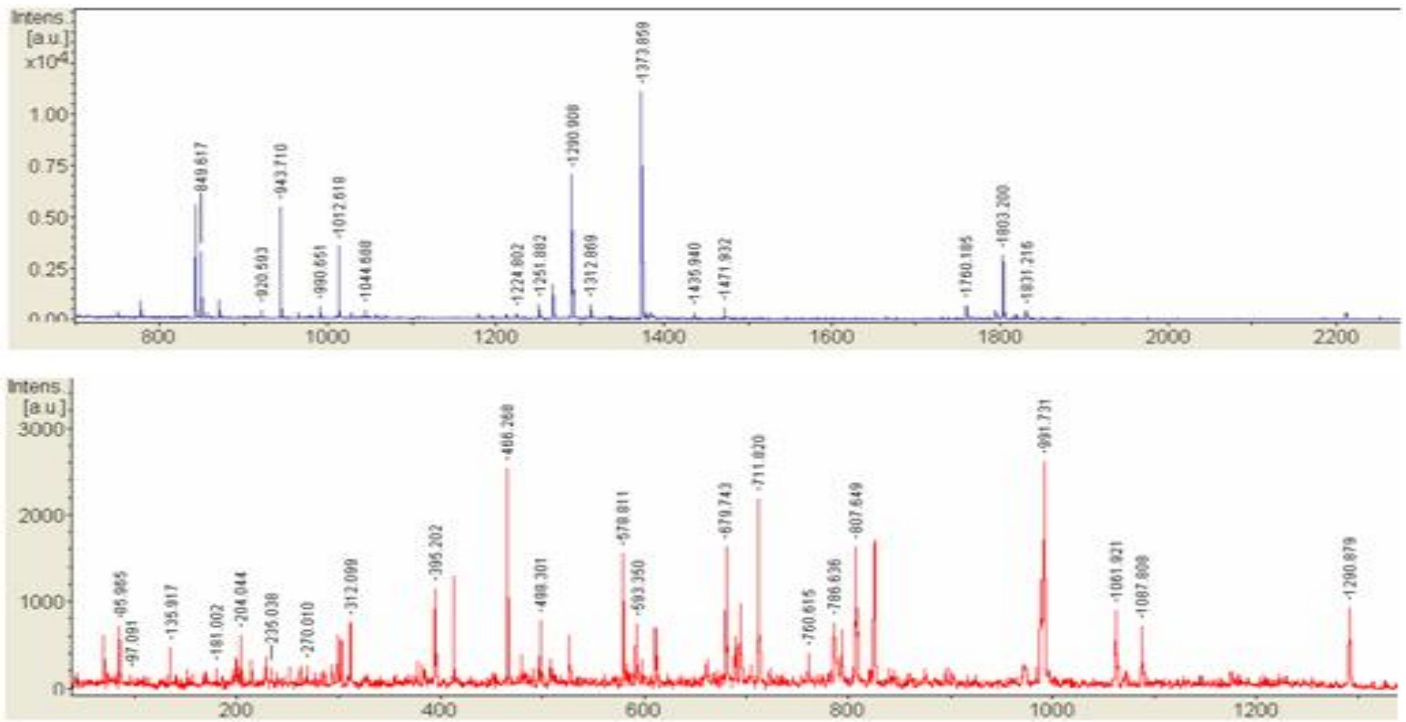
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|11131967 from Rattus norvegicus
gi|6562849 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **12**
Number of mass values matched: **9**
Sequence Coverage: **17%**

Matched peptides shown in **Bold Red**

```
     1 MGVLRLLVLC TLAACCVARS PPAPPLPNAP FAPLRPLGCN DSEVLAVAGF
    51 ALQNINRVQK DGYMLTLNRV HDARVHRQED MGSLFYLTLD VLETGCHVLS
   101 RKALKDCGPR IFYETVHGQC KAMFHVNKPR RVLYLPAYNC TLRPVSKRKI
   151 HSMCPDCPHP VDLSAPSVLE AATESLAKFN SENPSKQYAL VKVTKATTQW
   201 VVGPSYFVEY LIKESPCTQS QDSCSLQASD SEPVGLCQGS LIKSPGVPPQ
   251 RFKKTVTVSC EFFESQDQVP GGENPADTQD AKKLPQKNTA PTSSPSITAP
   301 RGSIQHLPEQ EEPEDSKGKS PEEPFPVQLD LTTNPQGDTL DVSFLYLEPE
   351 EKKLVVLPFP GKEQRSPECP GPEKQRTP
```

Haptoglobin



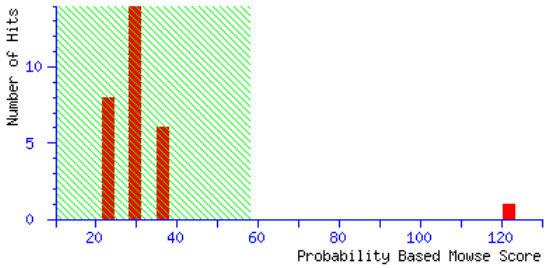
PMF of haptoglobin (top). MS/MS spectrum of its 1290.908 (bottom).

{MATRIX SCIENCE} Mascot Search Results

User : **katie**
Email :
Search title : **database search**
Database : **NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy : **Rattus (39483 sequences)**
Timestamp : **17 Dec 2008 at 04:02:11 GMT**
Top Score : **122 for gi|60097941, haptoglobin [Rattus norvegicus]**

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help
Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|60097941 **Mass:** 39052 **Score:** **122** **Expect:** 2.5e-008 **Queries matched:** 11
haptoglobin [Rattus norvegicus]

Protein View

Match to: **gi|60097941** Score: **122** Expect: **2.5e-008**
haptoglobin [Rattus norvegicus]

Nominal mass (M_r): **39052**; Calculated pI value: **6.10**
NCBI BLAST search of gi|60097941 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|124053462 from Rattus norvegicus
gi|59808182 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **22**
Number of mass values matched: **11**
Sequence Coverage: **27%**

Matched peptides shown in **Bold Red**

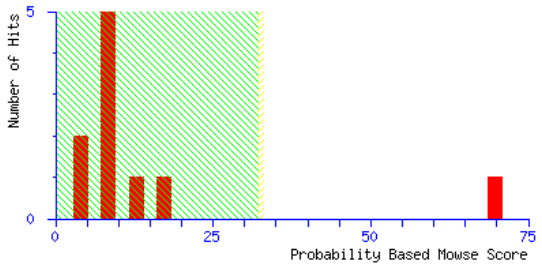
```
     1 MRALGAVVTL LLWGQLFAVE LGNDATDIED DSCPKPPEIA NGYVEHLVRY
    51 RCRQFYKLQT EGDGIYTLNS EKQWVNPAAG DKLPKCEAVC GKPKHPVDQV
   101 QRIIGGSMDA KGSFPWQAKM ISRHGLTTGA TLISDQWLLT TAQNLFLNHS
   151 ENATAKDIAP TLTLYVGKNQ LVEIEKVVLH PERSVVDIGL IKLKQKVLVT
   201 EKVMPICLPS KDYVAPGRMG YVSGWGRNVN FRFTERLKYV MLPVADQEKC
   251 ELHYEKSTVP EKKGAVSPVG VQPILNKHTF CAGLTKYEED TCYGDAGSAF
   301 AVHDTEEDTW YAAGILSFDK SCAVAEYGVY VRATDLKDWV QETMAKN
```

Mascot Search Results

User : **katie**
Email :
Search title : **database search**
MS data file : **DATA.TXT**
Database : **NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy : **Rattus (39483 sequences)**
Timestamp : **17 Dec 2008 at 04:05:26 GMT**
Protein hits : gi|204655 haptoglobin (Hp)
gi|109488177 PREDICTED: similar to zinc finger protein 21 isoform 1 [Ra
gi|16131445 predicted DNA-binding transcriptional repressor [Escherich
gi|109496224 PREDICTED: similar to putative pheromone receptor (Go-VN4)
gi|109487708 PREDICTED: similar to CTP synthase (UTP--ammonia ligase)
gi|62652280 PREDICTED: similar to NGFI-A-binding protein 2 (EGR-1-binc
gi|16128477 short chain dehydrogenase [Escherichia coli K12]
gi|20302105 a disintegrin and metallopeptidase domain 6 [Rattus norveg
gi|109489297 PREDICTED: similar to ATP-binding cassette, sub-family A
gi|109504605 PREDICTED: similar to Zinc finger CCHC domain-containing p

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 32 indicate peptides with significant homology.
Individual ions scores > 33 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|60097941** Score: **70**
haptoglobin [Rattus norvegicus]
Found in search of DATA.TXT

Nominal mass (M_r): **39052**; Calculated pI value: **6.10**
NCBI BLAST search of gi|60097941 against nr
Unformatted sequence string for pasting into other applications

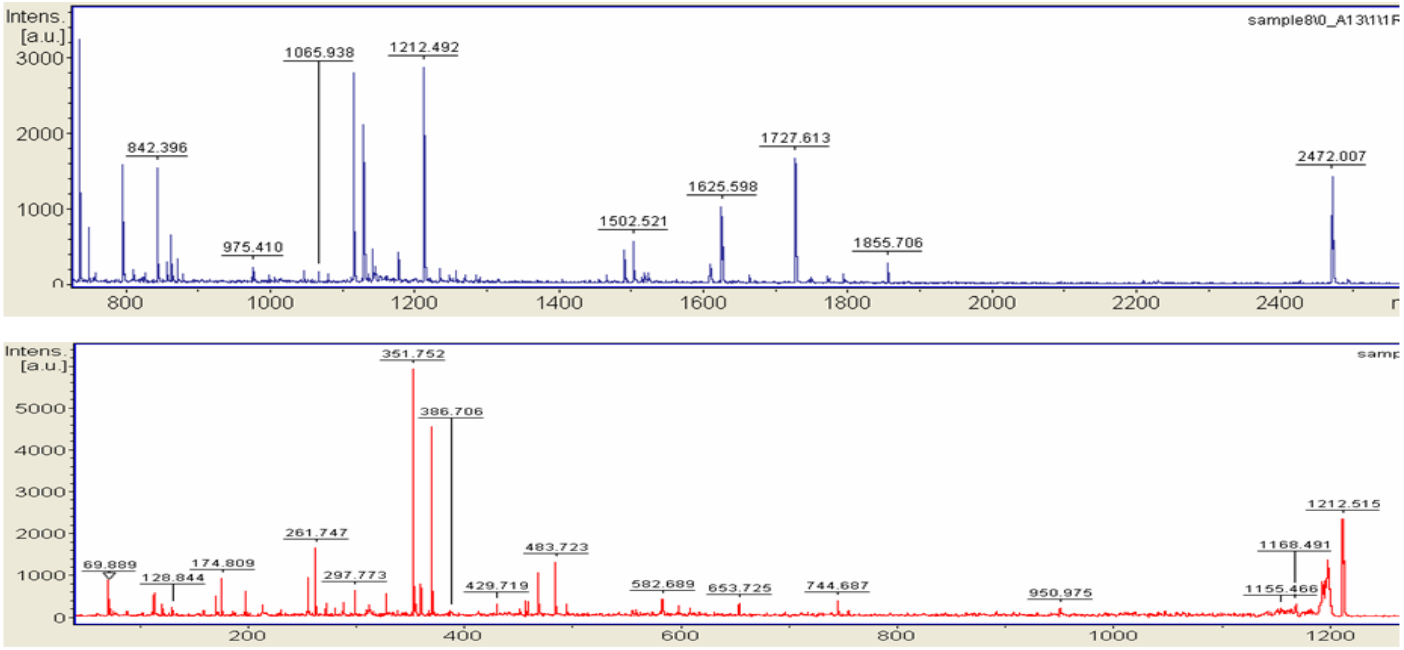
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|124053462 from Rattus norvegicus
gi|59808182 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **3%**

Matched peptides shown in **Bold Red**

```
     1 MRALGAVVTL LLWGQLFAVE LGNDATDIED DSCPKPPEIA NGYVEHLVRY
    51 RCRQFYKLQT EGDGIYTLNS EKQWVNPAAG DKLPKCEAVC GKPKHPVDQV
   101 QRIIGGSMDA KGSFPWQAKM ISRHGLTTGA TLISDQWLLT TAQNLFLNHS
   151 ENATAKDIAP TLTLYVGKNQ LVEIEKVVLH PERSVVDIGL IKLKQKVLVT
   201 EKVMPICLPS KDYVAPGRMG YVSGWGRNVN FRFTERLKYV MLPVADQEKC
   251 ELHYEKSTVP EKKGAVSPVG VQPILNKHTF CAGLTKYEED TCYGDAGSAF
   301 AVHDTEEDTW YAAGILSFDK SCAVAEYGVY VRATDLKDWV QETMAKN
```

Hemopexin precursor



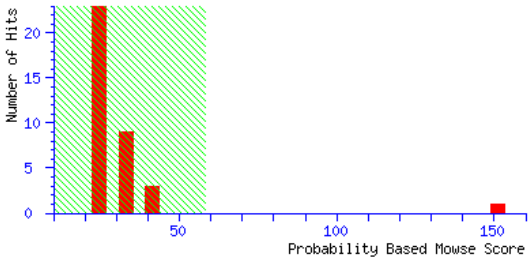
PMF of hemopexin precursor (top). MS/MS spectrum of its 1212.492 peptide (bottom).

Mascot Search Results

User : Katie
Email :
Search title : database search
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Rattus (39483 sequences)
Timestamp : 29 Oct 2008 at 02:32:13 GMT
Top Score : 151 for gi|122065203, Hemopexin precursor

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|122065203 **Mass:** 52060 **Score:** 151 **Expect:** 3.1e-011 **Queries matched:** 14
Hemopexin precursor
gi|16758014 **Mass:** 52000 **Score:** 135 **Expect:** 1.2e-009 **Queries matched:** 13
hemopexin [Rattus norvegicus]

Protein View

Match to: **gi|122065203** Score: **151** Expect: **3.1e-011**
Hemopexin precursor

Nominal mass (M_r): **52060**; Calculated pI value: **7.58**
NCBI BLAST search of gi|122065203 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|60688311 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **31**
Number of mass values matched: **14**
Sequence Coverage: **28%**

Matched peptides shown in **Bold Red**

```
  1 MARTVVALNI LVLLGLCWSL AVANPLPAAH ETVAKGENGT KPDSDVIEHC
 51 SDAWSFDATT MDHNGTMLFF KGEFVWRGHS GIRELISERW KNPVTSVDAA
101 FRGPDSVFLI KEDKVWVYPP EKKENGYPKL FQEEFPGIPY PPDAAVECHR
151 GECQSEGVLF FQGNRKWFWD FATRTQKERS WPAVGNCTAA LRWLERYYCF
201 QGNKFLRFNP VTGEVPPRYP LDARDYFISC PGRGHGKLRN GTAHGNSTHP
251 MHSRCNADPG LSALLSDHRG ATYAFSGSHY WRLDSSRDGW HSWPIAHHWP
301 QGPSAVDAAF SWDEKVYLIQ GTQVYVFLTK GGNNLVSGYP KRLEKELGSP
351 PGISLDTIDA AFSCPGSSKL YVTSGRRLWW LDLKSGAQAT WAELSWPHEK
401 VDGALCLEKS LGPYSCSSNG PNLFFIHGPN LYCYSSIDKL NAAKSLPQPQ
451 KVNSILGCSQ
```

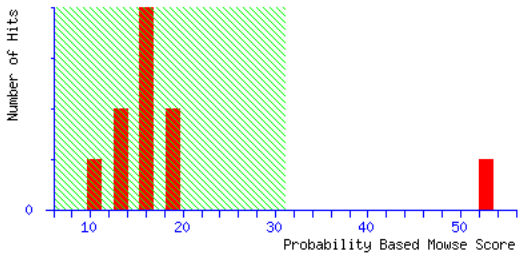
Mascot Search Results

User : **Katie**
Email :
Search title : **database search**
MS data file : **DATA.TXT**
Database : **NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy : **Rattus (39483 sequences)**
Timestamp : **29 Oct 2008 at 02:40:51 GMT**
Protein hits :

gi|16758014	hemopexin [Rattus norvegicus]
gi|62656231	PREDICTED: similar to Heparan sulfate glucosamine 3-O-sulfot
gi|396270	heat shock protein 70 [Rattus norvegicus]
gi|109505027	PREDICTED: similar to Zinc finger protein GLI3 [Rattus norve
gi|109510084	PREDICTED: hypothetical protein [Rattus norvegicus]
gi|16758760	mitogen activated protein kinase kinase kinase 1 [Rattus nor
gi|86129582	zinc finger, DHHC domain containing 14 [Rattus norvegicus]
gi|8170816	integrin beta 5 subunit [Rattus sp.]
gi|206515	quinone reductase (EC 1.6.99.2)
gi|109463105	PREDICTED: similar to gamma-aminobutyric acid (GABA-B) recep

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 31 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|16758014** Score: **53**
hemopexin [Rattus norvegicus]
Found in search of DATA.TXT

Nominal mass (M_r): **52000**; Calculated pI value: **7.58**
NCBI BLAST search of gi|16758014 against nr
Unformatted sequence string for pasting into other applications

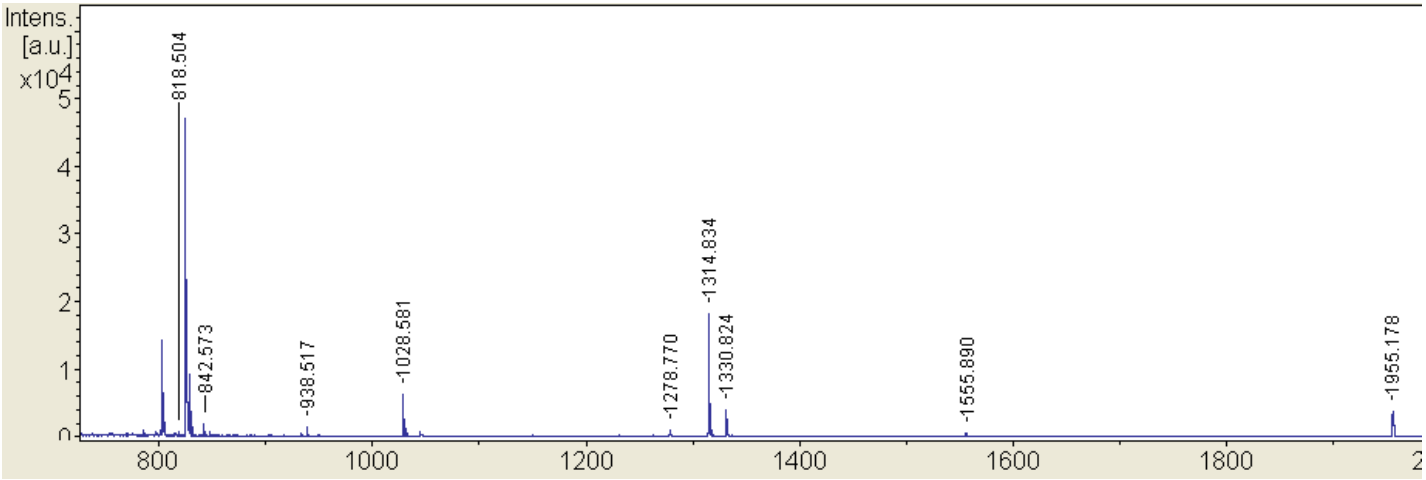
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|204621 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **2%**

Matched peptides shown in **Bold Red**

```
   1 MARTVVALNI LVLLGLCWSL AVANPLPAAH ETVAKGENGT KPDSDVIEHC
  51 SDAWSFDATT MDHNGTMLFF KGEFVWRGHS GIRELISERW KNPVTSVDAA
 101 FRGPDSVFLI KEDKVWVYPP EKKENGYPKL FQEESPGIPY PPDAAVECHR
 151 GECQSEGVLF FQGNRKWFWD FATRTQKERS WPAVGNCTAA LRWLERYYCF
 201 QGNKFLRFNP VTGEVPPRYP LDARDYFISC PGRGHGKLRN GTAHGNSTHP
 251 MHSRCNADPG LSALLSDHRG ATYAFSGSHY WRLDSSRDGW HSWPIAHHWP
 301 QGPSAVDAAF SWDEKVYLIQ GTQVYVFLTK GGNNLVSGYP KRLEKELGSP
 351 PGISLDTIDA AFSCPGSSKL YVTSGRRLWW LDLKSGAQAT WAELSWPHEK
 401 VDGALCLEKS LGPYSCSSNG PNLFFIHGPN LYCYSSIDKL NAAKSLPQPQ
 451 KVNSILGCSQ
```

Plasma glutathione peroxidase



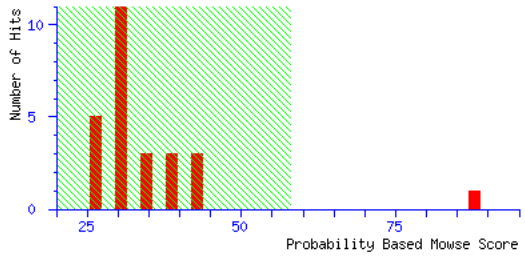
PMF of plasma glutathione peroxidase.

Mascot Search Results

```
User            : katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 26 Dec 2008 at 05:33:46 GMT
Top Score       : 88 for gi|6723180, plasma glutathione peroxidase precursor [Rattus norveg
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|6723180 **Mass:** 25653 **Score:** 88 **Expect:** 6.6e-005 **Queries matched:** 8
plasma glutathione peroxidase precursor [Rattus norvegicus]

Protein View

Match to: **gi|6723180** Score: **88** Expect: **6.6e-005**
plasma glutathione peroxidase precursor [Rattus norvegicus]

Nominal mass (M_r): **25653**; Calculated pI value: **8.26**
NCBI BLAST search of gi|6723180 against nr
Unformatted sequence string for pasting into other applications

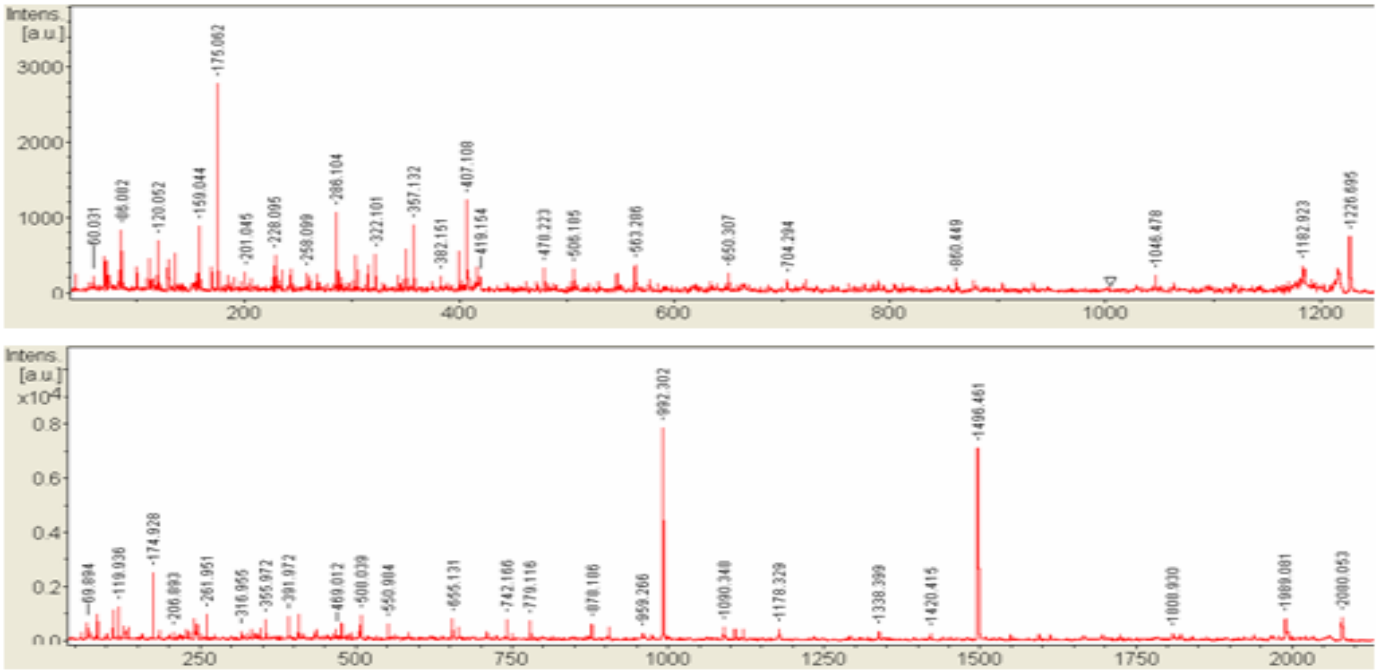
Taxonomy: Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **14**
Number of mass values matched: **8**
Sequence Coverage: **29%**

Matched peptides shown in **Bold Red**

```
     1 MSRILRASCL LSLLLAGFVP PGRGQEKSKT DCHGGMSGTI YEYGALTIDG
    51 EEYIPFKQYA GKYILFVNVA SYUGLTDQYL ELNALQEELG PFGLVILGFP
   101 CNQFGKQEPG ENSEILPSLK YVRPGGGFVP NFQLFEKGDV NGEKEQKFYT
   151 FLKNSCPPTA ELLGSPGRLF WEPMKIHDIR WNFEKFLVGP DGIPIMRWYH
   201 RTTVSNVKMD ILSYMRRQAA LGARGK
```

Retinol-binding protein



MS/MS spectra of retinol-binding protein 1227.393 peptides (top) and 2080.014 peptides (bottom).

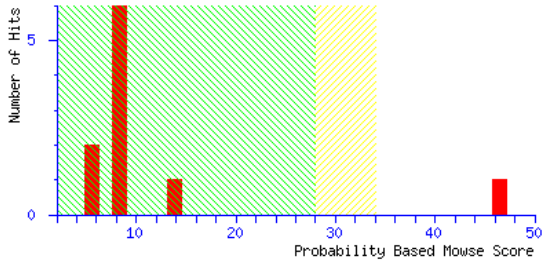
Mascot Search Results

User : **katie**
Email :
Search title : **database search**
MS data file : **DATA.TXT**
Database : **NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy : **Rattus (39483 sequences)**
Timestamp : **17 Dec 2008 at 04:16:40 GMT**
Protein hits :

gi|132407	Plasma retinol-binding protein precursor (PRBP) (RBP)
gi|203014	nonerythroid alpha-spectrin
gi|75905479	aldehyde dehydrogenase family 9, subfamily A1 [Rattus nor
gi|62078635	hypothetical protein LOC300663 [Rattus norvegicus]
gi|1705855	Voltage-dependent N-type calcium channel subunit alpha-1B
gi|109475883	PREDICTED: similar to tumor necrosis factor receptor supe
gi|62652886	PREDICTED: similar to tRNA (5-methylaminomethyl-2-thiouri
gi|25282449	high density lipoprotein binding protein [Rattus norvegic
gi|109468356	PREDICTED: similar to ceramide kinase-like isoform a [Rat
gi|109496648	PREDICTED: similar to caspase recruitment domain family,

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 28 indicate peptides with significant homology.
Individual ions scores > 34 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|132407** Score: **46**
Plasma retinol-binding protein precursor (PRBP) (RBP)
Found in search of DATA.TXT

Nominal mass (M_r): **23547**; Calculated pI value: **5.69**
NCBI BLAST search of gi|132407 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|206585 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **4%**

Matched peptides shown in **Bold Red**

```
    1 MEWVWALVLL AALGGGSAER DCRVSSFRVK ENFDKARFSG LWYAIAKKDP
   51 EGLFLQDNII AEFSVDEKGH MSATAKGRVR LLSNWEVCAD MVGTFTDTED
  101 PAKFKMKYWG VASFLQRGND DHWIIDTDYD TFALQYSCRL QNLDGTCADS
  151 YSFVFSRDPN GLTPETRRLV RQRQEELCLE RQYRWIEHNG YCQSRPSRNS
  201 L
```

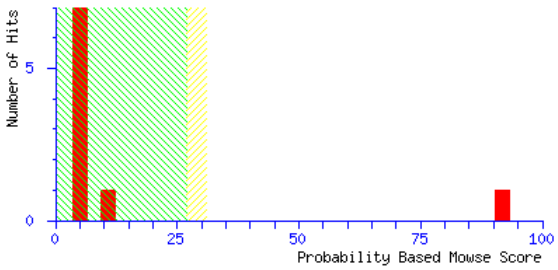
Mascot Search Results

User : **katie**
Email :
Search title : **database search**
MS data file : **DATA.TXT**
Database : **NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy : **Rattus (39483 sequences)**
Timestamp : **17 Dec 2008 at 04:18:56 GMT**
Protein hits :

gi|132407	Plasma retinol-binding protein precursor (PRBP) (RBP)
gi|109497883	PREDICTED: similar to pleckstrin homology domain contai
gi|75195093	COG1944: Uncharacterized conserved protein [Escherichia
gi|3122037	Dihydropyrimidinase-related protein 4 (DRP-4) (Collapsi
gi|57164117	apolipoprotein N [Rattus norvegicus]
gi|89107043	DNA-binding transcriptional regulator [Escherichia coli
gi|25282431	ELL associated factor 2 [Rattus norvegicus]
gi|112984212	hypothetical protein LOC503165 [Rattus norvegicus]
gi|24308494	zinc finger protein 14 (KOX 6) [Rattus norvegicus]

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 27 indicate peptides with significant homology.
Individual ions scores > 31 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Mascot Search Results

Protein View

Match to: **gi|132407** Score: **92**
Plasma retinol-binding protein precursor (PRBP) (RBP)
Found in search of DATA.TXT

Nominal mass (M_r): **23547**; Calculated pI value: **5.69**
NCBI BLAST search of gi|132407 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|206585 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **8%**

Matched peptides shown in **Bold Red**

```
   1 MEWVWALVLL AALGGGSAER DCRVSSFRVK ENFDKARFSG LWYAIAKKDP
  51 EGLFLQDNII AEFSVDEKGH MSATAKGRVR LLSNWEVCAD MVGTFTDTED
 101 PAKFKMKYWG VASFLQRGND DHWIIDTDYD TFALQYSCRL QNLDGTCADS
 151 YSFVFSRDPN GLTPETRRLV RQRQEELCLE RQYRWIEHNG YCQSRPSRNS
 201 L
```

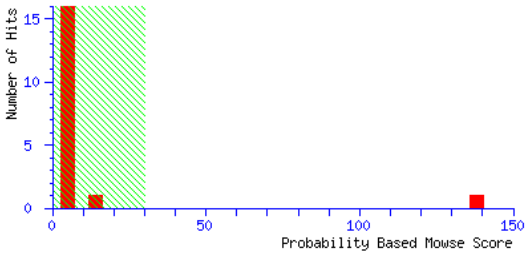
Mascot Search Results

User : **katie**
Email :
Search title : **database search**
MS data file : **DATA.TXT**
Database : **NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy : **Rattus (39483 sequences)**
Timestamp : **17 Dec 2008 at 04:23:06 GMT**
Protein hits :

gi|132407	Plasma retinol-binding protein precursor (PRBP) (RBP)
gi|203014	nonerythroid alpha-spectrin
gi|75905479	aldehyde dehydrogenase family 9, subfamily A1 [Rattus norv
gi|62078635	hypothetical protein LOC300663 [Rattus norvegicus]
gi|109497883	PREDICTED: similar to pleckstrin homology domain containin
gi|1705855	Voltage-dependent N-type calcium channel subunit alpha-1B
gi|109475883	PREDICTED: similar to tumor necrosis factor receptor super
gi|75195093	COG1944: Uncharacterized conserved protein [Escherichia co
gi|62652886	PREDICTED: similar to tRNA (5-methylaminomethyl-2-thiourid
gi|3122037	Dihydropyrimidinase-related protein 4 (DRP-4) (Collapsin r
gi|25282449	high density lipoprotein binding protein [Rattus norvegicu
gi|57164117	apolipoprotein N [Rattus norvegicus]
gi|89107043	DNA-binding transcriptional regulator [Escherichia coli W3
gi|109468356	PREDICTED: similar to ceramide kinase-like isoform a [Ratt
gi|25282431	ELL associated factor 2 [Rattus norvegicus]
gi|112984212	hypothetical protein LOC503165 [Rattus norvegicus]
gi|109496648	PREDICTED: similar to caspase recruitment domain family, m
gi|24308494	zinc finger protein 14 (KOX 6) [Rattus norvegicus]

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 30 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|132407** Score: **138**
Plasma retinol-binding protein precursor (PRBP) (RBP)
Found in search of DATA.TXT

Nominal mass (M_r): **23547**; Calculated pI value: **5.69**
NCBI BLAST search of gi|132407 against nr
Unformatted sequence string for pasting into other applications

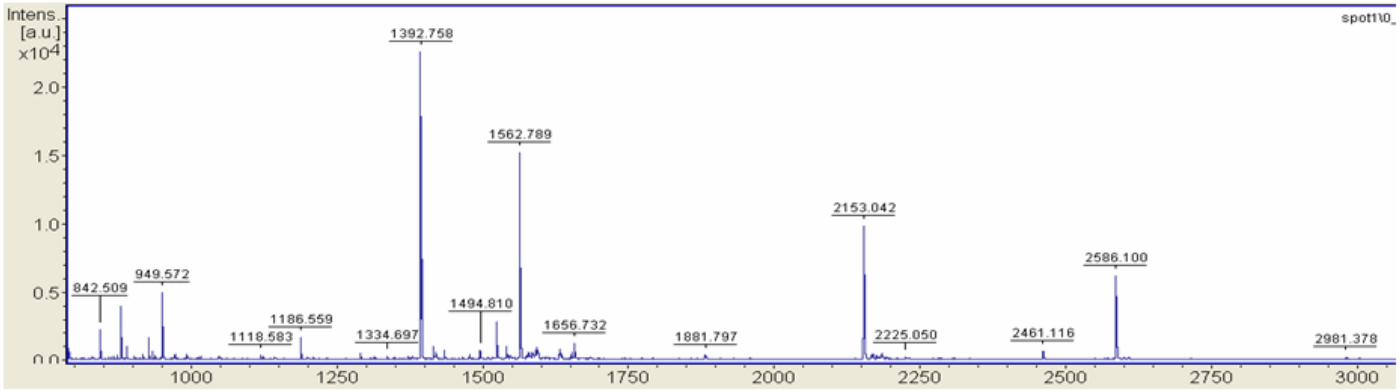
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|206585 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **13%**

Matched peptides shown in **Bold Red**

```
     1 MEWVWALVLL AALGGGSAER DCRVSSFRVK ENFDKARFSG LWYAIAKKDP
    51 EGLFLQDNII AEFSVDEKGH MSATAKGRVR LLSNWEVCAD MVGTFTDTED
   101 PAKFKMKYWG VASFLQRGND DHWIIDTDYD TFALQYSCRL QNLDGTCADS
   151 YSFVFSRDPN GLTPETRRLV RQRQEELCLE RQYRWIEHNG YCQSRPSRNS
   201 L
```

Transferrin



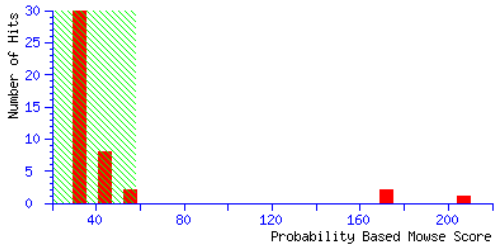
PMF of transferrin.

{MATRIX SCIENCE} Mascot Search Results

User : Yokibutt
Email :
Search title : database search
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Rattus (39483 sequences)
Timestamp : 29 Oct 2008 at 03:49:03 GMT
Top Score : 207 for gi|61556986, transferrin [Rattus norvegicus]

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help
Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|61556986 **Mass:** 78512 **Score:** 207 **Expect:** 7.9e-017 **Queries matched:** 25
transferrin [Rattus norvegicus]
gi|1854476 **Mass:** 78538 **Score:** 168 **Expect:** 6.3e-013 **Queries matched:** 22
transferrin [Rattus norvegicus]

Protein View

Match to: **gi|61556986** Score: **207** Expect: **7.9e-017**
transferrin [Rattus norvegicus]

Nominal mass (M_r): **78512**; Calculated pI value: **7.14**
NCBI BLAST search of gi|61556986 against nr
Unformatted sequence string for pasting into other applications

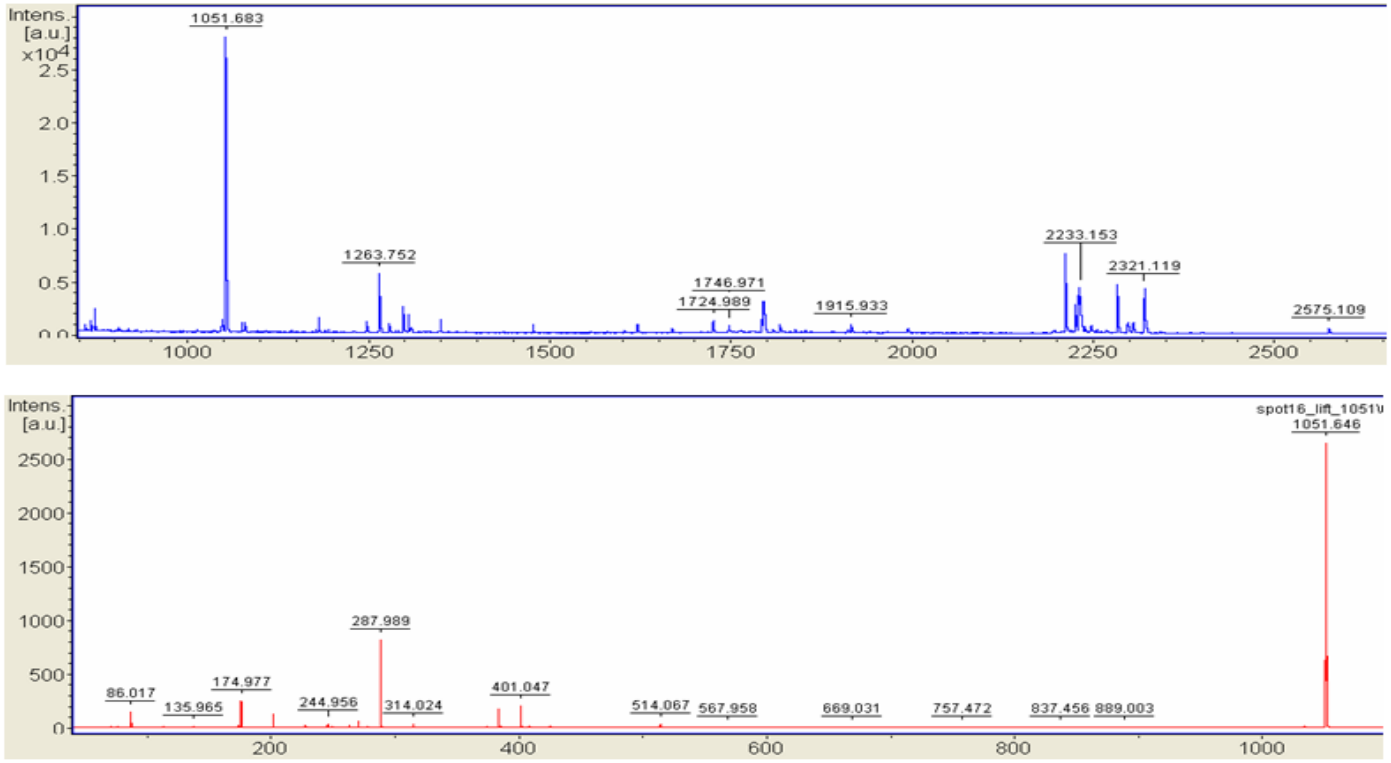
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|122066515 from Rattus norvegicus
gi|33187764 from Rattus norvegicus
gi|56540994 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **55**
Number of mass values matched: **25**
Sequence Coverage: **35%**

Matched peptides shown in **Bold Red**

```
     1 MRFAVGALLA CAALGLCLAV PDKTVKWCAV SEHENTKCIS FRDHMKTVLP
    51 ADGPRLACVK KTSYQDCIKA ISGGEADAIT LDGGWVYDAG LTPNNLKPVA
   101 AEFYGSLEHP QTHYLAVAVV KKGTDFQLNQ LQGKKSCHTG LGRSAGWIIP
   151 IGLLFCNLPE PRKPLEKAVA SFFSGSCVPC ADPVAFPQLC QLCPGCGCSP
   201 TQPFFGYVGA FKCLRDGGGD VAFVKHTTIF EVLPQKADRD QYELLCLDNT
   251 RKPVDQYEDC YLARIPSHAV VARNGDGKED LIWEILKVAQ EHFGKGKSKD
   301 FQLFGSPLGK DLLFKDSAFG LLRVPPRMDY RLYLGHSYVT AIRNQREGVC
   351 PEGSIDSAPV KWCALSHQER AKCDEWSVSS NGQIECESAE STEDCIDKIV
   401 NGEADAMSLD GGHAYIAGQC GLVPVMAENY DISSCTNPQS DVFPKGYYAV
   451 AVVKASDSSI NWNNLKGKKS CHTGVDRTAG WNIPMGLLFS RINHCKFDEF
   501 FSQGCAPGYK KNSTLCDLCI GPAKCAPNNR EGYNGYTGAF QCLVEKGDVA
   551 FVKHQTVLEN TNGKNTAAWA KDLKQEDFQL LCPDGTKKPV TEFATCHLAQ
   601 APNHVVVSRK EKAARVSTVL TAQKDLFWKG DKDCTGNFCL FRSSTKDLLF
   651 RDDTKCLTKL PEGTTYEEYL GAEYLQAVGN IRKCSTSRLL EACTFHKS
```

Vitamin-D binding protein



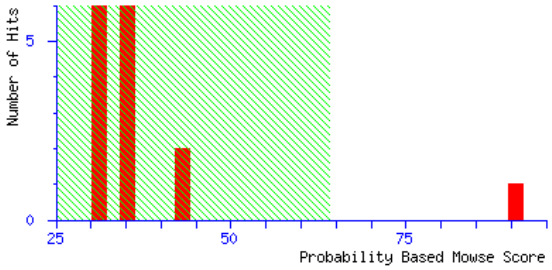
PMF of vitamin-D binding protein (top). MS/MS spectrum of its 1051.683 peptide (bottom).

Mascot Search Results

User : **Katie**
Email :
Search title : **database search**
Database : **NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy : **Rodentia (Rodents) (149892 sequences)**
Timestamp : **29 Oct 2008 at 04:29:58 GMT**
Top Score : **91 for gi|203941, vitamin D-binding protein precursor**

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 64 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help
Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|203941 **Mass:** 55079 **Score:** 91 **Expect:** 0.00013 **Queries matched:**
vitamin D-binding protein precursor

Protein View

Match to: **gi|203941** Score: **91** Expect: **0.00013**
vitamin D-binding protein precursor

Nominal mass (M_r): **55079**; Calculated pI value: **5.76**
NCBI BLAST search of gi|203941 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **16**
Number of mass values matched: **8**
Sequence Coverage: **24%**

Matched peptides shown in **Bold Red**

```
     1 MKRVLVLLLA LAFGHALERG RDYEKDKVCQ ELSTLGKDDF RSLSLILYSR
    51 KFPSSTFEQV SQLVKEVVSL TEECCAEGAD PNCYDTRTSE LSIKSCESDA
   101 PFPVHPGTSE CCTKEGLERK LCMAALSHQP QQFPAYVEPT NDEICEAFRK
   151 DPKGFADQFL FEYSSNYGQA PLPLLVGYTK SYLSMVGSCC TSAKPTVCFL
   201 KERLQMKQLS LLTTMSNRVC SQYAAYGKEK SRMSHLIKLA QKVPTANLED
   251 VLPLAEDLTE ILSRCCKSTS EDCMARELPE HTLKICGNLS KKNSKFEECC
   301 YETTPMGIFM CSYFMPTAEP LQLPAIKLPT SKDLCGQSAT QAMDQYTFEL
   351 SRRTQVPEVF LSKVLDTTLK TLRECCDTQD SVSCFSTQSP LMKRQLTSFI
   401 EKGQEMCADY SENTFTEYKK KLAERLRTKM PNASPEELAD MVAKHSDFAS
   451 KCCSINSPPR YCSSQIDAEM RDILQS
```

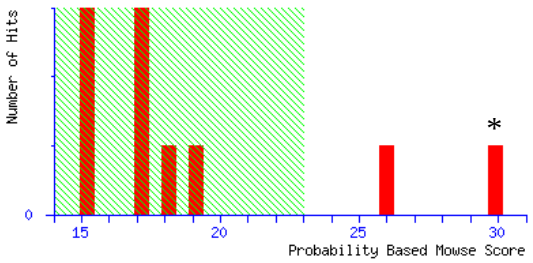
{MATRIX SCIENCE} Mascot Search Results

User : Katie
Email :
Search title : database search
MS data file : DATA.TXT
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Rattus (39483 sequences)
Timestamp : 29 Oct 2008 at 04:25:46 GMT
Protein hits :

gi|203927	vitamin D binding protein prepeptide
gi|89109544	hypothetical protein [Escherichia coli W3110]
gi|109483387	PREDICTED: similar to Casitas B-lineage lymphoma [Rattus :
gi|47059110	DOM-3 homolog Z [Rattus norvegicus]
gi|109477401	PREDICTED: similar to phosphatase and actin regulator 4 [:
gi|15804175	phosphopantetheine adenylyltransferase [Escherichia coli '
gi|16129049	23S rRNA pseudouridylate synthase [Escherichia coli K12]
gi|62664632	PREDICTED: similar to Coiled-coil domain-containing prote
gi|109463356	PREDICTED: hypothetical protein [Rattus norvegicus]
gi|117940025	RAD18 homolog [Rattus norvegicus]

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 23 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|203941** Score: **30**
vitamin D-binding protein precursor
Found in search of DATA.TXT

Nominal mass (M_r): **55079**; Calculated pI value: **5.76**
NCBI BLAST search of gi|203941 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **1%**

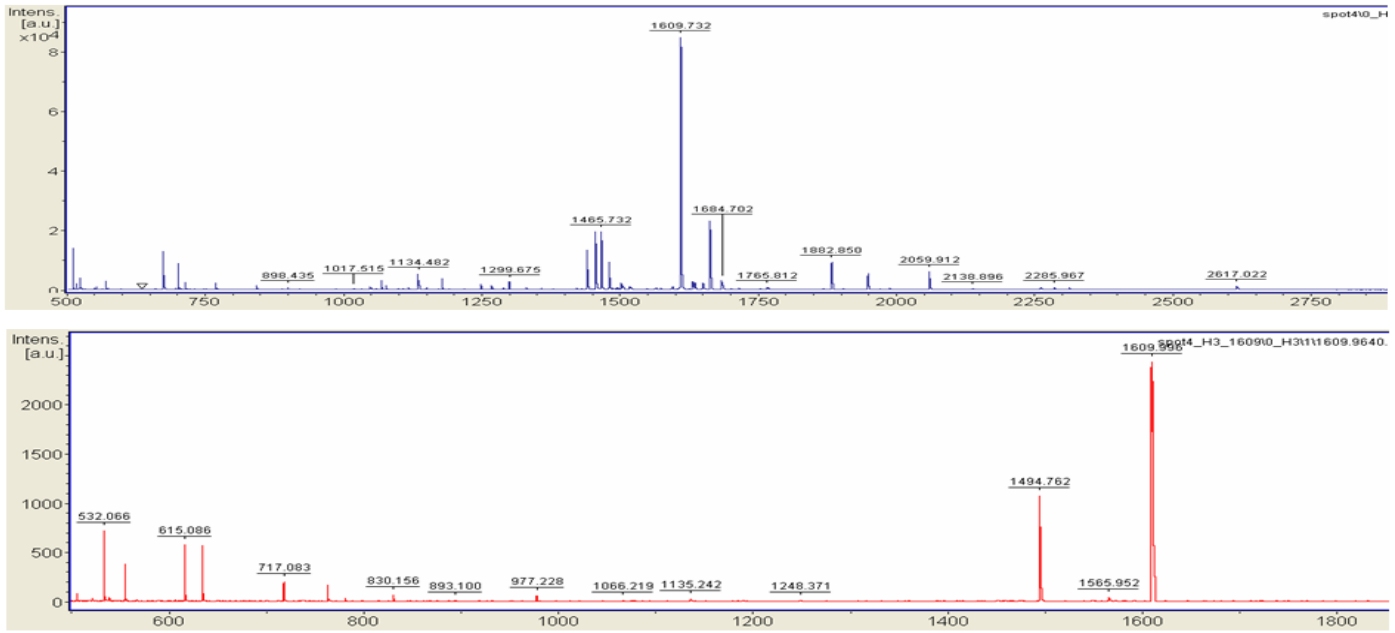
Matched peptides shown in **Bold Red**

```
     1 MKRVLVLLLA LAFGHALERG RDYEKDKVCQ ELSTLGKDDF RSLSLILYSR
    51 KFPSSTFEQV SQLVKEVVSL TEECCAEGAD PNCYDTRTSE LSIKSCESDA
   101 PFPVHPGTSE CCTKEGLERK LCMAALSHQP QQFPAYVEPT NDEICEAFRK
   151 DPKGFADQFL FEYSSNYGQA PLPLLVGYTK SYLSMVGSCC TSAKPTVCFL
   201 KERLQMKQLS LLTTMSNRVC SQYAAYGKEK SRMSHLIKLA QKVPTANLED
   251 VLPLAEDLTE ILSRCCKSTS EDCMARELPE HTLKICGNLS KKNSKFEECC
   301 YETTPMGIFM CSYFMPTAEP LQLPAIKLPT SKDLCGQSAT QAMDQYTFEL
   351 SRRTQVPEVF LSKVLDTTLK TLRECCDTQD SVSCFSTQSP LMKRQLTSFI
   401 EKGQEMCADY SENTFTEYKK KLAERLRTKM PNASPEELAD MVAKHSDFAS
   451 KCCSINSPPR YCSSQIDAEM RDILQS
```

*It matches the protein identified by MS analysis in the previous page.

Appendix II - Details of the sequence coverage and peptide mass finger prints of proteins identified in Chapter 5

Albumin



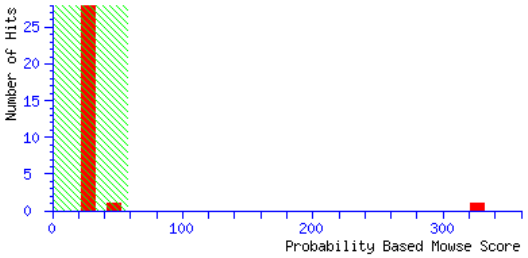
PMF of albumin (top). MS/MS spectrum of its 1609.732 peptide (bottom).

Mascot Search Results

User : Katie
Email :
Search title : database search
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Rattus (39483 sequences)
Timestamp : 22 Oct 2008 at 02:57:30 GMT
Top Score : 326 for gi|19705431, albumin [Rattus norvegicus]

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help
Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|19705431 **Mass:** 70670 **Score:** 326 **Expect:** 9.9e-029 **Queries matched:** 28
albumin [Rattus norvegicus]
gi|124028612 **Mass:** 70682 **Score:** 326 **Expect:** 9.9e-029 **Queries matched:** 28
Serum albumin precursor

Protein View

Match to: **gi|19705431** Score: **326** Expect: **9.9e-029**
albumin [Rattus norvegicus]

Nominal mass (M_r): **70670**; Calculated pI value: **6.09**
NCBI BLAST search of gi|19705431 against nr
Unformatted sequence string for pasting into other applications
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|55628 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **41**
Number of mass values matched: **28**
Sequence Coverage: **49%**

Matched peptides shown in **Bold Red**

```
     1 MKWVTFLLLL FISGSAFSRG VFRREAHKSE IAHRFKDLGE QHFKGLVLIA
    51 FSQYLQKCPY EEHIKLVQEV TDFAKTCVAD ESAENCDKSI HTLFGDKLCA
   101 IPKLRDNYGE LADCCAKQEP ERNECFLQHK DDNPNLPPFQ RPEAEAMCTS
   151 FQENPTSFLG HYLHEVARRH PYFYAPELLY YAEKYNEVLT QCCTESDKAA
   201 CLTPKLDAVK EKALVAAVRQ RMKCSSMQRF GERAFKAWAV ARMSQRFPNA
   251 EFAEITKLAT DVTKINKECC HGDLLECADD RAELAKYMCE NQATISSKLQ
   301 ACCDKPVLQK SQCLAETEHD NIPADLPSIA ADFVEDKEVC KNYAEAKDVF
   351 LGTFLYEYSR RHPDYSVSLL LRLAKKYEAT LEKCCAEGDP PACYGTVLAE
   401 FQPLVEEPKN LVKTNCELYE KLGEYGFQNA VLVRYTQKAP QVSTPTLVEA
   451 ARNLGRVGTK CCTLPEAQRL PCVEDYLSAI LNRLCVLHEK TPVSEKVTKC
   501 CSGSLVERRP CFSALTVDET YVPKEFKAET FTFHSDICTL PDKEKQIKKQ
   551 TALAELVKHK PKATEDQLKT VMGDFAQFVD KCCKAADKDN CFATEGPNLV
   601 ARSKEALA
```

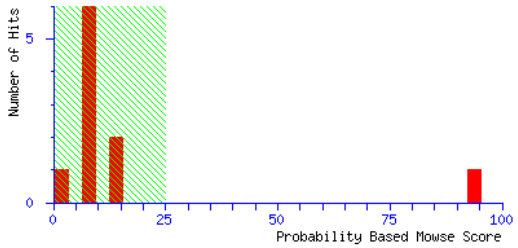
Mascot Search Results

User : Katie
Email :
Search title : database search
MS data file : DATA.TXT
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Rattus (39483 sequences)
Timestamp : 22 Oct 2008 at 03:09:19 GMT
Protein hits :

gi|19705431	albumin [Rattus norvegicus]
gi|109506635	PREDICTED: similar to MEGF10 protein [Rattus norvegicus]
gi|9845251	rapamycin and FKBP12 target-1 protein [Rattus norvegicus]
gi|115485	Calcitonin gene-related peptide 1 precursor (Calcitonin gene-r
gi|109474014	PREDICTED: similar to cell adhesion molecule with homology to
gi|16128844	conserved protein with nucleoside triphosphate hydrolase doma:
gi|34860298	PREDICTED: similar to matrix metalloproteinase 1a (interstiti:
gi|34878065	PREDICTED: similar to Putative pre-mRNA-splicing factor ATP-d:
gi|16128551	bacteriophage N4 receptor, outer membrane subunit [Escherichi:
gi|1710906	coatomer beta subunit [Rattus norvegicus]

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 25 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|19705431** Score: **94**
albumin [Rattus norvegicus]
Found in search of DATA.TXT

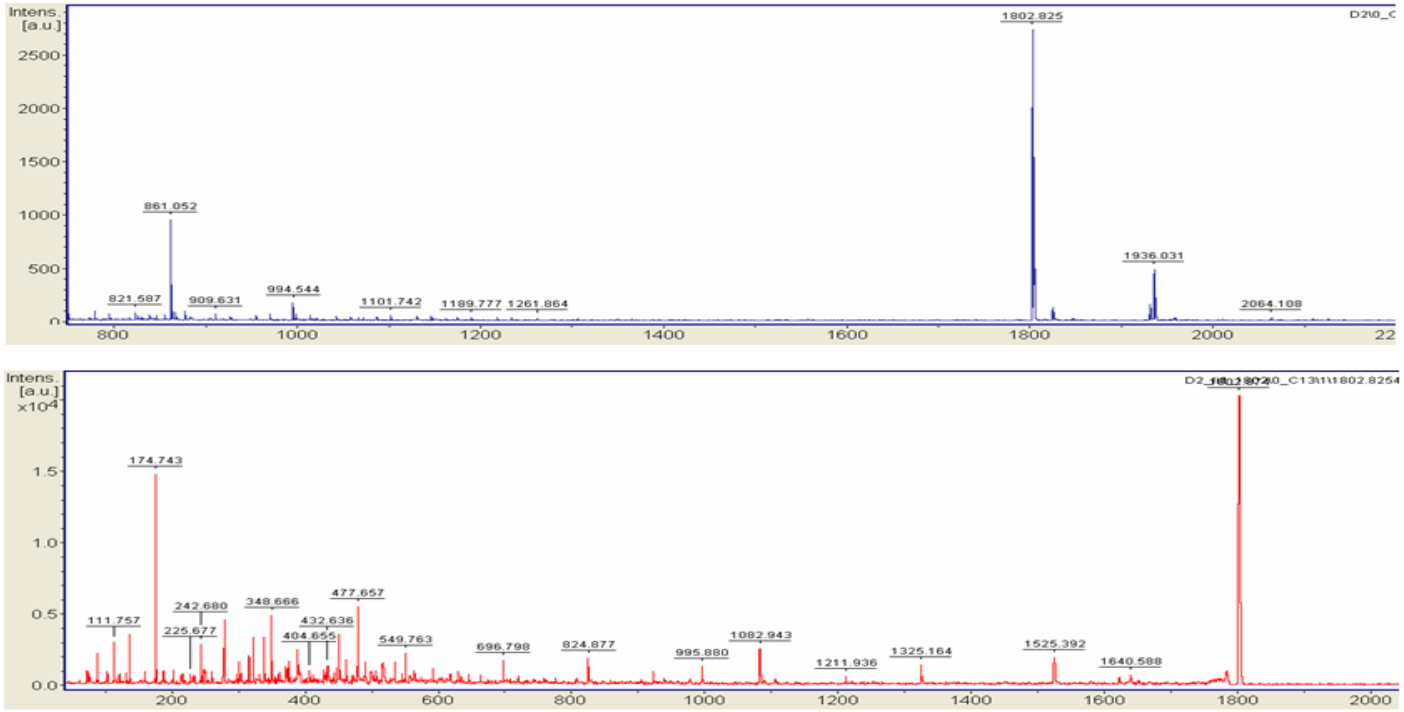
Nominal mass (M_r): **70670**; Calculated pI value: **6.09**
NCBI BLAST search of gi|19705431 against nr
Unformatted sequence string for pasting into other applications
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|55628 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **2%**

Matched peptides shown in **Bold Red**

```
    1 MKWVTFLLLL FISGSAFSRG VFRREAHKSE IAHRFKDLGE QHFKGLVLIA
   51 FSQYLQKCPY EEHIKLVQEV TDFAKTCVAD ESAENCDKSI HTLFGDKLCA
  101 IPKLRDNYGE LADCCAKQEP ERNECFLQHK DDNPNLPPFQ RPEAEAMCTS
  151 FQENPTSFLG HYLHEVARRH PYFYAPELLY YAEKYNEVLT QCCTESDKAA
  201 CLTPKLDAVK EKALVAAVRQ RMKCSSMQRF GERAFKAWAV ARMSQRFPNA
  251 EFAEITKLAT DVTKINKECC HGDLLECADD RAELAKYMCE NQATISSKLQ
  301 ACCDKPVLQK SQCLAETEHD NIPADLPSIA ADFVEDKEVC KNYAEAKDVF
  351 LGTFLYEYSR RHPDYSVSLL LRLAKKYEAT LEKCCAEGDP PACYGTVLAE
  401 FQPLVEEPKN LVKTNCELYE KLGEYGFQNA VLVRYTQKAP QVSTPTLVEA
  451 ARNLGRVGTK CCTLPEAQRL PCVEDYLSAI LNRLCVLHEK TPVSEKVTKC
  501 CSGSLVERRP CFSALTVDET YVPKEFKAET FTFHSDICTL PDKEKQIKKQ
  551 TALAELVKHK PKATEDQLKT VMGDFAQFVD KCCKAADKDN CFATEGPNLV
  601 ARSKEALA
```

T-kininogen



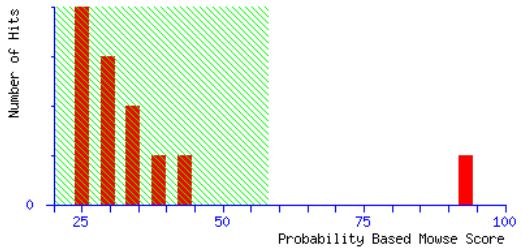
PMF of T-kininogen (top). MS/MS spectrum of its 1802.825 peptide (bottom).

Mascot Search Results

```
User            : Katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 22 Oct 2008 at 03:22:25 GMT
Top Score       : 93 for gi|60392582, T-kininogen 1 precursor (T-kininogen I) (Major acute pl
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 5

Re-Search All | Search Unmatched

1. gi|60392582 **Mass:** 48828 **Score:** 93 **Expect:** 2e-005 **Queries matched:** 7
T-kininogen 1 precursor (T-kininogen I) (Major acute phase protein) (Alpha-1-MAP) (Thi
gi|80861401 **Mass:** 48817 **Score:** 93 **Expect:** 2e-005 **Queries matched:** 7
kininogen 1 [Rattus norvegicus]

Protein View

Match to: **gi|60392582** Score: **93** Expect: **2e-005**
T-kininogen 1 precursor (T-kininogen I) (Major acute phase protein) (Alpha-1-MAP) (Thiostatin)

Nominal mass (M_r): **48828**; Calculated pI value: **6.08**
NCBI BLAST search of gi|60392582 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **25**
Number of mass values matched: **7**
Sequence Coverage: **19%**

Matched peptides shown in **Bold Red**

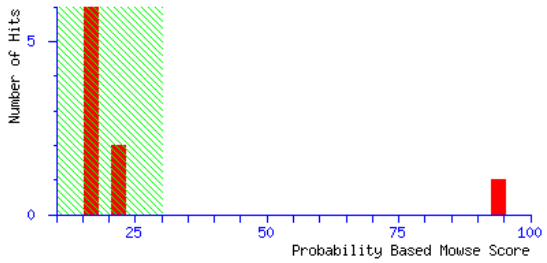
```
     1 MKLITILLLC SRLLPSLAQE EGAQELNCND ETVFQAVDTA LKKYNAELES
    51 GNQFVLYRVT EGTKKDGAET LYSFKYQIKE GNCSVQSGLT WQDCDFKDAE
   101 EAATGECTTT LGKKENKFSV ATQICNITPG KGPKKTEEDL CVGCFQPIPM
   151 DSSDLKPVLK HAVEHFNNNT KHTHLFALRE VKSAHSQVVA GMNYKIIYSI
   201 VQTNCSKEDF PSLREDCVPL PYGDHGECTG HTHVDIHNTI AGFSQSCDLY
   251 PGDDLFELLP KNCRGCPREI PVDSPELKEA LGHSIAQLNA QHNHIFYFKI
   301 DTVKKATSQV VAGVIYVIEF IARETNCSKQ SKTELTADCE TKHLGQSLNC
   351 NANVYMRPWE NKVVPTVRCQ ALDMMISRPP GFSPFRLVRV QETKEGTTRL
   401 LNSCEYKGRL SKARAGPAPD HQAEASTVTP
```

Mascot Search Results

User : Katie
Email :
Search title : database search
MS data file : DATA.TXT
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Rattus (39483 sequences)
Timestamp : 22 Oct 2008 at 03:29:42 GMT
Protein hits : gi|205085 LMW T-kininogen I precursor
gi|71043668 transcription elongation factor A (SII) 1 [Rattus norvegic
gi|109499578 PREDICTED: similar to UDP glucuronosyltransferase 2 family
gi|109489321 PREDICTED: hypothetical protein [Rattus norvegicus]
gi|23397411 cytochrome P450, family 4, subfamily b, polypeptide 1 [Rat

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 30 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|60392582** Score: **94**
T-kininogen 1 precursor (T-kininogen I) (Major acute phase protein) (Alpha-1-MAP) (Thiosta
Found in search of DATA.TXT

Nominal mass (M_r): **48828**; Calculated pI value: **6.08**
NCBI BLAST search of gi|60392582 against nr
Unformatted sequence string for pasting into other applications

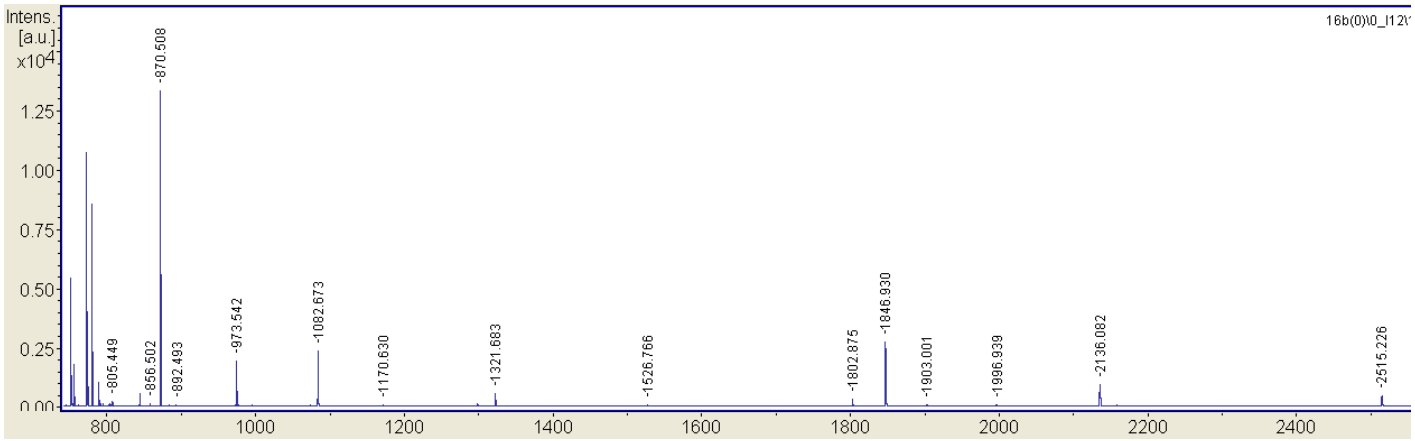
Taxonomy: Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: 3%

Matched peptides shown in **Bold Red**

```
     1 MKLITILLLC SRLLPSLAQE EGAQELNCND ETVFQAVDTA LKKYNAELES
    51 GNQFVLYRVT EGTKKDGAET LYSFKYQIKE GNCSVQSGLT WQDCDFKDAE
   101 EAATGECTTT LGKKENKFSV ATQICNITPG KGPKKTEEDL CVGCFQPIPM
   151 DSSDLKPVLK HAVEHFNNNT KHTHLFALRE VKSAHSQVVA GMNYKIIYSI
   201 VQTNCSKEDF PSLREDCVPL PYGDHGECTG HTHVDIHNTI AGFSQSCDLY
   251 PGDDLFELLP KNCRGCPREI PVDSPELKEA LGHSIAQLNA QHNHIFYFKI
   301 DTVKKATSQV VAGVIYVIEF IARETNCSKQ SKTELTADCE TKHLGQSLNC
   351 NANVYMRPWE NKVVPTVRCQ ALDMMISRPP GFSPFRLVRV QETKEGTTRL
   401 LNSCEYKGRL SKARAGPAPD HQAEASTVTP
```

α-2-HS glycoprotein



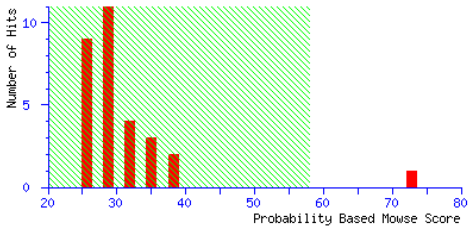
PMF of α-2-HS glycoprotein.

{MATRIX SCIENCE} Mascot Search Results

```
User            : Katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 16 Dec 2008 at 09:45:35 GMT
Top Score       : 73 for gi|231468, Alpha-2-HS-glycoprotein precursor (Fetuin-A) (Glycoprotein PP63)
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary — Help

Significance threshold p< 0.05 — Max. number of hits 500

Re-Search All | Search Unmatched

1. gi|231468 **Mass:** 38757 **Score:** 73 **Expect:** 0.0021 **Queries matched:** 6
Alpha-2-HS-glycoprotein precursor (Fetuin-A) (Glycoprotein PP63) (59 kDa bone sialic acid-cont
gi|6978477 **Mass:** 38778 **Score:** 56 **Expect:** 0.11 **Queries matched:** 5
alpha-2-HS-glycoprotein [Rattus norvegicus]

Protein View

Match to: **gi|231468** Score: **73** Expect: **0.0021**
Alpha-2-HS-glycoprotein precursor (Fetuin-A) (Glycoprotein PP63) (59 kDa bone sialic acid-cont

Nominal mass (M_r): **38757**; Calculated pI value: **6.05**
NCBI BLAST search of gi|231468 against nr
Unformatted sequence string for pasting into other applications

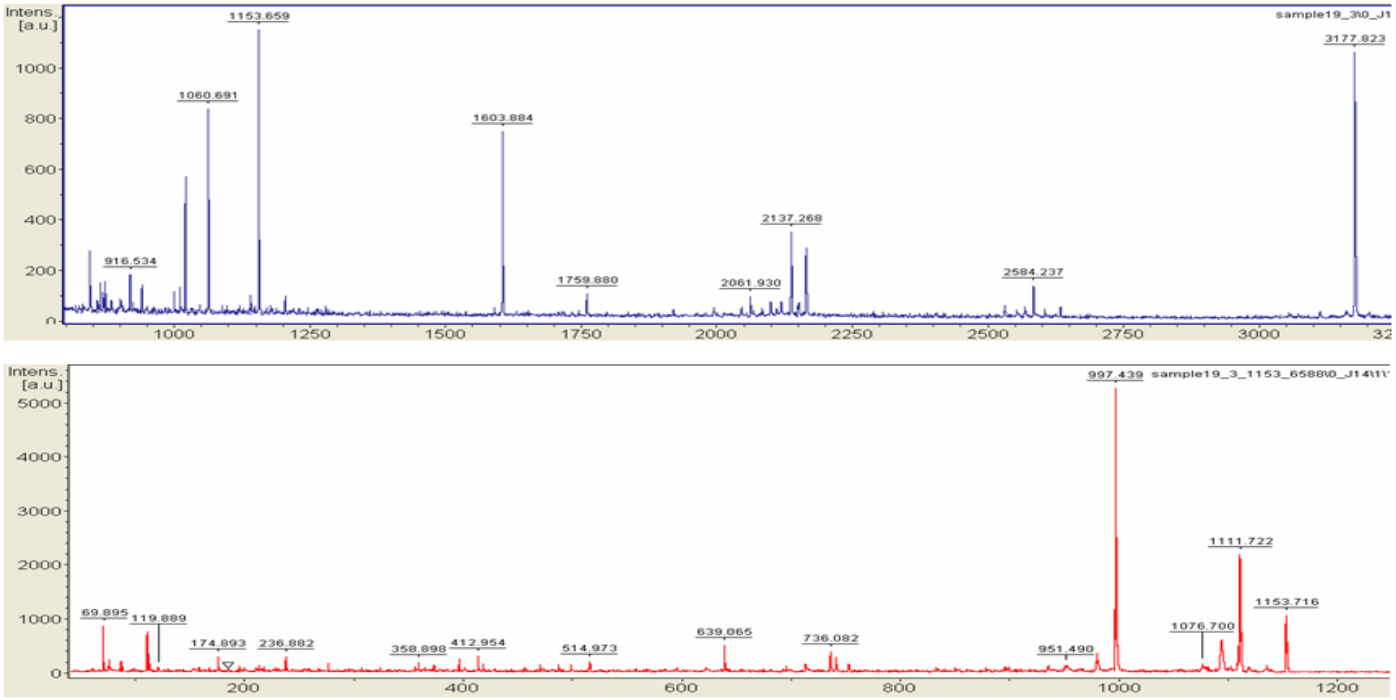
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|56140 from Rattus norvegicus
gi|220676 from Rattus norvegicus
gi|60552688 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **21**
Number of mass values matched: **6**
Sequence Coverage: **26%**

Matched peptides shown in **Bold Red**

```
     1 MKSLVLLLCF AQLWSCQSAP QGAGLGFREL ACDDPETEHV ALIAVDYLNK
    51 HLLQGFRQIL NQIDKVKVWS RRPFGEVYEL EIDTLETTCH ALDPTPLANC
   101 SVRQQAEHAV EGDCDFHILK QDGQFRVLHA QCHSTPDSAE DVRKFCPRCP
   151 ILIRFNDTNV VHTVKTALAA FNAQNNGTYF KLVEISRAQN VPFPVSTLVE
   201 FVIAATDCTG QEVTDPAKCN LLAEKQYGFC KATLIHRLGG EEVSVACKLF
   251 QTQPQPANAN PAGPAPTVGQ AAPVAPPAGP PESVVVGPVA VPLGLPDHRT
   301 HHDLRHAFSP VASVESASGE VLHSPKVGQP GDAGAAGPVA PLCPGRVRYF
   351 KI
```

α-1-antitrypsin



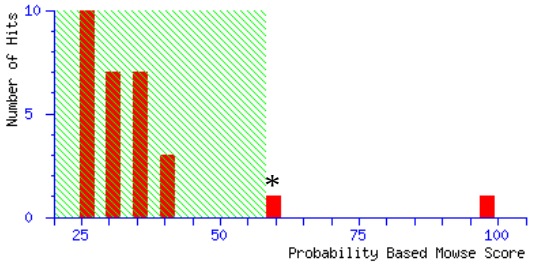
PMF spectrum of α-1-antitrypsin (top). MS/MS spectrum of its 1153.659 peptide (bottom).

Mascot Search Results

```
User            : Katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 22 Oct 2008 at 03:54:33 GMT
Top Score       : 98 for gi|203063, alpha-1-antitrypsin precursor
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 5

Re-Search All | Search Unmatched

1. gi|203063 **Mass:** 45978 **Score:** 98 **Expect:** 6.1e-006 **Queries matched:** 8
alpha-1-antitrypsin precursor

Protein View

Match to: **gi|203063** Score: **98** Expect: **6.1e-006**
alpha-1-antitrypsin precursor

Nominal mass (M_r): **45978**; Calculated pI value: **5.70**
NCBI BLAST search of gi|203063 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **13**
Number of mass values matched: **8**
Sequence Coverage: **24%**

Matched peptides shown in **Bold Red**

```
     1 SISRGLLLLA ALCCLAPSFL AEDAQETDTS QQDQSPTYRK ISSNLADFAF
    51 SLYRELVHQS NTSNIFFSPM SITTAFAMLS LGSKGDTRKQ ILEGLEFNLT
   101 QIPEADIHKA FHHLLQTLNR PDSELQLNTG NGLFVNKNLK LVEKFLEEVK
   151 NNYHSEAFSV NFADSEEAKK VINDYVEKGT QGKIVDLMKQ LDEDTVFALV
   201 NYIFFKGKWK RPFNPEHTRD ADFHVDKSTT VKVPMMNRLG MFDMHYCSTL
   251 SSWVLMMDYL GNATAIFLLP DDGKMQHLEQ TLTKDLISRF LLNRQTRSAI
   301 LYFPKLSISG TYNLKTLLSS LGITRVFNND ADLSGITEDA PLKLSQAVHK
   351 AVLTLDERGT EAAGATVVEA VPMSLPPQVK FDHPFIFMIV ESETQSPLFV
   401 GKVIDPTR
```

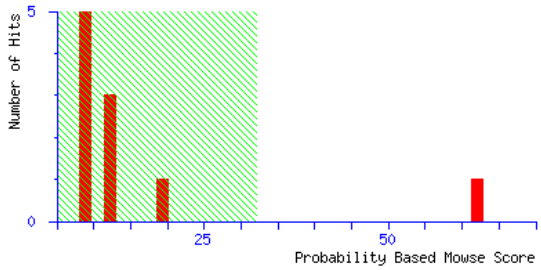
*This is due to the sequence homology, however, the MS/MS shown on the next page further confirms its identity.

{MATRIX SCIENCE} Mascot Search Results

User : Katie
Email :
Search title : database search
MS data file : DATA.TXT
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Rattus (39483 sequences)
Timestamp : 22 Oct 2008 at 03:53:57 GMT
Protein hits : gi|112889 Alpha-1-antiproteinase precursor (Alpha-1-antitrypsin)
gi|109505146 PREDICTED: hypothetical protein [Rattus norvegicus]
gi|109505115 PREDICTED: similar to MAP/microtubule affinity-regulati
gi|53850602 hypothetical protein LOC293343 [Rattus norvegicus]
gi|62955040 potassium channel modulatory factor 1 [Rattus norvegicu

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 32 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|203063** Score: **62**
alpha-1-antitrypsin precursor
Found in search of DATA.TXT

Nominal mass (M_r): **45978**; Calculated pI value: **5.70**
NCBI BLAST search of gi|203063 against nr
Unformatted sequence string for pasting into other applications

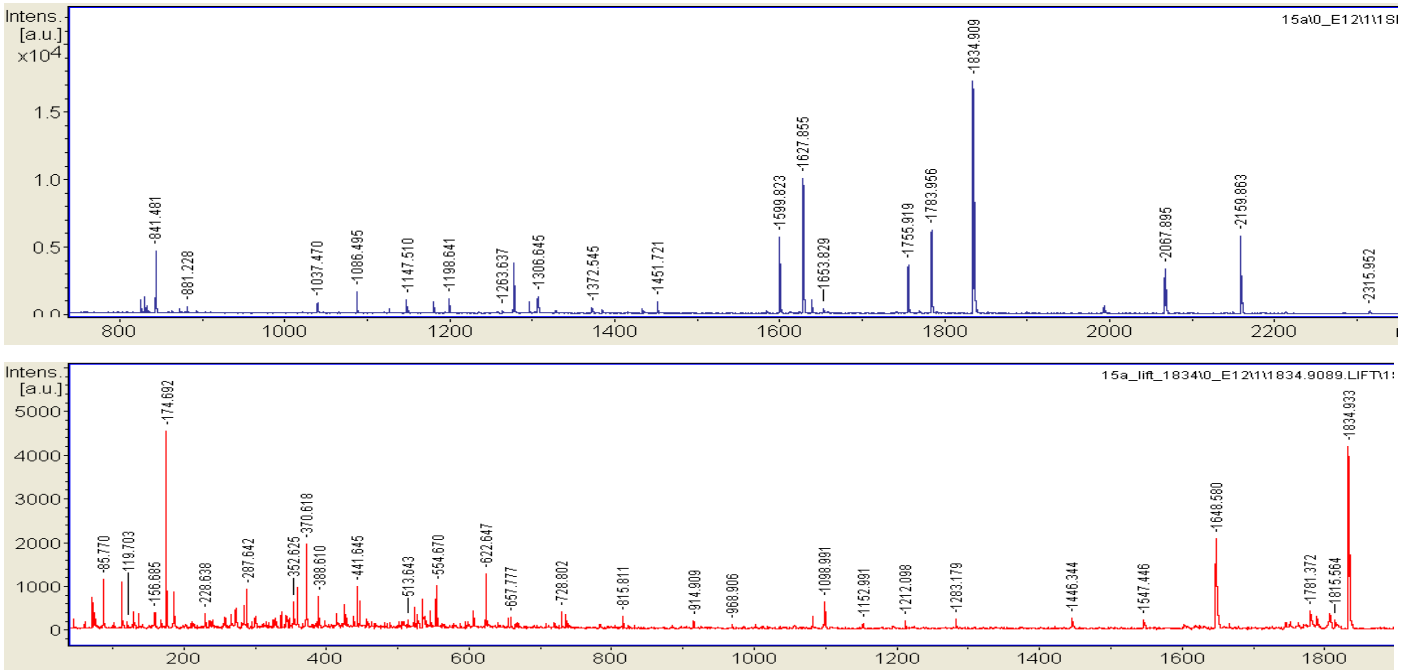
Taxonomy: Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **2%**

Matched peptides shown in **Bold Red**

```
   1 SISRGLLLLA ALCCLAPSFL AEDAQETDTS QQDQSPTYRK ISSNLADFAF
  51 SLYRELVHQS NTSNIFFSPM SITTAFAMLS LGSKGDTRKQ ILEGLEFNLT
 101 QIPEADIHKA FHHLLQTLNR PDSELQLNTG NGLFVNKNLK LVEKFLEEVK
 151 NNYHSEAFSV NFADSEEAKK VINDYVEKGT QGKIVDLMKQ LDEDTVFALV
 201 NYIFFKGKWK RPFNPEHTRD ADFHVDKSTT VKVPMMNRLG MFDMHYCSTL
 251 SSWVLMMDYL GNATAIFLLP DDGKMQHLEQ TLTKDLISRF LLNRQTRSAI
 301 LYFPKLSISG TYNLKTLLSS LGITRVFNND ADLSGITEDA PLKLSQAVHK
 351 AVLTLDERGT EAAGATVVEA VPMSLPPQVK FDHPFIFMIV ESETQSPLFV
 401 GKVIDPTR
```

Afamin protein



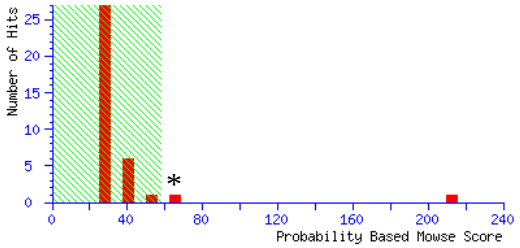
PMF of afamin (top). MS/MS spectrum of its 1834.933 peptide (bottom).

Mascot Search Results

```
User            : Katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 26 Nov 2008 at 10:49:31 GMT
Top Score       : 213 for gi|27229290, afamin [Rattus norvegicus]
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|27229290 **Mass:** 71172 **Score:** 213 **Expect:** 2e-017 **Queries matched:** 19
afamin [Rattus norvegicus]
gi|60688254 **Mass:** 54205 **Score:** 185 **Expect:** 1.2e-014 **Queries matched:** 16
Afm protein [Rattus norvegicus]

Protein View

Match to: **gi|27229290** Score: **213** Expect: **2e-017**
afamin [Rattus norvegicus]

Nominal mass (M_r): **71172**; Calculated pI value: **5.87**
NCBI BLAST search of gi|27229290 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|543793 from Rattus norvegicus
gi|456359 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **26**
Number of mass values matched: **19**
Sequence Coverage: **30%**

Matched peptides shown in **Bold Red**

```
    1 MRHLKLTGFI FFLLSLTESL ALPTKPQDVD HFNATQKFIN ENVAYLTIIA
   51 SAQYVQEASF EEVEMLVKVM LDYKDRCLAD STLPECSKIA NDAIQDMLCD
  101 MKGLPQKHNF SHCCRQAGFQ RRLCFFYNKK ANVGFLPPFP TLDPEEKCQA
  151 YKNNSESFLN LYMYEVARRN PFAFAPVLLN VAARFEEAAT TCCEQQQKAT
  201 YFQDKAAPIT QYLKALSSYQ RNVCGALLKF GPKTLNSINI AVFSKKFPKI
  251 GFEDLTSLLE DVSSMYDGCC EGDVVQCIRS QSQVMHHICS KQDSISSKIK
  301 ACCEKKLPER ADCIINANKD DRPEDLSLRT PKFTDSENVC QERDSEQDKF
  351 FAEFLYDYSR RHTELSTPEL LRITKVYKDL LEDCCNRKNP LSCYRHAEDK
  401 FNETTERSLA MVQQECKQFQ ELGKDALQRH FLVKFTKAAP QLPMEELVSL
  451 SKEMVAALAT CCTLSDEFAC VDNLADLVLG ELCGMNKNRT INPTVDHCCR
  501 ADFAFRRPCF EHLKADTTYA LPSVSALVSA LRADWCQPLK EDLQNKRHRF
  551 LVNLVKWMPE ITDEERLCLF TKFTAAGEEC GNIQKPEACF SPESSKTGDV
  601 SQDAEKQR
```

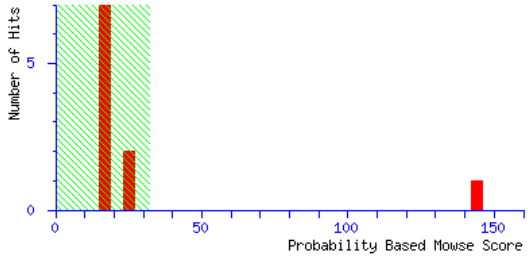
*This is due to the sequence homology, however, the MS/MS shown on the next page further confirms its identity.

Mascot Search Results

User	:	**Katie**
Email	:	
Search title	:	**database search**
MS data file	:	**DATA.TXT**
Database	:	**NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy	:	**Rattus (39483 sequences)**
Timestamp	:	**26 Nov 2008 at 10:37:26 GMT**
Protein hits	:	gi|27229290 afamin [Rattus norvegicus]
		gi|121308631 hypothetical protein [Rattus norvegicus]
		gi|16128840 HCP oxidoreductase, NADH-dependent [Escherichia coli K12]
		gi|16130533 heat shock protein [Escherichia coli K12]
		gi|51948390 dehydrogenase/reductase (SDR family) member 8 [Rattus norv
		gi|81866683 Usherin precursor (Usher syndrome type-2A protein homolog)
		gi|57012428 type I hair keratin KA36 [Rattus norvegicus]
		gi|11024672 kynurenine 3-monooxygenase [Rattus norvegicus]
		gi|19424166 f-box only protein 32 [Rattus norvegicus]
		gi|57527530 signal-induced proliferation-associated 1 like 2 [Rattus n

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 32 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|27229290** Score: **144**
afamin [Rattus norvegicus]
Found in search of DATA.TXT

Nominal mass (M_r): **71172**; Calculated pI value: **5.87**
NCBI BLAST search of gi|27229290 against nr
Unformatted sequence string for pasting into other applications

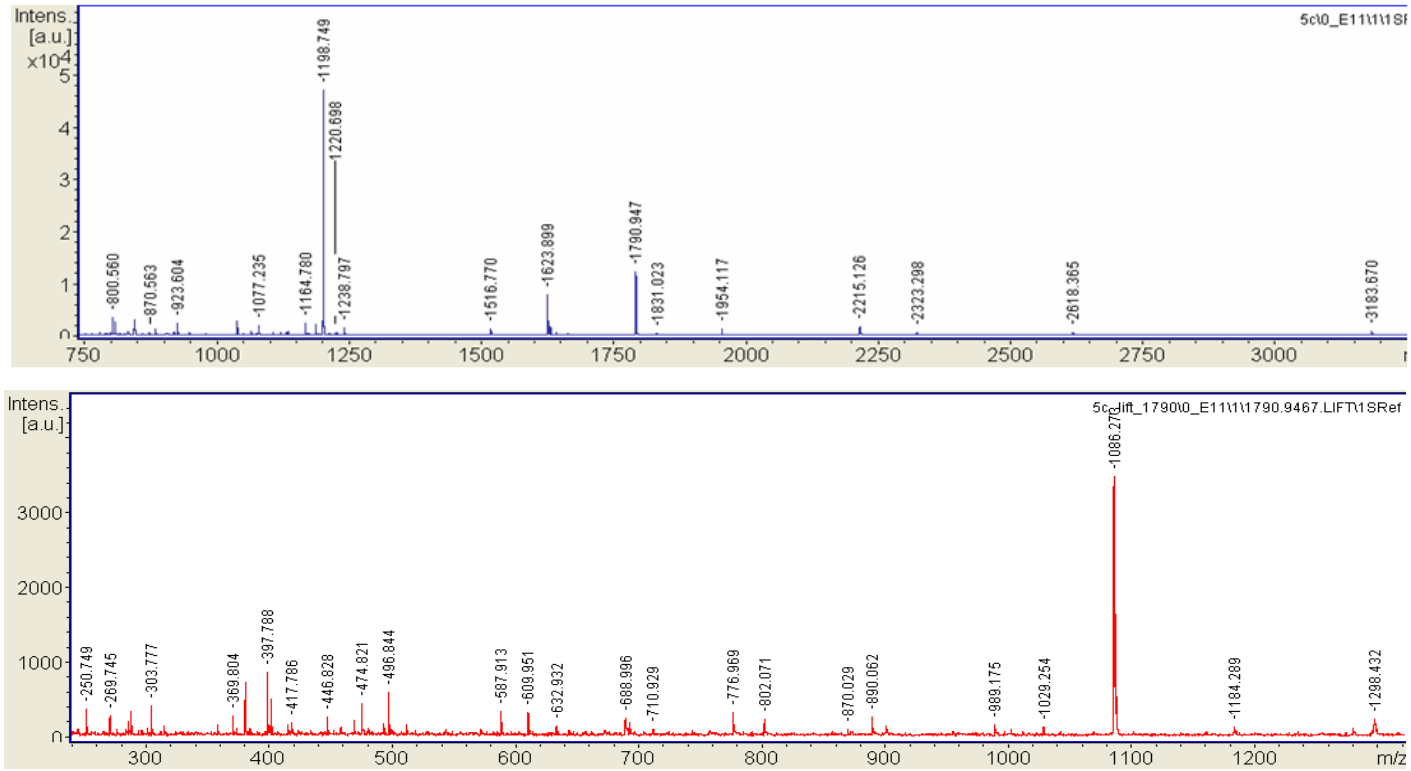
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|543793 from Rattus norvegicus
gi|456359 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **2%**

Matched peptides shown in **Bold Red**

```
     1 MRHLKLTGFI FFLLSLTESL ALPTKPQDVD HFNATQKFIN ENVAYLTIIA
    51 SAQYVQEASF EEVEMLVKVM LDYKDRCLAD STLPECSKIA NDAIQDMLCD
   101 MKGLPQKHNF SHCCRQAGFQ RRLCFFYNKK ANVGFLPPFP TLDPEEKCQA
   151 YKNNSESFLN LYMYEVARRN PFAFAPVLLN VAARFEEAAT TCCEQQQKAT
   201 YFQDKAAPIT QYLKALSSYQ RNVCGALLKF GPKTLNSINI AVFSKKFPKI
   251 GFEDLTSLLE DVSSMYDGCC EGDVVQCIRS QSQVMHHICS KQDSISSKIK
   301 ACCEKKLPER ADCIINANKD DRPEDLSLRT PKFTDSENVC QERDSEQDKF
   351 FAEFLYDYSR RHTELSTPEL LRITKVYKDL LEDCCNRKNP LSCYRHAEDK
   401 FNETTERSLA MVQQECKQFQ ELGKDALQRH FLVKFTKAAP QLPMEELVSL
   451 SKEMVAALAT CCTLSDEFAC VDNLADLVLG ELCGMNKNRT INPTVDHCCR
   501 ADFAFRRPCF EHLKADTTYA LPSVSALVSA LRADWCQPLK EDLQNKRHRF
   551 LVNLVKWMPE ITDEERLCLF TKFTAAGEEC GNIQKPEACF SPESSKTGDV
   601 SQDAEKQR
```

γ-actin



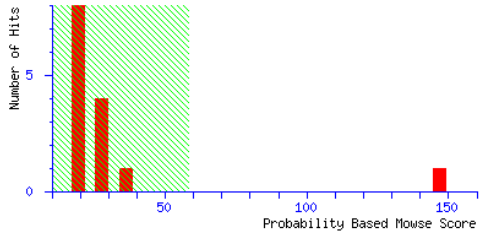
PMF of γ-actin (top). MS/MS spectrum of its1298.432 peptide (bottom).

{MATRIX SCIENCE} Mascot Search Results

```
User            : Katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 26 Nov 2008 at 10:05:35 GMT
Top Score       : 147 for gi|4501885, beta actin [Homo sapiens]
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|4501885 **Mass:** 42052 **Score:** 147 **Expect:** 7.9e-011 **Queries matched:** 13
beta actin [Homo sapiens]
gi|4501887 **Mass:** 42108 **Score:** 147 **Expect:** 7.9e-011 **Queries matched:** 13
actin, gamma 1 propeptide [Homo sapiens]
gi|109507063 **Mass:** 42109 **Score:** 147 **Expect:** 7.9e-011 **Queries matched:** 13
PREDICTED: similar to Actin, cytoplasmic 2 (Gamma-actin) [Rattus norvegicus]

Protein View

Match to: **gi|109507063** Score: **147** Expect: **7.9e-011**
PREDICTED: similar to Actin, cytoplasmic 2 (Gamma-actin) [Rattus norvegicus]

Nominal mass (M_r): **42109**; Calculated pI value: **5.31**
NCBI BLAST search of gi|109507063 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **35**
Number of mass values matched: **13**
Sequence Coverage: **46%**

Matched peptides shown in **Bold Red**

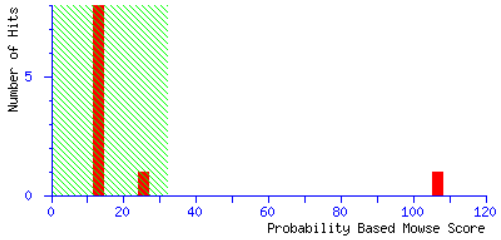
```
     1 MEEEIAALVI DNGSGMCKAG FAGDDAPRAV FPSIVGRPRH QGVMVGMGQK
    51 DSYVGDEAQS KRGILTLKYP NEHGIVTNWD DMEKIWHHTF YNELRVAPEE
   101 HPVLLTEAPL NPKANREKMT QIMFETFNTP AMYVAIQAVL SLYASGRTTG
   151 IVMDSGDGVT HTVPIYEGYA LPHAILRLDL AGRDLTDYLM KILTERGYSF
   201 TTTAEREIVR DIKEKLCYVA LDFEQEMATA ASSSSLEKSY ELPDGQVITI
   251 GNERFRCPEA LFQPSFLGME SCGIHETTFN SIMKCDVDIR KDLYANTVLS
   301 GGTTMYPGIA DRMQKEITAL APSTMKIKII APPERKYSVW IGGSILASLS
   351 TFQQMWISKQ EYDESGPSIV HRKCF
```

Mascot Search Results

User	:	**Katie**
Email	:	
Search title	:	**database search**
MS data file	:	**DATA.TXT**
Database	:	**NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy	:	**Rattus (39483 sequences)**
Timestamp	:	**26 Nov 2008 at 10:18:12 GMT**
Protein hits	:	gi|4501881 alpha 1 actin precursor [Homo sapiens]
		gi|68163565 secernin 1 [Rattus norvegicus]
		gi|16130922 sensory histidine kinase in two-component regulatory system with
		gi|1705907 Chloride channel protein 4 (ClC-4)
		gi|129894 Alpha platelet-derived growth factor receptor precursor (PDGF-R-a
		gi|47575951 olfactory receptor Olr1498 [Rattus norvegicus]
		gi|109458945 PREDICTED: similar to RIKEN cDNA 5730590G19-like [Rattus norvegic
		gi|11120710 collagen, type V, alpha 3 [Rattus norvegicus]
		gi|21930125 osteoclast stimulating factor [Rattus norvegicus]
		gi|118763771 Unknown (protein for MGC:156700) [Rattus norvegicus]

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 32 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|109507063** Score: **107**
PREDICTED: similar to Actin, cytoplasmic 2 (Gamma-actin) [Rattus norvegicus]
Found in search of DATA.TXT

Nominal mass (M_r): **42109**; Calculated pI value: **5.31**
NCBI BLAST search of gi|109507063 against nr
Unformatted sequence string for pasting into other applications

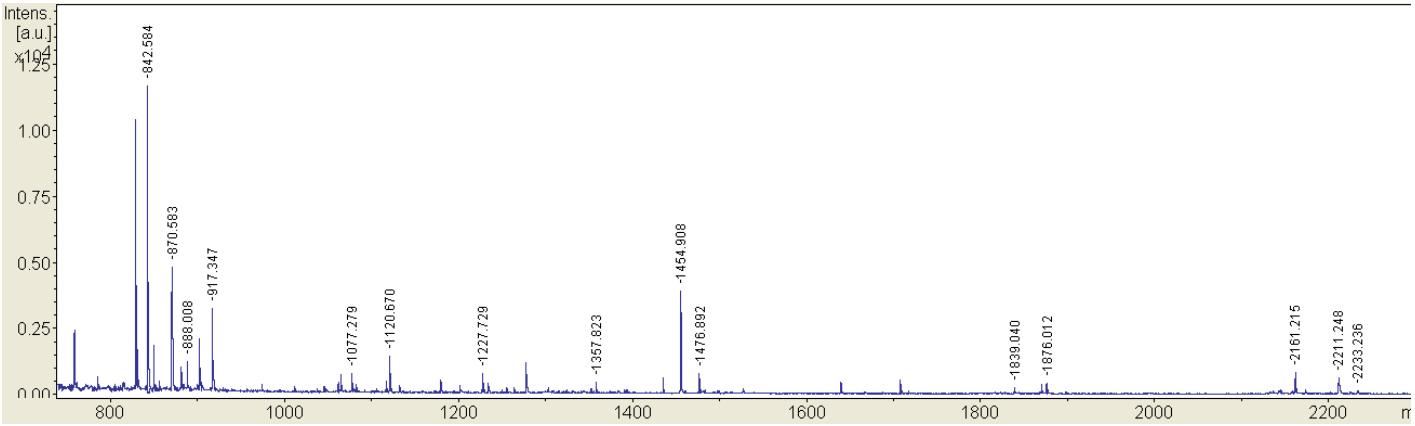
Taxonomy: Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **4%**

Matched peptides shown in **Bold Red**

```
     1 MEEEIAALVI DNGSGMCKAG FAGDDAPRAV FPSIVGRPRH QGVMVGMGQK
    51 DSYVGDEAQS KRGILTLKYP NEHGIVTNWD DMEKIWHHTF YNELRVAPEE
   101 HPVLLTEAPL NPKANREKMT QIMFETFNTP AMYVAIQAVL SLYASGRTTG
   151 IVMDSGDGVT HTVPIYEGYA LPHAILRLDL AGRDLTDYLM KILTERGYSF
   201 TTTAEREIVR DIKEKLCYVA LDFEQEMATA ASSSSLEKSY ELPDGQVITI
   251 GNERFRCPEA LFQPSFLGME SCGIHETTFN SIMKCDVDIR KDLYANTVLS
   301 GGTTMYPGIA DRMQKEITAL APSTMKIKII APPERKYSVW IGGSILASLS
   351 TFQQMWISKQ EYDESGPSIV HRKCF
```

Stress 70 protein



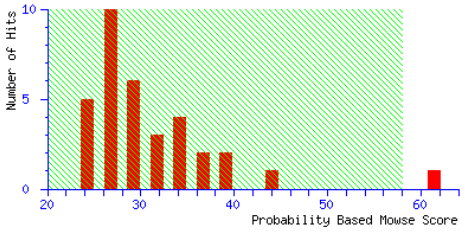
PMF of stress 70 protein.

Mascot Search Results

User : Katie
Email :
Search title : database search
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Rattus (39483 sequences)
Timestamp : 16 Dec 2008 at 09:24:10 GMT
Top Score : 61 for gi|62664205, PREDICTED: similar to Stress-70 protein, mitochondrial precursor

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help
Significance threshold p< 0.05 | Max. number of hits 50

Re-Search All | Search Unmatched

1. gi|62664205 **Mass:** 73984 **Score:** 61 **Expect:** 0.029 **Queries matched:** 7
PREDICTED: similar to Stress-70 protein, mitochondrial precursor (75 kDa glucose regulated prote

Protein View

Match to: **gi|62664205** Score: **61** Expect: **0.029**
PREDICTED: similar to Stress-70 protein, mitochondrial precursor (75 kDa glucose regulated protein)

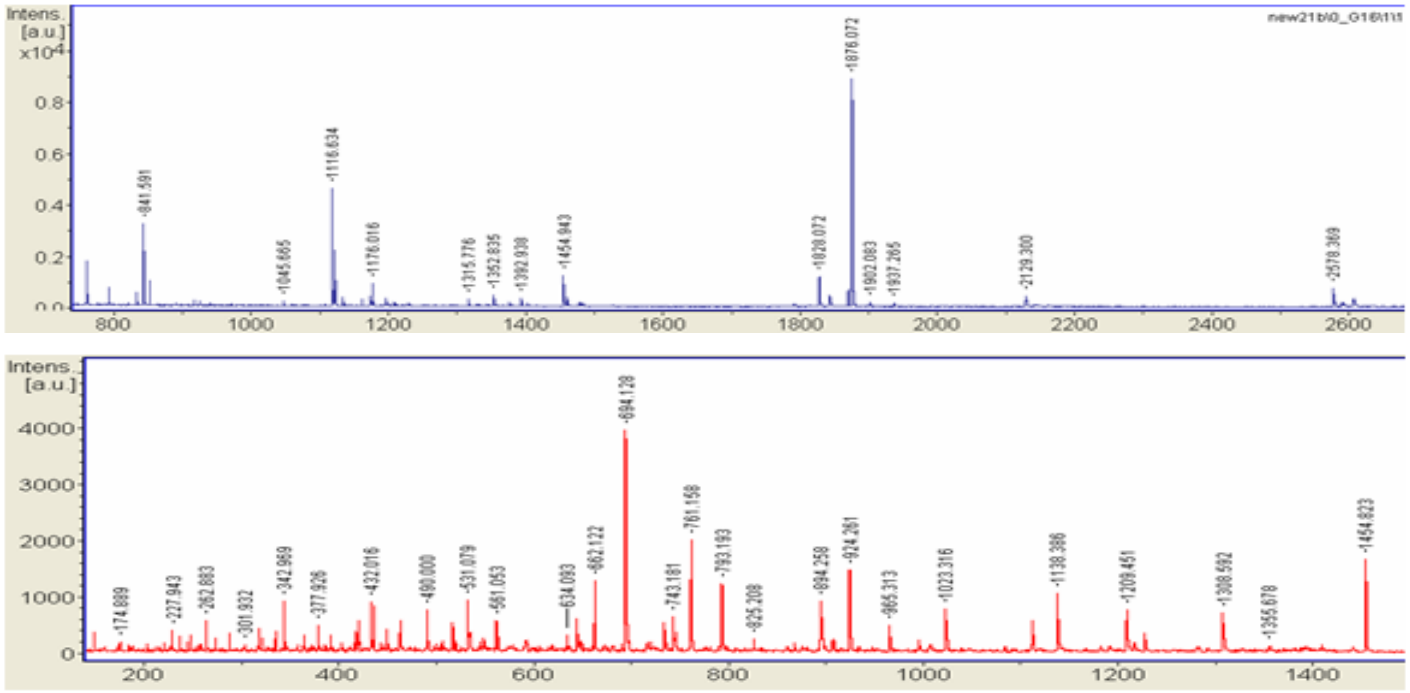
Nominal mass (M_r): **73984**; Calculated pI value: **5.87**
NCBI BLAST search of gi|62664205 against nr
Unformatted sequence string for pasting into other applications
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|109507099 (no taxonomy information for this entry)
gi|1000439 from Rattus sp

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **24**
Number of mass values matched: **7**
Sequence Coverage: **13%**

Matched peptides shown in **Bold Red**

```
    1 MISASRAAAA RLVGTTASRS PAAARHQDGW NGLSHEAFRF VSRRDYASEA
   51 IKGAVVGIDL GTTNSCVAVM EGKQAKVLEN AEGARTTPSV VAFTPDGERL
  101 VGMPAKRQAV TNPNNTFYAT KRLIGRRYDD PEVQKDTKNV PFKIVRASNG
  151 DAWVEAHGKL YSPSQIGAFV LMKMKETAEN YLGHTAKNAV ITVPAYFNDS
  201 QRQATKDAGQ ISGLNVLRVI NEPTAAALAY GLDKSEDKVI AVYDLGGGTF
  251 DISILEIQKG VFEVKSTNGD TFLGGEDFDQ ALLRHIVKEF KRETGVDLTK
  301 DNMALQRVRE AAEKAKCELS SSVQTDINLP YLTMDASGPK HLNMKLTRAQ
  351 FEGIVTDLIK RTIAPCQKAM QDAEVSKSDI GEVILVGGMT RMPKVQQTVQ
  401 DLFGRAPSKA VNPDEAVAIG AAIQGGVLAG DVTDVLLLDV TPLSLGIETL
  451 GGVFTKLINR NTTIPTKKSQ VFSTAADGQT QVEIKVCQGE REMAGDNKLL
  501 GQFTLIGIPP APRGVPQIEV TFDIDANGIV HVSAKDKGTG REQQIVIQSS
  551 GGLSKDDIEN MVKNAEKYAE EDRRKKERVE AVNMAEGIIH DTETKMEEFK
  601 DQLPADECNK LKEEISKMRE LLARKDSETG ENIRQAASSL QQASLKLFEM
  651 AYKKMASERE GSGSSSTGEQ KEDQKEEKQ
```

Apolipoprotein A-I



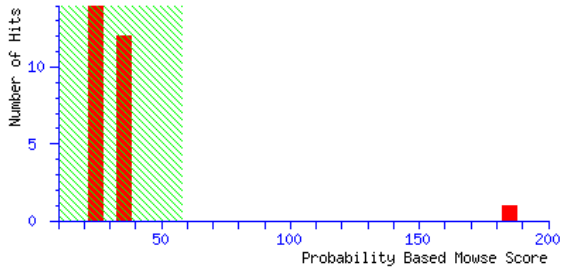
PMF of apolipoprotiein A-I (top). MS/MS spectrum of its 1454.943 peptide (bottom).

Mascot Search Results

User : Katie
Email :
Search title : database search
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Rattus (39483 sequences)
Timestamp : 26 Nov 2008 at 10:41:26 GMT
Top Score : 185 for gi|6978515, apolipoprotein A-I [Rattus norvegicus]

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

1. gi|6978515 **Mass:** 30100 **Score:** 185 **Expect:** 1.2e-014 **Queries matched:** 1
apolipoprotein A-I [Rattus norvegicus]

Protein View

Match to: **gi|6978515** Score: **185** Expect: **1.2e-014**
apolipoprotein A-I [Rattus norvegicus]

Nominal mass (M_r): **30100**; Calculated pI value: **5.52**
NCBI BLAST search of gi|6978515 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|202939 from Rattus norvegicus
gi|202945 from Rattus norvegicus
gi|59808388 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **26**
Number of mass values matched: **17**
Sequence Coverage: **50%**

Matched peptides shown in **Bold Red**

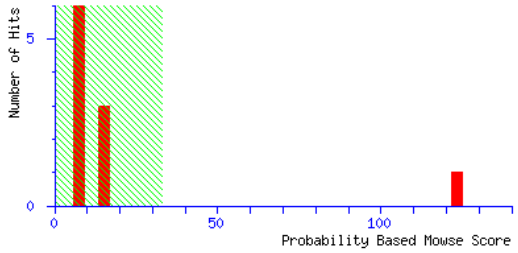
```
     1 MKAAVLAVAL VFLTGCQAWE FWQQDEPQSQ WDRVKDFATV YVDAVKDSGR
    51 DYVSQFESST LGKQLNLNLL DNWDTLGSTV GRLQEQLGPV TQEFWANLEK
   101 ETDWLRNEMN KDLENVKQKM QPHLDEFQEK WNEEVEAYRQ KLEPLGTELH
   151 KNAKEMQRHL KVVAEEFRDR MRVNADALRA KFGLYSDQMR ENLAQRLTEI
   201 KNHPTLIEYH TKASDHLKTL GEKAKPALDD LGQGLMPVLE AWKAKIMSMI
   251 DEAKKKLNA
```

Mascot Search Results

User	:	**Katie**
Email	:	
Search title	:	**database search**
MS data file	:	**DATA.TXT**
Database	:	**NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy	:	**Rattus (39483 sequences)**
Timestamp	:	**26 Nov 2008 at 11:00:47 GMT**
Protein hits	:	gi|113997 Apolipoprotein A-I precursor (Apo-AI) (ApoA-I)
		gi|18426862 proteasome 26S non-ATPase subunit 9 [Rattus norvegicus]
		gi|62079021 hypothetical protein LOC361301 [Rattus norvegicus]
		gi|109501906 PREDICTED: similar to F11C1.5a [Rattus norvegicus]
		gi|31542447 cysteine rich protein 61 [Rattus norvegicus]
		gi|75188767 COG2937: Glycerol-3-phosphate O-acyltransferase [Escherichia
		gi|47577843 olfactory receptor Olr119 [Rattus norvegicus]
		gi|109497906 PREDICTED: similar to UDP-N-acetylglucosamine: alpha-1,3-D-ma
		gi|16130617 anaerobic nitric oxide reductase flavorubredoxin [Escherichia
		gi|16128297 betaine aldehyde dehydrogenase, NAD-dependent [Escherichia co

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 33 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|6978515** Score: **123**
apolipoprotein A-I [Rattus norvegicus]
Found in search of DATA.TXT

Nominal mass (M_r): **30100**; Calculated pI value: **5.52**
NCBI BLAST search of gi|6978515 against nr
Unformatted sequence string for pasting into other applications

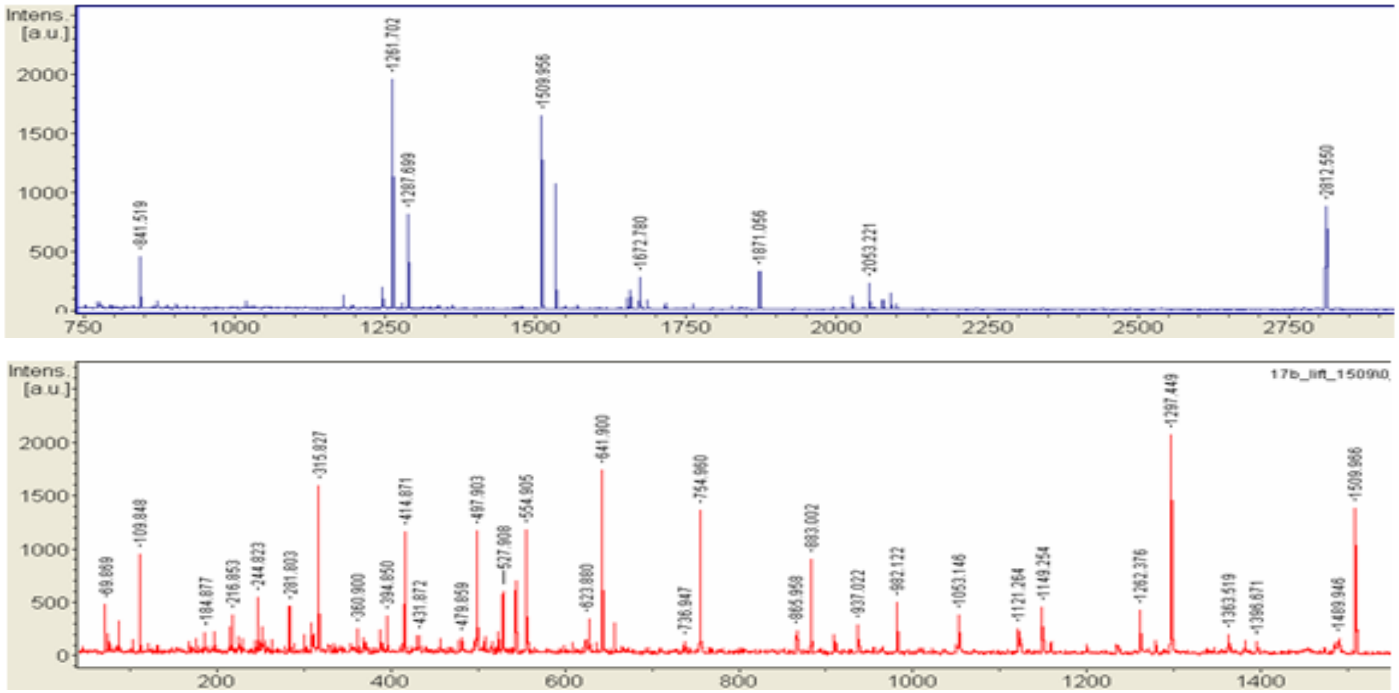
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|202939 from Rattus norvegicus
gi|202945 from Rattus norvegicus
gi|59808388 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **5%**

Matched peptides shown in **Bold Red**

```
   1 MKAAVLAVAL VFLTGCQAWE FWQQDEPQSQ WDRVKDFATV YVDAVKDSGR
  51 DYVSQFESST LGKQLNLNLL DNWDTLGSTV GRLQEQLGPV TQEFWANLEK
 101 ETDWLRNEMN KDLENVKQKM QPHLDEFQEK WNEEVEAYRQ KLEPLGTELH
 151 KNAKEMQRHL KVVAEEFRDR MRVNADALRA KFGLYSDQMR ENLAQRLTEI
 201 KNHPTLIEYH TKASDHLKTL GEKAKPALDD LGQGLMPVLE AWKAKIMSMI
 251 DEAKKKLNA
```

Apolipoprotein A-IV



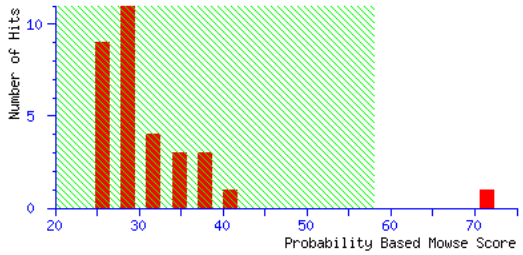
PMFof apolipoprotiein A-IV (top). MS/MS spectrum of its 1509.956 peptide (bottom).

Mascot Search Results

User : Katie
Email :
Search title : database search
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Rattus (39483 sequences)
Timestamp : 26 Nov 2008 at 11:29:31 GMT
Top Score : 71 for gi|114008, Apolipoprotein A-IV precursor (Apo-AIV) (ApoA-IV)

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help
Significance threshold p< 0.05 | Max. number of hits 500

Re-Search All | Search Unmatched

1. gi|114008 **Mass:** 44429 **Score:** 71 **Expect:** 0.0029 **Queries matched:** 7
Apolipoprotein A-IV precursor (Apo-AIV) (ApoA-IV)
gi|8392909 **Mass:** 44438 **Score:** 71 **Expect:** 0.0029 **Queries matched:** 7
apolipoprotein A-IV [Rattus norvegicus]

Protein View

Match to: **gi|114008** Score: **71** Expect: **0.0029**
Apolipoprotein A-IV precursor (Apo-AIV) (ApoA-IV)

Nominal mass (M_r): **44429**; Calculated pI value: **5.12**
NCBI BLAST search of gi|114008 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|202941 from Rattus norvegicus
gi|202943 from Rattus norvegicus
gi|60552712 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **17**
Number of mass values matched: **7**
Sequence Coverage: **24%**

Matched peptides shown in **Bold Red**

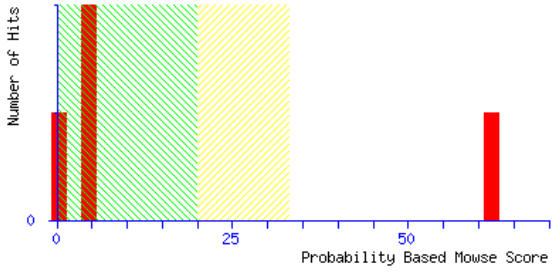
```
     1 MFLKAVVLTV ALVAITGTQA EVTSDQVANV MWDYFTQLSN NAKEAVEQLQ
    51 KTDVTQQLNT LFQDKLGNIN TYADDLQNKL VPFAVQLSGH LTKETERVRE
   101 EIQKELEDLR ANMMPHANKV SQMFGDNVQK LQEHLRPYAT DLQAQINAQT
   151 QDMKRQLTPY IQRMQTTIQD NVENLQSSMV PFANELKEKF NQNMEGLKGQ
   201 LTPRANELKA TIDQNLEDLR SRLAPLAEGV QEKLNHQMEG LAFQMKKNAE
   251 ELQTKVSTNI DQLQKNLAPL VEDVQSKLKG NTEGLQKSLE DLNKQLDQQV
   301 EVFRRAVEPL GDKFNMALVQ QMEKFRQQLG SDSGDVESHL SFLEKNLREK
   351 VSSFMSTLQK KGSPDQPLAL PLPEQVQEQV QEQVQPKPLE S
```

{MATRIX SCIENCE} Mascot Search Results

User : Katie
Email :
Search title : database search
MS data file : DATA.TXT
Database : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy : Rattus (39483 sequences)
Timestamp : 26 Nov 2008 at 11:33:14 GMT
Protein hits : gi|114008 Apolipoprotein A-IV precursor (Apo-AIV) (ApoA-IV)
gi|416805 Choline O-acetyltransferase (CHOACTase) (Choline acetylase) (ChAT)
gi|16131295 hypothetical protein b3421 [Escherichia coli K12]
gi|61557082 telomeric repeat binding factor 2, interacting protein [Rattus norvegicus]

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 20 indicate peptides with significant homology.
Individual ions scores > 33 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|114008** Score: **62**
Apolipoprotein A-IV precursor (Apo-AIV) (ApoA-IV)
Found in search of DATA.TXT

Nominal mass (M_r): **44429**; Calculated pI value: **5.12**
NCBI BLAST search of gi|114008 against nr
Unformatted sequence string for pasting into other applications

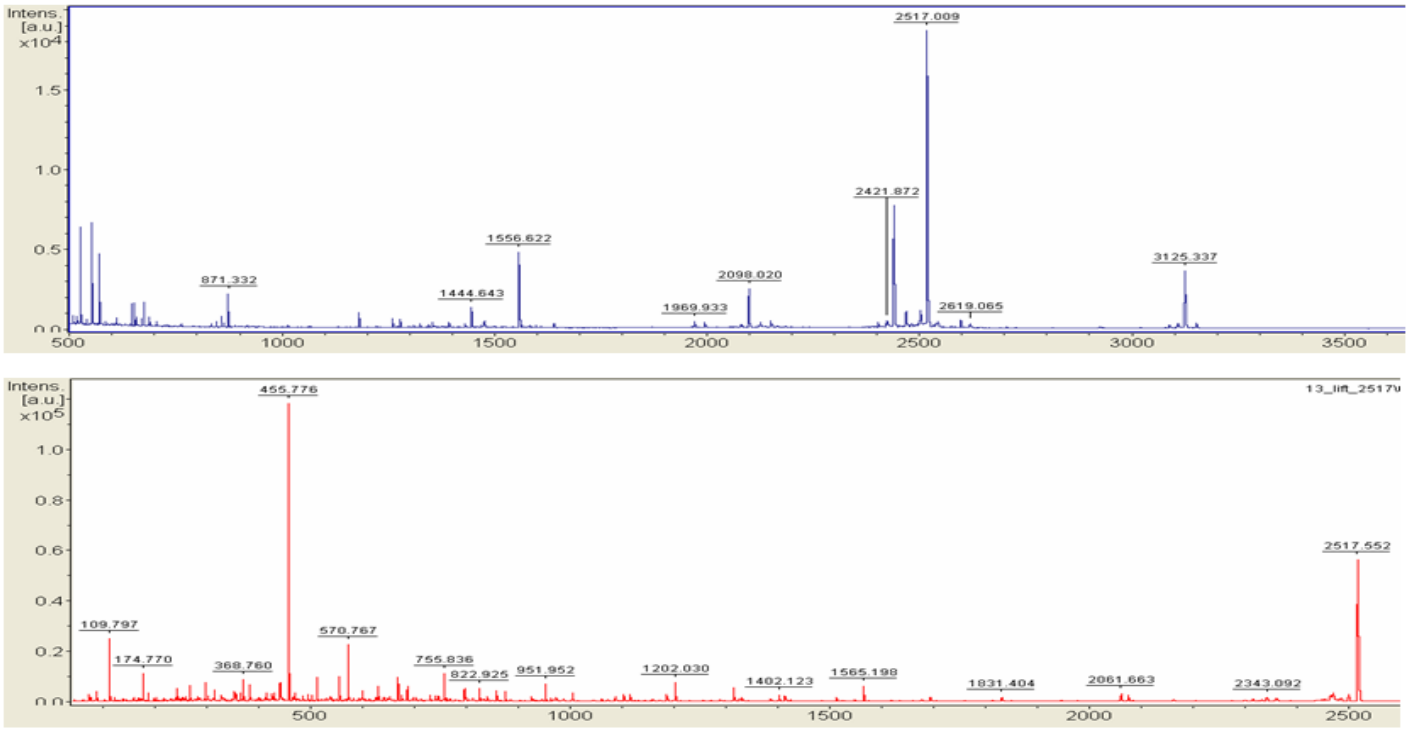
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|202941 from Rattus norvegicus
gi|202943 from Rattus norvegicus
gi|60552712 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: 3%

Matched peptides shown in **Bold Red**

```
     1 MFLKAVVLTV ALVAITGTQA EVTSDQVANV MWDYFTQLSN NAKEAVEQLQ
    51 KTDVTQQLNT LFQDKLGNIN TYADDLQNKL VPFAVQLSGH LTKETERVRE
   101 EIQKELEDLR ANMMPHANKV SQMFGDNVQK LQEHLRPYAT DLQAQINAQT
   151 QDMKRQLTPY IQRMQTTIQD NVENLQSSMV PFANELKEKF NQNMEGLKGQ
   201 LTPRANELKA TIDQNLEDLR SRLAPLAEGV QEKLNHQMEG LAFQMKKNAE
   251 ELQTKVSTNI DQLQKNLAPL VEDVQSKLKG NTEGLQKSLE DLNKQLDQQV
   301 EVFRRAVEPL GDKFNMALVQ QMEKFRQQLG SDSGDVESHL SFLEKNLREK
   351 VSSFMSTLQK KGSPDQPLAL PLPEQVQEQV QEQVQPKPLE S
```

Transthyretin



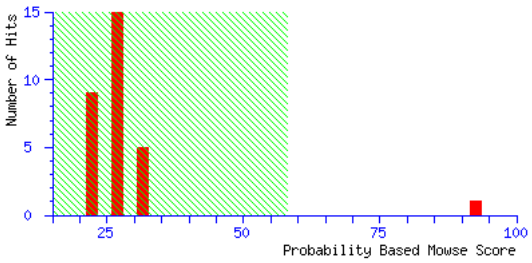
PMF of transthyretin (top). MS/MS spectrum of its 2517.009 peptide (bottom).

Mascot Search Results

```
User            : Katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 22 Oct 2008 at 04:25:10 GMT
Top Score       : 92 for gi|3212532, Chain A, Rat Transthyretin
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As: Concise Protein Summary — Help

Significance threshold p< 0.05 — Max. number of hits 5

Re-Search All | Search Unmatched

1. gi|3212532 **Mass:** 13122 **Score:** 92 **Expect:** 2.2e-005 **Queries matched:** 6
Chain A, Rat Transthyretin

Protein View

Match to: **gi|3212532** Score: **92** Expect: **2.2e-005**
Chain A, Rat Transthyretin

Nominal mass (M_r): **13122**; Calculated pI value: **6.04**
NCBI BLAST search of gi|3212532 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|3212533 from Rattus norvegicus
gi|3212534 from Rattus norvegicus
gi|3212535 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **30**
Number of mass values matched: **6**
Sequence Coverage: **73%**

Matched peptides shown in **Bold Red**

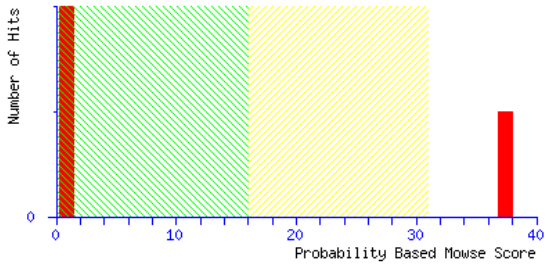
```
     1 SKCPLMVKVL DAVRGSPAVD VAVKVFKKTA DGSWEPFASG KTAESGELHG
    51 LTTDEKFTEG VYRVELDTKS YWKALGISPF HEYAEVVFTA NDSGHRHYTI
   101 AALLSPYSYS TTAVVSNPQN
```

{MATRIX SCIENCE} Mascot Search Results

User : **Katie**
Email :
Search title : **database search**
MS data file : **DATA.TXT**
Database : **NCBInr 20070216 (4626804 sequences; 1596079197 residues)**
Taxonomy : **Rattus (39483 sequences)**
Timestamp : **22 Oct 2008 at 04:28:30 GMT**
Protein hits : gi|136467 Transthyretin precursor (Prealbumin) (TBPA)
gi|47169510 TPA: mast cell protease 1-like 2 [Rattus norvegicus]
gi|109503443 PREDICTED: similar to deleted in liver cancer 1 isoform

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 16 indicate peptides with significant homology.
Individual ions scores > 31 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Protein View

Match to: **gi|3212532** Score: **37**
Chain A, Rat Transthyretin
Found in search of DATA.TXT

Nominal mass (M_r): **13122**; Calculated pI value: **6.04**
NCBI BLAST search of gi|3212532 against nr
Unformatted sequence string for pasting into other applications

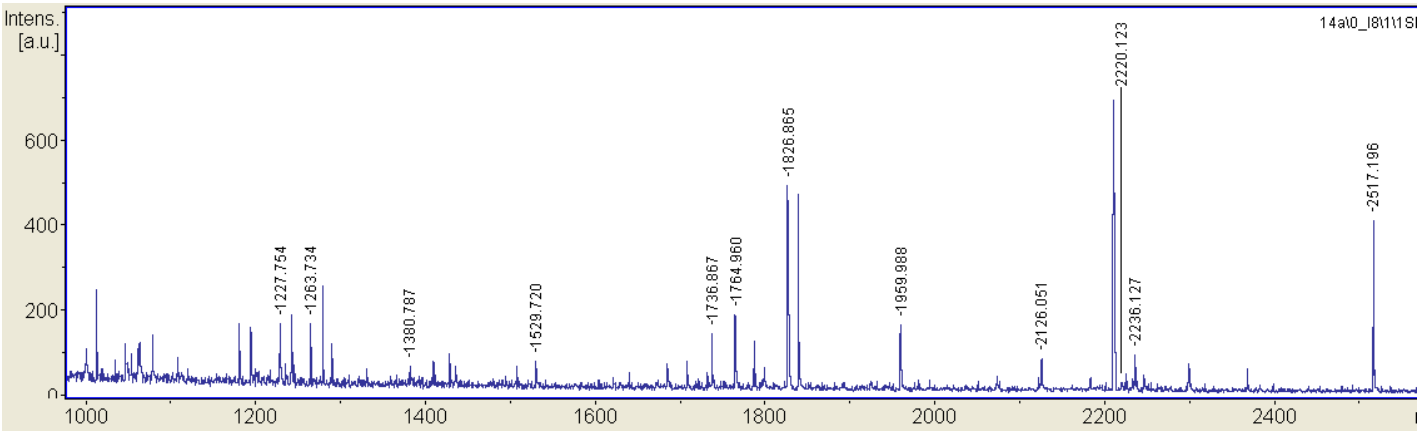
Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|3212533 from Rattus norvegicus
gi|3212534 from Rattus norvegicus
gi|3212535 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **19%**

Matched peptides shown in **Bold Red**

```
     1 SKCPLMVKVL DAVRGSPAVD VAVKVFKKTA DGSWEPFASG KTAESGELHG
    51 LTTDEKFTEG VYRVELDTKS YWKALGISPF HEYAEVVFTA NDSGHRHYTI
   101 AALLSPYSYS TTAVVSNPQN
```

Murinoglobulin



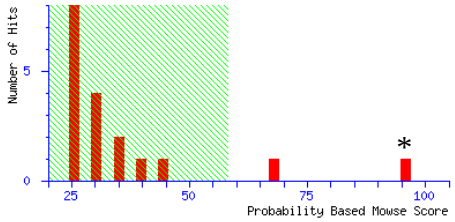
PMF of murinoglobulin.

Mascot Search Results

```
User            : Katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 26 Nov 2008 at 11:06:37 GMT
Top Score       : 96 for gi|12831225, Murinoglobulin 1 homolog [Rattus norvegicus]
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 500

Re-Search All | Search Unmatched

1. gi|12831225 **Mass:** 166590 **Score:** 96 **Expect:** 1e-005 **Queries matched:** 14
Murinoglobulin 1 homolog [Rattus norvegicus]

Protein View

Match to: **gi|12831225** Score: **96** Expect: **1e-005**
Murinoglobulin 1 homolog [Rattus norvegicus]

Nominal mass (M_r): **166590**; Calculated pI value: **5.68**
NCBI BLAST search of gi|12831225 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Rattus norvegicus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|81870614 from Rattus norvegicus
gi|55562 from Rattus norvegicus

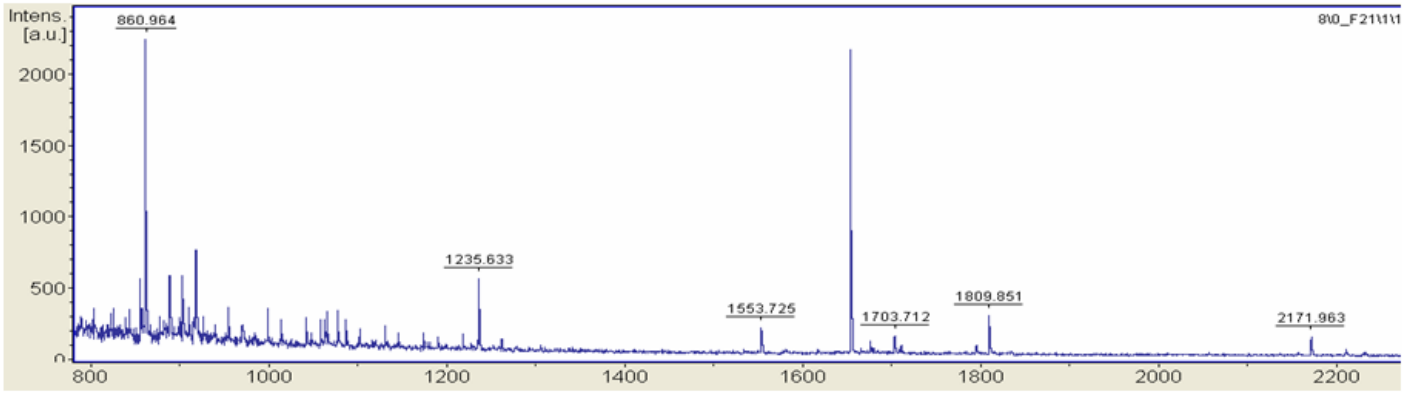
Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **18**
Number of mass values matched: **14**
Sequence Coverage: **9%**

Matched peptides shown in **Bold Red**

```
   1 MKKNREAQLC LFSALLAFLP FASLLNGNSK YMVLVPSQLY TETPEKICLH
  51 LYHLNETVTV TASLISQRGT RKLFDELVVD KDLFHCLSFT IPRLPSSEEE
 101 ESLDINIEGA KHKFSERRVV LVKNKESVVF VQTDKPVYKP GQSVKFRVVS
 151 MDKNLHPLNE LFPLAYIEDP KMNRIMQWQD IKTENGLKQL SFSLSAEPIQ
 201 GPYKIVILKQ SGVKEEHSFT VMEFVLPRFG VDVKVPNAIS VYDEIINVTA
 251 CAIYTYGKPV PGHVKISLCH GNPSFSSETK SACKEEDSEL DNNGCSTQEV
 301 NITEFQLKEN YLKMHQAFHV NATVTEEGTG SEFSGSGRIE VERTRNKFLF
 351 LKADSHFRHG IPFFVKIRLV DIKGDPIPNE QVFIKAQEAG YTNATTTDQH
 401 GLAKFSIDTS SISGYSLNIK VYHKEESSCI HSSCTAERHA EEHHTAYAVY
 451 SLSKSYIYLD TEAGVLPCNQ IHTVQAHFIL KGQVLGVLPQ IVFHYLVMAQ
 501 GSILQTGNHT HQVEPGVSQV QGNFALEIPV EFSMVPVAKM LIYTILPDGE
 551 VIADSVTFQV EKCLRNKVHL SFSPSQSLPA SQTHMRVTAS PQSLCGLRAV
 601 DQSVLLLKPE AELSPSLIYD LPGMQDSNFI PSSYHPFEDE YDCLMYQPRD
 651 TEELTYSVPY GREKDVYRYV RDMGLTAFTN LKIKHPTYCY EMNMVVLSAP
 701 AVESELSPRG GEFEMMPLGV NKSPLPKEPP RKDPPPKDPV IETIRNYFPE
 751 TWIWDLVTVN SSGVTEVEMT VPDTITEWKA GALCLSNDTG LGLSSVATLQ
 801 AFQPFFVELT MPYSVIRGEA FMLKATVMNY LPTSLPMAVQ LEASPDFTAV
 851 PVGNDQDSYC LGANGRHTSS WLVTPKSLGN VNFSVSVEAQ QSPELCGSQV
 901 ATVPETGRKD TVVKVLIVEP EGIKKEHTFS SLLCASDAEL SETLSLLLPP
 951 TVVKDSARAH FSVMGDILSS AIKNTQNLIQ MPYGCGEQNM VLFAPNIYVL
1001 KYLNETQQLT EKIKSKALGY LRAGYQRELN YKHKDGSYSA FGDHNGQGQG
1051 NTWLTAFVLK SFAQARAFIF IDESHITDAF TWLSKQQKDS GCFRSSGSLF
1101 NNAMKGGVDD EITLSAYITM ALLESSLPDT DPVVSKALGC LEASWETIEQ
1151 GRNGSFVYTK TLMAYAFALA GNQEKRNEIL KSLDKEAIRE DNSIHWERPQ
1201 KPTKSEGYLY TPQASSAEVE MSAYVVLARL TAQPAPSPED LALSMGTIKW
1251 LTKQQNSHGG FSSTQDTVVA LDALSKYGAA TFSKSQKTPL VTIQSSGSFS
1301 QKFQVDNSNR LLLQQVSLPD IPGNYTVSVS GEGCVYAQTT LRYNMPLEKQ
1351 QPAFALKVQT VPLTCNNPKG QNSFQISLEI SYTGSRPASN MVIADVKMLS
1401 GFIPLKPTVK KLERLEHVSR TEVTTNNVLL YLDQVTNQTL SFSFIIQQDI
1451 PVKNLQPAIV KVYDYYETDE VAFAEYSSPC SSDKQNV
```

*This identity was chosen as its molecular weight matches the gel.

Inter-α-inhibitor H4 heavy chain



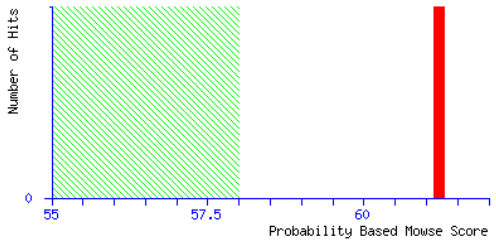
PMF of inter-α-inhibitor H4 heavy chain.

{MATRIX SCIENCE} Mascot Search Results

```
User            : Katie
Email           :
Search title    : database search
Database        : NCBInr 20070216 (4626804 sequences; 1596079197 residues)
Taxonomy        : Rattus (39483 sequences)
Timestamp       : 22 Oct 2008 at 04:15:25 GMT
Top Score       : 61 for gi|9506819, inter-alpha-inhibitor H4 heavy chain [Rattus norvegicus]
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 58 are significant (p<0.05).



Concise Protein Summary Report

Format As | Concise Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 5

Re-Search All | Search Unmatched

1. gi|9506819 **Mass:** 103885 **Score:** 61 **Expect:** 0.03 **Queries matched:** 5
inter-alpha-inhibitor H4 heavy chain [Rattus norvegicus]

Protein View

Match to: **gi|9506819** Score: **61** Expect: **0.03**
inter-alpha-inhibitor H4 heavy chain [Rattus norvegicus]

Nominal mass (M_r): **103885**; Calculated pI value: **6.08**
NCBI BLAST search of gi|9506819 against nr
Unformatted sequence string for pasting into other applications
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|2292988 from Rattus norvegicus

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Oxidation (M)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Number of mass values searched: **6**
Number of mass values matched: **5**
Sequence Coverage: **5%**

Matched peptides shown in **Bold Red**

```
     1 MKSPAPAHMW NIVLVLLSLL AVLPITTTEK NGIDIYSLTV DSRVSSRFAH
    51 TVVTSRVVNR ADTVQEATFQ VELPRKAFIT NFSMIIDGVT YPGLSKRRLK
   101 PEAIHCCCGR GESAGLVKTT GRKTEQFEVS VNVAPGSKTT FELIYQELLQ
   151 RRLGMYELLL KVRPEQLVKH LQMTSTSLSP QGISTLETES TFMTQELANA
   201 LTTSQNKTKA HIQFKPTLSQ QRKSQNEQDT VLDGDFTVRY DVDRSSTGGY
   251 LQIENGYFVH HFAPEDLPTM AKNVLFVIDK SGSMAGKKIQ QTREALIKIL
   301 KDLSTQDQFN IIVFSGEANQ WEQLLVQATE ENLNRAVDYA SKIPAQGGTN
   351 INKAVLSAVE LLDKSNQAEL LPSKSVSLII LLTDGEPTVG ETNPKIIQKN
   401 TQEAINGRYS LFCLGFGFDV NYPFLEKLAL DNGGLARRIY EDSDSALQLQ
   451 DFYQEVANPL LSSVTFEYPS NAVEDVTRYN FQHHFKGSEM VVAGKLRDQG
   501 PDVLLAKVSG QMHLQNITFQ TEASIAQQEK EFQGPKYIFH NFMERLWALL
   551 TIQQQLEQRI SASGAELEAL EAQVLNLSLK YNFVTPLTHM VVTKPEDQEQ
   601 FQVAEKPTEV DGGVWSILSA VQRHFKTPTT GSKLLTSRLR GNRFQTLSRL
   651 GDGLVGSRQY MPPPGLPGPP GLPGPPGPPG HPHFASSIDY GRQPSLGRVL
   701 DLPSLSSQDP AGPSLAMLPK VVEQEGTTPE ESPNPDHPRA PTIILPLPGS
   751 GVDQLCVDIL HSEKPMKLFV DINQGLEVVG KYEKNIGFSW IEVTILKPHL
   801 QVHATPERLV VTRGRKNSEY KWKKTLFSVL PGLKMTMDKT GLLQLSGPDK
   851 VTISLLSLDD PQRGLMLLLN DTHHFSNDIT GELGQFYQDI IWDDTKQTVR
   901 VLGIDYPATR ELKLSYQDGF PGTEISCWTV KI
```

Reference:

ABDUL-RAHMAN, P.S., LIM, B.K. and HASHIM, O.H. Expression of high-abundance proteins in sera of patients with endometrial and cervical cancers: analysis using 2-DE with silver staining and lectin detection methods. Electrophoresis. vol. 28, pp. 1989-1996 (2007)

ABE, M., YAMASHITA, S., KURAMOTO, T., HIRAYAMA, Y., TSUKAMOTO, T., OHTA, T., TATEMATSU, M., OHKI, M., TAKATO, T., SUGIMURA, T. and USHIJIMA, T. Global expression analysis of N-methyl-N'-nitro-N-nitrosoguanidine-induced rat stomach carcinomas using oligonucleotide microarrays. Carcinogenesis. vol. 24, pp. 861-867 (2003)

AHMED, N., BARKER, G., OLIVA, K., GARFIN, D., TALMADGE, K., GEORGIOU, H., QUINN, M. and RICE, G. An approach to remove albumin for the proteomic analysis of low abundance biomarkers in human serum. Proteomics. vol. 3, pp. 1980-1987 (2003)

AKITA, E.M. and NAKAI, S. Comparison of four purification methods for the production of immunoglobulins from eggs laid by hens immunized with an enterotoxigenic E. coli strain. Journal of Immunological Methods. vol. 160, pp. 207-214 (1993)

ALENCAR, H., FUNOVICS, M.A., FIGUEIREDO, J., SAWAYA, H., WEISSLEDER, R. and MAHMOOD, U. Colonic adenocarcinomas: near-infrared microcatheter imaging of smart probes for early detection--study in mice. Radiology. vol. 244, pp. 232-238 (2007)

ANDERSON, N.L. and ANDERSON, N.G. The human plasma proteome: history, character, and diagnostic prospects. Molecular and Cellular Proteomics. vol. 1, pp. 845-867 (2002)

BALKWILL, F. and MANTOVANI, A. Inflammation and cancer: back to Virchow? Lancet. vol. 357, pp. 539-545 (2001)

BARBER, M., FITZGERALD, R.C. and CALDAS, C. Familial gastric cancer - aetiology and pathogenesis. Best Practice and Research Clinical Gastroenterology. vol. 20, pp. 721-734 (2006)

BIZHANOV, G. and VYSHNIAUSKIS, G. A comparison of three methods for extracting IgY from the egg yolk of hens immunized with Sendai virus. Veterinary Research Communications. vol. 24, pp. 103-113 (2000)

BRALOW, S.P., GRUENSTEIN, M. and MERANZE, D.R. Host resistance to gastric adenocarcinomatosis in three strains of rats ingesting N-methyl-N'-nitro-N-nitrosoguanidine. Oncology. vol. 27, pp. 168-180 (1973)

CAMENISCH, G., TINI, M., CHILOV, D., KVIETIKOVA, I., SRINIVAS, V., CARO, J., SPIELMANN, P., WENGER, R.H. and GASSMANN, M. General applicability of chicken egg yolk antibodies: the performance of IgY immunoglobulins raised against the hypoxia-inducible factor 1alpha. FASEB Journal. vol. 13, pp. 81-88 (1999)

CAPUTO, E., LOMBARDI, M.L., LUONGO, V., MOHARRAM, R., TOMATORE, P., PIROZZI, G., GUARDIOLA, J. and MARTIN, B.M. Peptide profiling in epithelial tumor plasma by the emerging proteomic techniques. Journal of Chromatography B Analytical Technologies in the Biomedical and Life Sciences. vol. 819, pp. 59-66 (2005)

CARUSO, C., LIO, D., CAVALLONE, L. and FRANCESCHI, C. Aging, longevity, inflammation, and cancer. Annals of the New York Academy of Sciences. vol. 1028, pp. 1-13 (2004)

CHAN, D.C., CHEN, C.J., CHU, H.C., CHANG, W.K., YU, J.C., CHEN, Y.J., WEN, L.L., HUANG, S.C., KU, C.H., LIU, Y.C. and CHEN, J.H. Evaluation of serum amyloid A as a biomarker for gastric cancer. Annals of Surgical Oncology. vol. 14, pp. 84-93 (2007)

CHANG, S.H., AI, Y., BREVER, R.M., LANE, T.F. and HLA, T. The prostaglandin E2 receptor EP2 is required for cyclooxygenase 2-mediated mammary hyperplasia. Cancer Research. vol. 65, pp. 4496-4499 (2005)

CHEN, J., KAHNE, T., ROCKEN, C., GOTZE, T., YU, J., SUNG, J.J., CHEN, M., HU, P., MALFERTHEINER, P. and EBERT, M.P. Proteome analysis of gastric cancer metastasis by two-dimensional gel electrophoresis and matrix assisted laser desorption/ionization-mass spectrometry for identification of metastasis-related proteins. Journal of Proteome Research. vol. 3, pp. 1009-1016 (2004)

CHUNG, J.C., OH, M.J., CHOI, S.H. and BAE, C.D. Proteomic analysis to identify biomarker proteins in pancreatic ductal adenocarcinoma. ANZ Journal of Surgery. vol. 78, pp. 245-251 (2008)

CONCANNON, C.G., GORMAN, A.M. and SAMALI, A. On the role of Hsp27 in regulating apoptosis. Apoptosis. vol. 8, pp. 61-70 (2003)

CORREA, P. A human model of gastric carcinogenesis. Cancer Research. vol. 48, pp. 3554-3560 (1988)

CRADDOCK, V.M. The reaction of N-methyl-N'-nitro-N-nitrosoguanidine with deoxyribonucleic acid. Biochemical Journal. vol. 106, pp. 921-922 (1968)

DANNENBERG, A.J., LIPPMAN, S.M., MANN, J.R., SUBBARAMAIAH, K. and DUBOIS, R.N. Cyclooxygenase-2 and epidermal growth factor receptor: pharmacologic targets for chemoprevention. Journal of Clinical Oncology. vol. 23, pp. 254-266 (2005)

DIAMANDIS, E.P. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations. Molecular and Cellular Proteomics. vol. 3, pp. 367-378 (2004)

DOWLING, P., O'DRISCOLL, L., MELEADY, P., HENRY, M., ROY, S., BALLOT, J., MORIARTY, M., CROWN, J. and CLYNES, M. 2-D difference gel electrophoresis of the lung squamous cell carcinoma versus normal sera demonstrates consistent alterations in the levels of ten specific proteins. Electrophoresis. vol. 28, pp. 4302-4310 (2007)

DOWLING, P., WORMALD, R., MELEADY, P., HENRY, M., CURRAN, A. and CLYNES, M. Analysis of the saliva proteome from patients with head and neck squamous cell carcinoma reveals differences in abundance levels of proteins associated with tumour progression and metastasis. J Proteomics. vol. 71, pp. 168-175 (2008)

EBERT, M.P., KRUGER, S., FOGERON, M.L., LAMER, S., CHEN, J., PROSS, M., SCHULZ, H.U., LAGE, H., HEIM, S., ROESSNER, A., MALFERTHEINER, P. and ROCKEN, C. Overexpression of cathepsin B in gastric cancer identified by proteome analysis. Proteomics. vol. 5, pp. 1693-1704 (2005)

EBERT, M.P., NIEMEYER, D., DEININGER, S.O., WEX, T., KNIPPIG, C., HOFFMANN, J., SAUER, J., ALBRECHT, W., MALFERTHEINER, P. and ROCKEN, C. Identification and confirmation of increased fibrinopeptide a serum protein levels in gastric cancer sera by magnet bead assisted MALDI-TOF mass spectrometry. Journal of Proteome Research. vol. 5, pp. 2152-2158 (2006)

EL-AKAWI, Z.J., AL-HINDAWI, F.K. and BASHIR, N.A. Alpha-1 antitrypsin (alpha1-AT) plasma levels in lung, prostate and breast cancer patients. Neuroendocrinology Letters. vol. 29, pp. 482-484 (2008)

FELDMAN, A.L., ESPINA, V., PETRICOIN, E.F., 3RD, LIOTTA, L.A. and ROSENBLATT, K.P. Use of proteomic patterns to screen for gastrointestinal malignancies. Surgery. vol. 135, pp. 243-247 (2004)

FISCHER, M., HLINAK, A., MONTAG, T., CLAROS, M., SCHADE, R. and EBNER, D. Comparison of standard methods for the preparation of egg yolk antibodies. Tierarztl Prax. vol. 24, pp. 411-418 (1996)

GABRIJELCIC, D., SVETIC, B., SPAIC, D., SKRK, J., BUDIHNA, J. and TURK, V. Determination of cathepsins B, H, L and kininogen in breast cancer patients. Agents and Actions Supplements. vol. 38, no. 2, pp. 350-357 (1992)

GANAPATHY, E., PERAMAIYAN, R., RAJASEKARAN, D., VENKATARAMAN, M. and DHANAPAL, S. Modulatory effect of naringenin on N-methyl-N'-nitro-N-nitrosoguanidine- and saturated sodium chloride-induced gastric carcinogenesis in male Wistar rats. Clinical and Experimental Pharmacology and Physiology. vol. 35, pp. 1190-1196 (2008)

GARRIDO, C., FROMENTIN, A., BONNOTTE, B., FAVRE, N., MOUTET, M., ARRIGO, A.P., MEHLEN, P. and SOLARY, E. Heat shock protein 27 enhances the tumorigenicity of immunogenic rat colon carcinoma cell clones. Cancer Research. vol. 58, pp. 5495-5499 (1998)

GENTA, R.M. and RUGGE, M. Assessing risks for gastric cancer: new tools for pathologists. World Journal of Gastroenterology. vol. 12, pp. 5622-5627 (2006)

GIANAZZA, E. and ARNAUD, P. A general method for fractionation of plasma proteins. Dye-ligand affinity chromatography on immobilized Cibacron blue F3-GA. Biochemical Journal. vol. 201, pp. 129-136 (1982)

GORG, A., OBERMAIER, C., BOGUTH, G., HARDER, A., SCHEIBE, B., WILDGRUBER, R. and WEISS, W. The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis. vol. 21, pp. 1037-1053 (2000)

GRUNDMANN, E. Histologic types and possible initial stages in early gastric carcinoma. Beiträge zur Pathologie. vol. 154, pp. 256-280 (1975)

HATTA, H., TSUDA, K., AKACHI, S., KIM, M. and YAMAMOTO, T. Productivity and some properties of egg yolk antibody (IgY) against human rotavirus compared with rabbit IgG. Bioscience, Biotechnology, and Biochemistry. vol. 57, pp. 450-454 (1993)

HE, Q.Y., CHEUNG, Y.H., LEUNG, S.Y., YUEN, S.T., CHU, K.M. and CHIU, J.F. Diverse proteomic alterations in gastric adenocarcinoma. Proteomics. vol. 4, pp. 3276-3287 (2004)

HEO, S.H., LEE, S.J., RYOO, H.M., PARK, J.Y. and CHO, J.Y. Identification of putative serum glycoprotein biomarkers for human lung adenocarcinoma by multilectin affinity chromatography and LC-MS/MS. Proteomics. vol. 7, pp. 4292-4302 (2007)

HERVIEU, V., LOMBARD-BOHAS, C., DUMORTIER, J., BOILLOT, O. and SCOAZEC, J.Y. Primary acinar cell carcinoma of the liver. Virchows Archiv: an international journal of pathology. vol. 452, pp. 337-341 (2008)

HOFFMAN, S.A., JOO, W.A., ECHAN, L.A. and SPEICHER, D.W. Higher dimensional (Hi-D) separation strategies dramatically improve the potential for cancer biomarker detection in serum and plasma. Journal of Chromatography B Analytical Technologies in the Biomedical and Life Sciences. vol. 849, pp. 43-52 (2007)

HSU, P.I., CHEN, C.H., HSIEH, C.S., CHANG, W.C., LAI, K.H., LO, G.H., HSU, P.N., TSAY, F.W., CHEN, Y.S., HSIAO, M., CHEN, H.C. and LU, P.J. Alpha1-antitrypsin precursor in gastric juice is a novel biomarker for gastric cancer and ulcer. Clinical Cancer Research. vol. 13, pp. 876-883 (2007)

HUANG, L., HARVIE, G., FEITELSON, J.S., GRAMATIKOFF, K., HEROLD, D.A., ALLEN, D.L., AMUNNGAMA, R., HAGLER, R.A., PISANO, M.R., ZHANG, W.W. and FANG, X. Immunoaffinity separation of plasma proteins by IgY microbeads: meeting the needs of proteomic sample preparation and analysis. Proteomics. vol. 5, pp. 3314-3328 (2005)

JACKSON, D., CRAVEN, R.A., HUTSON, R.C., GRAZE, I., LUETH, P., TONGE, R.P., HARTLEY, J.L., NICKSON, J.A., RAYNER, S.J., JOHNSTON, C., DIEPLINGER, B., HUBALEK, M., WILKINSON, N., PERREN, T.J., KEHOE, S., HALL, G.D., DAXENBICHLER, G., DIEPLINGER, H., SELBY, P.J. and BANKS, R.E. Proteomic profiling identifies afamin as a potential biomarker for ovarian cancer. Clinical Cancer Research. vol. 13, pp. 7370-7379 (2007)

Japanese Gastric Cancer Association. Japanese Classification of Gastric Carcinoma - 2nd English Edition. Gastric Cancer. vol. 1, pp. 10-24 (1998)

JUAN, H.F., CHEN, J.H., HSU, W.T., HUANG, S.C., CHEN, S.T., YI-CHUNG LIN, J., CHANG, Y.W., CHIANG, C.Y., WEN, L.L., CHAN, D.C., LIU, Y.C. and CHEN, Y.J. Identification of tumor-associated plasma biomarkers using proteomic techniques: from mouse to human. Proteomics. vol. 4, pp. 2766-2775 (2004)

KALOGERAKI, A., GARBAGNATI, F., DARIVIANAKI, K., DELIDES, G.S., SANTINAMI, M., STATHOPOULOS, E.N. and ZORAS, O. HSP-70, C-myc and HLA-DR expression in patients with cutaneous malignant melanoma metastatic in lymph nodes. Anticancer Research. vol. 26, pp. 3551-3554 (2006)

KANDA, S., SUGIYAMA, K., TAKAHASHI, M., SHUMIYA, S., TOMINO, S. and NAGASE, S. Identification of a protein increasing in serum of Nagase analbuminemic rats bearing intestinal tumors as an isotype of T-kininogen. Japanese Journal of Cancer Research. vol. 81, pp. 63-68 (1990)

KANEKO, M., MORIMURA, K., NISHIKAWA, T., WANIBUCHI, H., TAKADA, N., OSUGI, H., KINOSHITA, H. and FUKUSHIMA, S. Different genetic alterations in rat forestomach tumors induced by genotoxic and non-genotoxic carcinogens. Carcinogenesis. vol. 23, pp. 1729-1735 (2002)

KAPRANOS, N., KOMINEA, A., KONSTANTINOPOULOS, P.A., SAVVA, S., ARTELARIS, S., VANDOROS, G., SOTIROPOULOU-BONIKOU, G. and PAPAVASSILIOU, A.G. Expression of the 27-kDa heat shock protein (HSP27) in gastric carcinomas and adjacent normal, metaplastic, and dysplastic gastric mucosa, and its prognostic significance. Journal of Cancer Research and Clinical Oncology. vol. 128, pp. 426-432 (2002)

KATO, M., HERZ, F., KATO, S. and HIRANO, A. Expression of stress-response (heat-shock) protein 27 in human brain tumors: an immunohistochemical study. Acta Neuropathol. vol. 83, pp. 420-422 (1992)
KAWAKAMI, T., HOSHIDA, Y., KANAI, F., TANAKA, Y., TATEISHI, K., IKENOUE, T., OBI, S., SATO, S., TERATANI, T., SHIINA, S., KAWABE, T., SUZUKI, T., HATANO, N., TANIGUCHI, H. and OMATA, M. Proteomic analysis of sera from hepatocellular carcinoma patients after radiofrequency ablation treatment. Proteomics. vol. 5, pp. 4287-4295 (2005)

KLEMPERER, F. Ueber natürliche Immunität und ihre Verwerthung für die Immunisirungstherapie. Archiv für Experimentelle Pathologie und Pharmakologie. vol. 31, pp. 356-382 (1893)

KLOSE, J. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. Humangenetik. vol. 26, pp. 231-243 (1975)

KOBORI, O., GEDIGK, P. and TOTOVIC, V. Early changes of glandular stomach in Wistar rats ingesting N-methyl-N'-nitro-N-nitrosoguanidine (MNNG): with special reference to light microscopic, electron microscopic, and enzyme histochemical study of the regenerating epithelium induced by MNNG. Zeitschrift für Krebsforschung und klinische Onkologie. Cancer research and clinical oncology. vol. 87, pp. 127-138 (1976)

KOBORI, O., GEDIGK, P. and TOTOVIC, V. Adenomatous changes and adenocarcinoma of glandular stomach in Wistar rats induced by N-methyl-N'-nitro-N-nitrosoguanidine. An electron microscopic and histochemical study. Virchows Archive A: Pathological Anatomy and Histology. vol. 373, pp. 37-54 (1977)

KOBORI, O., SHIMIZU, T., MAEDA, M., ATOMI, Y., WATANABE, J., SHOJI, M. and MORIOKA, Y. Enhancing effect of bile and bile acid on stomach tumorigenesis induced by N-methyl-N'-nitro-N-nitrosoguanidine in Wistar rats. Journal of the National Cancer Institute. vol. 73, pp. 853-861 (1984)

KUNZE, E., SCHAUER, A., EDER, M. and SEEFELDT, C. Early sequential lesions during development of experimental gastric cancer with special reference to dysplasias. Journal of Cancer Research and Clinical Oncology. vol. 95, pp. 247-264 (1979)

KWAK, J.Y., MA, T.Z., YOO, M.J., CHOI, B.H., KIM, H.G., KIM, S.R., YIM, C.Y. and KWAK, Y.G. The comparative analysis of serum proteomes for the discovery of biomarkers for acute myeloid leukemia. Experimental Hematology. vol. 32, pp. 836-842 (2004)

LARSSON, A., BALOW, R.M., LINDAHL, T.L. and FORSBERG, P.O. Chicken antibodies: taking advantage of evolution--a review. Poultry Science. vol. 72, pp. 1807-1812 (1993)

LAUREN, P. The Two Histological Main Types of Gastric Carcinoma: Diffuse and So-Called Intestinal-Type Carcinoma. An Attempt at a Histo-Clinical Classification. Acta Pathologica Microbiologica Scandinavia. vol. 64, pp. 31-49 (1965)

LAWLEY, P.D. and THATCHER, C.J. Methylation of deoxyribonucleic acid in cultured mammalian cells by N-methyl-N'-nitro-N-nitrosoguanidine. The influence of cellular thiol concentrations on the extent of methylation and the 6-oxygen atom of guanine as a site of methylation. Biochemical Journal. vol. 116, pp. 693-707 (1970)

LAWLEY, P.D. and SHAH, S.A. Methylation of ribonucleic acid by the carcinogens dimethyl sulphate, N-methyl-N-nitrosourea and N-methyl-N'-nitro-N-nitrosoguanidine. Comparisons of chemical analyses at the nucleoside and base levels. Biochemical Journal. vol. 128, pp. 117-132 (1972)

LEATHERBARROW, R.J. and DEAN, P.D. Studies on the mechanism of binding of serum albumins to immobilized cibacron blue F3G A. Biochemical Journal. vol. 189, pp. 27-34 (1980)

LEE, C.H., LUM, J.H., CHEUNG, B.P., WONG, M.S., BUTT, Y.K., TAM, M.F., CHAN, W.Y., CHOW, C., HUI, P.K., KWOK, F.S., LO, S.C. and FAN, D.M. Identification of the heterogeneous nuclear ribonucleoprotein A2/B1 as the antigen for the gastrointestinal cancer specific monoclonal antibody MG7. Proteomics. vol. 5, pp. 1160-1166 (2005)

LEE, K., KYE, M., JANG, J.S., LEE, O.J., KIM, T. and LIM, D. Proteomic analysis revealed a strong association of a high level of alpha1-antitrypsin in gastric juice with gastric cancer. Proteomics. vol. 4, pp. 3343-3352 (2004)

LEMIEUX, P., OESTERREICH, S., LAWRENCE, J.A., STEEG, P.S., HILSENBECK, S.G., HARVEY, J.M. and FUQUA, S.A. The small heat shock protein hsp27 increases invasiveness but decreases motility of breast cancer cells. Invasion and Metastasis. vol. 17, pp. 113-123 (1997)

LI, C. and LEE, K.H. Affinity depletion of albumin from human cerebrospinal fluid using Cibacron-blue-3G-A-derivatized photopatterned copolymer in a microfluidic device. Analytical Biochemistry. vol. 333, pp. 381-388 (2004) LEMIEUX, P., OESTERREICH, S., LAWRENCE, J.A., STEEG, P.S., HILSENBECK, S.G., HARVEY, J.M. and FUQUA, S.A. The small heat shock protein hsp27 increases invasiveness but decreases motility of breast cancer cells. Invasion and Metastasis. vol. 17, pp. 113-123 (1997)

LIU, L., LIU, J., DAI, S., WANG, X., WU, S., WANG, J., HUANG, L., XIAO, X. and HE, D. Reduced transthyretin expression in sera of lung cancer. Cancer Science. vol. 98, pp. 1617-1624 (2007a)

LIU, W., LIU, B., XIN, L., ZHANG, Y., CHEN, X., ZHU, Z. and LIN, Y. Down-regulated expression of complement factor I: a potential suppressive protein for gastric cancer identified by serum proteome analysis. Clinical Chimica Acta. vol. 377, pp. 119-126 (2007b)

MANIKANDAN, P., VIDJAYA LETCHOUMY, P., PRATHIBA, D. and NAGINI, S. Combinatorial chemopreventive effect of Azadirachta indica and Ocimum sanctum on oxidant-antioxidant status, cell proliferation, apoptosis and angiogenesis in a rat forestomach carcinogenesis model. Singapore Medical Journal. vol. 49, pp. 814-822 (2008)

MANTOVANI, A., ALLAVENA, P., SICA, A. and BALKWILL, F. Cancer-related inflammation. Nature. vol. 454, pp. 436-444 (2008)

MAO, J.T., CUI, X., RECKAMP, K., LIU, M., KRYSAN, K., DALWADI, H., SHARMA, S., HAZRA, S., STRIETER, R., GARDNER, B. and DUBINETT, S.M. Chemoprevention strategies with cyclooxygenase-2 inhibitors for lung cancer. Clinical Lung Cancer. vol. 7, pp. 30-39 (2005)

MARTIN, M.S., MARTIN, F., JUSTRABO, E., MICHIELS, R., BASTIEN, H. and KNOBEL, S. Susceptibility of inbred rats to gastric and duodenal carcinomas induced by N-methyl-N'-nitro-N-nitrosoguanidine. Journal of the National Cancer Institute. vol. 53, pp. 837-840 (1974)

MATSUKURA, N., KAWACHI, T., SASAJIMA, K., SANO, T., SUGIMURA, T. and HIROTA, T. Induction of intestinal metaplasia in the stomachs of rats by N-methyl-N'-nitro-N-nitrosoguanidine. Journal of the National Cancer Institute. vol. 61, pp. 141-144 (1978)

MC, K.A. and WRIGHT, G.F. The nitration products of 2-nitramino-delta2-1,3-diazacycloalkenes. Journal of the American Chemical Society. vol. 70, pp. 3990-3994 (1948)

MCCALLA, D.R. Reaction of N-methyl-N'-nitro-N-nitroguanidine and N-methyl-N-nitroso-p-toluenesulfonamide with DNA in vitro. Biochimica et Biophysica Acta. vol. 155, pp. 114-120 (1968)

MING, S.C. Dysplasia of gastric epithelium. Frontiers of Gastrointestinal Research. vol. 4, pp. 164-172 (1979)

MORSON, B.C., SOBIN, L.H., GRUNDMANN, E., JOHANSEN, A., NAGAYO, T. and SERCK-HANSSEN, A. Precancerous conditions and epithelial dysplasia in the stomach. Journal of Clinical Pathology. vol. 33, pp. 711-721 (1980)

NAGAI, T., PFEIFFER, C.J., FUJIMURA, M., HATTORI, T. and TOBE, T. Susceptibility of healed gastric ulcers to chemical carcinogenesis in rats and implications of cellular kinetic changes. Cancer Research. vol. 44, pp. 5828-5835 (1984)

NELSON, W.G., DE MARZO, A.M., DEWEESE, T.L. and ISAACS, W.B. The role of inflammation in the pathogenesis of prostate cancer. Journal of Urology. vol. 172, pp. 6-12 (2004)

NEWBERNE, P.M., CHARNLEY, G., ADAMS, K., CANTOR, M., SUPHAKARN, V., ROTH, D. and SCHRAGER, T.F. Gastric carcinogenesis: a model for the identification of risk factors. Cancer Letters. vol. 38, pp. 149-163 (1987)

NISHIGAKI, R., OSAKI, M., HIRATSUKA, M., TODA, T., MURAKAMI, K., JEANG, K.T., ITO, H., INOUE, T. and OSHIMURA, M. Proteomic identification of differentially-expressed genes in human gastric carcinomas. Proteomics. vol. 5, pp. 3205-3213 (2005)

NIWA, T., YAMASHITA, S., TSUKAMOTO, T., KURAMOTO, T., NOMOTO, T., WAKAZONO, K., FUJITA, H., MATSUSHIMA, T., TATEMATSU, M., SUGIMURA, T. and USHIJIMA, T. Whole-genome analyses of loss of heterozygosity and methylation analysis of four tumor-suppressor genes in N-methyl-N'-nitro-N-nitrosoguanidine-induced rat stomach carcinomas. Cancer Science. vol. 96, pp. 409-413 (2005)

O'FARRELL, P.H. High resolution two-dimensional electrophoresis of proteins. Journal of Biological Chemistry. vol. 250, pp. 4007-4021 (1975)

OHGAKI, H., KAWACHI, T., MATSUKURA, N., MORINO, K., MIYAMOTO, M. and SUGIMURA, T. Genetic control of susceptibility of rats to gastric carcinoma. Cancer Research. vol. 43, pp. 3663-3667 (1983)

OHGAKI, H., KATO, T., MORINO, K., MATSUKURA, N., SATO, S., TAKAYAMA, S. and SUGIMURA, T. Study of the promoting effect of sodium chloride on gastric carcinogenesis by N-methyl-N'-nitro-N-nitrosoguanidine in inbred Wistar rats. Gann. vol. 75, pp. 1053-1057 (1984)

OHGAKI, H., KLEIHUES, P. and SUGIMURA, T. Experimental gastric cancer. Italian Journal of Gastroenterology. vol. 23, pp. 371-377 (1991)

OIEN, K.A., MCGREGOR, F., BUTLER, S., FERRIER, R.K., DOWNIE, I., BRYCE, S., BURNS, S. and KEITH, W.N. Gastrokine 1 is abundantly and specifically expressed in superficial gastric epithelium, down-regulated in gastric carcinoma, and shows high evolutionary conservation. Journal of Pathology. vol. 203, pp. 789-797 (2004)

PATZ, E.F., JR., CAMPA, M.J., GOTTLIN, E.B., KUSMARTSEVA, I., GUAN, X.R. and HERNDON, J.E. Panel of serum biomarkers for the diagnosis of lung cancer. Journal of Clinical Oncology. vol. 25, pp. 5578-5583 (2007)

PAWLIK, T.M., HAWKE, D.H., LIU, Y., KRISHNAMURTHY, S., FRITSCHE, H., HUNT, K.K. and KUERER, H.M. Proteomic analysis of nipple aspirate fluid from women with early-stage breast cancer using isotope-coded affinity tags and tandem mass spectrometry reveals differential expression of vitamin D binding protein. BMC Cancer. vol. 6, pp. 68 (2006)

PETRICOIN, E.F., ORNSTEIN, D.K., PAWELETZ, C.P., ARDEKANI, A., HACKETT, P.S., HITT, B.A., VELASSCO, A., TRUCCO, C., WIEGAND, L., WOOD, K., SIMONE, C.B., LEVINE, P.J., LINEHAN, W.M., EMMERT-BUCK, M.R., STEINBERG, S.M., KOHN, E.C. and LIOTTA, L.A. Serum proteomic patterns for detection of prostate cancer. Journal of the National Cancer Institute. vol. 94, pp. 1576-1578 (2002)

RADWANSKA, A., BACZYNSKA, D., NOWAK, D., BREZILLON, S., POPOW, A., MAQUART, F.X., WEGROWSKI, Y. and MALICKA-BLASZKIEWICZ, M. Lumican affects actin cytoskeletal organization in human melanoma A375 cells. Life Sciences. vol. 83, pp. 651-660 (2008)

RAMP, U., MAHOTKA, C., HEIKAUS, S., SHIBATA, T., GRIMM, M.O., WILLERS, R. and GABBERT, H.E. Expression of heat shock protein 70 in renal cell carcinoma and its relation to tumor progression and prognosis. Histology and Histopathology. vol. 22, pp. 1099-1107 (2007)

REGLER, R., SICKINGER, S. and SCHWEIZER, M. Differential regulation of the two mRNA species of the rodent negative acute phase protein alpha 1-inhibitor 3. FEBS Letters. vol. 282, pp. 368-372 (1991)

ROSE, M.E., ORLANS, E. and BUTTRESS, N. Immunoglobulin classes in the hen's egg: their segregation in yolk and white. European Journal of Immunology. vol. 4, pp. 521-523 (1974)

RUAN, G.P., MA, L., HE, X.W., MENG, M.J., ZHU, Y., ZHOU, M.Q., HU, Z.M. and WANG, X.N. Efficient production, purification, and application of egg yolk antibodies against human HLA-A*0201 heavy chain and light chain (beta2m). Protein Expression and Purification. vol. 44, pp. 45-51 (2005)

RYU, J.W., KIM, H.J., LEE, Y.S., MYONG, N.H., HWANG, C.H., LEE, G.S. and YOM, H.C. The proteomics approach to find biomarkers in gastric cancer. Journal of Korean Medical Science. vol. 18, pp. 505-509 (2003)

SAITO, T., INOKUCHI, K., TAKAYAMA, S. and SUGIMURA, T. Sequential morphological changes in N-methyl-N'-nitro-N-nitrosoguanidine carcinogenesis in the glandular stomach of rats. Journal of the National Cancer Institute. vol. 44, pp. 769-783 (1970)

SANDOVAL, J.A., TURNER, K.E., HOELZ, D.J., RESCORLA, F.J., HICKEY, R.J. and MALKAS, L.H. Serum protein profiling to identify high-risk neuroblastoma: preclinical relevance of blood-based biomarkers. Journal of Surgical Research. vol. 142, pp. 268-274 (2007)

SCHADE, R., HENKLEIN, P., HLINAK, A., DE VENTE, J. and STEINBUSCH, H. Specificity of Chicken (IgY) versus Rabbit (IgG) Antibodies Raised against Cholecystokinin Octapeptide (CCK-8). ALTEX. vol. 13, pp. 80-85 (1996)

SCHADE, R., CALZADO, E.G., SARMIENTO, R., CHACANA, P.A., PORANKIEWICZ-ASPLUND, J. and TERZOLO, H.R. Chicken egg yolk antibodies (IgY-technology): a review of progress in production and use in research and human and veterinary medicine. Alternatives to Laboratory Animals. vol. 33, pp. 129-154 (2005)

SCHEIMAN, J.M. and CUTLER, A.F. Helicobacter pylori and gastric cancer. American Journal of Medicine. vol. 106, pp. 222-226 (1999)

SCHWAB, G.P., WETSCHER, G.J., KLINGER, A., KRECZY, A., OFNER, C., BERRESHEIM, U. and GADENSTATTER, M. Is there a dysplasia-carcinoma sequence in rat gastric remnant? Digestive Diseases and Sciences. vol. 42, pp. 608-615 (1997)

SEIBERT, V., EBERT, M.P. and BUSCHMANN, T. Advances in clinical cancer proteomics: SELDI-ToF-mass spectrometry and biomarker discovery. Briefings in Functional Genomics and Proteomics. vol. 4, pp. 16-26 (2005)

SINGER, B., FRAENKEL-CONRAT, H., GREENBERG, J. and MICHELSON, A.M. Reaction of nitrosoguanidine (N-methyl-N'-nitro-N-nitrosoguanidine) with tobacco mosaic virus and its RNA. Science. vol. 160, pp. 1235-1237 (1968)

SINHA, A.A., MORGAN, J.L., WOOD, N., BETRE, K., REDDY, A., WILSON, M.J. and RAMANANI, D.M. Heterogeneity of cathepsin B and stefin A expression in Gleason pattern 3+3 (score 6) prostate cancer needle biopsies. Anticancer Research. vol. 27, pp. 1407-1413 (2007)

SONG, C.S., YU, J.H., BAI, D.H., HESTER, P.Y. and KIM, K.H. Antibodies to the alpha-subunit of insulin receptor from eggs of immunized hens. Journal of Immunology. vol. 135, pp. 3354-3359 (1985)

SUGIMURA, T. and FUJIMURA, S. Tumour production in glandular stomach of rat by N-methyl-N'-nitro-N-nitrosoguanidine. Nature. vol. 216, pp. 943-944 (1967)

SUGIYAMA, K., IZUMI, S., TOMINO, S. and NAGASE, S. Tumor-associated expression of a serum protein, termed aX protein (alpha 1-inhibitor III), and its mRNA in rat liver. Japanese Journal of Cancer Research. vol. 80, pp. 759-764 (1989)

SUZUKI, H., NAGATA, H., SHIMADA, Y. and KONNO, A. Decrease in gamma-actin expression, disruption of actin microfilaments and alterations in cell adhesion systems associated with acquisition of metastatic capacity in human salivary gland adenocarcinoma cell clones. International Journal of Oncology. vol. 12, pp. 1079-1084 (1998)

TAHARA, E., HAIZUKA, S. and TOKUOKA, S. An ultrastructural study on precancerous and cancerous lesions of the glandular stomach in rats induced by N-methyl-N'-nitro-N-nitrosoguanidine. Acta Pathologica Japonica. vol. 26, pp. 1-15 (1976)

TAKAHASHI, M., KOKUBO, T., FURUKAWA, F., KUROKAWA, Y., TATEMATSU, M. and HAYASHI, Y. Effect of high salt diet on rat gastric carcinogenesis induced by N-methyl-N'-nitro-N-nitrosoguanidine. Gann. vol. 74, pp. 28-34 (1983)

TAKAHASHI, M., KOKUBO, T., FURUKAWA, F., KUROKAWA, Y. and HAYASHI, Y. Effects of sodium chloride, saccharin, phenobarbital and aspirin on gastric carcinogenesis in rats after initiation with N-methyl-N'-nitro-N-nitrosoguanidine. Gann. vol. 75, pp. 494-501 (1984)

TAKIKAWA, M., AKIYAMA, Y., MARUYAMA, K., SUZUKI, A., LIU, F., TAI, S., OHSHITA, C., KAWAGUCHI, Y., BANDOU, E., YONEMURA, Y. and YAMAGUCHI, K. Proteomic analysis of a highly metastatic gastric cancer cell line using two-dimensional differential gel electrophoresis. Oncology Reports. vol. 16, pp. 705-711 (2006)

TAN, Y.K. and FIELDING, J.W. Early diagnosis of early gastric cancer. European Journal of Gastroenterology and Hepatology. vol. 18, pp. 821-829 (2006)

TINI, M., JEWELL, U.R., CAMENISCH, G., CHILOV, D. and GASSMANN, M. Generation and application of chicken egg-yolk antibodies. Comparative Biochemistry and Physiology. Part A, Molecular and Integrative Physiology. vol. 131, pp. 569-574 (2002)

TONGE, R., SHAW, J., MIDDLETON, B., ROWLINSON, R., RAYNER, S., YOUNG, J., POGNAN, F., HAWKINS, E., CURRIE, I. and DAVISON, M. Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. Proteomics. vol. 1, pp. 377-396 (2001)

TSUKAMOTO, T., MIZOSHITA, T. and TATEMATSU, M. Animal models of stomach carcinogenesis. Toxicologic Pathology. vol. 35, pp. 636-648 (2007)

TURASHVILI, G., BOUCHAL, J., BURKADZE, G. and KOLAR, Z. Differentiation of tumours of ductal and lobular origin: II. Genomics of invasive ductal and lobular breast carcinomas. Biomedical Papers of the Medical Faculty of the University Palacky, Olomouc, Czechoslovakia. vol. 149, pp. 63-68 (2005)

TURCK, N., RICHERT, S., GENDRY, P., STUTZMANN, J., KEDINGER, M., LEIZE, E., SIMON-ASSMANN, P., VAN DORSSELAER, A. and LAUNAY, J.F. Proteomic analysis of nuclear proteins from proliferative and differentiated human colonic intestinal epithelial cells. Proteomics. vol. 4, pp. 93-105 (2004)

VILLAR-GAREA, A., GRIESE, M. and IMHOF, A. Biomarker discovery from body fluids using mass spectrometry. Journal of Chromatography B Analytical Technologies in the Biomedical and Life Sciences. vol. 849, pp. 105-114 (2007)

WHITEAKER, J.R., ZHANG, H., ENG, J.K., FANG, R., PIENING, B.D., FENG, L.C., LORENTZEN, T.D., SCHOENHERR, R.M., KEANE, J.F., HOLZMAN, T., FITZGIBBON, M., LIN, C., COOKE, K., LIU, T., CAMP, D.G., ANDERSON, L., WATTS, J., SMITH, R.D., MCINTOSH, M.W. and PAULOVICH, A.G. Head-to-head comparison of serum fractionation techniques. Journal of Proteome Research. vol. 6, pp. 828-836 (2007)

WIESTLER, O., VON DEIMLING, A., KOBORI, O. and KLEIHUES, P. Location of N-methyl-N'-nitro-N-nitrosoguanidine-induced gastrointestinal tumors correlates with thiol distribution. Carcinogenesis. vol. 4, pp. 879-883 (1983)

WITKIN, S.S. Heat shock protein expression and immunity: relevance to gynecologic oncology. European Journal of Gynaecological Oncology. vol. 22, pp. 249-256 (2001)

WOOLLEY, J.A. and LANDON, J. Comparison of antibody production to human interleukin-6 (IL-6) by sheep and chickens. Journal of Immunological Methods. vol. 178, pp. 253-265 (1995)

YAMASHITA, S., WAKAZONO, K., SUGIMURA, T. and USHIJIMA, T. Profiling and selection of genes differentially expressed in the pylorus of rat strains with different proliferative responses and stomach cancer susceptibility. Carcinogenesis. vol. 23, pp. 923-928 (2002)

YAN, L., TONACK, S., SMITH, R., DODD, S., JENKINS, R.E., KITTERINGHAM, N., GREENHALF, W., GHANEH, P., NEOPTOLEMOS, J.P. and COSTELLO, E. Confounding Effect of Obstructive Jaundice in the Interpretation of Proteomic Plasma Profiling Data for Pancreatic Cancer. Journal of Proteome Research. vol. 8, no. 1, pp. 142-148 (2008)

YOSHIHARA, T., KADOTA, Y., YOSHIMURA, Y., TATANO, Y., TAKEUCHI, N., OKITSU, H., UMEMOTO, A., YAMAUCHI, T. and ITOH, K. Proteomic alteration in gastic adenocarcinomas from Japanese patients. Molecular Cancer. vol. 5, pp. 75 (2006)

YOU, W.C., BLOT, W.J., LI, J.Y., CHANG, Y.S., JIN, M.L., KNELLER, R., ZHANG, L., HAN, Z.X., ZENG, X.R. and LIU, W.D. Precancerous gastric lesions in a population at high risk of stomach cancer. Cancer Research. vol. 53, pp. 1317-1321 (1993)

YU, K.H., RUSTGI, A.K. and BLAIR, I.A. Characterization of proteins in human pancreatic cancer serum using differential gel electrophoresis and tandem mass spectrometry. Journal of Proteome Research. vol. 4, pp. 1742-1751 (2005)

YU, L.R., ZENG, R., SHAO, X.X., WANG, N., XU, Y.H. and XIA, Q.C. Identification of differentially expressed proteins between human hepatoma and normal liver cell lines by two-dimensional electrophoresis and liquid chromatography-ion trap mass spectrometry. Electrophoresis. vol. 21, pp. 3058-3068 (2000)

YUASA, H., HIRANO, K., KODAMA, H., NAKANISHI, H., IMAI, T., TSUDA, H., IMAIDA, K. and TATEMATSU, M. Immunohistochemical demonstration of intestinal-type alkaline phosphatase in stomach tumors induced by N-methyl-N'-nitro-N-nitrosoguanidine in rats. Japanese Journal of Cancer Research. vol. 85, pp. 897-903 (1994)

ZECH, V.F., DLASKA, M., TZANKOV, A. and HILBE, W. Prognostic and diagnostic relevance of hnRNP A2/B1, hnRNP B1 and S100 A2 in non-small cell lung cancer. Cancer Detection and Prevention. vol. 30, pp. 395-402 (2006)

ZHANG, J., KANG, B., TAN, X., BAI, Z., LIANG, Y., XING, R., SHAO, J., XU, N., WANG, R., LIU, S. and LU, Y. Comparative analysis of the protein profiles from primary gastric tumors and their adjacent regions: MAWBP could be a new protein candidate involved in gastric cancer. Journal of Proteome Research. vol. 6, pp. 4423-4432 (2007)

ZHANG, Y., HU, M.Y., CHEN, B.H., WANG, Z.J., WU, W.Z., ZHOU, K. and LIU, K.D. The effects of ezrin on the proliferation and the invasiveness of cells of different hepatocellular carcinoma cell lines. Zhonghua Gan Zang Bing Za Zhi. vol. 14, pp. 489-494 (2006)

REVIEW

Discovery of diagnostic serum biomarkers of gastric cancer using proteomics

Katie Wing-kei Lam[1] *and Samuel Chun-lap Lo*[1, 2]

[1] The Proteomic Task Force, Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hong Kong, China
[2] State Key Laboratory of Chinese Medicine and Molecular Pharmacology, Shenzhen, China

Gastric cancer has significant morbidity and mortality worldwide and locally. Good prognosis relies on an early diagnosis. However, this remains a challenge due to the lack of specific and sensitive serum biomarkers for early detection. Hence, there is a constant search for these biomarkers for screening purposes. Proteomic profiling enables a new approach to the discovery of biomarkers in disease. This review presents recent attempts in search of gastric cancer serum biomarker using proteomics. Different methodologies and different types of samples were employed by different groups of researchers. Major difficulties were encountered in the discovery processes, including interference from abundant proteins and continuous changing serum proteomes from different individuals.

Received: July 14, 2007
Revised: October 31, 2007
Accepted: November 1, 2007

Keywords:
Biomarker / Gastric cancer

1 Introduction

Gastric cancer is the fourth most common cancer and the second most cause of cancer-related death worldwide. Gastric cancer accounts for nearly 1000 000 new cases and over 850 000 deaths annually [1]. In Hong Kong, stomach cancer ranks as the sixth most common cancer (Figs. 1 and 2). According to the official statistics in the year 2003, there were 1005 new stomach cancer cases and 680 deaths in that year. Among them, there were 635 new cases in men and 416 deaths in men, and 370 new cases in women and 264 deaths in women. The mortality: incidence ratio of stomach cancer of men and women were 0.66 and 0.71, respectively. This ratio is higher than the overall mortality: incidence ratio of all cancers, which were 0.61 and 0.45 for men and women, respectively. Therefore, in the context of Hong Kong, these figures show patients with stomach cancer have a poor prognosis and thus a higher chance to die from it as compared to the other kinds of cancers [2].

In the United States, it is estimated that there were 22 280 new cases of stomach cancer and 11 430 stomach cancer related deaths in 2006 (American Cancer Society. Cancer Facts and Figures 2006. Atlanta, Ga: American Cancer Society 2006. Available online (http://www.cancer.org/docroot/PRO/content/PRO_1_1_Cancer_Statistics_2006_presentation.asp. Last accessed 3rd July 2007)). It ranks 14th in incidence among major types of cancer malignancies. According to the American Cancer Society, the overall survival rate in these patients at 5 years ranges from almost no survival for patients with disseminated disease to almost 50% survival for patients with localized distal gastric cancers confined to respectable regional diseases. However, even with the apparent localized disease, the 5 years survival

Correspondence: Professor Samuel Chun-lap Lo, The Proteomic Task Force, The Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR, China
E-mail: bcsamlo@inet.polyu.edu.hk
Fax: +852-23649932

Abbreviations: AMP-18, 18 kDa antrum mucosal protein; **CFI,** complement factor I; **FOV,** foveolin; **HMG-1,** high-mobility group-1; **SAA,** serum amyloid A

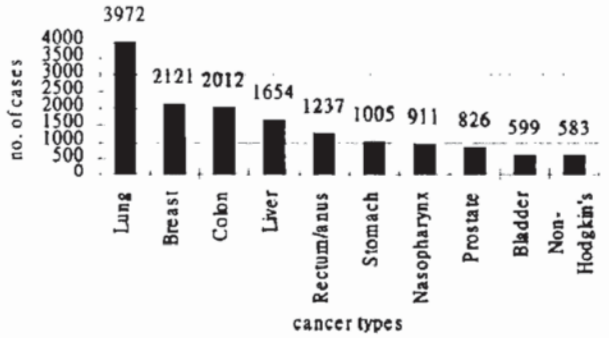


Figure 1. Leading cancer cases in Hong Kong 2003.

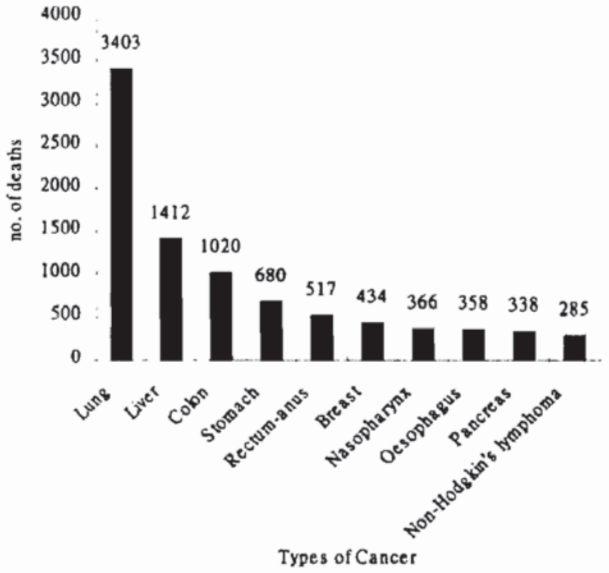


Figure 2. The ten major causes of cancer deaths in Hong Kong 2003.

rate of patients with proximal gastric cancer is only 10–15%. Although the treatment of patients with disseminated gastric cancer may result in palliation of symptoms and some prolongation of survival, long remissions are uncommon (American Cancer Society. Cancer Facts and Figures 2006. Atlanta, GA: American Cancer Society 2006. Available online (http://www.cancer.org/docroot/PRO/content/PRO_1_1_-Cancer_Statistics_2006_presentation.asp. Last accessed 3rd July 2007)).

Gastric carcinogenesis is a continuous process leading from nonatrophic gastritis to glandular atrophy, to metaplasia and dysplasia, and finally to adenocarcinoma [3]. There are a number of factors that are acknowledged to increase the risk of having gastric cancer [4–6]: helicobacter pylori gastric infection; advanced age; male gender; diet low in fruits and vegetables; diet high in salted, smoked, or preserved foods; chronic atrophic gastritis; intestinal metaplasia; pernicious anemia; gastric adenomatous polyps; family history of gastric cancer; cigarette smoking; Menetrier's disease (giant hypertrophic gastritis); and familial adenomatous polyposis. Gastric cancer is often asymptomatic or causes only nonspecific symptoms in its early stages. By the time when more severe symptoms occur, the cancer has usually metastasized to other parts of the body, which counters for its poor prognosis. Early detection and accurate preoperative staging of an early gastric cancer offer the best prognosis and are essential for planning optimal therapy. However, there is currently no specific and sensitive biomarker for its early detection. Thus, there is a need for finding specific biomarkers for diagnostic screening purposes.

Nowadays, endoscopic evaluation is frequently used and is the golden standard for diagnosis of gastrointestinal neoplasm. However, endoscopy is an invasive procedure. The cost effectiveness of this aggressive approach remains questionable. Further, this invasive method is time-consuming and is not suitable for large-scale screening purposes. These catalyze the need for finding suitable gastric cancer biomarkers and preferably in blood samples.

Studies using genomics approach to screen for cancer-related biomarkers have been launched for decades. However, there are still questions about the correlation between the expression levels of mRNA and the corresponding changes in expression levels of proteins expressed. One gene may express multiple proteins, with multiple biological functions. The proteins expressed from the genes may have different isoforms and may undergo a variety of PTMs which may be important in the carcinogenic processes. That is why only a handful of genes were approved by FDA for cancer diagnosis purposes. In USA and Hong Kong, there is no gene that had been approved for diagnosis of stomach cancer. Currently in our hospital systems in Hong Kong, there is no specific biomarker available for large-scale screening of gastric cancer either. The obvious alternative is the proteome approach. Being the final form of gene products, proteins are most directly related to biological functions, and more responsive to physiological and disease states.

2 Proteomics approach in finding disease biomarkers

The most frequently used gastric tumor markers are carcinoembryonic antigen (CEA), CA 19-9, and α-fetoprotein antigen (AFP). However, they are not sensitive biomarkers as only a modest proportion of patients with gastric cancer have elevated levels of these proteins. As elaborated earlier, there is a strong need to find specific biomarkers for gastric cancer for large-scale screening purposes. (The proteomic approach in finding biomarkers have some success in cancers of the bladder [7], prostate [7, 8], ovary [9], breast [10], and liver malignancies [11]). However, it is slow in the gastric cancer front.

3 Perspectives of proteomic approach in the discovery of serum biomarkers of gastric cancer

Serum, being the best protein sample representative of the whole body condition, is a favorite sample for investigation. It is easily accessible with low degree of invasiveness. The possible presence of serum biomarkers rest on the hypothesis that blood/plasma/serum contains tens of thousands of different proteins, along with their cleaved or modified forms, is a reflection of ongoing physiological and pathological events in the body. Specifically, when blood and/or lymph perfuse the carcinogenic tissues, it will carry "neoplasm-related" proteins and protein fragments, passively or actively, into the systemic circulation. The complex chemistry of the tumor-host microenvironment and/or reactions of the body to the neoplasm should generate unique signatures in the blood microenvironment. Therefore, discovery of these specific patterns – specific serum biomarkers, can be used to screen patients with early gastric cancers.

Some researchers compared serum proteomes from patients having cancers with those from controls. Any differentially expressed proteins will be taken as possible candidates of gastric cancer serum biomarkers. Some other researchers compared proteomes from gastric tumor tissues with those from matching samples either taken from the same patients or with sex- and age-matched controls. Quantities of differentially expressed proteins identified in the gastric tumor samples will then be measured in the patients' sera. It is usually hoped that these proteins will also be increased in blood/serum. Generally, in this comparative proteomic approach, it is hoped that the proteins found are either secretory proteins from the tumor so that they will go out into blood, or the proteins are specific surface proteins in the tumor which will leak out into blood. However, most of these studies were retrospective in nature and the patient samples were from those who have already been diagnosed of having gastric cancers. Therefore, the diagnostic values of these serum biomarkers found had to be validated.

4 Serum sample analysis

Because of historical reasons, the classical approach for analyzing serum proteomes is 2-DE followed by MALDI-TOF MS. Ebert *et al.* [12] used this combination to screen for potential biomarkers for gastric cancer. They screened gastrectomy specimens obtained from ten consecutive patients with gastric cancer. After enriching an epithelial cell fraction of these samples, these cells were resolved by 2-DE. They compared the cancer-proteome to that obtained from matched normal nonmalignant mucosal preparation. One hundred ninety-one differentially expressed protein spots were found by 2-DE and identified by MALDI-TOF MS. Among the 191 differentially expressed protein spots that were identified, cathepsin B was found to be overexpressed in 60% of cancer patients (The other proteins that were listed in the publication are shown in Table 1). Western blotting confirmed that the active form of cathepsin B was overexpressed, while immunohistochemistry showed strong cytoplasmic staining in cancer tissues of 98% patients. The serum level of cathepsin B was then determined. Among 149 patients tested, 148 patients had cathepsin B in their serum. Further, the serum levels of cathepsin B were significantly higher in gastric cancer patients (mean concentration 129.41 ± 79.55 pmol/L) compared with control patients without evidence of a malignant disease (mean 56.92 ± 67.4 pmol/L; $p = 0.0026$). Besides, there was also a significant association with the size of the primary tumor and the presence of distant metastases. Furthermore, high cathepsin B serum levels were also associated with a poor survival in patients undergoing cancer surgery [8]. However, as this study is retrospective in nature, upregulated proteins such as cathepsin B are not early diagnostic biomarkers. They are rather prognosis biomarkers. Further, cathepsin B is also overexpressed in other cancer patients, *e.g.*, with prostate cancer [13], colon cancer [14], breast cancer [15], *etc.* Its potential as a specific biomarker for gastric cancer is yet to be confirmed.

Because of limitations over availability of human samples (*e.g.*, date and time of collection, drugs administered, heterogeneity due to difference in diet, race, age, *etc.*), some researchers resorted to use animal models. Juan *et al.* [16] used the cancer-cell-nude mice injection model. They screened sera samples of nude mice that had received xenotransplants of human gastric cancer cell line SC-M1. One month after inoculation of tumor cells, plasma was collected for proteomic analysis using 2-DE and MALDI-TOF MS. The plasma proteomes were compared with those obtained using plasma from the untreated mice. Various acute phase proteins such as haptoglobin and serum amyloid A (SAA) were found to be differentially upregulated in tumor-bearing mice. Furthermore, by using immunohistochemistry, SAA and haptoglobin were found to originate from the mouse hosts and not from the human cancer cell line donors. Upregulation of SAA was further confirmed in the serum samples obtained from patients with gastric cancer. One of the concerns of this study, however, is the fact that injection of PBS into these nude mice was taken as controls to those nude mice receiving the human cancer cells. Instead of using PBS, injection of some noncancerous cell lines should have been used as control. SAA is a known acute phase protein that is related to inflammation. Whether SAA production in the SC-M1-injected nude mice is due to rudiment activities of their much impaired immune system is currently unknown. Further, there was no absorption procedure to remove abundant serum proteins such as albumin. Therefore, major abundant proteins including SAA, were seen overexpressed in the 2-DE gels of these patients. Despite of the above concerns, in a follow-up study by the same group, Chan *et al.* [17] attempted to reiterate that SAA level is really increased in gastric cancer patients by measuring serum

Table 1. Proteins found to have an altered expression level in patients with gastric cancers

Reference	Protein	Expression[a)]
Ebert *et al.* [12]	Cathepsin B	↑
	Myosin light chain alkali, nonmuscle isoform/myosin light chain alkali, smooth-muscle	↑
	Proteasome activator hPA28 β-subunit	↑
	Enhancer of rudimentary homolog	↑
	PDZ and LIM domain protein 1	↑
	Transitional ER ATPase	↑
	Gastricsin precursor	↓
	CA11 protein	↓
	Peroxiredoxin 2	↓
	Cathepsin E precursor	↓
	BC007716) unknown/(AF058954) GTPspecific succinyl-CoA synthetase β-subunit	↓
	ATP synthase D chain	↓
	Complement component1, Q subcomponent binding protein	↓
	Dihydrolipoamide dehydrogenase, mitochondrial precursor	↓
	Electron transfer flavoprotein-ubiquinone oxidoreductase	↓
	Fumarate hydratase, mitochondrial precursor	↓
	NADH-ubiquinone oxidoreductase 42 kDa subunit	↓
	Succinate dehydrogenase (ubiquinone) flavoprotein	↓
	150 kDa oxygen-regulated protein precursor	↓
Juan *et al.* [16]	Haptoglobin	↑
	SAA	↑
Chan *et al.* [17]	SAA	↑
Liu *et al.* [19]	Complement C4-B precursor	↑
	CFI	↓
	Haptoglobin	↓
Ryu *et al.* [20]	NPS3	↑
	Transgelin	↑
	Prohibitin	↑
	HSP27 and variant	↑
	Protein disulfide isomerase A3	↑
	Glucose-regulated protein	↑
	Apolipoprotein A-1	↓
	P20	↓
	Nucleoside diphophate isomerase A	↓
	α-1-Antitrypsin	↓
	Desmin	↓
	Serum albumin	↓
	Serotransferrin	↓
He *et al.* [21]	Cytokeratin 8	↑
	Tropomyosin isoform	↑
	Enolase 1	↑
	Trisoephosphate isomerase	↑
	Phosphoglycerate mutase 1	↑
	Pyruvate kinase	↑
	Chaperonin containing TCP1	↑
	HSP60	↑
	Heat shock cognate 71 protein	↑
	Protein disulfide isomerase	↑
	Translation elongation factor EF-Tu	↑
	Prohibitin	↑
	Cytokeratin 20	↓
	α-1-Antitrypsin	↓
	Apolipoprotein A-1	↓
	GMP reductase 2	↓

Table 1. Continued

Reference	Protein	Expression[a)]
	Creatine kinase B	↓
	Selenium binding protein 1	↓
	Carbonic anhydrase I	↓
	Carbonic anhydrase II	↓
	AMP-18 (Gastrokine 1)	↓
Oien *et al.* [22]	Gastrokine 1	↓
Yoshihara *et al.* [23]	Manganese superoxide dismutase	↑
	Nonhistone chromosomal protein HMG-1	↑
	Carbonic anhydrase I	↓
	Carbonic anhydrase II	↓
	Gastrokine 1	↓
	Aspartate aminotransferase 2 precursor	↓
	GST	↓
Chen *et al.* [24]	Catecholamine-*O*-methyltransferase	↑
	Protein disulfide isomerase	↑
	α-1-Antitrypsin precursor	↑
	Heat shock 27 kDa protein	↑
	Haptoglobin precursor	↑
	Desmin	↑
	Albumin	↑
	3-mercaptopyruvate sulfurtransferase	↓
	Calreticulin	↓
	ATP synthase β-subunit	↓
Nishigaki *et al.* [32]	Annexin 5	↑
	Bisphosphate 3-nucleotidase	↑
	Microtubule-associated protein, RP/EB family, member 1	↑
	GST	↑
	Heat shock 27 kDa protein 1	↑
	MADS box transcription enhancer factor 2, polypeptide C	↑
	Nicotinamide *N*-methyltransferase	↑
	UDP-glucuronosyltransferase	↑
	ApoA-I binding protein	↓
	Serum amyloid P component precursor	↓
	CA11 protein	↓
	ATP-dependent proteolytic subunit *Escherichia coli* homolog	↓
	Cytochrome *c* oxidase subunit Va	↓
	Tumor RMS cell line RD specific product	↓
	Enoyl Coenzyme A hydratase 1, mitochondrial and peroxisomal	↓
	Serine proteinase inhibitor, clade B (ovalbumin), member 1	↓
	Ferritin, heavy chain	↓
	Cytosolic inorganic pyrophosphatase	↓
	Mitotic checkpoint protein isoform MAD1a	↓
	Mannose-6-phosphate isomerase	↓
Melle *et al.* [33]	Pepsinogen C	↓
Takikawa *et al.* [35]	Interferon-induced Mx protein	↑
	Glycyl-tRNA synthetase	↑
	Flavoprotein unbunit of complex II	↑
	Ts11 cell cycle protein	↑
	Tyrosyl-tRNAsynthetase	↑
	Keratin 5	↑
	Transfer RNA-Trp synthetase	↑
	Adenylate kinase 2	↑
	Pyruvate kinase	↓
	Cytokeratin 8	↓

Table 1. Continued

Reference	Protein	Expression[a)]
	Dihydrodiol dehydrogenase isoform DD1	↓
	Annexin I	↓
	Annexin A2, isoform 2	↓
	Carbonic anhydrase II	↓
Lee *et al.* [39]	α-1-Antitrypsin	↑

a) ↑ overexpression, ↓ underexpression.

levels of SAA in these patients using an ELISA assay. Two groups of patients: (a) ninety-six gastric cancer patients before and after curative gastrectomy and (b) 32 patients with gastric ulcers and 52 healthy subjects as controls, were tested. These authors found that the mean SAA concentration was higher in gastric cancer patients (88.54 ± 50.44 mg/L) than in healthy subjects (3.36 ± 2.29 mg/L) and gastric ulcer patients (10.48 ± 8.97 mg/L) ($p<0.05$). The SAA concentration was associated with tumor stage ($p = 0.0244$) and location ($p = 0.0016$) but not with Lauren's histological type ($p = 0.839$). With 2-DE and subsequent MALDI-TOF MS, they also found that SAA is overexpressed in serum samples. During patient follow-up studies, the mean SAA concentration was found to increase significantly in 24 patients with tumor recurrence ($p<0.05$) but did not change in 77 patients without recurrence. In the survival analysis, patients with SAA levels >97 mg/L had a nearly four-fold increase in risk of death. Validation of the upregulation of SAA was done by immunohistochemical staining which showed visible immunoreactivity on resected tumor specimens from patients with gastric cancer. They demonstrated that SAA was useful in predicting survival of patients with gastric cancer, and was a valuable tool for postoperative follow-up [17]. They also recommended that as the sensitivity of SAA was higher than that reported using conventional tumor markers for gastric cancer, SAA might be a useful new biomarker for detecting gastric cancer [17].

Another group, Liang *et al.* [18] implemented the use of SELDI on serum samples aiming to discriminate between gastric cancer patients and healthy controls with peptide mass fingerprints. They found three peak masses that were differentially expressed in patients with gastric cancer. The intensities of peaks at 5084 and 5910 Da were significantly higher in sera from patients with gastric cancer. Conversely, the peak at 8691 Da in sera from patients with gastric cancer was markedly reduced in intensity when compared with that of controls. Although the findings are interesting, application value of these data is limited unless identities of these peak masses can be found.

Back to the criticism that only abundant proteins in the plasma proteome were found, Liu *et al.* [19] employed an immunoaffinity removal method to clear most of the highly abundant proteins from serum samples from gastric cancer patients. Subsequently, three proteins including complement C_4-B precursor, complement factor I (CFI) precursor, and haptoglobin precursor were found to be differentially expressed in sera from gastric cancer patients when compared with that of healthy controls. Further validation by Western blotting showed that CFI precursor was underexpressed in gastric cancer sera when compared to normal sera. The diagnostic value of these protein biomarkers have to be validated from large-scale longitudinal studies.

5 Tissue sample and fluid analysis

Human gastric tissues are also widely used in the search for differential protein expression. As mentioned earlier, it is generally envisioned that some gastric cancer specific proteins that were found in tissue samples may also leak into blood. These proteins or fragments of these proteins can then be used as serum biomarkers for gastric cancer. Without any abundant proteins removal methodology in place, Ryu *et al.* [20] used 2-DE to analyze human cancer tissues and 140 protein spots were identified. A comparison of stomach cancer tissue with normal tissue from the same patients showed that seven upregulated proteins, namely: NPS3, transgelin, prohibitin, heat shock protein (HSP) 27 and variant, protein disulfide isomerase A3, an unnamed protein product, and glucose-regulated protein were overexpressed. In addition, another seven proteins were underexpressed, namely: apolipoprotein A-1, p20, nucleoside diphophate isomerase A, α-1-antitrypsin, desmin, serum albumin, and serotransferrin. It is uncertain why levels of serum albumin detected were also decreased. Further, the authors neither discuss the possible reasons why specific proteins are up- or downregulated nor suggesting the use of any of the differentially expressed proteins as specific biomarkers. However, it should be stressed among the proteins found to be differentially expressed, increased expression of HSP27 is also found by others (see below).

Another group, He *et al.* [21] also employed 2-DE to analyze primary gastric cancer tissues. Compared to nontumor tissues, multiple protein alterations were observed in tumor tissues. These included variations in the expression of cyto-

skeleton proteins, including an increase in cytokeratin 8, and a tropomyosin isoform as well as a decrease in cytokeratin 20. Co-upregulations of HSPs and glycolytic enzymes were also observed in tumor tissues. According to the authors, these upregulation indicate self-protective efforts of cells and the growing energy requirement during malignant transformation. Downregulations were also seen with proteins involved in cell proliferation and differentiation, such as GMP reductase 2, creatine kinase B, and proteins bearing potential tumor suppressor activities, including prohibitin and selenium binding protein 1. More interestingly, a human stomach-specific protein, 18 kDa antrum mucosal protein (AMP-18 with NCBI Accession number = 26000208, AAN75447 and experimental M_r/pI = 20 kDa/6.0), was found to be dramatically underexpressed in cancer tissues. According to the authors, the decreased AMP-18 production implicates a possible special pathological role for this protein in gastric carcinogenesis [21]. In the same year, Oien *et al.* [22] reported that this AMP-18 was also downregulated in their cohort of gastric cancer patients. According to these authors, the protein was renamed as gastrokine 1 by the Human Gene Nomenclature Committee in 2003. The authors found that gastrokine 1 was highly expressed in normal stomach, where it was located in the superficial/foveolar gastric epithelium, but was absent from gastric carcinomas. Further, gastrokine 1 was found only in epithelia showing gastric metaplasia, *e.g.*, Barrett's oesophagus, the ulcer-associated cell lineage and ovarian mucinous neoplasms. Although there is no solid proof, these authors speculated that gastrokine 1 (also called CA11, AMP-18, or foveolin (FOV)) has a role in mucosal protection. Further research is required to see if gastrokine 1 could be used as a specific cancer biomarker.

Yoshihara *et al.* [23] also tried to minimize individual variations by studying protein alternations between tumor and nontumor tissues from the same gastric cancer patients. It was found that a nonhistone chromosomal protein called high-mobility group-1 protein (HMG-1) was elevated in tumor tissue. HMG-1 is a DNA-binding protein that regulates the transcription of various genes and the structural stabilization of chromosomes. It was reported to be associated with carcinogenesis and metastasis in colorectal and breast cancer. On the other hand, these authors reported a remarkable decrease in the level of FOV precursor (FOV, Accession number = 38488935; AAR21211 with M_r/pI = 20 kDa/5.7). This is the same protein that was called gastrokine 1 in the report by Oien *et al.* [22]. RT-PCR was performed and the results revealed significant downregulation of FOV mRNA expression in tumor tissues [23]. Hence, results of Yoshihara *et al.* reinforced observations made by He *et al.* [21] as well as Oien *et al.* [22]. Therefore, a decrease in gastrokine 1/FOV may be used as a specific phenomenon for the occurrence of gastric cancer. However, how is gastrokine 1 compared to pepsinogen C (see below) in terms of being better gastric cancer specific biomarker is currently unknown.

Chen *et al.* [24] used a well described *N*-methyl-*N'*-Nitro-*N*-Nitrosoguanidine (MNNG) intoxicated Wistar rat gastric cancer model to investigate the protein profiling of carcinogenesis and metastasis in the induced gastric cancer tissues. After 2-DE and MALDI-TOF MS analysis of the cancer tissues and matching normal tissues, they found that 11 proteins were upregulated and two proteins were downregulated (Table 1). In particular, the overexpression of HSP27 in gastric cancer was confirmed by immunohistochemical analysis of human gastric cancer specimens. HSP27 is a known molecular chaperone with an ability to interact with a large number of proteins [24, 25]. It was said that its production is accelerated in nonphysiological conditions and it will aid cell survival. Thus, it was speculated that overproduction of HSPs could protect malignantly transformed cells from apoptotic cell death and fosters resistance to chemotherapeutic agents and irradiation [26, 27]. Its expression is associated with poor prognosis [28]. However, higher expression of HSP27 protein had been previously reported in liver, breast, colon, and gastric cancer [20, 28–31]. Therefore, HSP27 may not be a specific biomarker for the occurrence of gastric cancer.

After proteomic analysis of 14 paired samples of gastric carcinomas and corresponding noncancerous gastric mucosae, Nishigaki *et al.* [32] found nine proteins with an increased expression and thirteen with a decreased expression in gastric cancer. The two most notable groups included proteins involved in mitotic checkpoint (MAD1L1 and EB1) and mitochondrial functions (CLPP, COX5A, and ECH1). These results suggested that there were links between dysfunctions in these processes and gastric carcinogenesis. More importantly, immunohistochemical analyses confirmed that the levels of expression of MAD1L1 and CYR61 were decreased in gastric carcinoma tissues while HSP27 was increased in gastric carcinoma tissues. On the other hand, Melle *et al.* [33] applied laser capture microdissection and SELDI-TOF/MS on tissue samples. They analyzed 74 cryostat sections of central gastric tumor, tumor margin, and normal gastric epithelium. Pepsinogen C was the one peak that was found to be significantly downregulated in tumor tissue. It should be stressed that although it is generally accepted that pepsinogen A and C reflect the functional and morphological status of the gastric mucosa and serum pepsinogen test had been used for screening of the occurrence of gastric cancer; pepsinogen C level is normally constant. After an extensive review of the data in (Japan, Miki [34]) concluded that the pepsinogen test method can be used as a screening test for identifying high-risk subjects, rather than as a tool for screening for cancer itself!

Given the retrospective nature of most gastric cancer biomarkers studies discussed above, researchers had tried to improve quality of the data with samples from gastric cancer cell lines. Takikawa *et al.* [35] used 2-D-DIGE coupled with MALDI-TOF MS to identify specific proteins differentially expressed between a highly metastatic stomach cancer cell line MKN-45-P and its parental cell line MKN-45. It was

hoped that the results could be informative and be applied in large-scale validation studies. Their results revealed upregulation of eight proteins (IFN-induced Mx protein, Gly-, Tyr-, Trp-tRNA synthetase, the flavoprotein subunit of complex II, the ts11 cell cycle protein, keratin 5, and adenylate kinase) and downregulation of five proteins (pyruvate kinase, cytokeratin 8, didydrodiol dehydrogenase, annexin I, and carbonic anhydrase II). It was found that these 13 proteins were mainly involved in protein synthesis, metabolism, receptor and signal transduction, the cytoskeleton and cell cycling. On the other hand, Lee *et al.* [36] tried to identify the target antigen of a reportedly stomach cancer specific antibody - the MG7 mAb using KATO III and MKN-45 gastric carcinoma cell lines. MG7 antigen was found to be decreased in gastric cancer patients who underwent gastrectomy. This allowed MG7 detection to be used in diagnosis of the presence of gastrointestinal cancer as well as to evaluate the effectiveness and treatment outcome after cancer therapy. With the 2-DE and Western blotting, two proteins of 35 kDa were consistently detected by the MG7 antibody. Followed by MALDI-TOF MS, these MG7 immunoreactive proteins were identified as the herterogeneous nuclear ribonucleoprotein A2/B1 (hnRNP A2/B1). Four out of the five gastrectomy samples from patients with gastric cancer showed positive results in Western blotting with commercially available hnRNAP A2/B1 antibodies. However, hnRNAP upregulation is also found in other types of cancer [37, 38].

Lastly, Lee *et al.* [39] utilized gastric juice to perform 2-DE aiming to obtain disease-specific protein expression patterns. In healthy subjects, pepsin A, pepsin B, and gastric lipase were the major proteins, but were not detected in 60% of gastric cancer cases. Conversely, an extraordinary amount of α-1-Antitrypsin was observed in gastric cancer patients, while it is only detected in 5% of the healthy group [39].

In the studies discussed above, several proteins were repeatedly found to be differentially expressed in serum or tissue samples from patients with gastric cancer compared to nonneoplastic serum/tissue samples. They are apolipoprotein A-1, α-1-antitrypsin, and HSP27. However, we knew that apolipoprotein A-1 and α-1-antitrypsin are abundant proteins that present in human blood. Therefore, their roles as specific biomarkers for gastric cancer have to be verified. However, minor proteins such as HSP27 and hnRNP A2/B1 are not specifically related to gastric cancer. Taken overall, we are yet to have a specific (set of) early diagnostic protein biomarker for large-scale screening of gastric cancer.

6 Difficulties in proteomics in the discovery of biomarkers

Although researchers are working hard on biomarker discovery, finding sensitive and specific serum biomarkers for early diagnosis of gastric cancer still remains a daunting challenge. The biggest problem is the presence of the high abundant proteins which interfere with the exhibition of some low abundant proteins that may be very important. It is customary to believe that most of the clinically important biomarkers are rarely exceeding in amount. A few proteins such as albumin, immunoglobulins, haptoglobin, antitrypsin, and transferrin constitute around 90% of all protein masses present in serum samples [40]. Serum samples should therefore be treated to concentrate these relatively rare biomarkers for the discovery process. Multiple approaches for simplifying the serum proteome have been done and compared. The serum fractionation schemes include *N*-linked glycopeptide enrichment, cysteinyl-peptide enrichment, magnetic bead separation (C3, C8, and WCX), size fractionation, protein A/G depletion, and immunoaffinity column depletion of abundant serum proteins. After comparing with the results obtained from the unfractionated human serum, it was found that the immunoaffinity subtraction is the most effective means for simplifying the serum proteome while maintaining reasonable sample throughput [41].

Another difficulty for biomarker discovery will be the intrinsic variation of serum samples within individuals [16]. The proteome pattern of serum depends on a number of factors, including the genetic background, sex, age, nutritional status, lifestyle, medical treatment, bed rest, *etc.* The methods of samples collection, protocols for preparation, and storage of samples are also extremely sensitive to the proteome patterns.

Lastly, given the significant research efforts that are in place worldwide aiming to discover useful cancer biomarkers, the process is slow and most of the time it is difficult for biomarkers to translate from discovery to viable clinical screening tools. This is because the validation of a biomarker is more difficult and time-consuming than its discovery.

7 Conclusion

Conventional serum tumor markers are not suitable for the early diagnosis of gastric cancers. Modern proteomic technologies have been applied as a powerful tool for discovery of biomarkers for early diagnosis of gastric cancer. Yet daunting challenges still have to be overcome before these biomarkers can be found.

K. W.-K. L. is supported by a postgraduate studentship from the Research Committee of The Hong Kong Polytechnic University. This study is also supported by a grant from The Hong Kong Polytechnic University (account no. G-YF69).

The authors have declared no conflict of interest.

8 References

[1] Pisani, P., Parkin, D. M., Bray, F., Ferlay, J., Estimates of the worldwide mortality from 25 cancers in 1990. *Int. J. Cancer* 1999, *8*, 18–29.

[2] Hong Kong Cancer Stat 2003, Hong Kong Cancer Registry, Hospital Authority, 2006.

[3] Correa, P., A human model of gastric carcinogenesis. *Cancer Res.* 1988, *48*, 3554–3560.

[4] Kurtz, R. C., Sherlock, P., The diagnosis of gastric cancer. *Semin. Oncol.* 1985, *12*, 11–18.

[5] Scheiman, J. M., Cutler, A. F., *Helicobacter pylori* and gastric cancer. *Am. J. Med.* 1999, *106*, 222–226.

[6] Fenoglio-Preiser, C. M., Noffsinger, A. E., Belli, J., Stemmermann, G. N., Pathologic and phenotypic features of gastric cancer. *Semin. Oncol.* 1996, *23*, 292–306.

[7] Adam, B. L., Vlahou, A., Semmes, O. J., Wright, G. L., Jr., Proteomic approaches to biomarker discovery in prostate and bladder cancers. *Proteomics* 2001, *1*, 1264–1270.

[8] Petricoin, E. F., Ornstein, D. K., Paweletz, C. P., Ardekani, A. *et al.*, Serum proteomic patterns for detection of prostate cancer. *J. Natl. Cancer Inst.* 2002, *94*, 1576–1578.

[9] Petricoin, E. F., Ardekani, A. M., Hitt, B. E., Levine, P. J. *et al.*, Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 2002, *359*, 572–577.

[10] Li, J., Zhang, Z. S., Rosenzweig, J., Wang, Y. Y., Chan, D. W., Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. *Clin. Chem.* 2002, *48*, 1296–1304.

[11] Feng, J. T., Liu, Y. K., Song, H. Y., Dai, Z. *et al.*, Heat-shock protein 27: A potential biomarker for hepatocellular carcinoma identified by serum proteome analysis. *Proteomics* 2005, *5*, 4581–4588.

[12] Ebert, M. P., Krüger, S., Fogeron, M. L., Lamer, S. *et al.*, Overexpression of cathepsin B in gastric cancer identified by proteome analysis. *Proteomics* 2005, *5*, 1693–1704.

[13] Sinha, A. A., Morgan, J. L., Wood, N., Betre, K. *et al.*, Heterogeneity of cathepsin B and stefin A expression in Gleason pattern 3+3 (score 6) prostate cancer needle biopsies. *Anticancer Res.* 2007, *27*, 1407–1413.

[14] Alencar, H., Funovics, M. A., Figueiredo, J., Sawaya, H. *et al.*, Colonic adenocarcinomas: Near-infrared microcatheter imaging of smart probes for early detection–study in mice. *Radiology* 2007, *244*, 232–238.

[15] Turashvili, G., Bouchal, J., Burkadze, G., Kolár, Z., Differentiation of tumours of ductal and lobular origin: II. Genomics of invasive ductal and lobular breast carcinomas. *Biomed. Pap. Med. Fac. Univ. Palacky Olomouc Czech. Repub.* 2005, *149*, 63–68.

[16] Juan, H. F., Chen, J. H., Hsu, W. T., Huang, S. C. *et al.*, Identification of tumor-associated plasma biomarkers using proteomic techniques: From mouse to human. *Proteomics* 2004, *4*, 2766–2775.

[17] Chan, D. C., Chen, C. J., Chu, H. C., Chang, W. K. *et al.*, Evaluation of serum amyloid A as a biomarker for gastric cancer. *Ann. Surg. Oncol.* 2007, *1*, 84–93.

[18] Liang, Y., Fang, M., Li, J., Liu, C. B. *et al.*, Serum proteomic patterns for gastric lesions as revealed by SELDI mass spectrometry. *Exp. Mol. Pathol.* 2006, *2*, 176–180.

[19] Liu, W., Liu, B., Xin, L., Zhang, Y. *et al.*, Down-regulated expression of complement factor I: A potential suppressive protein for gastric cancer identified by serum proteome analysis. *Clin. Chim. Acta* 2007, *1*, 119–126.

[20] Ryu, J. W., Kim, H. J., Lee, Y. S., Myong, N. H. *et al.*, The proteomics approach to find biomarkers in gastric cancer. *J. Korean Med. Sci.* 2003, *18*, 505–509.

[21] He, Q. Y., Cheung, Y. H., Leung, S. Y., Yuen, S. T. *et al.*, Diverse proteomic alterations in gastric adenocarcinoma. *Proteomics* 2004, *4*, 3276–3287.

[22] Oien, K. A., McGregor, F., Butler, S., Ferrier, R. K. *et al.*, Gastrokine 1 is abundantly and specifically expressed in superficial gastric epithelium, down-regulated in gastric carcinoma, and shows high evolutionary conservation. *J. Pathol.* 2004, *203*, 789–797.

[23] Yoshihara, T., Kadota, Y., Yoshimura, Y., Tatano, Y. *et al.*, Proteomic alteration in gastric adenocarcinomas from Japanese patients. *Mol. Cancer* 2006, *5*, 75.

[24] Chen, J., Kähne, T., Röcken, C., Götze, T. *et al.*, Proteome analysis of gastric cancer metastasis by two-dimensional gel electrophoresis and matrix assisted laser desorption/ionisation-mass spectrometry for identification of metastasis-related proteins. *J. Proteome Res.* 2004, *5*, 1009–1016.

[25] Kato, M., Herz, F., Kato, S., Hirano, A., Expression of stress-response (heat-shock) protein 27 in human brain tumors: An immunohistochemical study. *Acta Neuropathol. (Berl).* 1992, *83*, 420–422.

[26] Concannon, C. G., Gorman, A. M., Samali, A., On the role of Hsp27 in regulating apoptosis. *Apoptosis* 2003, *8*, 61–70.

[27] Witkin, S. S., Heat shock protein expression and immunity: Relevance to gynecologic oncology. *Eur. J. Gynaecol. Oncol.* 2001, *22*, 249–256.

[28] Kapranos, N., Kominea, A., Konstantinopoulos, P. A., Savva, S. *et al.*, Expression of the 27 kDa heat shock protein (HSP27) in gastric carcinomas and adjacent normal, metaplastic, and dysplastic gastric mucosa, and its prognostic significance. *J. Cancer Res. Clin. Oncol.* 2002, *128*, 426–432.

[29] Yu, L. R., Zeng, R., Shao, X. X., Wang, N. *et al.*, Identification of differentially expressed proteins between human hepatoma and normal liver cell lines by two-dimensional electrophoresis and liquid chromatography-ion trap mass spectrometry. *Electrophoresis* 2000, *21*, 3058–3068.

[30] Garrido, C., Fromentin, A., Bonnotte, B., Favre, N. *et al.*, Heat shock protein 27 enhances the tumorigenicity of immunogenic rat colon carcinoma cell clones. *Cancer Res.* 1998, *1*, 5495–5499.

[31] Lemieux, P., Oesterreich, S., Lawrence, J. A., Steeg, P. S. *et al.*, The small heat shock protein hsp27 increases invasiveness but decreases motility of breast cancer cells. *Invasion Metastasis* 1997, *17*, 113–123.

[32] Nishigaki, R., Osaki, M., Hiratsuka, M., Toda, T. *et al.*, Proteomic identification of differentially expressed genes in human gastric carcinomas. *Proteomics* 2005, *5*, 3205–3213.

[33] Melle, C., Ernst, G., Schimmel, B., Bleul, A. *et al.*, Characterization of pepsinogen C as a potential biomarker for gastric cancer using a histo-proteomic approach. *J. Proteome Res.* 2005, *4*, 1799–1804.

[34] Miki, K., Gastric cancer screening using the serum pepsinogen test method. *Gastric Cancer* 2006, *9*, 245–253.

[35] Takikawa, M., Akiyama, Y., Maruyama, K., Suzuki, A. *et al.*, Proteomic analysis of a highly metastatic gastric cancer cell line using two-dimensional differential gel electrophoresis. *Oncol. Rep.* 2006, *16*, 705–711.

[36] Lee, C. H., Lum, J. H., Cheung, B. P., Wong, M. S. *et al.*, Identification of the heterogeneous nuclear ribonucleoprotein A2/B1 as the antigen for the gastrointestinal cancer specific monoclonal antibody MG7. *Proteomics* 2005, *5*, 1160–1166.

[37] Zech, V. F., Dlaska, M., Tzankov, A., Hilbe, W., Prognostic and diagnostic relevance of hnRNP A2/B1, hnRNP B1 and S100 A2 in nonsmall cell lung cancer. *Cancer Detect. Prev.* 2006, *30*, 395–402.

[38] Turck, N., Richert, S., Gendry, P., Stutzmann, J. *et al.*, Proteomic analysis of nuclear proteins from proliferative and differentiated human colonic intestinal epithelial cells. *Proteomics* 2004, *4*, 93–105.

[39] Lee, K., Kye, M., Jang, J. S., Lee, O. J. *et al.*, Proteomic analysis revealed a strong association of a high level of alpha-1-antitrypsin in gastric juice with gastric cancer. *Proteomics* 2004, *4*, 3343–3352.

[40] Yu, K. H., Rustgi, A. K., Blair, I. A., Characterization of proteins in human pancreatic cancer serum using differential gel electrophoresis and tandem mass spectrometry. *J. Proteome Res.* 2005, *5*, 1742–1751.

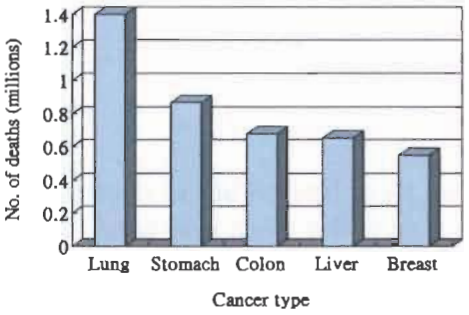
[41] Whiteaker, J. R., Zhang, H., Eng, J. K., Fang, R. *et al.*, Head-to-head comparison of serum fractionation echniques. *J. Proteome Res.* 2007, *6*, 828–836.

Protein biomarkers associated with experimentally induce gastric cancer in rats

Katie Wing-Kei Lam and Samuel Chun-Lap Lo
The Proteomic Task Force, Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hung Hom, Hong Kong
Email address: bcsamlo@inet.polyu.edu.hk

Introduction

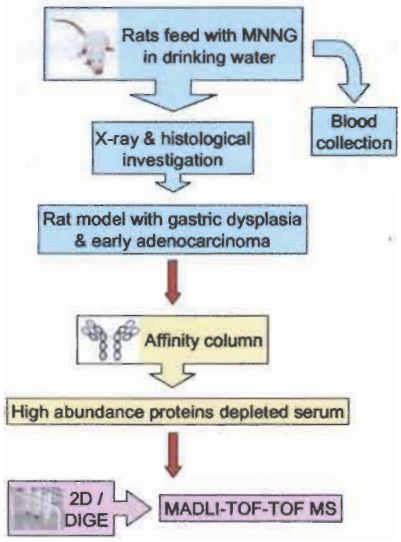
Gastric cancer has significant morbidity and mortality worldwide. According to the World Health Organization, stomach cancer is the second most cause of cancer-related death world-wide in 2007, which accounts for 866,000 deaths (1).



Good prognosis relies on early diagnosis. The most frequently used gastric tumor markers are carcinoembryonic antigen (CEA), cancer antigen 19-9 (CA 19-9), alpha-fetoprotein antigen (AFP) and human chorionic gonadotropin (hCG). However, they are not sensitive biomarkers as only a modest proportion of patients with gastric cancer have elevated levels of these proteins. Therefore, there is an urgent need for specific biomarkers indicative of early gastric cancer (2). Major hindrances of biomarker discovery are the presence of high abundance proteins and the availability of the early stage gastric cancer samples. Here we attempted to remove the high abundance proteins in serial serum samples of a well established experimental gastric cancer model in rats by a customized affinity column. Differentially expressed proteins were pinpointed by 2DE with subsequent image analysis. These proteins were than identified by MALDI-TOF-TOF mass spectrometry.

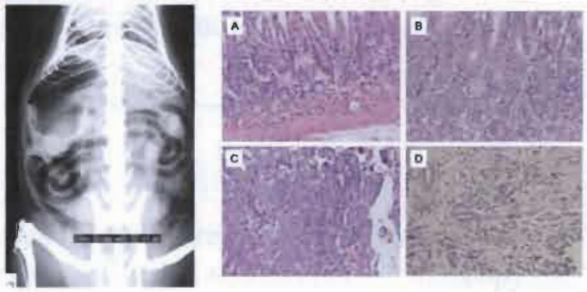
Methods

The project consisted of three stages: establishment of animal model, depletion of high abundance proteins, and identification of differential expressed proteins.

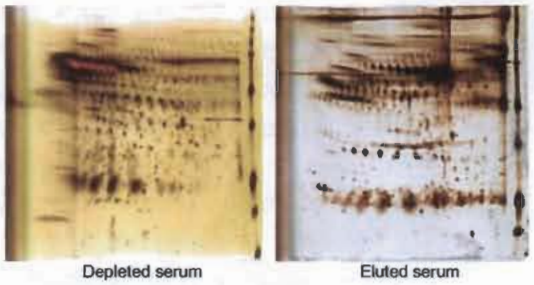


Results

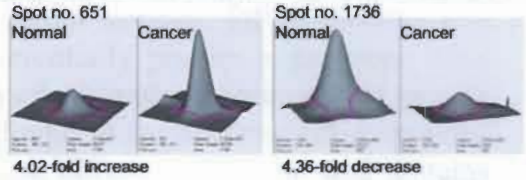
Gastric cancer had occurred in rats. The X-ray photos were taken and histology study was done.



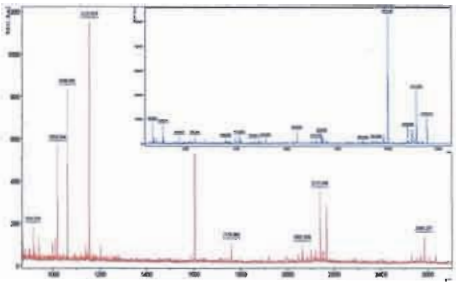
The affinity column successfully depletes 99.8% of the high abundance proteins and therefore, generates a new serum proteome for better visualization of the low abundance proteins.



From the results obtained by DIGE, several proteins were found to express differently between normal and cancer stages. Among the differential proteins, most of them were up-regulated, while a few were down-regulated.



One of the up-regulated proteins, α-1-antitrypsin, was identified by the MALDI-TOF-TOF MS.



Conclusion

Novel protein biomarkers can be used to detect gastric cancer at the early stage for screening purpose.

Acknowledgement
This work is supported by

References

(1)World Health Organization, Fact Sheet No297, July 2008
(2)K.W.K Lam, S.C.L Lo, Discovery of diagnostic serum biomarkers of gastric cancer using proteomics; Proteomics Clin. Appl. 2008, 2, 219-228

Putative protein biomarkers found in experimentally induced gastric cancer in rats

Katie W.K. Lam[1], Yau-Ming Lai[2], Phoebe S.T. Chan[2], Pak-Kwan Hui[3], Samuel C.L. Lo[1]
[1]The Proteomic Task Force, Department of Applied Biology and Chemical Technology, The Hong Kong Polytechnic University, Hung Hom, HK
[2]Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hung Hom, HK
[3]Department of Pathology, Kwong Wah Hospital, HK

Abstract

Gastric cancer is the second most common causes of cancer-related death world-wide in 2007. Good prognosis relies on early diagnosis. Clinical diagnosis is by endoscopy and biopsy, which are not suitable for large-scale screening. There is a lack of specific and sensitive serum biomarkers indicative of gastric cancer. Major hindrances of biomarker discovery are the availability of the early stage gastric cancer samples and the presence of high abundance proteins (HAPs) in the serum samples. In this project, gastric dysplasia, an early stage of gastric carcinogenesis, was induced in Wistar rats using a chemical carcinogen, MNNG. Possible occurrence of gastric lesion was examined by high resolution X-ray mammographic unit with double contrast enhancement, and confirmed by histopathological examinations on the gastric tissues. Serial rat serum samples were collected during the induction process. On another front, antibodies to normal rat serum were raised in chickens before being purified by established technologies. An affinity column was generated using these antibodies. Subsequently, serial rat serum samples of interest were loaded on to the customized affinity column in order to deplete HAP in the serum samples. DIGE labeling was performed on the depleted serum samples. Differentially protein expressions were compared using serum samples from rats with dysplasia and adenocarcinoma as well as normal control. A total of 16 proteins were found to be differentially expressed (differed at least by 5 folds). Identification of these differentially expressed proteins by MALDI-TOF mass spectrometry is currently in progress.