# Analysis of the effects of BrdU on DLKP Human Lung Cancer Cells by two-dimensional difference gel electrophoresis and mass spectrometry

A thesis submitted for the degree of M.Sc.

Dublin City University

By

Michael Henry

This research work described in this thesis was performed under the supervision of Professor Martin Clynes and Dr. Paula Meleady

National Institute for Cellular Biotechnology

2007

*I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of M.Sc. is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged with the text of my own work.*

Signed: Michael Henry Date: 27.04.07

*This thesis is dedicated to my father Anthony.*

*Always want to make you proud.*

**Acknowledgements:**

I would like to sincerely thank Professor Martin Clynes for giving me the opportunity to carry out the research work presented here.

Thanks also to those at the N.I.C.B. (N.C.T.C.C.), both past and present who helped me in so many ways. Dr. Paula Meleady for letting me join her proteomics group and starting me off down the route of protein identification by Mass Spectrometry. Dr. Andrew Dowd for all his help with the DIGE work, Dr. Scott McMillen for all the MALDI help and Maurice Burke for all the trouble shooting with the MDLC-LTQ.

I would also like to thank Dr. Liz Moran for starting me off in scientific research as an immunologist, Dr. Annemarie Larkin and Dr. John Milne for all the work on that troublesome antibody that never made it into this thesis.

A big thanks to my coffee and lunch gangs over the years and especially the current ones of Paudie, Paddy, Niall, Finbar (thanks for being my intern), Paula and Derek (man you've so changed). Also, thanks to the people behind the scene in the centre Joe, Carol and Yvonne who quietly keep it ticking over.

Friends and family are the most important part of my life. I would like to thank the lads Shea, John and Jim for the distraction from work that is PE soccer. The girls who I have the pleasure of their company but only if I cook (only joking) that is Eadaoin, Rachel and Cora. Long term and life time friends that I've made who are Róisín and Paula, better friends I could not ask for.
A very special thanks to my family. My dad whom I miss so much (I know I've made you proud), my mam who is a tower of strength to me, my brother Jim who has always looked out for me, my sister Catherine who I will make into a better chef and newest members Dom and Neil.

An' though the course may change sometimes,
Rivers always reach the sea.
(Page/Plant)

**Abbreviations**

| | | |
|---|---|---|
| ACN | - | Acetonitrile |
| ATCC | - | American Type Culture Collection |
| BrdU | - | Bromodeoxyuridine |
| CA | - | Carrier Ampholytes |
| Da | - | Dalton |
| 2-DE | - | Two-Dimensional Electrophoresis |
| EMBL | - | European Molecular Biology Laboratory |
| ESI | - | Electrospray Ionisation |
| FA | - | Formic Acid |
| HPLC | - | High-performance liquid chromatography |
| IEF | - | Isoelectric Focusing |
| IPG | - | Immobilised pH Gradient |
| MALDI | - | Matrix Assisted Laser Desorption Ionisation |
| MDLC | - | Multi-Dimensional Liquid Chromatography |
| MH+ | - | Protonated molecular ion |
| min | - | Minute |
| mL | - | Milliliter |
| $M_r$ | - | Molecular Weight |
| MS | - | Mass Spectrometry |
| m/z | - | Mass-to-charge ratio |
| NCBI | - | National Centre for Biotechnology Information (USA) |
| PAGE | - | Polyacrylamide gel electrophoresis |
| p*I* | - | Isoelectric Point |
| PMF | - | Peptide Mass Fingerprinting |

| | | |
|---|---|---|
| RPC | - | Reverse phase chromatography |
| PTM | - | Post Translational Modification |
| SD | - | Standard deviation |
| SDS | - | Sodium-dodecyl sulphate |
| TFA | - | Tri-Fluoric Acid |
| ToF | - | Time of Flight |
| v/v | - | volume/volume |
| w/v | - | weight/volume |

**Abstract**

Bromodeoxyuridine (BrdU) is a thymidine analogue that incorporates into DNA of dividing cells during the S-phase of the cell cycle. Previous work in laboratories reported that treatment with 10μM BrdU in the human lung carcinoma cell line (DLKP) resulted in increased expression of the cytoskeletal proteins Keratin 8 and 18 and the cell adhesion proteins $\alpha_2$ and $\beta_1$ integrin.

This study investigated protein expression changes in differentiating DLKP cells following exposure to 10μM BrdU. DLKP cells were grown in culture flasks and harvested after 7 days exposure to BrdU. Two-dimensional gel electrophoresis was used to investigate BrdU specific changes in the proteome of DLKP BrdU treated and control cells.

Cy3-labeled DLKP control were combined with Cy-5 labeled BrdU DLKP treated proteins and separated on the same 2-D gel along with a Cy-2 labelled mixture of both samples as an internal standard. Using DIGE technology, the statistically significant comparisons of each protein abundance was made over three biological replicates. 43 protein spots were identified as differentially regulated. Among the 43 protein spots, 25 were found to be up-regulated and 18 were found to be down-regulated.

Gel plugs of the differentially regulated proteins spots were subjected to optimised in-gel digestion/sample preparation methods. Samples were analysed by Matrix-assisted laser desorption/ionisation-time of flight (MALDI) and tandem (ion trap) mass spectrometry. Optimisation parameters included sample clean up and concentration, internal calibration and re-calibration methods, increased volume gel plug picking and liquid chromatography run times.

Use of the optimised methods allowed the identification of 23 differentially regulated proteins. The identified proteins are involved in important biological processes such as cytoskeletal/structural organisation, cell growth, maintenance, metabolism and immune response

**Table of Contents**

# 1.0 Introduction

## 1.1 General Introduction to Proteomics

The importance of changes in protein expression and/or function has long been appreciated in environmental chemical-induced toxicity and cancer. For many years, however, profiling protein targets of toxic chemicals has been a nearly impossible task. The development of both matrix-assisted laser desorption-ionisation (MALDI) (Karas *et al.,* 1988) and electrospray ionisation (ESI) (Fenn *et al.,* 1989) has now made it possible for high throughput analysis and identification of proteins. Proteomics focuses on the identification of a large number of proteins from cellular extracts. Biological samples normally contain numerous proteins, and the complexity of these materials requires separation steps. The most common method involves the following steps: separation of proteins from a given biological sample by gel electrophoresis (Blackstock *et al.,* 1999), excision of spots from the gel and digestion by trypsin (Quadroni *et al.,* 1999), and extraction from the gel followed by analysis by mass spectrometry (Yates *et al.,* 2000).

Two-Dimensional Electrophoresis (2-DE) combined with protein identification by mass spectrometry (MS) is currently one of the main workhorses for proteomics. Proteome analysis is technically challenging, as the number of different proteins expressed at a given time under defined biological conditions is likely to be in the range of several thousand for simple prokaryotic organisms and up to at least 10, 000 in eukaryotic cell extracts. Moreover, current proteomic studies have revealed that the majority of identified proteins are abundant housekeeping proteins that are present in numbers of $10^5$ to $10^6$ copies *per* cell, whereas proteins such as receptor molecules that are present in much lower concentrations (typically < 100 molecules

*per* cell) are usually not detected. Improved methods for enrichment of low-abundance proteins (i.e. prefractionation) as well as more sensitive detection and quantitation methods are required.

2-DE couples Isoelectric Focusing (IEF) in the first dimension with SDS-PAGE in the second dimension which enables the separation of complex mixtures of proteins according to their Isoelectric Point (*pI*), Molecular weight ($M_r$), solubility and relative abundance. 2-DE can routinely resolve ~2000 proteins simultaneously and detect < 1 ng of protein *per* spot. It also delivers a map of intact proteins, which reflects changes in protein expression level, isoforms or Post Translation Modification (PTM). This is in contrast to LC-MS/MS based methods, which perform analysis on peptides, where $M_r$ and *p*I information is lost and where stable isotope labelling is required for quantitative analysis. One of the greatest strengths of 2-DE is its capability to study proteins that have undergone some of PTM (such as phophorylation, glycosylation or limited proteolysis) and which can, in many instances, be readily located in 2-DE gels as they appear as distinct spot trains in the horizontal and/or vertical axis of the 2-DE gel. In addition, 2-DE not only provides information on protein modifications and/or changes in their expression levels, but also permits the isolation of proteins for further structural analyses by MALDI-TOF MS, ESI-MS, NMR (Rantalainen *et al.,* 2006) or Edman Microsequencing (Edman, *et al.,* 1967). The major steps of the 2-DE-MS workflow include (1) sample preparation and protein solubilisation; (2) protein separation by 2-DE; (3) protein detection and quantification; (4) computer assisted analysis of 2-DE patterns; (5) protein identification and characterisation (Dunn *et al.,* 1993).

## 1.2 Sample preparation

To take advantage of the high resolution of 2-DE, proteins of the sample have to be denatured, disaggregrated, reduced and solubilised to achieve complete disruption of the molecular interactions and to ensure that each spot represents an individual polypeptide. The major problems concerning the visualisation of proteins from total cell line extracts is the high dynamic range of proteins with respect to $M_r$, p*I* and solubility. Although a one-step procedure for protein extraction would be highly desireable with regard to simplicity and reproducibility, there is no single method of sample preparation that can be universally applied to all kinds of samples analyzed by 2-DE (Dunn *et al.,* 1993). Sample preparation should be as simple as possible to increase reproducibility and protein modifications during sample preparation must be minimised, because they might result in artifactual spots on 2-D gels. In particular, proteolytic enzymes in the sample must be inactivated. Samples containing urea must not be heated, in order to avoid charge heterogeneities caused by carbamylation of the proteins by isocyanate formed in the decomposition of urea (Dunn *et al.,* 1993). The three fundamental steps in sample preparation are cell disruption, inactivation or removal of interfering substances, and solubilisation of the proteins (Dunn *et al.,* 2004).

Cell disruption can be achieved by detergent lysis, osmotic lysis, freeze thaw cycling, sonication, grinding with or without liquid nitrogen, nitrogen cavitation or a rotating blade homogeniser. Typically microbial cells or plant tissues require harsh conditions for cell lysis due to their cell walls while more gentle methods can be applied to mammalian cells. During or after cell lysis, interfering compounds

such as proteolytic enzymes, salts, lipids, nucleic acids and/or highly abundant proteins have to removed or inactivated. The two most important parameters are salt and proteolysis.

Proteases must be inactivated to prevent protein degradation that otherwise may result in artifactual spots and loss of high $M_r$ proteins. Protease inhibitors are usually added but they may modify proteins and cause charge artifacts (Dunn *et al.*, 1993). Schuchard *et al.*, (2005) described isoform profile changes in plasma proteins on 2D-gels with the inclusion of 4-(2-Aminoethyl) Benzenesulfonyl Fluoride (AENSF) (serine protease inhibitor) which covalently modifies serine residues. Other remedies are boiling the sample in SDS-buffer (without urea) or inactivating proteases by low pH (ice-cold TCA precipitation).

Salt ions may interfere with electrophoretic separation and should be removed if their concentration is too high (>100mM). This is particularly a problem when samples are applied by sample in-gel rehydration. Higher salt concentrations can be tolerated by cup loading, however low voltages (approximately 150V) have to be applied for several hours otherwise proteins may precipitate at the site of sample application. Salt will increase the conductivity of the Isoelectric Focusing (IEF) gel thus prolonging the time required to reach the steady-state. Salt removal can be achieved by spin-dialysis or precipitation (TCA or cold acetone), 2-D clean-up kits (e.g. from GE Healthcare) or dilution of the sample below a critical salt concentration followed by application of a larger sample volume onto the Immobilised pH Gradient (IPG) gel. The sample is then desalted in the gel by applying low voltages (100V) at the beginning of the run for up to several hours

and replacing the filter paper pads beneath the electrodes (where the salt ions have collected) several times (Görg *et al.*, 2000). High amounts of lipids may interact with membrane proteins and consume detergents. High speed centrifugation of the sample and subsequent removal of the lipid layer avoids this problem.

Polysaccarides and nucleic acids can interact with carrier ampholytes and proteins resulting in streaky 2-D patterns. These macromolecules may also increase the viscosity of the solution and obstruct the pores of the polyacrylamide gels. The nucleic acids can be removed by their digestion with RNAses and DNAses.

Sometimes highly abundant proteins present problems since they impair separation and detection of lower abundant proteins by limiting the amount of these proteins to be loaded onto the 2-D gel and/or masking them on the 2-D pattern. Albumin constitutes up to 60% of the bulk protein in plasma and is a major problem (Pieper *et al.*, 2003). There are commercially available albumin removal kits but one has to be aware that these kits remove proteins other than albumin (Simpson *et al.*, 2004).

### 1.2.1 Protein solubilisation

After cell disruption and/or removal of interfering compounds, the individual polypeptides must be denatured and reduced to disrupt intra- and intermolecular interactions and solubilised while maintaining the inherent charge properties. Sample solubilisation is usually carried out in a buffer containing chaotropes (e.g. urea and/or thiourea), non-ionic and/or zwitterionic detergents (e.g. CHAPS),

reducing agents, carrier ampholyte (CA) and depending on the sample, protease inhibitors. The most popular sample solubilisation buffer is based on O'Farrell's lysis buffer ( O'Farrell *et al.,* 1975) and modifications thereof (9M Urea, 2-4% CHAPS, 1%DTT and 2% v/v CA).

Urea is quite effective in disrupting hydrogen bonds, leading to protein unfolding and denaturation. However, urea lysis buffers are not ideal for the solubilisation of all protein classes, particularly for membrane or other hydrophobic proteins. Improvement in the solubilisation of hydrophobic proteins has come with the use of thiourea (Rabilloud *et al.,* 1998) and new zwitteronic detergents such as sulfobetaines (Luche *et al.,* 2003). Thiourea is better suited for breaking hydrophobic interactions but its usefulness is limited due to its poor solubility in water. However it is more soluble in concentrated urea solutions. Currently the best solution for solubilisation is a combination of 5-7M urea and 2M thiourea. The major problem associated with urea in aqueous solutions is that urea exists in equilibrium with ammonium (iso)cyanate, which can react with the α-amino groups of the N-terminus and the ε-amino groups of lysine residues, thereby forming artifacts such as blocking the N-terminus and introducing charge heterogeneities. To prevent this carbamylation reaction temperatures above 37°C have to be avoided under all circumstances and CA (2%v/v), which act as cyanate scavengers should be included in the urea solution. The zwitterionic detergent CHAPS is effective in solubilising very hydrophobic proteins and in combination with urea and thiourea have been shown to solubilise at least several integral membrane proteins (Molloy *et al.,* 2000).

Reduction and prevention of re-oxidation of disulfide bonds is a critical step of the sample preparation procedure. Reducing agents are necessary for cleavage of intra- and intermolecular disulphide bonds to achieve complete protein unfolding. The most commonly used reducant is DTT and is applied in excess, i.e. concentrations up to 100mM. Unfortunately DTT is a weak acid with pKvalue between 8.5-9, which means it will ionise at basic pH and therefore will not allow complete resolution in the alkaline gel area due to migration to the anode during IEF.

### 1.2.2 Prefractionation procedures

Since there is no amplification step for proteins analogous to the PCR method for amplifying nucleic acids, and due to the high dynamic range and diversity of expressed proteins, it is often preferable to carry out a prefractionation step to reduce the complexity of the sample, enrich for certain proteins such as low-copy number proteins or alkaline proteins and to get some information on the topology of the proteins. This can be accomplished by (1) isolation of specific cell types from a tissue, e.g. laser capture micro dissection (LCM) (Zhou *et al.,* 2002), or fluorescence activated cell sorting (FACS) (Li W *et al.,* 2006), (2) isolation of cell compartments and/or organelles e.g. by sucrose gradient centrifugation (Zhang *et al.,* 2006), (3) selective precipitation of certain protein classes e.g. TCA/acetone precipitation for ribosomal proteins (Jiang *et al.,* 2004), (4) sequential extraction procedures with increasingly powerful solubilisation buffers e.g. aqueous buffers, organic solvents (ethanol or chloroform/methanol) (Ruth *et al.,* 2006) and detergent-based extraction solutions (Williams *et al.,* 2007), (5) chromatographic

separation methods e.g phosphoprotein enrichment using Phosphoprotein Protein Purification kit (Qiagen) (Laugesen *et al.*, 2006). One major disadvantage of most prefractionation procedures lies in cross-contamination between individual fractions and in the fact that they are either time consuming, complicated to handle, require concentration steps due to elution/dilution procedures and/or do not allow more than a few samples in parallel to be processed. Another disadvantage with fractionation procedures for the depletion of high abundant proteins for example in biomarker discovery can result in the loss of many low abundant proteins too. Zhou *et al* (2004) described when albumin is immunodepleted, another 63 proteins bound to albumin are co-depleted as well which can introduce variability in results and possible loss of information.

## 1.3 2-Dimensional Electrophoresis (2-DE)

The prerequisite, but also the challenge for proteome analysis is to separate proteins from complex biological samples with high reproducibility and high resolution. During 2-DE, proteins in complex mixtures are electrophoretically separated according to their isoelectric point (pI) and subsequently according to their molecular weight by a sequential combination of isoelectric focusing (IEF) and sodium-dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE). 2-DE has been used for large-scale protein separation since 1975 (Klose 1975) and allows for purification, identification as well as quantification of proteins. 2-DE analysis provides several types of information, including molecular weight, pI and quantity.

### 1.3.1 First dimension: IEF

Differences in proteins isoelectric points are the basis of separations by isoelectric focusing (IEF). The pI is the pH at which a protein will not migrate in an electric field and is determined by the charged groups in the protein. Proteins can carry positive, negative or zero charge depending on their local pH, and for every protein there is a specific pH at which its net charge is zero (this is its pI).

IEF represents the first dimension of two-dimensional (2D) electrophoresis, and immobilised pH gradient (IPG) strips facilitate this analysis.

Linear or nonlinear (Bjellqvist *et al.*, 1993) wide pH range (e.g. IPG 3-12), medium (e.g. IPG 4-7), narrow (e.g. 4.5-5.5) and/or ultra narrow (e.g. IPG 4.9-5.3).

IPG's can vary in lengths to accommodate various gel sizes, usually from 7-24cm. Besides laboratory-made IPG gels, a variety of commercially available IPG dry strips can now be purchased from different suppliers. Ready-made IPG dry strips are increasingly popular due to easier handling, better comparability of results and exchange of data, and have significantly contributed to the widespread application of 2-DE in proteomics.

### 1.3.2 IPG strip rehydration and sample application

Prior to IEF, the IPG dry strips must be rehydrated (usually overnight) to their original thickness of 0.5mm with a rehydration buffer containing either 8M urea or 2M thiourea and 6M urea, 0.5-4% non-ionic or zwitterionic detergents (e.g. 2% CHAPS), a reductant (typically 0.4% DTT) and 0.5% v/v CA (e.g. IPG buffer or Pharmalyte 3-10). IPG dry strips are either rehydrated with the sample already dissolved in rehydration buffer (sample in-gel rehydration), or with rehydration buffer without sample, followed by sample application by cup-loading. For sample in-gel rehydration (Rabilloud *et al.*, 1994), the sample (1-10mg protein/ml) is directly solubilised in a defined volume of rehydration buffer. Sample in-gel rehydration is not recommended for samples containing very high $M_r$, very alkaline and/or very hydrophobic proteins since these are taken up into the gel only with difficulty. For cup-loading, the IPG dry strips are reswollen in rehydration buffer without sample. After IPG strip rehydration, samples dissolved in lysis buffer are applied into disposable plastic cups placed onto the surface of the IPG strip. Best results are obtained when the samples are applied at the pH extremes.

The choice of pH gradient primarily depends on the complexity of the sample. Wide or medium range IPGs such a IPGs 3-12, 4-9, or 4-7, are typically used to analyze simple proteomes or to get an overview of a more complex proteome. Settings are usually limited to 50μA per IPG stip and 150V to avoid the protein denaturing effects caused by Joule heating, because conductivity is initially high due to salts. As the run proceeds, the salt ion migrate to the electrodes, resulting in decreased conductivity and allowing high voltages to be applied

### 1.3.3 Equilibration of IPG gel strips

Following isoelectric focusing, IPG strips are equilibrated to further denature proteins, maintain solubility and establish an appropriate pH for the second dimension SDS-PAGE analysis. Due to the observation that the focused proteins bind more strongly to the fixed charged groups of the IPG gel matrix than to CA gels, relatively long equilibration times (10-15 min), as well as urea and glycerol to reduce electroendosmotic effects are required to improve protein transfer from the first to the second dimension (Görg *et al.,* 1988). Traditionally the best protocol is to incubate the IPG strips for 10-15mins with 50mM Tris-HCL (pH 8.8), containing 2% w/v SDS, 1% w/v DTT, 6M urea and 30% w/v glycerol followed by a further 10-15 minute equilibration in the same solution containing 4% (w/v) iodacetamide instead of DTT (Görg *et al.,* 1988). The latter step is used to alkylate any free DTT, as otherwise it migrates through the second-simension SDS-PAGE gel, resulting in an artefact known as point streaking that can be observed after silver staining (Görg *et al.,* 1995). More importantly, the iodacetamide alkylates sulfhydryl groups and

prevents their reoxidation; this step is highly recommended for subsequent spot identification by MS. After equilibration, the IPG strips are applied onto the surface of the second-dimensional SDS-PAGE gel.

### 1.3.4 Second dimension: SDS-PAGE

SDS-PAGE can be performed on horizontal or vertical systems (Görg *et al.,* 1995). Horizontal setups are ideally suited for ready-made gels (e.g. ExcelGelSDS; GE Healthcare), whereas vertical systems are preferred for multiple runs in parallel, in particular for large-scale proteome analysis which usually requires simultaneous electrophoresis of batches of second-dimension SDS-PAGE gels for higher through-put and maximal reproducibility (Anderson *et al.,* 1978). The most commonly used buffer for the second dimension of 2-DE is the discontinous buffer system of Laemmli (Laemmli *et al.,* 1970). The analysis of low $M_r$ (<15kDa) and high $M_r$ (150kDa) proteins is somewhat intricate since there is no standard 2-DE system which effectively allows separation of proteins over the entire $M_r$ range between 5kDa and 500kDa. A common approach is to combine several gels optimised for the appropriate $M_r$ ranges 5-30kDa, 15-200kDa and >150kDa.

## 1.4 Protein detection and quantitation

After 2-DE, the separated proteins have to be visualised, either by universal or by specific staining methods. Concentrations of individual proteins in a single cell differ between six or seven orders of magnitude, ranging from several millions of copies/cell for some highly abundant proteins to a few copies/cell for low abundant proteins. These enormous variations in protein concentrations are a major challenge for almost all currently available protein detection methods (Corthals *et al.,* 2000). The most important properties of protein visualisation methods are high sensitivity, high dynamic range, reproducibility and compatibility with post-electrophoretic protein identification procedures such as MS. Universal detection methods of proteins on 2-D gels include staining with anionic dyes (e.g. Coomassie Blue), negative staining with metal cations (e.g. zinc imidazole), Silver-staining, fluorescence staining or labelling, and radioactive isotopes, using autoradiography , fluorography, or Phosphor-imaging. For most of these staining procedures, the resolved polypeptides have to be fixed in solutions such as ethanol/acetic acid/$H_20$ for at least several hours before staining to remove any compounds (e.g. CA detergents) that might interfere with detection. Specific staining methods for detection of Post-Translational Modification (PTM) (glycosylation, phosphorylation etc.) are employed either directly in the 2-DE gel or more frequently after transfer (blotting) onto an immobilised membrane. The blotted proteins can be probed with specific antibodies or with lectins.

### 1.4.1 Universal protein detection and quantitation methods

Countless staining procedures have been reported in the literature ever since the pioneering work of Ornstein and Davis (Ornstein *et al.,* 1964; Davis *et al.,* 1964) established a high resolution zone electrophoretic method in polyacrylamide gels, called disc-electrophoresis on account of several discontinuities established along the migration path for sharpening the analyte zones. They can be roughly divided into the following categories: (1) organic dyes; (2) silver stains; (3) negative stains/reverse staining; (4) fluorescent stains.

The Coomassie stains (organic dyes) involve dye-dye interactions among Coomassie dyes that are ionically bound to, or in hydrophobic association with protein molecules (Wilson *et al.,* 1979) Colloidal Brilliant Blue (CBB) staining methods have found widespread used for the detection of proteins on 2-DE gels, because of their low price, ease of use and compatibility with most subsequent protein analysis and characterisation methods such as MS. The limitation of CBB stains lies in their insufficient sensitivity (detection limit in the range of 200-500ng protein per spot), which does not permit the detection of low abundance proteins.

Reverse staining exploits the fact that protein-bound metal cations (e.g. zinc, copper or potassium) are usually less reactive than the free salt in the gel. Thus, the speed of precipitation of free or only weakly bound ions to form an insoluble salt is slower on sites occupied by proteins than in the protein-free background. This generates transparent protein zones or spots while the gel background becomes opaque due to the precipitated, unsoluble salt. The advantages of reverse staining is

they act rapidly (15mins for most applications) and allow easy recovery of proteins from gel matrices, since they do not require a fixation step. The zinc stain method is quick and is more sensitive than CBB (Adams *et al.,* 1990). The detection limit is 20-50ng of protein per spot and is compatible with subsequent protein identification by MS. The major disadvantage of zinc staining is its restricted linear dynamic range making this staining method unsuitable for detecting quantitative differences on 2-DE gels.

Silver staining methods (Merril *et al.,* 1981) are far more sensitive than CBB or zinc stains. Silver staining is reported to increase the sensitivity of protein detection over Coomassie Blue staining by 2000-fold, from tenths of a microgram to tenths of a nanogram (Merril *et al.,* 1981). All silver stain protocols depend on the reduction of ionic to metallic silver onto the protein surface. Although silver staining is decidedly the most sensitive stain, silvering protocols are highly erratic due to a number of parameters such as temperature, quality of solvents, developing times. Moreover, a large number of silvering techniques do not allow for further processing of proteins since aldehydes irreversibly cross-link polypeptide chains, although aldehyde-free silver protocols are available, albeit with lower sensitivities (Sinha *et al.,* 2001).

Better and more reproducible results in terms of sensitivity and linear dynamic range of detection are obtained by protein detection methods relying on fluorescent compounds or radiolabelling of proteins combined with highly sensitive electronic detection methods. Prior to the advent of highly sensitive silver staining, detection of proteins labelled with radioisotopes was the only method of sensitive detection

for proteins separated on 2-DE gels. Radiolabelling can be accomplished by incorporating radioactive isotopes in proteins. *In vivo* metabolic radiolabelling of samples by the incorporation of radioactive amino has been extensively used for proteome analysis of microorganisms and cell culture systems. The radiolabelled 2-DE separated proteins can be detected by autoradiography of fluorography using X-ray films which are exposed to the dried gels and which can be quantified by densitometry. However, these film based techniques require long exposure times (up to several weeks) if high sensitivity is desired. To overcome the limitations of X-ray film-based autoradiography, several electronic methods for the detection of radiolabelled proteins in 2-D gels have been developed e.g. phosphor-imaging. The advantages are the possibility to detect very low levels of radioactivity in a considerably shorter time and the higher dynamic range (up to five orders of magnitude) (Patterson *et al.,* 1993). The major disadvantages of radiolabelling technology are the well known general shortcomings associated with radiolabelling, such as hazardous and expensive radiochemicals, waste disposal, safety considerations and the high costs for equipment (phosphor-imager, imaging screens).

Due to the shortcomings of organic dyes, silver staining or radiolabelling for visualisation and quantitation of proteins, fluorescent detection of proteins has gained popularity for proteome analysis. The best known example of pre-electrophoretic fluorescent labels is the flourophore monobromobimane (Urwin *et al.,* 1993) and the cyanine-based dyes (Unlu *et al.,* 1997) that react with cysteine and lysine residues, respectively. The latter dyes are commercially available as CyDyes (GE Healthcare). The major problem of pre-electrophoretic labelling is the

occurrence of protein size and/or protein charge modifications which may result in altered protein mobilities alongside the $M_r$ and/or p*I* axis. Alternatively, proteins can be stained with a fluorescent dye molecule after the electrophoretic separation has been completed. The most prominent example is the ruthenium-based dye SYPRO Ruby (Berggren *et al.,* 2002). Staining is accomplished within a few hours in a single step procedure which may be easily adapted for use with automated instrumentation. The detection limit is approximately 1-2ng protein/spot, and the linear dynamic range of quantitation is about three orders of magnitude. A cost effective alternative to SYPRO Ruby staining, which is based on Ruthenium II tris (bathophenanthroline disulfonate) has been developed by Rabilloud *et al.,* (2001). In conclusion, protein detection and quantitation methods based on fluorescent staining and/or labelling are rather promising. They have a comparatively wide linear dynamic range ($>10^3$) and are relatively easy to use. Furthermore, most fluorescent staining procedures are compatible with subsequent protein identification methods such as MS. The major limitation of most fluorescent staining methods is their lowered sensitivity compared to electronic detection methods of radiolabelled proteins. Typically, only proteins expressed at greater than $10^3$ copies/cell can be detected on a standard 2-DE gels by using fluorescent dye technologies, whereas at least in theory less than a dozen copies of a protein/cell can be visualised with the most sensitive electronic detection methods for radiolabelled proteins.

## 1.5 Methods for the analysis of protein (Post Translational Modifications) PTMs

Through genome sequencing no information can be gained on post-translational modifications PTMs of proteins. PTM is when the primary structure of the protein is altered after the protein has been translated and is already folded. Proteins can display a broad range of PTMs, the most common of which are summarised in table 1.1. These modifications act on individual residues either by cleavage at specific points, deletions, additions or having side chains converted or modified.

| **The main post-translational modifications associated with proteins** |
|---|
| Acetylation, acylation, ADP-ribosylation |
| Amidation |
| γ-Carboxylation an β-hydroxylation |
| Disulfide bond formation |
| Glycosylation |
| Phosphorylation |
| Proteolytic processing |
| Sulfation |

**Table 1.1 The main post-translational modifications associated with proteins**

Protein phosphorylation is a key PTMs crucial in the control of numerous regulatory pathways, enzyme activites and degradation of proteins, whereas glycosylation is associated with biochemical alterations, developmental changes and pathogenesis, e.g. tumorgenesis. Hence detection and characterisation of PTMs are a major task in proteomics. Phosphorylation results in a shift of the protein p*I* to more acidic values. One of the strengths of 2-DE is its capability to readily locate post-translationally modified proteins, as they frequently appear as distinct rows of spots in the horizontal and/or vertical axis of the 2-DE gel.

Phosphoproteins can be detected on 2-DE gels by autoradiography of phosphor-imaging after *in-vivo* incorporation of $^{32}P$ or $^{33}P$ orthophosphates into proteins. An alternative method for phosphoprotein detection is immunostaining with phosphoamino acid-specific poly- or monoclonal antibodies after transfer of the 2-DE separated proteins onto an immobilising membrane. Fluorescent detection of phosphoproteins separated on 2-DE using Pro-Q Diamond phosphoprotein dye (Molecular Probes) which are MS compatible is also possible. Phosphoserine-, phosphothreonine-, and phosphotyrosine- containing proteins can be detected using Pro-Q Diamond (Steinberg *et al.,* 2003). The detection limit is 1-2ng. However the specificity of the stain, in particular with complex protein samples has been questioned since highly abundant, non-phosphorylated proteins may also be stained, albeit less intense than the phosphorylated ones.

The most classical procedure for glycoprotein staining is periodic acid-Schiff (PAS). When treated with periodic acid, vicinal diols are oxidised to aldehydes, able to react with pararosaniline, a triphenylmethane derivative, and to sodium

metabisulfite (together, the Schiff's reagent) and to form an adduct, pink to magenta in colour. The procedure has been adapted to glycoprotein detection in gels (Zacharius *et al.,* 1969). The sensitivity is low, 25-100ng carbohydrate or 1-10μ highly glycosylated protein. Recently, Pro-Q emerald 488, a glycoprotein stain that reacts with periodic acid oxidised carbohydrate groups and which generates a green fluorescent signal on glycoproteins has been described (Hart *et al.,* 2003). This stain permits detection of 5-20ng of glycoprotein per spot, depending on the nature and degree of glycoprotein glycosylation. The second principle for glycoprotein detection is based on sugar binding proteins so-called lectins. A wide range of lectins with different carbohydrate specificities are commercially available. However for the detailed analysis of the saccharide composition of glycoproteins, HPLC or MS-based methods are usually preferred.

## 1.6 Difference In-Gel Electrophoresis (DIGE)

A bottleneck for high throughput proteomic studies is image analysis. In conventional 2-D methodology, protein samples are separated on individual gels, stained and quantified, followed by image comparison with computer aided image analysis programs. Because multistep 2-DE technology often prohibits different images from being perfectly superimposed, image analysis is frequently very time consuming. To shorten this laborious procedure Ünlü *et al.*, (1997) have developed a method called DIGE, in which samples are labelled *in vitro* using different fluorescent cyanine minimal dyes (GE Healthcare) differing in their excitation and emission wavelengths. Three CyDyes are commonly used, namely Cy2, Cy3 and Cy5, which allow labelling of two samples of proteins plus an internal standard (Alban *et al.*, 2003). This internal standard, typically a pooled mixture of all the samples in the experiment is used for normalisation of data between gels thereby minimising experimental variation and increasing the confidence in matching and quantitation of different gels in complex experimental designs. These different fluorescent cyanine minimal dyes differ in their excitation and emission wavelengths and are mixed with sample prior to IEF and separation on a single 2D gel. After consecutive excitation with all three wavelengths, the images are overlaid and subtracted (normalised), whereby only differences (e.g. up- or down regulated, and/or PTM proteins) between the samples are visualised. Due to the comigration of the sample, methodological variations in spot positions and protein abundance are excluded and consequently image analysis is facilitated considerably. Nevertheless, this approach is still dependent on accurate matching and comparison of large sets of 2-DE gels in order to generate meaningful data on differential

protein expression between sets of samples (Dowsey *et al.,* 2003). Applications that profit from DIGE include samples generated under various pre specified conditions, the comparison of extracts and the analysis of biological variance. The p*I*s of the cyanine dye labelled proteins remain unaffected, because the cyanine dyes compensate for the loss of the positive charge of the lysine residues. However, the $M_r$ increases by 434-464 Da (depending on the dye molecule) per labelled lysine residue. Consequently, labelling of more than one lysine reside per protein molecule must be avoided otherwise labelling with CyDyes would result in multiple spots in the vertical axis of the 2-DE gel. In practice, approximately 3-5% of protein is labelled (referred as minimal labelling). Since the bulk of the protein remains unaltered, the slight increase in $M_r$ sometimes presents a problem for spot excision for MS analysis, particularly with lower $M_r$ proteins (Shaw *et al.,* 2003). This off-set has to be taken into account when automatic spot pickers are used. One alternative is to stain the separated proteins additionally with SYPRO Ruby or CBB prior to spot picking from microperparative gels. Another is so-called saturation labelling with similar cyanine dyes with label cysteine (instead of lysine) residues to saturation (Shaw *et al.,* 2003). An additional advantage of saturation is the greater sensitivity of the stain with a detection limit of approximately 0.1ng of protein compared to 1ng with minimal labelling Van den Bergh *et al.,* (2004). Saturation labelling has several drawbacks, (1) the labelling reaction must be carried out at a defined protein/dye ration to obtain an optimal 2-DE spot pattern with a minimal number of spot trains in the vertical dimension, (2) the dye ratio must be assessed for different types of samples, depending on the percentage of cysteines, (3) care must be taken to avoid side-chain reactions, such as lysine labelling, (4) up to 25% of the protein material may precipitate during the labelling

reaction due to the introduction of the hydrophobic dye molecule and (5) the 2-DE spot pattern is significantly altered compared to that of unlabelled or minimal-labelled proteins (Shaw *et al.*, 2003).

## 1.7 Computerised 2-D image analysis

One of the key objectives of proteomics is to identify the differential expression between control and experimental samples run on a series of 2-D gels. That is, the protein spots that have been inhibited, induced or have changed abundance (increased or decreased in size and intensity). Once these gel features have been found, the proteins of interest can be identified using MS. This goal is usually accomplished with the help of computerised image analysis systems (Dowsey *et al.*, 2003). The first step in computerised image analysis of 2-DE protein patterns is capture of the gel image in a digital format. A range of devices, including modified document scanners, laser densitometers, CCD cameras, and fluorescent scanners, phosphor imagers are available for the acquisition of 2-D gel images. The saved images are then subjected to computer assisted image analysis. The traditional workflow for 2-DE software package is (1) preprocessing of the gel images i.e. normalisation, cropping and background subtraction, (2) spot segmentation, detection and expression quantification, (3) landmarking i.e. an initial user guided pairing of a few spots between the reference and sample gels. The sample gel is warped to align the landmarks, (4) matching i.e. automatic pairing of the rest of the spots, (5) identification of differentially expressed spots and (6) data presentation and interpretation (Dunn *et al.*, 1992). Currently, several 2-D image analysis

software packages are commercially available (Progenesis SameSpot-Nonlinear Dynamics; DeCyder™-GE Healthcare). These programs have been continuously improved and enhanced over the years in terms of faster matching algorithms with lesser manual intervention, and with focus on automation and better integration of data from various sources. The digitised image is first subjected to several clean up steps to reduce background smear and to remove horizontal and/or vertical streaks. This procedure is usually fast and does not require much user interaction. The individual spots on the 2-D pattern are then detected and quantified. This step is also performed automatically. Regrettably, most image analysis programs do not identify all spots correctly, particularly when the overall quality of the electrophoretic separation is low (e.g. when crowded areas and overlapping spots due to improper sample preparation or insufficient spatial resolution are present on the gel). Hence, manual spot editing with reference to the original stained 2-DE gel (or image) is usually necessary. Depending on the number of spots, quality of the 2-DE separation and the algorithms used for spot detection, this process may be quite laborious and time-consuming. After spot editing, each spot on one 2-DE gel must be matched by means of a reference (master) gel. For this, in most computer aided 2-D image analysis programs, several landmark spots (which should be evenly distributed over the entire gel) must be manually identified by the operator. Starting from these landmark spots, the program proceeds to match the other spots automatically. Again, mismatches must be carefully checked and edited manually. Typically, at least two 2-D gel patterns are matched (e.g. control versus disease) and then compared to each other with respect to qualitative and quantitiative differences between the 2-D patterns.

## 1.8 Protein Identification

MS has become the technique of choice for identification of proteins from excised 2-D gel spots. Despite the outstanding properties and widespread application of 2-DE, almost 20 years went by before systematic use of 2D PAGE gels became an integral part of the revived field now termed proteomics. 2D PAGE gels were far ahead of their time because no method was available for the rapid, routine identification of the protein spots on the gels. In the past decade, two crucial developments have changed the situation. First, whole genome sequencing for an increasing number of organisms has defined at the gene level all proteins that exist in that organism. This has meant that, rather than a complete covalent characterization being necessary to identify a protein, partial information on fragments or sequence is adequate to identify the protein at the gene level within a genome sequence database. Second, mass spectrometric methods, largely based on MALDI-TOF and/or electrospray ionization with subsequent fragmentation of peptides, can now supply the needed information to identify the protein at sensitivities routinely below 1 pM (Cañas *et al.,* 2006). Two mass spectrometric methods for rapid identification of proteins at the gene level with samples of the amounts readily available on 2D gels are now in widespread use. One of these methods uses the characteristic distribution of peptide masses obtained by enzymatic fragmentation of proteins (Henzel *et al.,* 1993). A number of computer programs are available for using the observed peptide masses to search gene sequence databases for proteins that fit the mass fingerprint (Henzel *et al.,* 1993, Yates et al., 1993; Clauser *et al.,* 1995; Wilkins *et al.,* 1997). The other method for protein identification uses sequence tags (Mann & Wilm, 1994; Mann, 1996); i.e.,

partial amino acid sequence information. Searching of gene sequence databases is again used to identify the protein at the gene level.

Mass spectrometric measurements are carried out in the gas phase on ionised analytes. By definition, a mass spectrometer consists of an ion source, a mass analyser that measures the mass-to-charge ratio (m/z) on the ionised analytes and a detector that registers the number of ions at each m/z value. Matrix-assisted laser desorption ionisation (MALDI) and electrospray ionisation (ESI) are the two techniques most commonly used to volatise and ionise the proteins or peptides for mass spectrometric analysis. MALDI sublimates and ionises the samples out of a dry, crystalline matrix via laser pulses. MALDI-MS is normally used to analyse relatively simple peptide mixtures. ESI ionises the analytes out of a solution and is therefore readily couple to a liquid-based separation tools. Integrated liquid-chromatography ESI-MS systems (LC-MS) are generally preferred for the analysis of complex samples.

The mass analyser is central to the technology. In the context of proteomics, its key parameters are sensitivity, resolution, mass accuracy and the ability to generate information-rich ion mass accuracy spectra from peptide fragments. There are four basic types of mass analyser currently in proteomics research. These are time of flight (TOF), ion trap, quadrupole and Fourier transform ion cyclotron (FR-MS) analysers. They are very different in design and performance, each with its own strength and weakness.

### 1.8.1 Matrix Assisted Laser Desorption Ionisation (MALDI) – Time of Flight (TOF)

The matrix assister laser desorption/ionisation technique was developed in 1987 (Karas *et al.,* 1987) and has increased the upper mass limit of mass spectrometric analyses of large biomolecules to over 300,000Da. It has also enabled the analyses of biomolecules by mass spectrometry to become easier and more sensitive. The time-of-flight (TOF) mass spectrometer operates on the principle that when a temporally and spacially well defined group of ions of differing mass/charge (m/z) ratios are subjected to the same electric field and allowed to drift in a region of constant electric field, they will traverse this region in a time which depends on their m/z ratios. Reflectron MS offers improves mass resolution by compensating for the difference in flight times of the same m/z ions of slightly different kinetic energies by means of an ion reflector. This results in focusing the ion packets in space and time at the detector.

The carbon, hydrogen, nitrogen, oxygen, and sulphur in proteins, as elsewhere, are mixtures of isotopes. About one carbon atom in 90 is $^{13}C$ instead of the usual $^{12}C$, so a peptide with a mass of 1900, containing about 90 carbon atoms, will on average contain one $^{13}C$ atom, but may contain none, one, two, or more. Thus such a peptide will not give a single sharp peak, but a multiplet, as observed in Figure 1.1 which is the autodigestion trypsin III peak from a typical MALDI-TOF MS spectrum of a protein digested with trypsin.

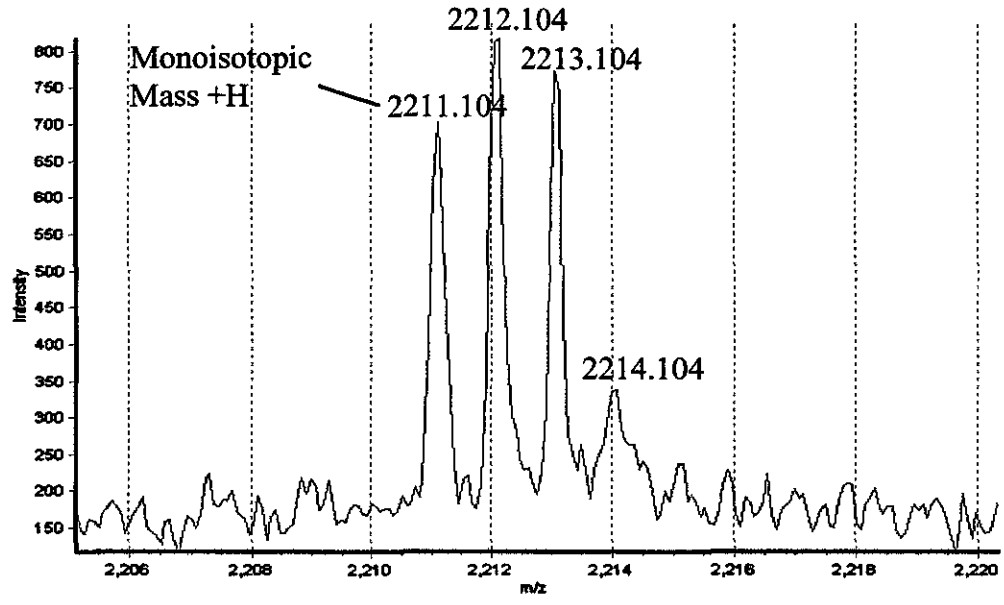



**Figure 1.1 MALDI ToF analysis of a single Charge Trypsin peak at 2211.104m/z**

The first peak seen in the cluster is the monoisotopic peak consisting of the most common basic element isotopes ($^{12}C$). The area under the peaks is proportional to the statistical chance of having a molecule with 0,1,2, or more $^{13}C$ atoms.
The difference between the isotopes shows the charge state of the ion. As shown in the spectrum above, if the m/z difference is one atomic mass unit then the charge is +1. A charge of +2 would give an m/z difference of 0.5, however in MALDI mass spectrometry this is fairly uncommon.

One of the most important aspects of MALDI-TOF-MS is the preparation of a good sample (Cohen *et al.,* 1996; Beavis *et al.,* 1996). Many variables influence the integrity of a good homogeneous MALDI sample and may include the concentration of matrix and analyte, choice of matrix (see table 1.2), analyte sample history (i.e. exposure to salts, strong ionic detergents).

| Matrix | Application |
|---|---|
| α-Cyano-4-hydroxycinnamic acid (CHCA) | Peptides, proteins, lipids and oligonucleotides |
| 3,5-Dimethoxy-4-hydrocinnamic acid (sinapinic acid) | Peptides, proteins and glycoproteins |
| 2,50-Dihydroxybenzoic acid (DHB) | Peptide, proteins, lipids and oligoasccarides |

**Table 1.2 Example of common MALDI matrices at 337nm**

MALDI uses laser pulses to ionise and volatilise proteins and peptides that have been embedded in the dry crystalline organic matrix. MALDI is usually coupled to TOF analysers that measure the mass of intact peptides, because there is a fixed flight distance to the detector and ions with different masses will travel at different speeds. Peptide ions with the same m/z will exhibit equal times of flight. The acquired spectra by MALDI-TOF mass spectrometry allows mass fingerprinting (PMF). PMF is typically the primary tool for protein identification. This technique, which is user friendly and quite fast, based on the finding that a set of peptide masses obtained by MS analysis of a protein digest (usually trypsin) provides a characteristic mass fingerprint of that protein. The protein is then identified by comparison of the experimental mass fingerprint with the theoretical peptide masses generated *in silico* using protein and nucleotide databases. Because mass mapping requires an essentially purified target protein, the technique is commonly used in conjunction with 2-DE. This approach proves very effective when trying to identify proteins from species whose genomes are relatively small, completely

sequenced and well annotated. PMF is not so reliable for organisms whose genomes have not been completely sequenced. Protein spots from 2-DE gels frequently require concentrating and desalting (salt ions suppress ionisation) prior to MALDI MS. The peptides can trapped with Reverse Phase material (ZipTip, Millipore), and after desalting washes with water, they can be eluted with 50%ACN containing matrix. The method results in highly purified, salt-free peptide solutions that can be directly applied to MALDI PMF database searching.

A second problem is to identify proteins that are extensively post-translationally modified, since the peptides generated from these proteins may not match with the unmodified protein in the database (Graves *et al., 1999*). A third problem is that PMF does not work as well if several different proteins are present in the spot. Search engines such as ProFound (Zhang, *et al.,* 2000) have been developed that enable identification of the correct protein(s) even when the data quality is relatively low or when the sample consists of a simple mixture of proteins. In cases proving impossible to identify by PMF alone it is essential to obtain amino acid sequence information. This is most readily accomplished using either MALDI-MS with Post Source Decay (PSD) or by MALDI-TOF Instruments capable of fragmenting MALDI generated precursor ions e.g. MALDI-TOF-TOF and hybrid quadrupole TOF Instruments.

To allow the fragmentation of MALDI-generated precursor ions, MALDI ion sources can be coupled to quadrupole ion-trap mass spectrometers (Krutchinsky *et al.,* 2001) and to two types of TOF instruments. In the first, two TOF sections are separated by a collision cell (TOF-TOF instrument), whereas in the second, the

hybrid quadrupole TOF instrument, the collision cell is placed between a quadrupole mass filter and a TOF analyser (Loboda, *et al.*, 2006).

### 1.8.2 Ion Trap

In ion-trap analysers, the ions are first captured or "trapped" for a certain time interval and are then subjected to MS or MS/MS analysis. Ions created by electrospray (ESI) are focused using electrostatic lensing system into the ion trap. An electrostatic ion gate pulses open and closed to inject ions into the trap. The time during which ions are allowed into the trap, termed the "ionisation period" is set to maximise signal while minimising space-charge effects. The ion trap is typically filled with helium. Collisions with helium dampens the kinetic energy of the ions and serve to quickly contract trajectories toward the center of the ion trap, enabling trapping of injected ions. Trapped ions are further focused toward the center of the trap through the use of an oscillating potential, called the fundamental *rf*, applied to the ring electrode. An ion will be stably trapped depending upon the values for the mass and charge of the ion, the size of the ion trap (r), the oscillating frequency of the fundamental rf (w), and the amplitude of the voltage on the ring electrode (*V*). Ion traps are robust and sensitive. The linear or two-dimensional ion trap (Hager *et al.*, 2003 and Schwartz *et al.*, 2002) are recent developments where ions are stored in a cylindrical volume that is considerably larger than that of the traditional, three dimensional ion traps allowing increased sensitivity, resolution and mass accuracy.

The Fourier transform ion cyclotron resonance mass spectrometer (FTICR-MS) instrument is also a trapping mass spectrometer. They measure mass indirectly by oscillating ions in a strong magnetic field. Because these ions will oscillate as a function of their m/z, measuring the frequency of these oscillations allows inference of m/z using a Fourier transform (FT). These instruments provide the highest combination of mass accuracy, resolution and intrascan dynamic range (Khalsa-Moyers *et al.,* 2006). They offer the highest resolution performance and is the most potential protein analyser for protein characterisation and analysis of PTM. Its strengths are high sensitivity, mass accuracy, resolution and dynamic range (Martin *et al.,* 2005). In spite of the enormous potential, the expense of FTIR-MS instruments has limited their routine use in proteomics research. The instruments are expensive, need permanent special experience with high-field magnets, their capacity is still limited, and hence only special analyses can be performed and high throughput seems not be feasible at present (Wittmann-Liebold *et al.,* 2006).

Recently Orbitrap Mass Analysers can deliver accurate mass and high resolution similar to those achievable with FT-MS instrumentation but are more affordable and have less running costs to the latter (Hardman *et al.,* 2003). Orbitrap is an ion trap where the ions are trapped in an electrostatic field (Hardman *et al.,* 2003). A Fourier Transform is employed to obtain oscillation frequencies for ions with different masses resulting in an accurate reading of their m/z.
Electropray Ionisation has mostly been coupled to ion traps and used to generate fragment ion spectra (collision-induced (CID) spectra) of selected precursor ions (Albersold *et al.,* 2001).

### 1.8.3 Tandem Mass Spectrometry

A Tandem mass spectrometry (MS/MS) consists of two mass analysers separated by an ion-activation device. This makes it possible to separate an ion of interest using the first analyser, induce the ion to dissociate by activation and then analyse the dissociation products in the second analyser. These dissociation products provide a "structural fingerprint" for that particular ion and thereby provide detailed structural information (e.g. amino acid sequence on it). The sequencing of peptides and proteins by various ion-sequencing techniques has been demonstrated, including low- and high energy collision –induced dissociation (CID) (Hunt *et al.*, 1996), surface-induced activation (SID) (Biemann *et al.*, 1988) and photoinduced dissociation (PID) (Beussman *et al.*, 1995). Currently, the most commonly used ion-activation technique for the dissociation of peptide ions is low-energy CID (which involves the collision of peptide ions at low speeds with an inert gas). The fragment ions produced upon peptide-ion activation are analysed by a second mass analyser. Under these low energy gas phase collision conditions, peptide ions mostly fragment at the amide (peptide) bonds along the backbone, generating a ladder of sequence ions. There are three different types of bonds that can fragment along the amino acid backbone: the NH-CH, CH-CO and CO-NH bonds. Each bond breakage gives rise to two species, one neutral and the other one charged, and only the charged species is monitored by the mass spectrometer. The charge can stay on either of the two fragments depending on the chemistry and relative proton affinity of the two species. Hence there are six possible fragment ions for each amino acid residue and these are labelled in Figure 1.2, with the a, b, and c ions having the charge retained on the N-terminal fragment and the x, y, z ions having

the charge retained on the C-terminal fragment. The most common cleavage sites are at the CO-NH bonds which give rise to the b and/or the y ions (Figure 1.2). The mass difference between two adjacent b ions, or y ions is indicative of a particular amino acid residue (Biemann *et al.,* 1990) (see table 1.3).

Immonium ions (labelled "i" appear in the very low m/z range of the MS-MS spectrum. Each amino acid residue leads to a diagnostic immonium ion, with the exception of the two pairs leucine (L) and iso-leucine (I), and lysine (K) and glutamine (Q), which produce immonium ions with the same m/z ratio, i.e. m/z 86 for I and L, m/z 101 for K and Q. The immonium ions are useful for detecting and confirming many of the amino acid residues in a peptide, although no information regarding the position of these amino acid residues in the peptide sequence can be ascertained from the immonium ions.

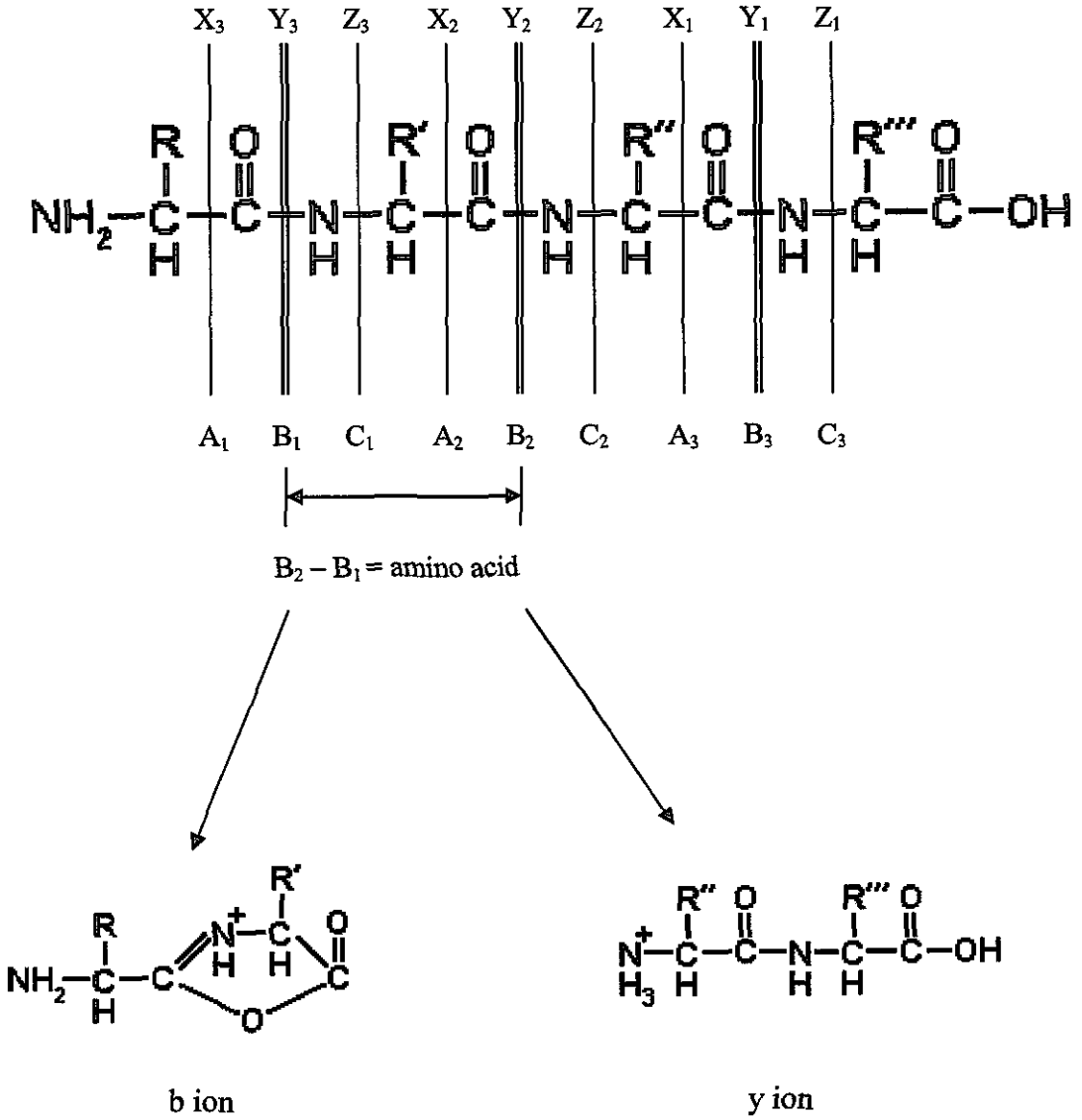



**Figure 1.2 Example of an amino acid backbone showing where the six possible fragments ions could be produced labelled as A, B, C ions having the charge retained on the N-terminal fragement and X, Y, Z ions having the charge retain on the C-terminal fragment. The mass difference between two adjacent b or y ions is indicative of a particular amino acid. Also, an example of the most common cleavage site producing b and y ions.**

| Letter | Name | Mass, Da |
|---|---|---|
| G | Glycine | 57.02 |
| A | Alanine | 71.04 |
| S | Serine | 87.03 |
| P | Proline | 97.05 |
| V | Valine | 99.07 |
| T | Threonine | 101.05 |
| C | Cysteine | 103.01 |
| I | Isoleucine | 113.08 |
| L | Leucine | 113.08 |
| N | Asparagines | 114.04 |
| D | aspartic acid | 115.03 |
| Q | Glutamine | 128.06 |
| K | Lysine | 128.09 |
| E | glutamic acid | 129.04 |
| M | Methionine | 131.04 |
| H | Histidine | 137.06 |
| F | Phenylalanine | 147.07 |
| R | Arginine | 156.10 |
| Y | Tyrosine | 163.06 |
| W | Tryptophan | 186.08 |
| | carboxymethyl cysteine | 161.05 |
| | carbamidated cysteine | 160.03 |
| | oxidised methionine | 147.04 |

**Table1.3 The amino acids and their respective Masses (Da).**

An example of an MS/MS product ions is illustrated in Figure 1.3. The molecular mass of the peptide was measured using the LTQ Linear Ion Trap in Full MS mode and found to be 667.6Da, the dominant ions in the MS spectrum being the protonated molecular ions ($M+H^+$) at m/z 668.6. These ions were selected for transmission through the first analyser to produce a MS/MS spectrum. The sequence (amino acid backbone) ions have been identified, the peptide fragmented predominantly at the CO-NH bonds and have both b and y ions. The b series ions are labelled with red vertical lines and the y series have been labelled with blue vertical lines. The mass difference between adjacent members of a series can be calculated e.g. y8 – y7 = 860.5 - 713.5 = 147 Da which is equivalent to a Phenylalanine (F) residue; and similarily b9 – b8 = 1005.6 – 908.5 = 97.1 Da which is equivalent to Proline (P) residue. In this way, using either the b series or the y series (y series reads from right to left and b from left to right), the amino acid sequence of the peptide can be determined and was found to be IMNTFSVVPSPK from the protein Beta Tubulin.

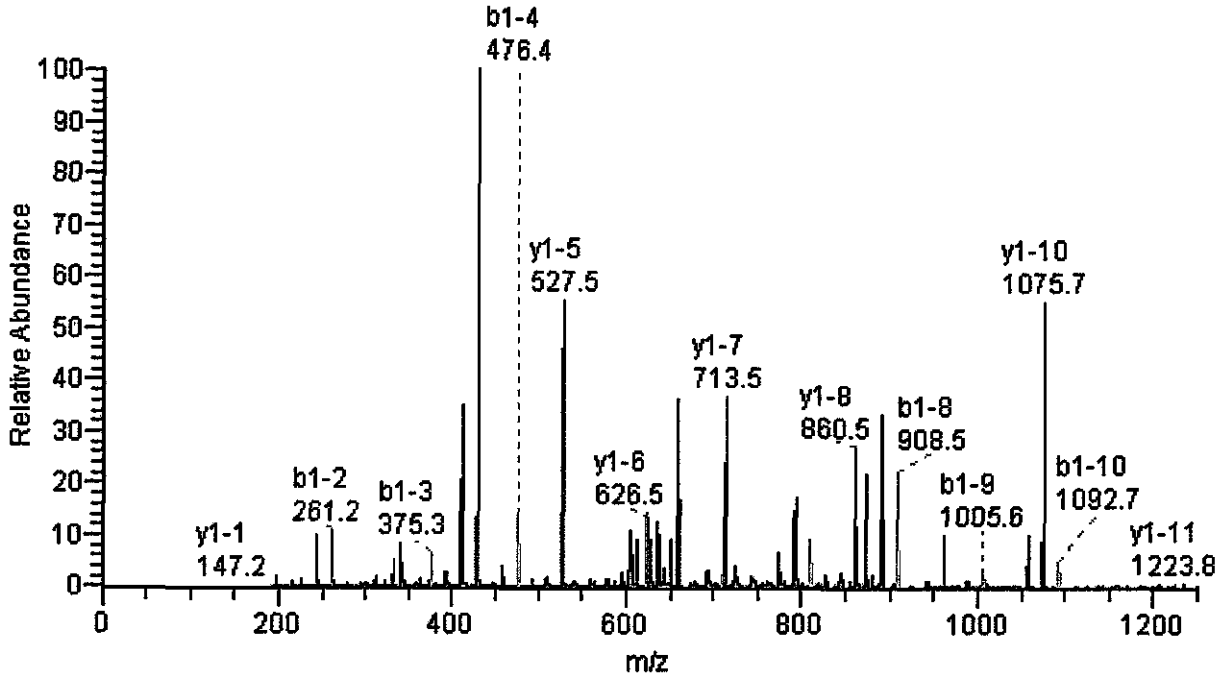



**Figure 1.3 MS/MS product ion spectrum using the LTQ Linear Ion Trap after fragmentation of a peptide taken from a tryptically digested protein.**

## 1.9 Combining liquid chromatography with mass spectrometry using ESI

Mass spectrometry is inherently well-suited for mixture analysis, and can be further augmented by electrospray ionisation, which enables liquid chromatographic techniques to be combined with tandem mass spectrometers, making it possible to assign structural identity to individual species in a complex mixture by chromatographic separation on their chemical properties followed by a separation based on mass/charge (m/z) ratios and subsequent structural characterisation in the MS/MS mode.

Low flow-rate electrospray (nano-LC-electrospray) can efficiently spray small volumes of aqueous solutions using pressure assisted flow rates of 200nl/min with fused silica capillaries. For most studies on peptides, the low flow rates used by small-bore liquid chromatography (LC) columns offer the greatest sensitivity for analysis (Griffin *et al.,* 1991). Consider a 100 femtomole ($10^{-15}$ mol) of protein isolated from a 2-DE gel. For simple analysis by ESI-MS/MS, the typical minimum molar concentration must range between 0.1 to 1.0μM. This means that the total sample must be contained within a volume no larger than 1μL as it flows through the end of the ESI emitter into the mass spectrometer. A 1mm ID microbore column, having a typical column volume of 200μL, flowing at 50μL/min, yields an elution volume of approximately 12μL for a fifteen second wide peak. Contrasted with a 75μm ID nanobore column, having a typical column volume of 0.5μL flowing at 200nL/min which yields an elution volume of approximately 50nL for a 15 second wide peak. Calculated on a per volume basis, a 75μm ID nanobore column yields a relative concentration improvement of over 3,700 fold compared to a conventional 3.6mm ID column (Tomer *et al.,* 2001).

In order to obtain primary structural information for peptides, it is necessary to acquire tandem-mass-spectral data on the chromatographic peaks that are eluted from the microcolumn LC onto the mass spectrometer. Two approaches are generally employed to acquire tandem-mass-spectral data for peptides. The first approach requires premeasurement of m/z values for all the components of the complex mixture, followed by a separate analysis of the same mixture to obtain tandem mass spectra. The shortcomings of this approach include high sample consumption and a requirement for considerable manual intervention. The second approach involved automated data acquisition by computer control of the instrument, a process that allows both MS and MS/MS data to be acquired in a single analysis.

A typical computer algorithm for the automated acquisition of tandem mass spectra would involve the following steps.

(1) scan a mass analyser to record m/z values of ions entering a mass analyser

(2) search this mass spectrum for an m/z value based on certain preset guidelines

(3) select the searched m/z value and subject it to MS/MS conditions that are calculated from the m/z value of the selected precursor ion

(4) acquire tandem mass spectra

(5) reset the instrument back to step 1

With these ionization methods mass accuracies below 10ppm (0.01Da) for a peptide with a mass of 1000 Da can be achieved.

## 1.10 *De Novo* sequencing using MS/MS spectra

It is possible to carry out *de novo* sequencing using the MS/MS spectra as the sole reference for deducing the sequence of the peptides they represent. All the information necessary to reconstruct a peptide sequence could possibly exist in the ions generated, although spectra of high enough quality are relatively rare. The advantages of using *de novo* methods are that they may be used when a reference database is simply not available, as in the case of organisms with an as yet unsequenced genome. *De novo* sequencing may also be helpful when attempting to identify peptides that have been post-translationally modified (Reinders *et al.*, 2004), which can prevent mass matching when searching a database. The primary challenge in *de novo* sequencing using MS/MS is to identify the type of ion that a peak in the spectrum represents. Building the spectrum is straightforward if each y ion peak can be confidently identified. Unfortunately often a complete enough set of reference ions need to create a full sequence is not always present. This may be due to ion suppression or poor signal-to-noise ratio in the spectrum.

## 1.11 Database searching to identify proteins

The peptide data (PMF and partial sequence data) are compared with the entire protein database (Swiss-Prot, TrEMBL, NCBI, ExPaSy) or with the deduced protein sequences from the gene database. The proteomics Tools of the Swiss Institiute of Bioinformatics (SIB) are dedicated to the analysis of protein sequence, structure and 2-DE, thereby all known protein sequences of cells and tissues of

various organisms are accessible. This search routine enables the unambiguous identification of proteins from 2-DE if the protein is known and listed in the 2-DE database. If only limited sequence coverage was achieved the protein identification becomes less safe. With unknown proteins, the identification may be derived from a comparison with homologous sequences from other organisms. If this yields ambiguous results more efforts to identify the protein has to be made, e.g. by preparing nucleotide probes for the identification of the corresponding gene or by complete protein sequence analysis.

The choice of protein cleavage agent affects the distribution of size of the resulting peptide fragment according to the specificity of its cleavage. Trypsin, which cleaves after arginine and lysine residues, unless followed by proline (Gattiker, *et al.,* 2002), produces peptides ranging from one residue long to theoretically any length. PMF is highly used and useful method for protein identification, but the lack of access to sequence data from the spectra makes its use in high-throughput peptide identification limited as mass is not a powerful enough discriminator when searching eukaryotic databases (Shadforth *et al.,* 2006). The use of MS/MS, however provides access to sequence data, which can be used to more confidently identify peptides. In all variants of MS/MS-based experiment, once the precursor ions have been selected, they are subject to collision, thus fragmenting them into a range of daughter ions as the amide backbone is broken. For any unique peptide, the complete spectrum of daughter ions is also unique. The series of y-ions that could be expected represents a chain of increasing m/z from the C-terminus, with each ion differing excluding N- and C-terminal groups at that position as successive amide linkages are broken, b-ions are complementary to these on the N-terminal side of the peptide ion upon fragmentation. Other ion types may also be created

through loss of water, $NH_3$, CO or immonium ion formation (Papayannopoulos *et al.*, 1992). Searching a protein sequence database with MS/MS spectra proceeds initially in a similar manner to PMF, with matches to the precursor ion mass within a potentially wide tolerance window being selected from the database. Following this, the experimental tandem spectrum is compared to a theoretical spectrum generated for each candidate peptide. Various scoring methods are then used to judge the validity of each match. The most used and widely known probability-based scoring algorithms are MASCOT and SEQUEST. MASCOT uses a probability-based scoring algorithm based upon MOWSE (Pappin *et al.*, 2003) PMF scoring method, but adapted to incorporate additional information such as signal intensity and consecutive fragment heuristics, provided by the tandem mass spectrum (Colinge *et al.*, 2003). SEQUEST uses a measure of cross-correlation, between the top 500 candidate theoretical spectra, chosen using essentially PMF identification and the experimental spectrum (Yates *et al.*, 1995).

## 1.12 Other methods to identify proteins

### 1.12.1 Edman Sequencing

Sequencing by Edman degradation. The Edman sequencing technique allows the unambiguously identification of proteins from 2-DE blots. 2-DE gels are blotted onto PVDF membranes that are developed with CBB for N-terminal Edman sequencing. In the blot procedure glycine and Tris buffers should be replaced by sodium borate buffers that result in cleaner Edman degradation cycles. Edman degradation takes place in a sequencer that automatically degrades proteins and peptides from their N-terminal end by a step-wise chemical procedure (Edman, *et al.*, 1967).

The reagent phenylisothiocyanate (PITC) is attached to the N-terminal end group of the peptide forming a PTC-peptide, which strongly acidic conditions cleaves off the first amino acid as aminoathiazolinone derivative (ATZ-amino acid) from the remaining peptide. The released amino acid is converted to the more stable phenylthiohydantoin (PTH-amino acid). Products are normally separated and identified by HPLC in a mirco RP C18 column. Blots are applied to the blot reaction cartridge of the sequencer. Peptides or proteins that are prone to be washed out during Edman degradation may be successfully sequenced by the addition of polymers (e.g Biobrene).

### 1.12.2 Internal sequencing

If the N-terminal amino acid is blocked then Edman degradation results in no sequence information. Therefore, it is recommended to perform internal sequencing of the protein spot of interest. Internal sequencing can be achieved by several different approaches. (1) the protein blot or gel piece is proteolytically digested by trypsin, which cleaves after lysines and arginines. The peptide fragments are eluted and separated by HPLC and then sequenced in the Edman sequencer; or (2) alternatively, the generated peptides are examined by MS. The advantage of the MS approach is that the peptide mixture can be applied directly to the MS system. Yet, application of Edman N-terminal or internal sequencing is useful, (1) when the identity and purification of a selected protein is questionable; (2) for identifying protein domains, obtained after limited proteolysis. Edman sequencing allows a safe identification of the N-terminal region(s) of protein domains in such experiments. Edman sequencing is highly valuable also for *de novo* sequencing of a protein if no sequencing information is available in databases or cannot be deduced from genetic data.

## 1.13 Proteomic Automated procedures

Due to the large number of samples which have to be analysed in high-throughput proteomics studies, there is an increasingly urgent need for automated procedures (Westermeier *et al.*, 2004). Despite earlier improvements such as the possibility to run twelve second dimension SDS gels in parallel, 2-DE has long been a laborious and difficult to automate procedure. Only recently, the situation has changed due to (1) the availability of ready-made gels (IPG DryStrips) on stable plastic supports, (2) the introduction of the IPGphor and similar devices for automated first dimensional IEF, (3) DIGE, (4) simplified image comparison and improved quantitation due to the possibility of including an internal standard, (5) improved algorithms and better computer programs for easier gel image analysis, (6) automation of spot excision and protein digestion and (7) high mass resolution and high mass accuracy of modern MS.

## 1.14 Other proteomic methodologies

### 1.14.1 Quantitative Proteomics using Isotope-coded affinity tag (ICAT)

ICAT reagents consist of conjugate molecules containing an affinity tag, an isotope tag, and a protein-reactive group such as iodacetamide for attachement to cysteine SG groups in proteins. A non-deuterated light (d0) and a deuterated heavy (d8) pair of ICAT reagents are used to determine the relative levels of proteins, e.g. in a complex protein mixture. See figure 1.4 for the chemical structure of the tag used for ICAT strategies (image from www.appliedbiosystems.com).

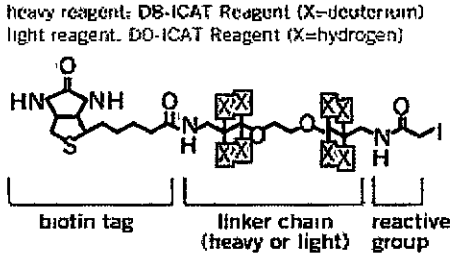



**Figure 1.4 Chemical structure of the tag used for ICAT**

Using labelling with ICAT reagents in combination with 2-DE it is possible to differentiate two cell states (Aebersold *et al., 2000).* After electrophoresis, the spots are excised and treated with trypsin. The tryptic peptides are analysed by ESI-MS/MS. Those peptides that contain both the normal mass and the isotope mass are derived from proteins that differ in the different situations. Comparison of the relative abundance of protonated versus deuterated ICAT peptide conjugates results in a mass difference of 8 Da between the two samples, allowing to select proteins

that are differentially expressed in cancer cells compared with normal (Smolka *et al.*, 2001). See Figure 1.5 for an example of a strategy for quantifying differential protein expression (image from Aebersold *et al., 1999).*

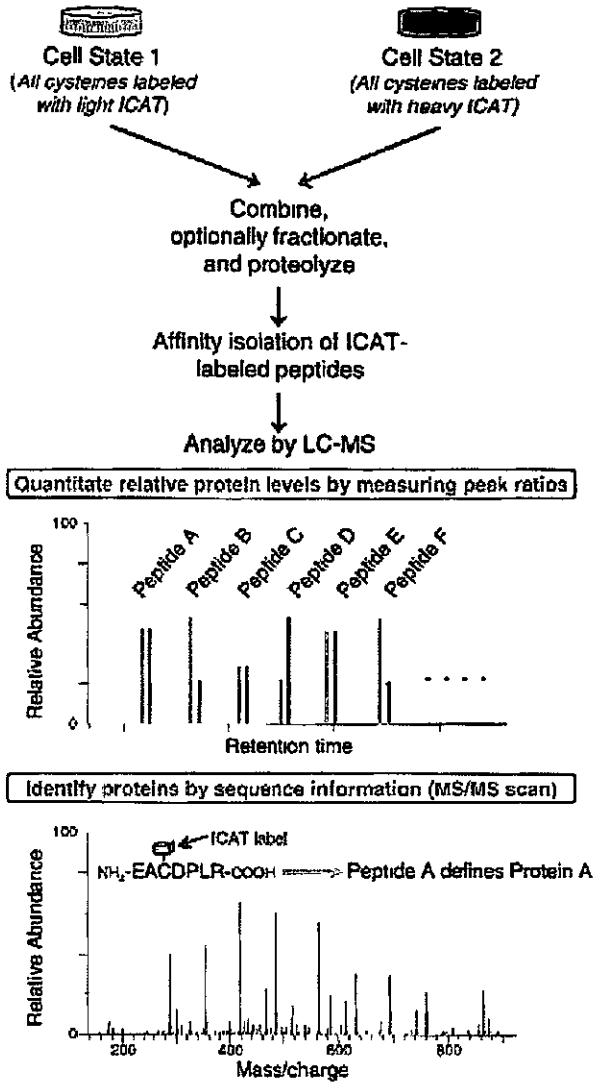



**Figure 1.5 Isotope-coded affinity tag has 3 elements, an affinity tag (biotin) that is used to isolate ICAT-labelled peptides, a linker that can incorporate stable isotopes and a reactive group with specificity toward thiol groups (cysteines). The reagent exists in 2 forms, a heavy (d8) and light (d0). ICAT strategy for quantifying differential protein expression. Two protein mixtures representing 2 different cell states are treated with d0 and d8 ICAT reagents respectively. ICAT reagent is covalently attached to each cysteinly residue in every protein. Protein mixtures are combined and proteolyzed to peptides. ICAT-labelled peptides are isolated by use of the biotin tag. These peptides are separated by microcapillary high-performance liquid chromatography. A pair of ICAT-labelled peptides is chemically identical and easily visualised because it coelutes. An 8- Da difference is measured by scanning mass spectrometer. Ratios of the original amounts of protein from the 2 cell states are strictly maintained in the peptide fragments. Relative quantification is determined by the ratio of the ion currents for the d0 and d8 tagged peptide pairs. Every other scan is devoted to fragmenting and then recording sequence information for the eluting peptide.**

A second generation of ICAT reagents have been developed such as the cleavable reagents (cICAT) and the visible-ICAT reagents (iTRAQ AppliedBiosystems) that react with amino-groups and contain isobaric tags with reported and balancer groups (Weise *et al.*, 2006). During CID-MS the reporter is cleaved off the peptide and releases an isotope series representing the quantity of a single peptide on known mass. This technology can be applied to compare total cell lysates without any 2-DE which are digested prior to ESI-MS/MS. It should be noted that quantification is limited if (1) the protein occurs in many species, (2) the reaction with the reagent tag is not quantitiative and (3) peptides of identical sequence occur from different proteins in the digest.

### 1.14.2 Surface Enhanced Laser Desorption Ionisation (SELDI) -ToF MS

A strategy for Biomarker discovery is the analysis of peptides and proteins present in plasma or blood serum by mass spectrometry. SELDI-ToF technology involves incubating samples on chips whose surface is coated with a protein-fractionating resin (e.g. C18 or strong cationic exchange). The unbound compounds are washed away, the chips are overlaid with an energy-absorbing matrix and finally spectra are required using laser ionisation and ToF separation mass spectrometry (Simpkins *et al.*, 2005). The resulting peak intensities are supposed to correlate with the concentrations of diverse peptides in the sample. Biomarker wizard software can analyses the spectral map and detects differentially expressed protein/peptides with statistical significance. Zhukov *et al.*, (2003) used SELDI to identify lung cancer biomarkers. Although SELDI technology is well documented for its ability of rapid

profiling of minute samples, the ability to identify proteins from SELDI remains challenging and conventional peptide mapping must be employed.

### 1.14.3 Multidimensional Protein Identification Technology (MuDPIT)

Multidimensional Protein Identification technology (MuDPIT) is a non-gel approach using multidimensional high-pressure liquid chromatography separation, tandem mass spectrometry and database searching (Koller *et al.,* 2002). MudPIT allows a rapid and simultaneous separation and identification of proteins and peptides in a complex mixture without the need for pre- or post- separation labelling which is not possible in 2D electrophoresis, ICAT and iTRAQ (Koller *et al.,* 2002). Following protein digestion the peptide fragments are allowed to separate in parallel with two-dimensional liquid chromatography using a strong cation exchange (SCX) column that separate peptides based on charge, and then by a reverse phase (RP) column based on hydrophobicity. The peptides are eluted directly into a tandem mass spectrometry for fragmentation followed by protein identification. One of the major weaknesses of MudPIT is in identifying quantitative differences in protein expression across protein mixtures (Washburn *et al.,* 2003). Durr *et al.,* (2004) observed that two replicate MuDPIT analyses produced two sets of protein identifications with ~65% overlap.

## 1.15 Bromodeoxyuridine (BrdU)

BrdU is a thymidine analogue capable of inducing epitheloid morphology and altering the expression of neuroendocrine markers in small cell lung carcinoma cell lines (Feyes *et al.,* 1991). The ability of BrdU to alter differentiation in neuronal, muscle and haematopoietic lineages is also well established (Esumi *et al.*, 1989) (Tapscott *et al.*, 1989) and (Yen *et al.,* 1987). BrdU competes with thymidine for incorporation in DNA, a property which has led to its widespread use in cell proliferation assays (Dolbeare *et al.*, 1996). While it appears that BrdU must be incorporated in DNA in order to modulate differentiation (Harding *et al.,* 1978), relatively little is known regarding the mechanism(s) of action involved. There is evidence that in some systems altered conformation of the BrdU-containing DNA may affect interactions with specific transcription factors, leading to either inhibition or induction of differentiation, depending on cell type (Lin *et al.,* 1972).

DLKP is a poorly differentiated squamous carcinoma cell line which does not express keratin proteins or other epithelial markers such as epithelial-specific antigen or desmosomal proteins: despite being of epithelial origins (McBride, *et al.,* 1998). DLKP cells may represent a stem cell like cell line of lung epithelial lineage which has the potential for both proliferation and differentiation. Also, the cells appear to be at a very early stage of differentiation, shown by the lack of expression of the simple epithelial keratins 8, 18 and 19, and can theoretically progress towards several different phenotypes. Treatment of DLKP with 10μM BrdU significantly altered the morphology of the cells (McBride *et al.,* 1999). In this study it was noted that within 4 days of treatment of DLKP with 10μM BrdU the cells showed an increase in cell

surface area when compared to untreated cells. There was an approximate 2 fold increase in cell size with the cells appearing stretched and flattened.

We have demonstrated in our laboratory that BrdU treatment of DLKP induces differentiation. McBride *et al* (1999) showed immunohistochemically, that treatment with 10μM BrdU over 7 days induced Keratin 8 and 18 protein (but with no change in keratin mRNA levels).

Meleady and Clynes (2001) demonstrated that BrdU treatment of DLKP resulted in increases of integrin $\alpha_2$ and $\beta_1$ subunit protein.

BrdU also alters the adhesive properties of DLKP with an increased adhesion to collagen Type IV, laminin, fibronectin and ECM gel (Meleady and Clynes 2001).

Walsh *et al* (2003) showed significant increased levels of eIF4E following exposure of DLKP to BrdU. The treatment also resulted in an increase in phosphorylated eIF4E.

## 1.16 Aims of Thesis

Previous work conducted in this laboratory has shown that treatment of the poorly differentiated lung cell line DLKP with the differentiation modulating agent BrdU caused increased expression of growth and differentiation related proteins. Protein expression of cytokeratins 8 and 18 and integrins $\alpha_2$ and $\beta_1$ were induced following BrdU exposure. No previous studies have been undertaken however, to examine global changes in protein profiles in this system.

The main aims of this thesis were as follows:

1. To optimise methods of protein identification from 2D colloidal stained gels using MALDI-ToF and nanoLC-Tandem Mass Spectrometry to achieve reliable high confidence protein identifications.
2. To examine the effects of BrdU on DLKP using 2-D DIGE technology for statistically significant comparisons of protein abundance using three biological replicates.
3. Using optimised Mass Spectrometry based protein identification methods to identify any differentially regulated proteins obtained by 2D-DIGE after BrdU exposure on DLKP.

## 2.0 Materials and Methods

### 2.1 Water

Ultrapure water was used in the preparation of all media and solutions. This water was purified by reverse osmosis (Millipore Milli-RO 10 Plus) and the further purified using the Elga Maxima Life Science to a standard of 18MΩ/cm resistance.

### 2.2 Glassware

Solutions pertaining to cell culture and maintenance were prepared and stored in sterile glass bottles. Bottles (and lids) and all other glassware used for any cell-related work were prepared as follows: all glassware and lids were soaked in a 2% (v/v) solution of RBS-25 (AGB Scientific) for at least 1 hour. This is a de-proteinising agent which removes proteinaceous materials from glassware. Following scrubbing and several rinses in tap water was then washed by machine using Neodisher GK (Dr.Weigert) detergent, an organic, phosphate-based acid detergent. Glassware was then rinsed twice with ultrapure water and sterilised by autoclaving.

## 2.3 Sterilisation

Water, glassware and all thermostable solutions were sterilised by autoclaving at 121°C for 20 minutes (min) under pressure of 15psi. Thermolabile solutions were filtered through 0.22μm sterile filters (Millipore, Millex SLGV-25BS).

## 2.4 Basal Media preparation

Medium was routinely prepared and sterility checked by Joe Carey in the NICB. The basal media used during routine cell culture were prepared according to the formulations shown in Table 2.1 10x media were added to sterile ultrapure water, buffered with HEPES and $Na_2HCO_3$ and adjusted to a pH of 7.45-7.55 using sterile 1.5M NaOH and 1.5M HCL. The media was then filtered through sterile 0.22μm bell filters (PALL 121-58), given a three month expiry date and stored in 500ml sterile bottles at 4°C. Sterility checks were carried out on each 500ml bottle of medium as described in Section 2.5.5.

Prior to use, 100ml aliquots of basal media were supplemented with 2mM L-glutamine (Gibco 25030-024) and 5% foetal calf serum (FCS) (Gibco) and this was used as routine culture medium. This was stored for up to 2 weeks at 4°C, after which time, fresh culture medium was prepared.

**Table 2.1 Preparation of basal media**

| | **Dulbecco's Modified Eagle Medium Nutrient Mixture F-12 Ham with L-Glutamineand 15mM HEPES*** (Sigma D8900) |
|---|---|
| **10x Medium** | Powder |
| **UHP Water** | 4700ml |
| **7.5% $NaHCO_3$ (BDH, 30151)** | 45ml |

* HEPES = N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)

Dulbecco's Modified Eagle Medium Nutrient Mixture F-12 Ham was supplemented with 5% FCS.

## 2.5 Cell lines

All cell culture was carried out in a class II down-flow re-circulation laminar flow cabinet (Nuaire Biological Cabinet) and any work which involved toxic compounds was carried out in a cytoguard (Holten). Strict aseptic techniques were adhered to at all times. The laminar cabinet was swabbed with 70% industrial methylated spirits (IMS) before and after use, as were all items used in the cabinet. Each cell line was assigned specific media and waste bottles and only one cell line was worked with at a time in the cabinet which was allowed to clear for 15min between different cell lines. The cabinet itself was cleaned each week with industrial detergents (Virkon, Antec, International; TEGO, TH.Goldschmidt Ltd.), as were the incubators.

The cell lines used during the course of this study, their sources and basal media requirements are listed in Table 2.5.1. Cell lines were maintained in 75 $cm^2$ flasks non-vented (Costar;3075) and 175 $cm^2$ flasks non-vented (NUNC;156502)

### 2.5.1. Subculture of Adherent Lines

During routine subculturing or harvesting of adherent lines, cells were removed from their flasks by enzymatic detachment.

Waste medium was removed from the flasks and rinsed with a pre-warmed (37°C) trypsin/EDTA (TV) solution (0.25% trypsin (Gibco, 25090-028), 0.01% EDTA (Sigma, EDS) solution in Phosphate Buffered Saline (Dulbecco A) (Oxoid, BR0014G). The purpose of this was to inhibit any naturally occurring trypsin inhibitor which

would be present in residual serum. Fresh TV was then placed on cells (4ml/75cm$^2$ flask or 10ml/175cm$^2$flask) and the flasks incubated at 37 °C until the cells were seen to have detached (5-10min). The trypsin was deactivated by addition of an equal volume of growth medium (*i.e.* containing 5% serum. The entire solution was transferred to a 30ml sterile universal tube (Sterilin 128a) and centrifuged at 1,000rpm for 5min. The resulting cell pellet was resuspended in pre-warmed (37°C) fresh growth medium, counted (Section 2.5.3.) and used to re-seed a flask at the required cell density or to set up an assay.

**Table 2.2** Preparation of basal media

| Cell Line | Basal Medium | Cell Type |
|---|---|---|
| | | |
| DLKP | ATCC | Human non-small-cell lung carcinoma |
| | | |

### 2.5.2 Cell Counting

Cell counting and viability determinations were carried out using a trypan blue (Gibco, 1520-012) dye exclusion technique.

An aliquout of trypan blue was added to a sample from a single cell suspension in a ratio of 1:5. After 3 min incubation at room temperature, a sample of this mixture was applied to the chamber of a haemocytometer over which a glass coverslip had been placed. Cells in the 16 squares of the four outer corner grids of the chamber were counted microscopically, an average per corner grid was calculated with the dilution factor being taken into account and final cell numbers were multiplied by $10^4$ to determine the number of cells per ml (volume occupied by sample in chamber is 0.1cm x 0.01cm *i.e.* 0.0001$cm^3$ therefore cell number x
$10^4$ is equivalent to cell per ml). Non-viable cells were those which stained blue while viable cells excluded the trypan blue and remained unstained.

### 2.5.3 Cell Freezing

To allow long term storage of cell stocks, cells were frozen and cryo-preserved in liquid nitrogen at temperatures below -180°C. Once frozen properly, such stocks should last indefinitely.

Cells to be frozen were harvested in the log phase of growth (*i.e.* actively growing and approximately 50-70% confluent) and counted as described in Sections 2.5.2. Pelleted cells were re-suspended in serum and an equal volume of a DMSO/Serum (1:9, v/v). This solution was slowly added dropwise to the cell suspension to give a final concentration of at least $5x10^6$ cells/ml. This step was very important as DMSO is toxic to cells. When added slowly the cells had a period of time to adapt to the presence of the DMSO, otherwise cells may have lysed. The suspension was then aliquoted into cryovials (Greiner, 122278) which were then quickly placed in the vapour phase of liquid nitrogen containers (approximately -80°C). After 2.5 to 3.5 hours, the cryovials were lowered down into the liquid nitrogen where they were stored until required.

### 2.5.4 Cell thawing

Immediately prior to the removal of a cryovial from the liquid nitrogen stores for thawing, a sterile universal containing growth medium was prepared for the rapid transfer and dilution of thawed cells to reduce their exposure time to DMSO freezing solution which is toxic at room temperature. The suspension was centrifuged at 1,000 rpm for 5 min, the DMSO-containing supernatant removed and the pellet re-suspended in fresh growth medium. A viability count was carried out (Section 2.5.2.) to determine the efficacy of the freezing/thawing procedures. Thawed cells were placed into tissue culture flasks with the appropriate volume of medium and allowed to attach overnight. After 24 hours, the cells were re-fed with fresh medium to remove any residual traces of DMSO.

### 2.5.5 Sterility Checks

Sterility checks were routinely carried out on all media, supplements and trypsin used for cell culture. Sample of basal media were inoculated into Tryptone Soya Broth (Oxoid CM0129) and Thioglycollate Broth (Oxoid CM173) and were incubated at 25°C and 35°C respectively. Tryptone Soya Broth when incubated at 25°C will detect for the presence of fungus while Thioglycollate Broth when incubated at 35°C will detect for the presence of aerobic and anaerobic bacteria. Growth media (*i.e.* supplemented with serum and L-Glutamine) were sterility checked at least 2 days prior to use and were subsequently examined for turbidity and other indications of contamination.

## 2.6 *Mycoplasma* Analysis

*Mycoplasma* examinations were carried out routinely (at least every three months) on all cell lines used in this study.

### 2.6.1 Indirect staining Procedure

In this procedure, *Mycoplasma*-negative NRK cells (a normal rat kidney fibroblast line) were used as indicator cells. As such, these cells were incubate with supernatant from test cell lines and then examined for *Mycoplasma* contamination. NRK cells were used for this procedure because cell integrity is well maintained during fixation. A fluorescent Hoescht stain (Sigma H6204) was utilised which bind specifically to DNA and so will stain the nucleus of the cell in addition to any *Mycoplasma* DNA present. A *Mycoplasma* infection would thus be seen as small fluorescent bodies in the cytoplasm of the NRK cells and sometime outside the cells.

NRK cells were seeded onto sterile coverslips in sterile Petri dishes at a cell density of $2x10^3$ cells per ml and allowed to attach overnight at 37°C in a 5% $CO_2$ humidified incubator. 1ml of cell-free (cleared by centrifugation at 1,000rpm for 5 min) supernatant from each test cell line was then incubated onto a NRK Petri dish and incubated as before until the cells reached 20-50% confluency (4-5 days). After this time, the waste medium was removed from the Petri dishes, the coverslips washed

twice with sterile PBS, once with a cold (0-4°C spark proof fridge for 15min) PBS/Carnoys (50/50) solution and fixed with 2ml of chilled (-20°C spark proof freezer for 15min) Carnoys solution (glacial acetic acid:methanol-1:3) for 10 min. The fixative was then removed and after air drying, the coverslips were washed twice in UHP water and stained with 2ml of Hoechst for 10 min.

From this point on, work proceeded in the dark to limit quenching of the fluorescent stain.

The coverslips were rinsed three times in PBS. They were then mounted in 50% (v/v) glycerol in 0.05M citric acid and 0.1M disodium phosphate and examined using a fluorescent microscope with a UV filter.

## 2.7 Proteomic techniques

### 2.7.1 BrdU experiment

DLKP cells (passage 32) were inoculated into 175cm$^2$ flasks at a density of 5x10$^5$ cells per flask and were incubated for two days at 37°C. BrdU (sigma, B5002) powder was reconsitiuted in UHP water to a stock concentration of 10mM and the resultant solution was filter sterilised through a sterile 0.22μm filter (Millipore SLGV), aliquotted into sterile vials and stored at -20°C. BrdU at a concentration of 10μm was then added to the cells. The medium was replaced with fresh drug containing medium 4 days later. After 7 days of exposure to BrdU the cells were harvested.

### 2.7.2 Protein Preparation for 2D-electrophoresis

The cells were removed from adherent culture by trypsinisation. Trypsin was neutralized by addition of media, and cells were subsequently washed twice in cold PBS and twice with cold sucrose-tris (10mM sucrose, 100mM Tris, pH 8.0) buffer. Cells were lysed using a buffer containing (4% (w/v) CHAPS, 7M Urea, 2M Thiourea, 10mM Tris-HCL, 5mM Magnesium Acetate pH 8.5 plus DNase) at a concentration of 3 x 10$^7$ cells per ml. Cells were passed through a sterile 18 gauge syringe until the solution became homogenous and non-viscous and was then incubated at room

temperature for 1 hour. Insoluble material was removed by centrifugation at 14,000rpm for 20mins at 4°C. Protein concentration was determined using the Bradford protein assay kit (BioRad) as per manufactures instructions. The protein lysates were stored at -80°C until used.

### 2.7.3 Cy-Dye labeling

Cell lysates prepared as described in section 3.10.2 were labeled with N-Hydroxy succinimidyl ester-derivatives of the cyamine dyes Cy2, Cy3 and Cy5 as follows. Typically, samples yielded 2-10μg/μl protein and 50μg of protein from total cell lysate were minimally labeled with 1μl solution containing either 200pmol of either Cy3 or Cy5 per comparison on the same 2D gel. Labelling reactions were performed on ice in the dark for 30mins and then quenched with a 1μl of 10mM free lysine per 200pmol of Cy Dye for 10mins on ice. A pool of all samples was also prepared and labeled with Cy2 to be used as a standard on all gels to aid image matching and cross-gel statistical analysis. The Cy3 and Cy5 labelling reactions (50μg of each) from each lysate were mixed and run on the same gels with an equal amount (50μg) of Cy2-labeled standard (Alban *et al,* 2003). In each 2D-DIGE experiment a total of 3 biological replicates were used for each group. Each biological replicate was run as technical duplicates.

### 2.7.4 Protein Separation by 2-DE

Immobilized 18cm linear pH gradient (IPG) strips, pH4-7, were reyhdrated in rehydration buffer (7M Urea, 2M Thiourea, 4% CHAPS, 0.5% IPG Buffer, 50mM DTT) overnight, according to the manufactures guidelines. Isoelectric focusing was performed using an IPGphor apparatus (GE Healthcare) for a total of 48kV/h at 20°C, 50mA. Strips were equilibrated for 20mins in 50mM Tris-HCL, pH8.8, 6M Urea, 30% (v/v) glycerol, 1% (w/v) SDS containing 65mM DTT and then for 20mins in the same buffer containing 240mM Iodacetamide. Equilibrated IPG strips were transferred onto 18cm x 20cm 12.5% uniform polyacrlamide gels poured between low fluorescence glass plates. Strips were overlaid with 0.5% (w/v) low melting point agarose prepared in 2X concentration SDS running buffer containing bromophenol blue. Gels were run using the Ettan Dalt 12 (GE Healthcare) at 2.5W/gel for 30mins then 100W at 10°C until the dye front had run off the bottom of the gels. All images were collected on a Typhoon 9400 Variable Mode Imager (GE Healthcare).

### 2.7.5 Image Acquisition and Analysis

The Cy-dye labeled gels were visualised using a Typhoon 9410 (GE Healthcare). Excitation and emission wavelengths were chosen specifically for each of the dyes according to manufacturer's recommendations. Images were preprocessed to remove areas extraneous to those of interest using ImageQuant v6.5 (GE Healthcare). Intra-gel analysis was performed using DeCyder DIA (Difference In-gel Analysis) v6.5 (GE Healthcare). Inter-gel matching and statistical analysis were performed using DeCyder BVA (Biological Variance Analysis) v5.0 (GE Healthcare). The DeCyder DIA program uses a normal distribution model to determine the differentially expressed spots. Each protein in the individual sample is represented in the pooled standard, thus a comparison between the test sample and the identical protein in the standard can be used to generate a ratio of relative expression.

### 2.7.6 Preparation of plates for spot picking

Low fluorescent plates were polished with lint free wipes soaked in methanol and allowed to dry at room temperature for 10mins. During this time the Bind Silane solution (8mls Ethanol, 200µl Glacial Acetic Acid, 10µl Bind-Silane, 1.8mls UHP) was prepared. 1.5mls of the Bind Silane solution was applied evenly using a lint free wipe to one side of each glass plate. The glass plates were placed in a rack for 90mins

taking note of which side Bind Silane had been applied and were covered with aluminium foil to prevent dust particulates adhering to the surface of the plates. Plates were again polished with a lint free wipe. The Bind Silane plates were place together with their accompanying space plate such that the Bind-Silane surface is in contact with the spacers on the spacer plate. For preparative gels proteins samples were prepared as described in section 3.10.2. A total of 300μg of protein was loaded on to each preparative gel. Loading of protein samples onto IEF strips was the same as with loading labeled protein samples as described in section 3.10.4.

### 2.7.7 Silver staining of 2D electrophoresis gels

Cell samples for silver staining were prepared as for DIGE analysis but are not labelled with Cy dyes. Samples are separated by 2D Gel Electrophoresis as for Cy Dye labelled proteins and processed as follows:

(1) Fixing: gels are placed in fixing solution for a minimum of 30 minutes.

(2) Washing: Using the outlet on the gel box, the fixing solution was carefully drained out. 150ml of water was added to rinse and the gel was placed on a belly dancer for five minutes. The gel was rinsed a total of three times. Sensitizing: After the third wash, the gel box was drained and approximately 200ml of sensitizing solution was added before returning to the belly dancer for 30 minutes.

(3) Washing: Then using the method outlined above, three 10 minute washes were carried out with distilled water. Silver Reaction: Following the silver step, a further two washes were carried out. Each wash was a minimum of 5 minutes.

(4) Developing: 200ml of developer was added to the gel and allowed to develop on the belly dancer until protein spots appeared, something in the region of 20 minutes.

(5) Stopping: When the desired amount of spots had appeared on the gel, the developing was stopped with the addition of 200ml of stopping solution and the gels returned to the belly dancer for 10min.

(6) Storage: For medium to long-term storage the gels were kept in distilled water. All equipment used in the procedure was cleaned using warm water and detergent before rinsing with distilled water.

See table 2.3 below for reagent preparations.

**Table 2.3** Preparation of silver staining required for 6 gels

| Step | Reagent | Duration |
|---|---|---|
| Fixer | 400ml methanol / 100ml acetic acid / 500ml distilled water | > 1 hour |
| Sensitiser | 300ml of Ethanol / 2g of Sodium Thiosulphate / 68g Sodium Acetate / 700ml distilled water | 30 min |
| Silver | 0.1% Silver nitrate / 0.5ml (37%) Formaldehyde / 1000ml distilled water | 30 min |
| Wash | Distilled water | 3 x 10 min |
| Developer | 25g of Sodium Carbonate / 1000ml distilled water | ~ 20 min |
| Stopper | 14.6g $Na_2EDTA$ / 1000ml distilled water | 10 min |

### 2.7.8 Colloidal staining of 2D-Gels, Imaging and Spot matching to BVA

Preparative 2D-gels were prepared as described in section 3.10.4 were removed from the DALT 12. The spacer plates were removed from the plate/gel sandwich using wedging action with the "wonder wedge" (GE Healthcare). The spacer plate was removed from cleaning later. The polyacrylamide gel fixed to the low fluorescent plate was allowed to fix in a solution of 7% Glacial Acetic Acid, 40% Methanol for at least 60mins in a plastic container under agitation. Brilliant Blue G-Colloidal concentrate (Sigma B2025) was diluted in UHP water at a ratio of 1:4. Immediately before staining the colloidals were activated with Methanol. This solution was mixed with severe agitation for 30 seconds as it is important in order to activate the colloidals. The fixing solution was removed from the gels and activated colloidal solution was poured on each gel. Gels were allowed to stain for 1-4 hours under vigorous agitation. Staining was complete when the gels appeared dark blue and highly abundant protein spots are clearly visible. Decant colloidal staining solution to a waste bottle and rinse the gel and container with UHP to remove as much of the remaining stain as possible. Destaining solution composed of 10% Acetic Acid, 25% Methanol and 65% UHP (v/v) was poured onto each gel and was agitated for 30 seconds. The gels were rinsed with 25% Methanol and allowed to destain in 25% Methanol for up to 24 hours. Destaining was stopped as soon as the gel background became clear. Two reference marker stickers were placed on the back of each plate and then plate was scanned using Lab Scan Software. The image was imported into Image Master and spots were detected. The location of the reference markers were assigned (IR1 for left reference and IR2 for

right reference marker). A single DIA file was created for the detection of spots and reference marker locations. Spot map was exported and imported into the BVA file created for the DIGE experiment for which the preparative gel was part of. Matching mode was selected in the BVA experiment file and landmark mode was activated. Obvious highly abundant spots present in distinct protein patterns were matched manually in this mode. The software was then instructed to match image to experiment using these landmarks. Pick spots were displayed in match mode gel images by alteration of image preferences. Matching of these protein spots between DIGE images prep gel images were checked individually and proteins that were matched incorrectly were corrected. Where the correct spot could not be found the spot was left unmatched. A pick list was then exported to the Spotter.

## 2.8 Proteomics Automated Workflow

The automated proteomic workflow used in our laboratory includes a logical sequence of picking, digestion, spotting, MS analysis and automatic protein database searching. The robots used for the first three parts of this process are ETTAN robots (GE Healthcare) the output files provided at each stage are directly usable with each robot however they are not in a format directly usable for the Ettan Pro MALDI-TOF MS or nanoLC ESI.

Automated spotting of peptide sample and matrix was carried out using the ETTAN spotter with the dried-droplet method. The peptide sample was applied to the target slide and was mixed with an equal volume of matrix.

The result of automation is the fact that experiments are conducted in exactly the same way each time, samples are handled in the same conditions, and the results are therefore more consistent.

The Ettan Pro MALDI-TOF has a high degree of automation allowing efficient detection of peptide fragment ions that can automatically calibrated and inputted into ProFound database search engine.

### 2.8.1 Spot picking of protein spot gel plugs from 2D-gels

Preparative gels stained as described in section were placed in a plate holder on the Ettan spot picker robot (GE Healthcare). Plates were positioned so that the reference markers would appear between the two white strips on the plate holder. The plate was secured in position by positioning of clamps on the edges of the plates and tightened with screws. The solution inlet pipe was removed from the 20% Ethanol storage bottle and placed in LC-MS grade water bottle (Sigma). The line was primed 20 times to replace all 20% Ethanol in the line. The pick list generated was imported into the spot picker software and the robot was instructed to automatically detect the reference markers. Upon detection of the reference markers the instrument instructs the user to insert an appropriate number of microtitre plates on the plate holder stand. A directory file is then created allowing for the tracking of spots along the processes of picking, digestion and spotting. Microtitre plates containing gel plugs were sealed with foil wrapping and stored at 4°C.

### 2.8.2 Destaining of gel plugs and protein digestion

Plates ready for digestion were placed in appropriate positions in the Robot Digestor (GE Healthcare). The files corresponding to the microtitre plate placed in the robot were uploaded onto the Digestor software. The water surrounding the gel plugs were removed and the gel plugs were destained with three twenty minute washes with 50% Methanol containing 20mM Ammonium Bicarbonate (SIGMA). The gel plugs were then given two fifteen minute washes with 70% Acetonitrile (SIGMA). The gel plugs were allowed to air dry for a minimum of 60 minutes. A working stock of Sequence grade trypsin (Promega V5280) was prepared. The trypsin was reconstituted by the addition of 200µl of Acetic Acid buffer supplied by the manufacturer. Working stock of this trypsin solution was made by the addition of 1.4ml of a 40mM Ammonium Bicarbonate, 10% Acetonitrile solution in LC-MS grade water resulting in a 10µg/ml trypsin solution. The robot was instructed to add 10µl of this 10µg/ml trypsin solution to each gel plug. The microtitre plate was wrapped in foil and incubated at 37°C overnight.

The following day the microtitre plate was removed from the incubator and returned to the digestor robot. A new microtitre plate was placed directly below the exisiting plate and the robot was instructed to carry out to two twenty minute peptide extractions with 50% Acetonitrile, 0.1% Trifluoroacetic Acid (Sigma). Upon transfer of the protein digests to the new microtitre plate, the plate was then placed in a SpeedVac Concentrator (Thermo) for 60 minutes to desiccate the peptides. Once the plate was fully dried the plates were stored at -20°C.

## 2.9 Preparation for MALDI-ToF Analysis

### 2.9.1 Direct sample spotting

Dried peptides were reconstituted in 3µl 0.1% TFA and 0.4µl was spotted onto MALDI ToF target slides with an equal volume of matrix (saturated α-cyano-4-hydroxycinnamic acid (LaserBio Labs) in 50%Acetonitrile/0.1%TFA) using Ettan Spotter (GE Healthcare).

All samples were allowed to air-dry at room temperature.

### 2.9.2 Sample desalting and concentration using $C_{18}$ Zip Tip (Millipore)

Peptides were resuspended by adding 10µL of wash solution (0.1%TFA) to each well. After attaching a $ZipTip_{\mu\text{-}c18}$ pipette tip onto a 10µL pipettor the tips were prewet by aspirating the wetting solution (100% Acetonitrile) two times and this was dispensed to waste. The tip was wetted again with a second wetting solution (50% Acetonitrile) two times and this was dispensed to waste. Th ZipTips were then equilibrated by aspirating wash solution (0.1%TFA) through the tip a total of ten times. Peptides were bound to ZipTips by aspirating and dispensing (cycling) the resuspended digests through the tips a total of five times. The tips were then washed by aspirating wash solution (0.1%TFA) and dispensing the eluate to waste a total of five times. The purified peptide digests were eluted from the ZipTips by aspirating 1µL of elution buffer

(50%Acetonitrile+0.1%TFA containing saturated α-cyano-4-hydroxyxinnamic acid (LaserBio Labs)). MALDI matrix solution and this solution was then applied directly onto a MALDI target. A total of three elutions were carried out on all peptide samples bound to ZipTips with the final elution mixed with a commercially available peptide standard mixture (Pep4 Laser BioLabs).

## 2.10 MALDI TOF MS

Peptide-mass fingerprinting was performed using Ettan MALDI-ToF Pro. The instrument setting was reflector mode with 175 ns delay extraction time, 60-65% grid voltage and 20kV accelerating voltage. Laser shots at 250 per spectrum were used to acquire the spectra with mass range from 600-3000Da with 8-shot mode. A total of three spectra were generated for each sample analysed. MALDI-generated mass spectra were internally calibrated using trypsin peaks or commercially available peptide standards (Pep4 Laser BioLabs). The peptide masses were searched against the National Centre for Biotechnology Information (NCBI) *Homo-sapiens* database using ProFound$^{TM}$. One missed cleavage per peptide was allowed and an initial mass tolerance of 100ppm was used in all searches. Fixed carbamidomethylation of cysteines and partial oxidation for methionine was assumed.

MALDI-time of flight (ToF) MS was carried out using Ettan pro MALDI-ToF Mass Spectrophotometer (GE Healthcare). Peak lists were generated and calibrated against internal trypsin autolysis peaks (m/z 842.501 and 2211.104) and or commercially available peptide standards (Pep 4 LaserBio Lab). PMF identifies a protein by measuring the molecular mass of all major trypsin products and matching these molecular masses using the NCBI database with ProFound. Proteins were identified using ProFound with a mass tolerance of +/- 0.2 m/z, a maximum of 1 missed cleavage, fixed carbamidomethylation of cysteines and partial oxidation of methionines.

## 2.11 nanoLC-MS/MS

Tryptically digested samples were resuspended in 20μL of a LC-MS grade water containing 0.1%TFA and analysed by one-dimensional LC-MS using the Ettan™ MDLC system (GE Healthcare) in highthroughput configuration directly connected to a Finnigan™ LTQ™(Thermo Electron). Samples were concentrated and desalted on RPC trap columns (Zorbax™ 300SB C18, 0.3mm x 5mm, Agilent Technologies) and the peptides were separated on a nano-RPC column (Zorbax™ 300SB C18, 0.075mm x 100mm, Agilent Technologies) using a linear acetonitrile gradient from 0-65% Acetonitrile (Riedel-de Haën LC-MS grade) over 180 minutes directly into the LTQ via a 10μm nanoESI emitter (Presearch FS360-20-10-CE-20). Thermo linear ion trap mass spectrometer was used for MS/MS. A scan time of ~0.15 s (one microscans with a maximum ion injection time of 10ms) over an m/z range of 300-2000 was used followed by MS/MS analysis of the 3 most abundant peaks from each scan which were then excluded for the next 60 seconds followed by MS/MS of the next next three abundant peaks which in turn were excluded for 60 seconds and so on. A "collision energy"setting of 35% was applied for ion fragmentation and dynamic exclusion was used to discriminate against previously analysed ions (data dependent analysis).
All buffers used for nanoLC separations contained 0.1% Formic Acid (Fluka) as the ion pairing reagent. Full scan mass spectra were recorded in profile mode and tandem mass spectra in centroid mode. The peptides were identified using the information in the tandem mass spectra by searching against SWISS PROT database using SEQUEST™ using an Xcorr value of 1.5 for singly charged peptides, 2.0 for doubly charged peptide and 2.5 for triply charged peptides.

## 3.0 Results

Previous studies in this laboratory have shown that 10μM 5,2'-Bromodeoxyuridine (BrdU) induces cytokeratin 8, 18 and 19 protein expression in the DLKP cell line (McBride *et al,* 1999). Induced $\alpha_2$-,$\beta_1$-integrin has also been observed (Meleady and Clynes 2001) and elevated levels of eIF4E (Walsh *et al,* 2003) .

In this study we ultilised 2-dimensional DIGE gel electrophoresis followed by Mass Spectrometry (MS) analysis to identify and quantify specific protein expression changes in human lung carcinoma cell line DLKP following exposure to the differentiation-modulating agent 5-bromodeoxyuridine.

### 3.1 5 day treatment of DLKP with 10μM BrdU

DLKP cells (passage 32) were seeded in 175cm$^2$ flasks at a concentration of 5x10$^5$ cells per flask. Three biological replicates consisting of 2 control (untreated) and 8 (treated) flasks per replicate were set up in three separate 37°C incubators on three consecutive days labeled A, B and C. Two days after each replicate set up, 10μM BrdU was added to each of the 8 treated labeled flasks but not to the 2 control labeled flasks. The medium in each flask was replaced with fresh10μM BrdU four

days later in the case of the treated flask while the control had its media renewed. After 7 days the cells were harvested (materials and methods 2.7.1.).
Following harvesting, the cells were prepared for 2DE analysis (materials and methods 2.7.2) and stored at -80°C until use.

## 3.2 Protein Separation by 2-DE

2-DE protein separation was carried out on the control and BrdU treated samples. Sample were separated in the 1st dimension (IEF) over the pH range of 4-7 and then in the 2nd dimension (materials and methods 2.7.4)

## 3.3 Quality control of protein lysates

Prior to analysis by DIGE, all protein extracts were analysed by 2D electrophoresis followed by silver staining (materials and methods section 2.7.7) to ensure extracts used were good quality and had not been adversely affected by protein extraction (figure 3.1)

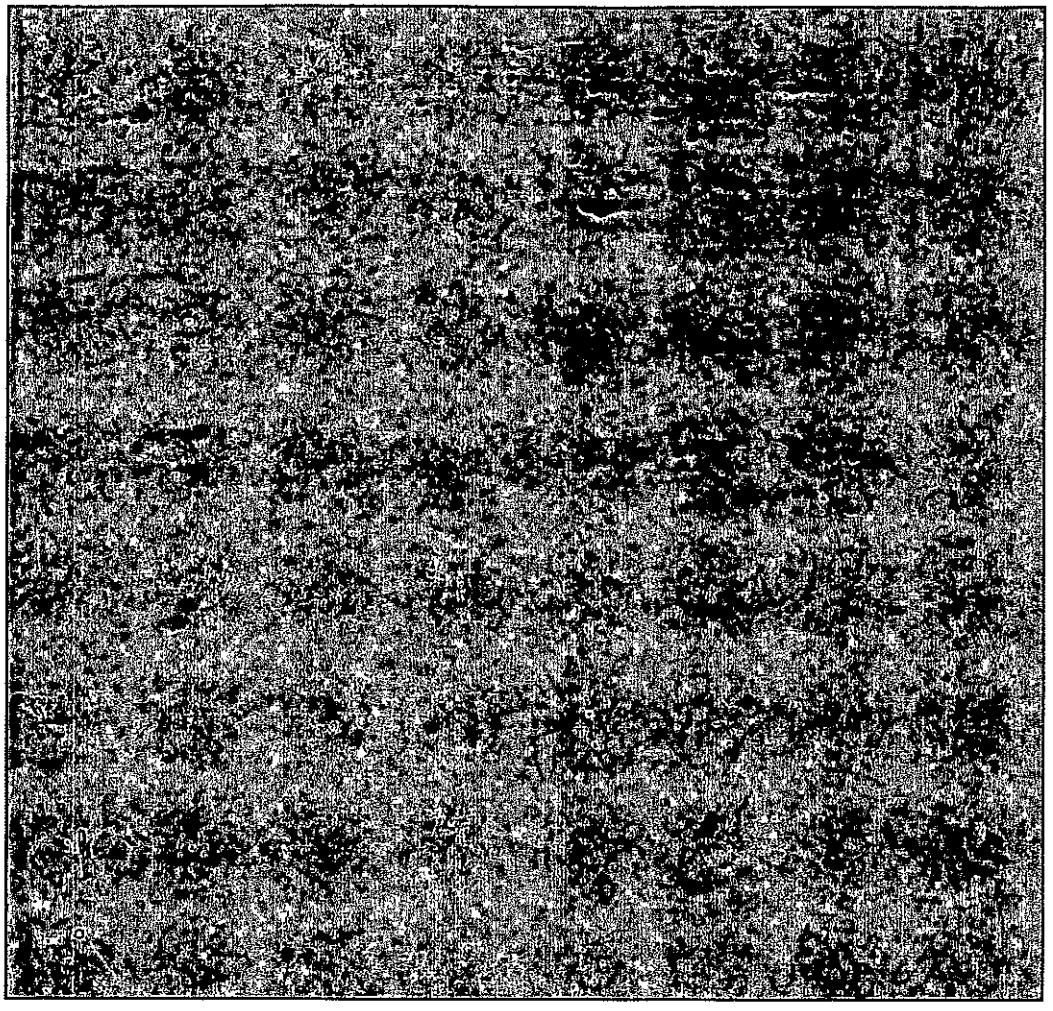


**Figure 3.1 All protein lysates for DIGE analysis were separated by IEF over the p*I* range 4-7 and run on 2D electrophoresis gels and visualised by silver staining to ensure the quality of protein samples. An example of a good quality sample showing clear well defined protein spots with no sign of breakdown is shown above**

## 3.4 DIGE Analysis of Control versus BrdU Treated DLKP

The experimental design for DIGE analysis of DLKP control versus BrdU treated is outlined in table 1. Three biological replicates of each control and test samples were used for the DIGE analysis and labeled with Cy-Dyes (see materials and methods 2.7.3.). Each gel contained 50µg of control, 50µg test sample and 50µg of internal control. Each gel was run in duplicate, resulting in a total of 6 gels in the experiment (table 3.1).

Stock Cy dye at 1nM/µl

Working stock at 200pm/µl

| Gel | Cy2 Internal Standard | Cy 3 Control | Cy 5 test |
|---|---|---|---|
| 1 | Pool | Control A | Test A |
| 2 | Pool | Control B | Test B |
| 3 | Pool | Control C | Test C |
| 4 | Pool | Control A | Test A |
| 5 | Pool | Control B | Test B |
| 6 | Pool | Control C | Test C |

**Table 3.1 Experimental Design for 2D-DIGE comparison of control versus BrdU-treated DLKP.**

## 3.5 Image Analysis

DIGE gels were scanned with a Typhoon 9400 scanner (GE Healthcare) producing a Cy2, Cy3 and Cy5 image for each gel at excitation wavelengths of 488, 532 and 633nm and emission wavelengths of 520, 580 and 670 nm respectively. Three images corresponding to three samples labeled with three different CyDyes are generated in one gel (figure 3.2). Image analysis was carried out with DeCyder 6.5 software (GE Healthcare) using the Batch Processor facility. The Batch Processor automatically analysed the scanned images first using the Differential In-gel Analysis (DIA) module where the Cy2, Cy3 and Cy5 image from each gel was normalised and the Cy3 and Cy5 samples compared with the Cy2 internal standard. Histograms are generated by DeCyder DIA software, plotting spot frequency (left y-axis) against the log volume ratio (x-axis) and a normal distribution model was fitted to the main peak of the frequency distribution. A 1.5 standard deviation model was set as threshold. The threshold values are represented by vertical black lines in the histogram, and the differentially expressed spot features are shown as blue spots and red spots (figure 3.3). The DIA is only allowed for single pairwise comparisons and cannot be used to analyse the data generated by all gels for group statistical analysis. The number of increased expressed spot, decreased expressed spot, similar expressed spot, and threshold value of control and treated are shown in table 3.2. To identify differentially expressed protein spots across these 6 gels, the results from the intra-gel comparison (6 DIA files) were imported in the BVA module of DeCyder software. One of the Cy2 images containing the most detected protein

spots was selected as the Master, and the other 5 internal standard images were matched sequentially to it (materials and methods 2.7.5.).

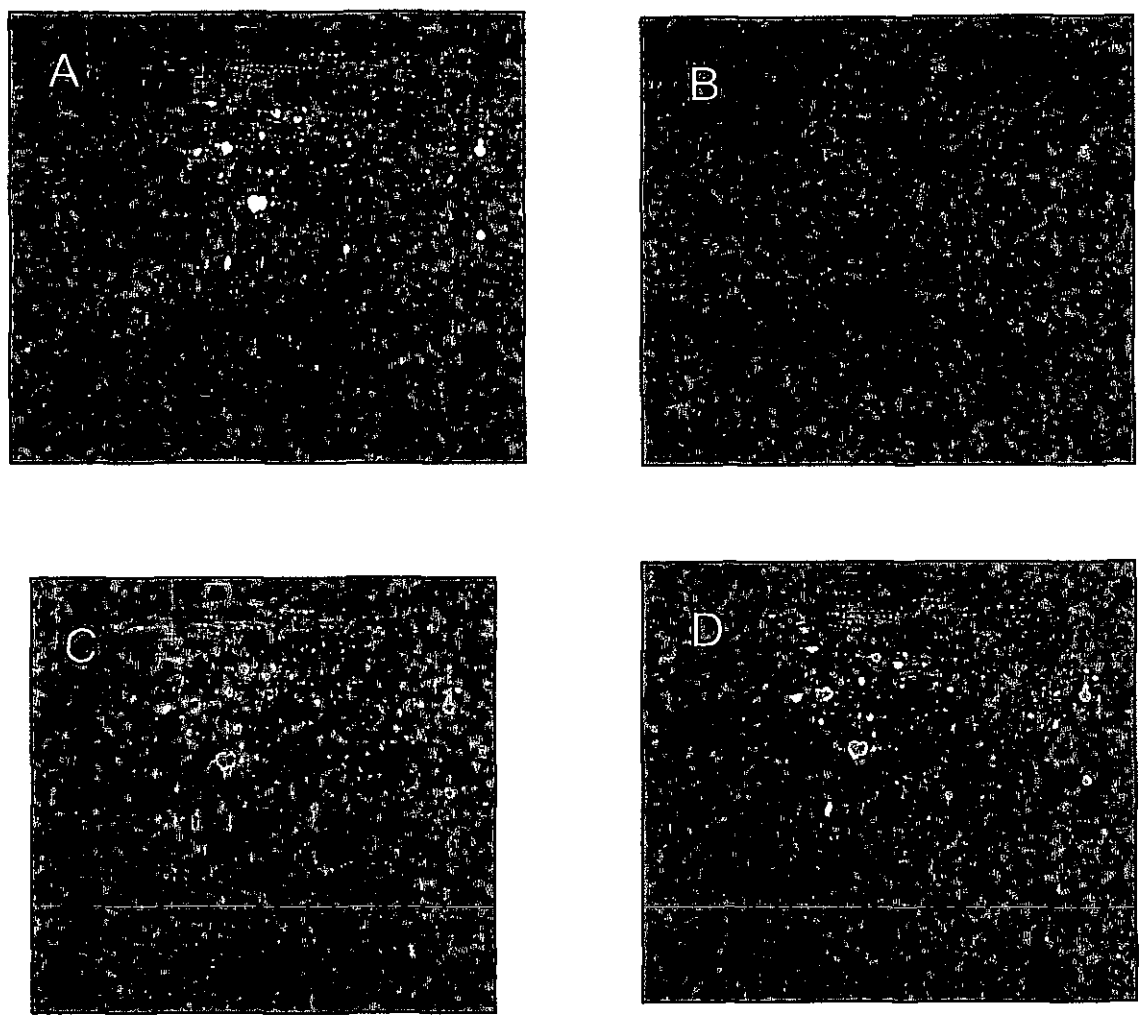



**Figure 3.2. Representative example of 2-D DIGE analytical gels: (A) colour overlay of 2-D DIGE images of DLKP control and BrdU treated samples labeled with Cy3 and Cy5. (B) 50µg of Cy2-labeled pooled standard (equal amount of control and treated sample); (C) 50µg of Cy3-labeled control samples; (D) 50µg of Cy5-labeled BrdU treated sample.**

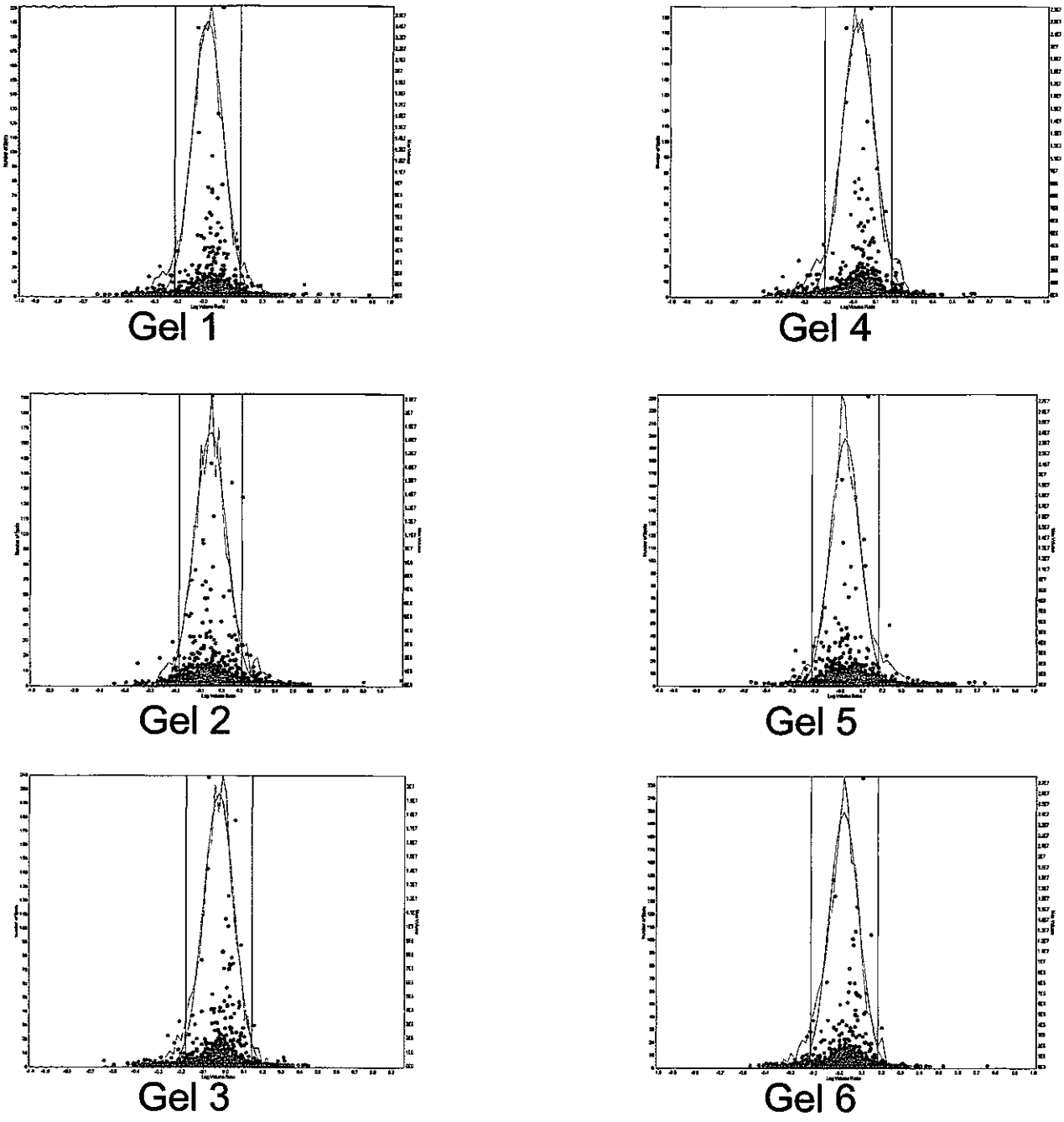



**Figure 3.3. Determining the differentially expressed proteins in the 6 gels. The histograms were generated by DeCyder DIA software, plotting spot frequency (left y-axis) against log volume ratio (x-axis) and a normal distribution model (red line) was fitted to the main peak of the frequency distribution histogram. A 1.5 fold up/down regulation was set as threshold. The threshold values represented by vertical black lines in the histogram and the differentially expressed spot features are shown as blue spots and red spots.**

| Gel 1 | Gel 2 | Gel 3 |
|---|---|---|
| Threshold value 1.5 | Threshold value 1.5 | Threshold value 1.5 |
| Decreased: 127 (5.9%) | Decreased: 97 (4.8%) | Decreased: 120 (5.6%) |
| Increased: 98 (4.6%) | Increased: 145 (7.1%) | Increased: 62 (2.9%) |
| Similar: 1919 (89.5) | Similar: 1787 (88.1%) | Similar 1976 (91.5%) |

| Gel 4 | Gel 5 | Gel 6 |
|---|---|---|
| Threshold value 1.5 | Threshold value 1.5 | Threshold value 1.5 |
| Decreased: 147 (7.1%) | Decreased: 75 (3.5%) | Decreased: 129 (6.1%) |
| Increased: 99 (4.8%) | Increased: 140 (6.6%) | Increased: 62 (2.9%) |
| Similar 1831 (88.2%) | Similar: 1914 (89.9% | Similar: 1929 (91%) |

**Table 3.2. Result of Difference In-Gel Analysis of Control vs Treated BrdU DLKP**

To identify differentially expressed spots across all the 6 gels, the results from the intra-gel comparision (6 DIA fies) were imported into the Biological Variation Analysis (BVA) module of DeCyder software.

This paired ratio data was then utilised in the Biological Variation Analysis (BVA) module when the Cy2 internal standard images from the gels were matched, allowing the spot ratios between the gels to be compared. A list was thus generated containing the protein spot ratio and T-test value for each spot (table 3.3).
The protein spots were assigned a Master Number. The location of the Master Number to its assigned protein spot can be seen in their treated and control gel in figure 3.4 and 3.5.

Relevant spots were selected by narrowing the filter settings to $\geq \pm 1.5$ with a T-test confidence of $\geq 95\%$ ($P \leq 0.05$) and visually confirmed not to be artifacts.
43 protein spots displayed statistical significant differences in their level of expression between DLKP control and BrdU Treatment (table 3.3). 25 protein spots were up-regulated and 18 were down-regulated.

Two examples are shown using the DeCyder BVA results to visualise up and down differentially regulated protein spots (figure 3.6+3.7). The expanded DIGE spot profile of spot 2213 which is 3-fold up regulated with corresponding 3D view and graphical view of the normalized log abundance of the protein spot relative to the standard for each biological replicate gel. The DIGE spot profile of spot 1511 which is 1.5-fold down regulated with corresponding 3D view and graphical view of the normalized log abundance of the protein spot relative to the standard for each biological replicate gel.

### 3.3 BVA results of Control versus Treated DLKP

| Position | Master Number | T-test <=0.05 | Av. Ratio>±1.5-fold difference |
|---|---|---|---|
| 1 | 2113 | 2.00E-09 | 3.09 |
| 2 | 2100 | 4.40E-10 | 2.46 |
| 3 | 2104 | 7.10E-07 | 2.20 |
| 4 | 160 | 2.90E-07 | 1.82 |
| 5 | 360 | 0.00097 | 1.80 |
| 6 | 389 | 5.90E-07 | 1.78 |
| 7 | 426 | 2.20E-06 | 1.76 |
| 8 | 402 | 7.10E-07 | 1.75 |
| 9 | 1468 | 0.0032 | 1.75 |
| 10 | 2090 | 3.20E-07 | 1.73 |
| 11 | 1003 | 0.019 | 1.68 |
| 12 | 151 | 0.00072 | 1.67 |
| 13 | 1058 | 0.03 | 1.64 |
| 14 | 1826 | 2.00E-06 | 1.63 |
| 15 | 810 | 0.037 | 1.62 |
| 16 | 141 | 7.10E-05 | 1.62 |
| 17 | 337 | 1.20E-05 | 1.61 |
| 18 | 1059 | 0.021 | 1.59 |
| 19 | 368 | 1.20E-06 | 1.59 |
| 20 | 1678 | 0.00019 | 1.58 |
| 21 | 138 | 5.90E-09 | 1.58 |
| 22 | 2117 | 2.20E-06 | 1.57 |
| 23 | 674 | 4.80E-05 | 1.54 |
| 24 | 622 | 9.60E-09 | 1.54 |
| 25 | 140 | 9.40E-07 | 1.54 |
| 26 | 511 | 0.0028 | -1.51 |
| 27 | 780 | 0.0046 | -1.52 |
| 28 | 697 | 0.0017 | -1.52 |
| 29 | 385 | 0.0067 | -1.53 |
| 30 | 550 | 0.0028 | -1.55 |
| 31 | 1511 | 0.0023 | -1.56 |
| 32 | 1432 | 0.00031 | -1.56 |
| 33 | 552 | 0.00021 | -1.57 |
| 34 | 1412 | 4.90E-06 | -1.63 |
| 35 | 893 | 0.0012 | -1.63 |
| 36 | 1518 | 0.002 | -1.65 |
| 37 | 234 | 0.0013 | -1.65 |
| 38 | 1425 | 6.10E-06 | -1.73 |
| 39 | 551 | 0.00033 | -1.73 |
| 40 | 1423 | 0.0056 | -1.66 |
| 41 | 76 | 0.0015 | -1.74 |
| 42 | 731 | 3.90E-07 | -1.75 |
| 43 | 1973 | 0.00014 | -2.34 |

**Table 3.3 Differentially regulated DLKP proteins following BrdU treatment.**

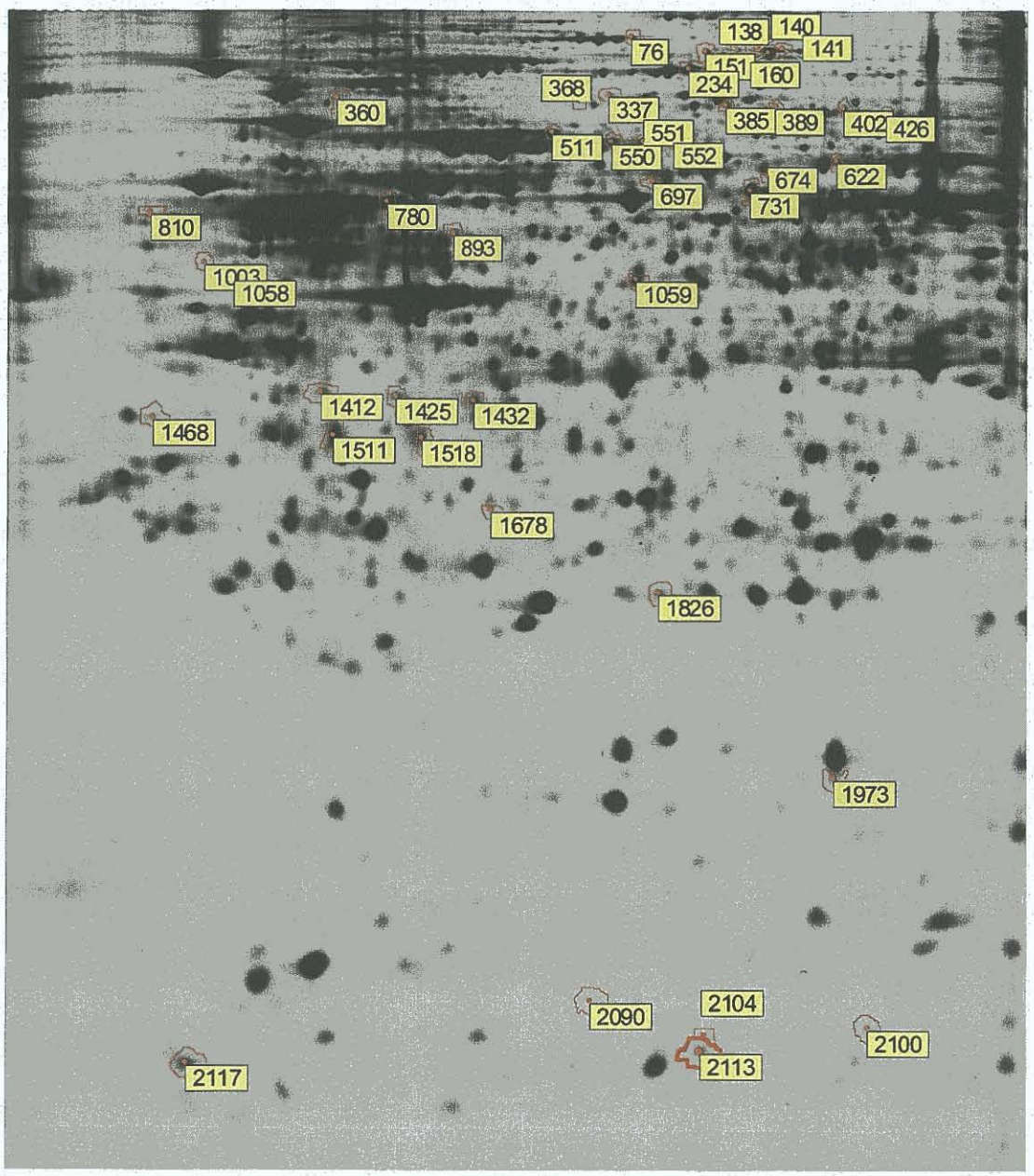



**Figure 3.4. 2D-DIGE scan of DLKP Control sample. 43 Protein spots were identified by their master number as confirmed differentially-regulated spots.**

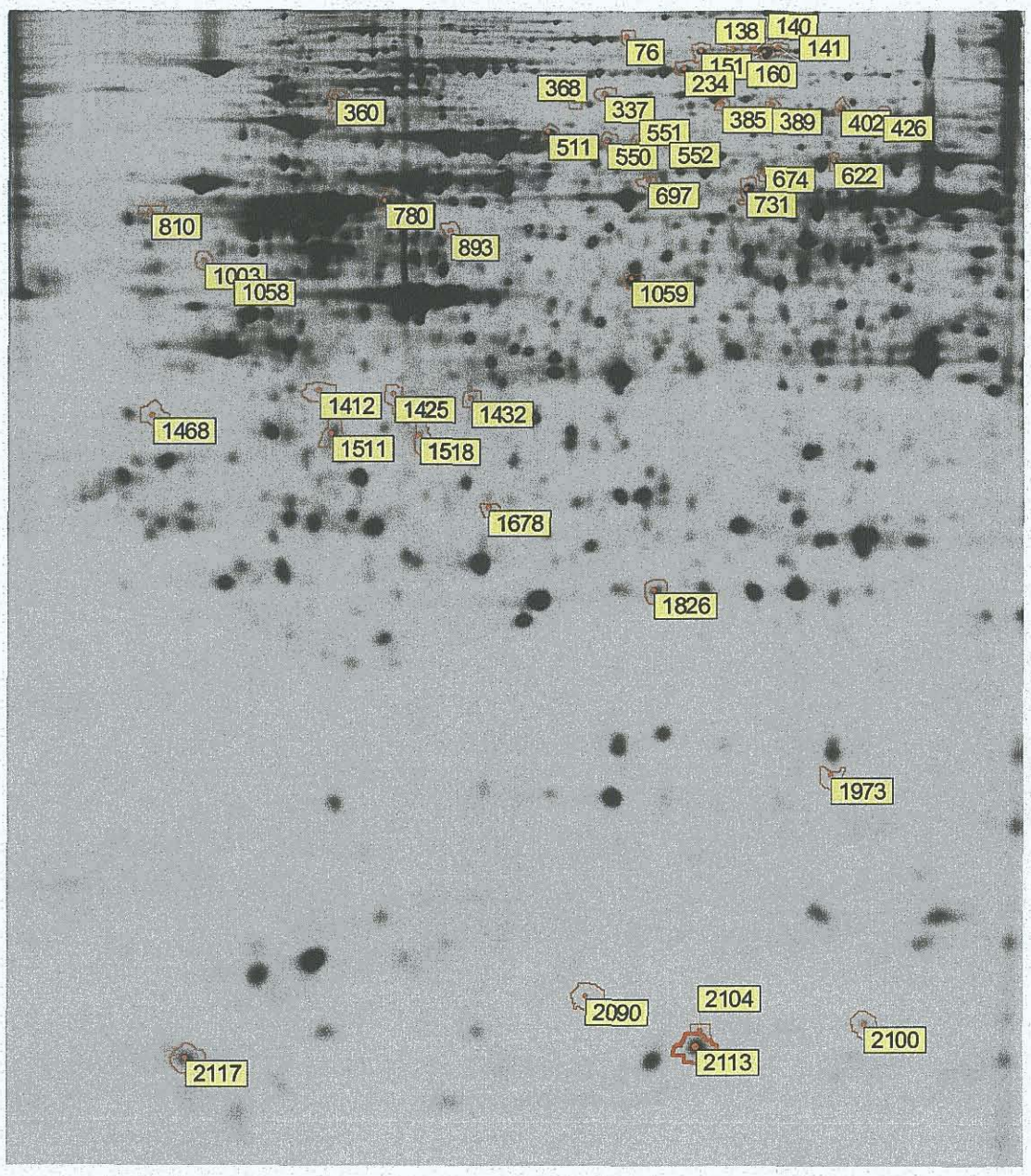



**Figure 3.5. 2D-DIGE scan of DLKP BrdU-Treated sample. 43 Protein spots were identified by their master number as confirmed differentially-regulated spots.**

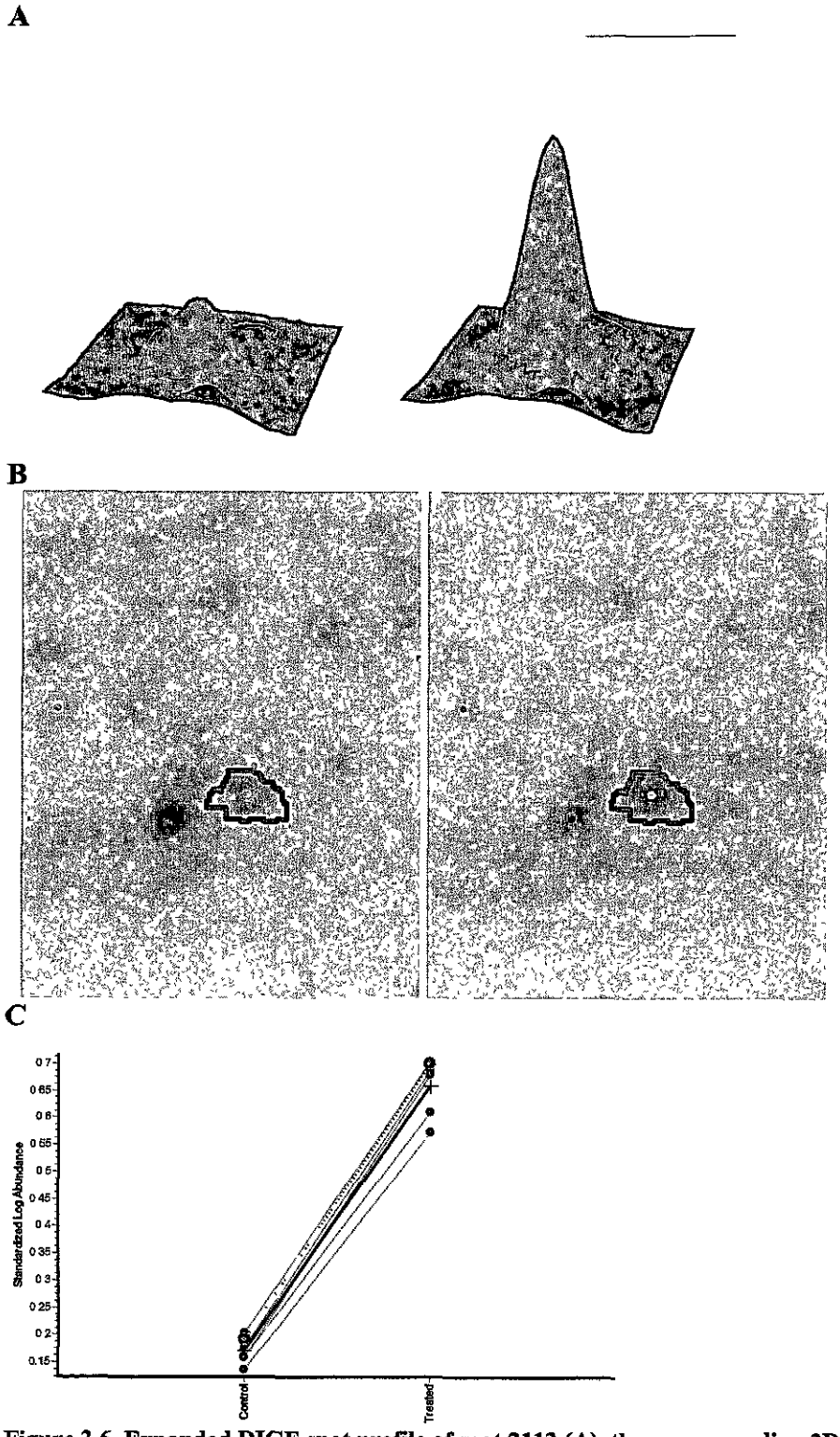



**Figure 3.6. Expanded DIGE spot profile of spot 2113 (A), the corresponding 3D view (B) and a graphical view of the normalized log abundance of the protein spot relative to the standard for each biological replicate gel (C).**

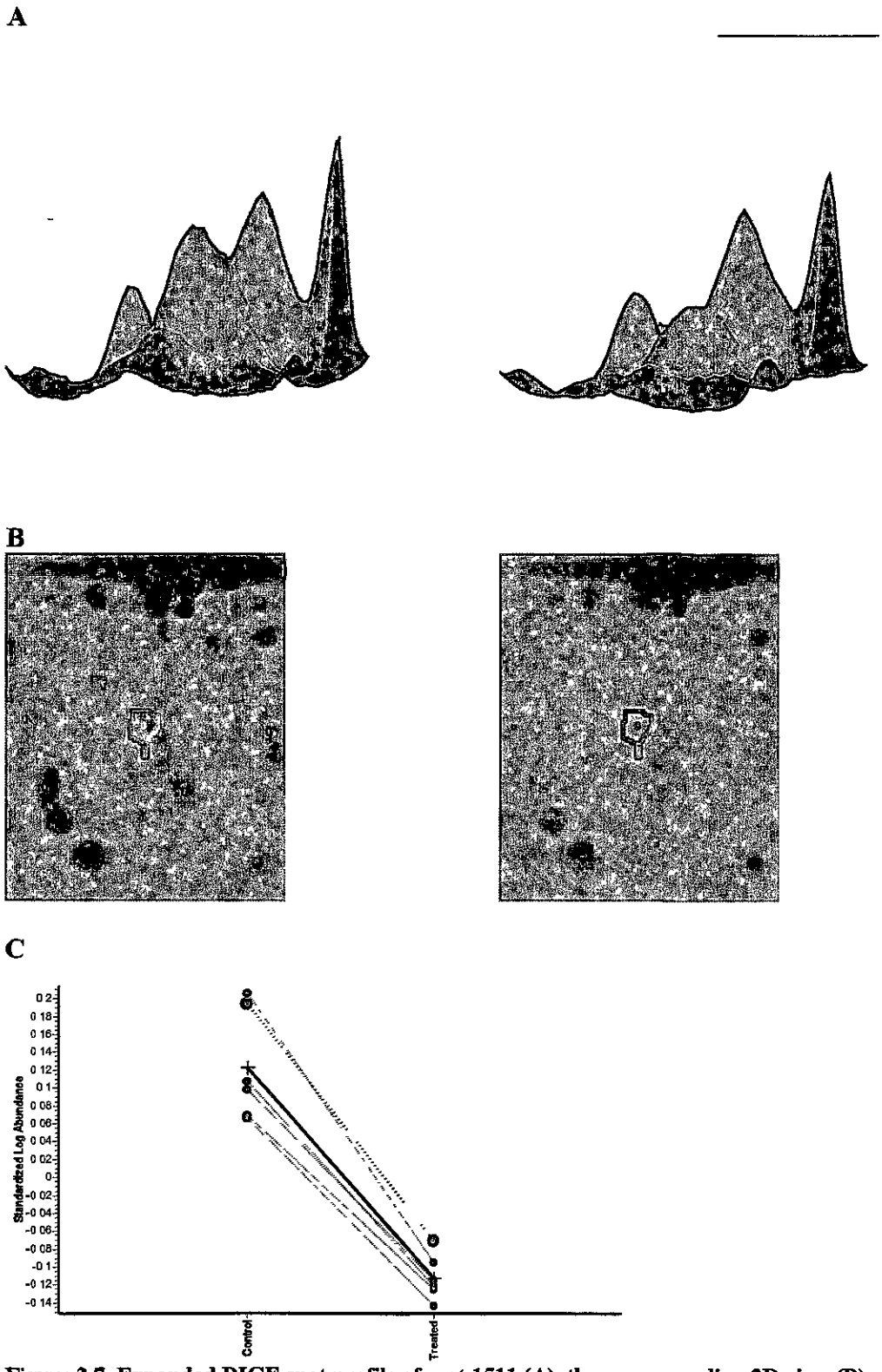



**Figure 3.7. Expanded DIGE spot profile of spot 1511 (A), the corresponding 3D view (B) and a graphical view of the normalized log abundance of the protein spot relative to the standard for each biological replicate gel (C).**

## 3.6 Preparative 2D gels for protein Identification

To identify the differentially expressed proteins preparative DLKP control and BrdU treated 2D gels were run (material and methods 2.7.8). Briefy, preparative 2D gels were prepared by loading 350µg total protein per gel. The gels were stained with Collodial Coomassie Blue (Sigma) and scanned using Labscan software. Scanned images were imported into *Imagemaster version 5 (GE Healthcare)*. Differentially regulated proteins from Table 3.3 were then matched to the DLKP control and treated preparative gels (figures 3.8 + 3.9) and a protein pick list was prepared using the spot picker software (GE Healthcare).

Using the Ettan Spot Picker (GE Healthcare) the 43 differentially regulated proteins were picked and placed into a 96 well plate (materials and methods 2.8.1).

*Destaining and digestion of the protein plugs was carried out using the* Ettan Digestor (GE Healthcare) and overnight digestion using sequence grade trypsin (Promega) at 37°C was carried out (materials and methods 2.8.2).

The digested proteins were either spotted onto MALDI-Target slides for identification by MALDI-ToF MS or using MS/MS in conjunction with nanoelectrospray (MDLC-LTQ) (materials and methods 2.9 and 2.11 respectively). Figure 3.10 below for a flow chart to illustrate the flowpath involved in gel electrophoresis to protein identification.

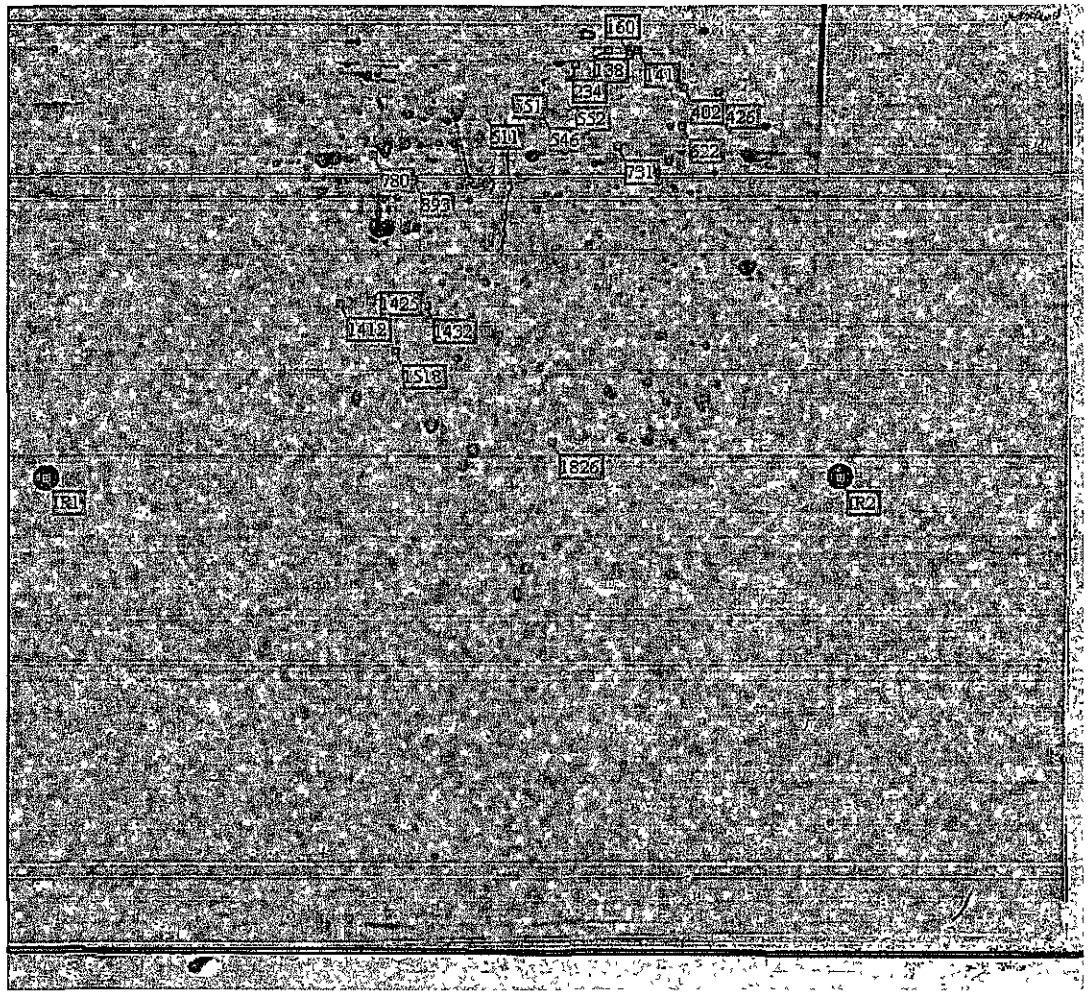



**Figure 3.8. Representative 2D image of the preparative DLKP Control gel. 350µg of protein sample from DLKP was first separated by IEF using Immobiline DryStrip pH 4-7, followed by second dimension separation with a 12.5% SDS-PAGE. The gel was post stained with colloidal coomassie blue. Arrows indicated differentially expressed proteins labelled with their master number, and were excised from the gel and subjected to in-gel digestion.**

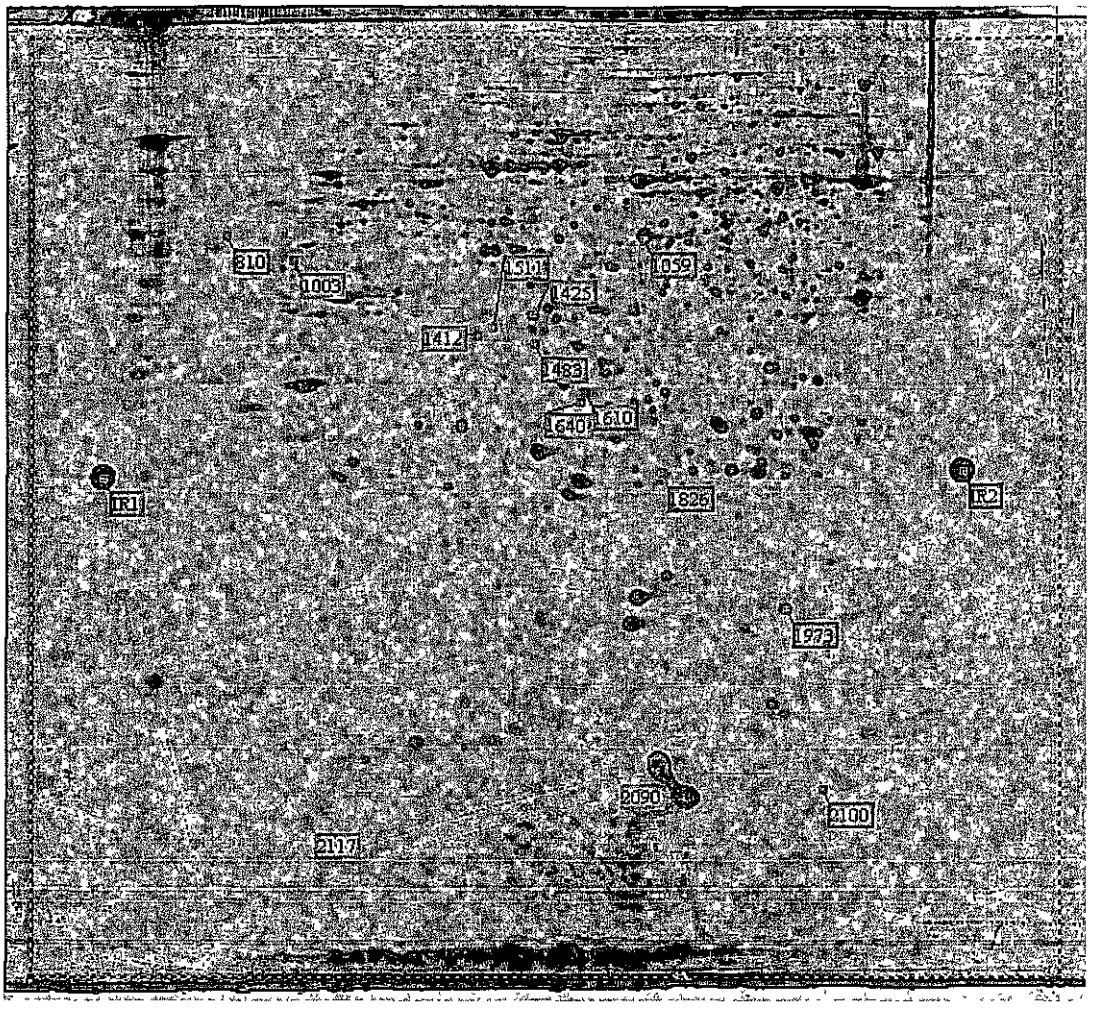



***Figure 3.9. Representative 2D image of the preparative* DLKP BrdU Treated gel. 350µg of protein sample from DLKP was first separated by IEF using Immobiline DryStrip pH 4-7, followed by second dimension separation with a 12.5% SDS-PAGE. The gel was post stained with colloidal Coomassie blue. Arrows indicated differentially expressed proteins labelled with their master number, and were excised from the gel and subjected to in-gel digestion.**

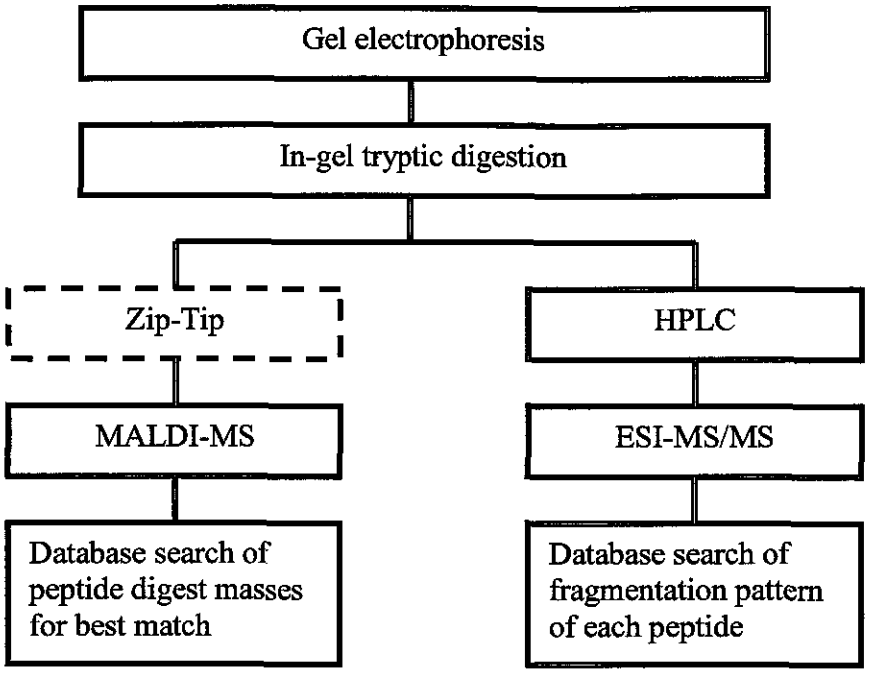



**Figure 3.10. A flow chart to illustrate the steps involved in protein identification from 2-D gels by MALDI and tandem mass spectrometry.**

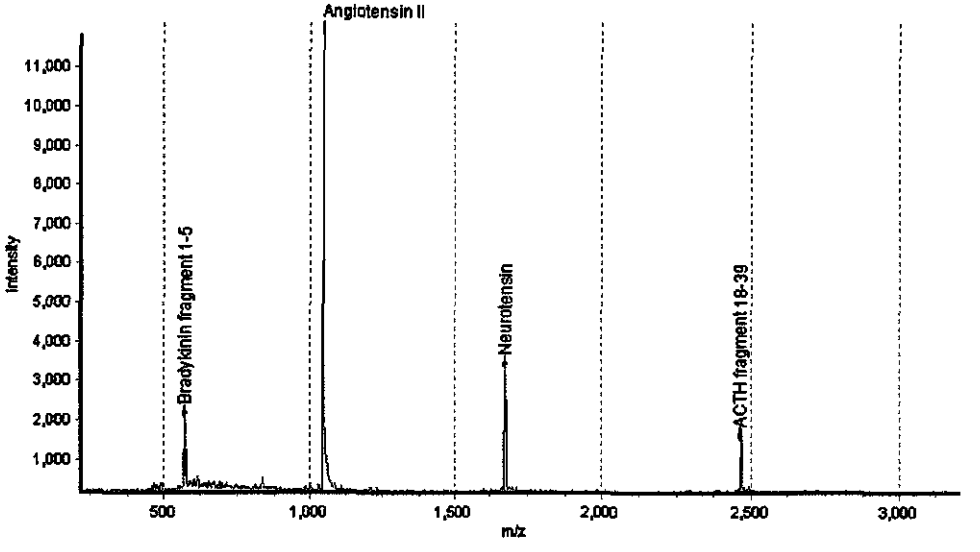
## 3.7 Matrix-Assisted Laser Desorption Ionization (MALDI) Time of Flight (ToF)

In Matrix-Assisted Laser Desorption Ionization (MALDI), the digested peptide sample is embedded in a low molecular weight, UV-absorbing matrix that enhances sample ionization called α-cyano-4-hydroxycinnamic acid (CHCA). The role of the matrix is to facilitate intact desorption and ionization of the sample. The matrix is present in vast excess of sample, and therefore isolates individual sample molecule. The protein can be identified from its peptide mass map or fingerprint, if the peptides from a sufficiently unique pattern when checked against the NCBI-PROFOUND database. At least an expected value of 0.01 (99% confident) or less was considered as a successful identification.

Mass spectrometry can be carried out automatically with the Ettan Pro MALDI-TOF with the laser intensity and slide position automatically adjusting to give the best ionization of the peptides present in each sample (materials and methods figure 2.10).

The use of an external calibrant in the first position on the MALDI target to calibrate the target slide is routinely used in our laboratory (figure 3.11 A) however during this study it was observed that peptide mass accuracy obtained using this method was poor leading to ambiguous protein identifications following database searches. The peptides masses from Spot 1678 were externally calibrated using the pep4 mix standard spotted on the first position on the side. Spot 1678 had been spotted on position twelve (24 positions in total on each slide) on the slide. The peptide list

generated was searched using Profound. An ambiguous protein identification was obtained (figure 3.11 B). The same Mass spectra was then internally calibrated using the trypsin autodigestion peaks, Trypsin I (842.509m/z) and Trypsin III (2211.104 m/z) which were present. Following a ProFound search, the protein designated "The Molecular Basis For The Local Conformational Rearrangement Of Human Phosphoserine Phosphatase" was identified with very high confidence as being spot 1768 (figure 3.11 C). The mass differences between the externally calibrated and internally calibrated peptides from spot 1678 can be seen in figure 3.11 D below.




**Spectrum processing method**

General | Filter | Calibrants | Peak detection | Mass list adjustments

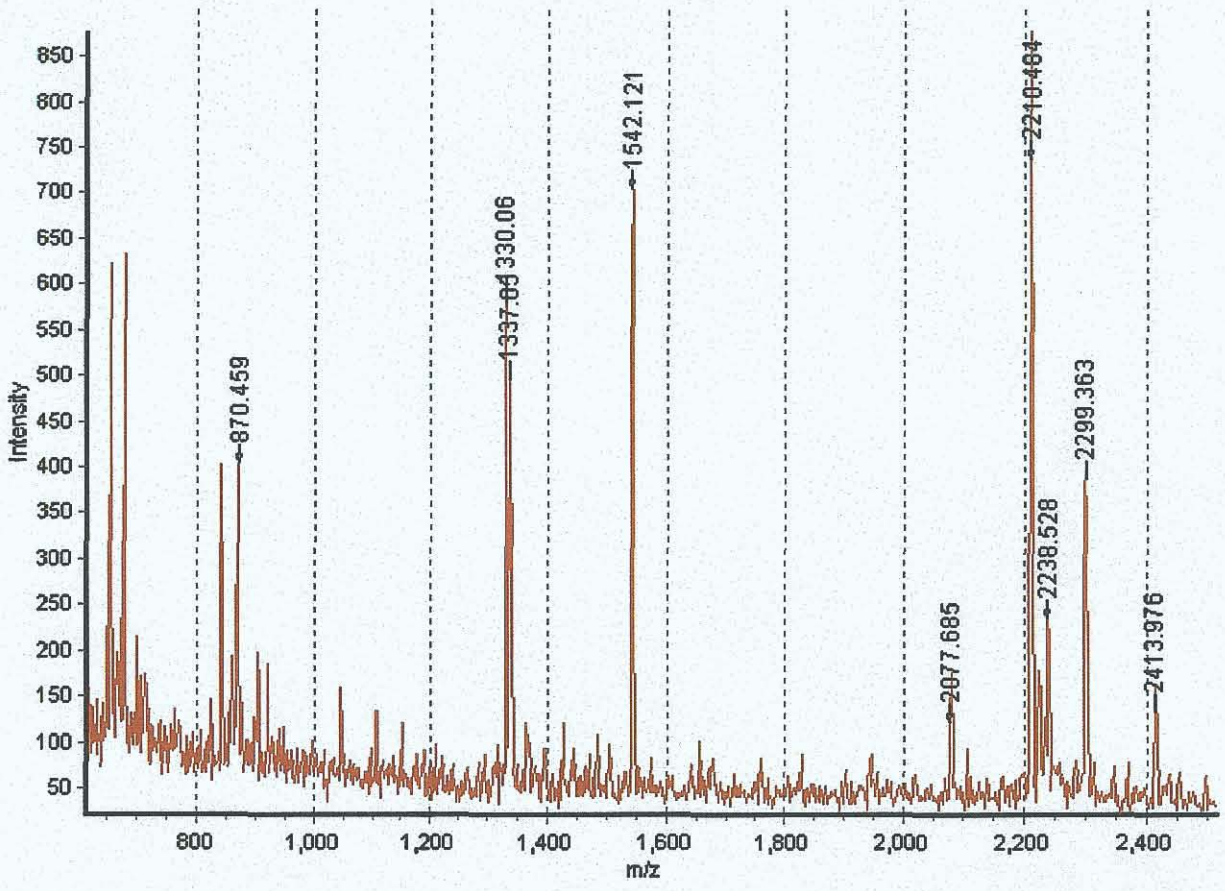
Tolerance: 2 m/z  Min. s/n: 2.5  Min. quality value: 0.60  Apply

New m/z: 573.315 m/z  Label: Bradykinin fragment 1-5

| | Old m/z | New m/z | Ave / Isotopic | Label |
|---|---|---|---|---|
| ▶ | 573.315 | 573.315 | Mono | Bradykinin fragment 1-5 |
| | 1046.542 | 1046.542 | Mono | Angiotensin II |
| | 1672.918 | 1672.918 | Mono | Neurotensin |
| | 2465.199 | 2465.199 | Mono | ACTH fragment 18-39 |

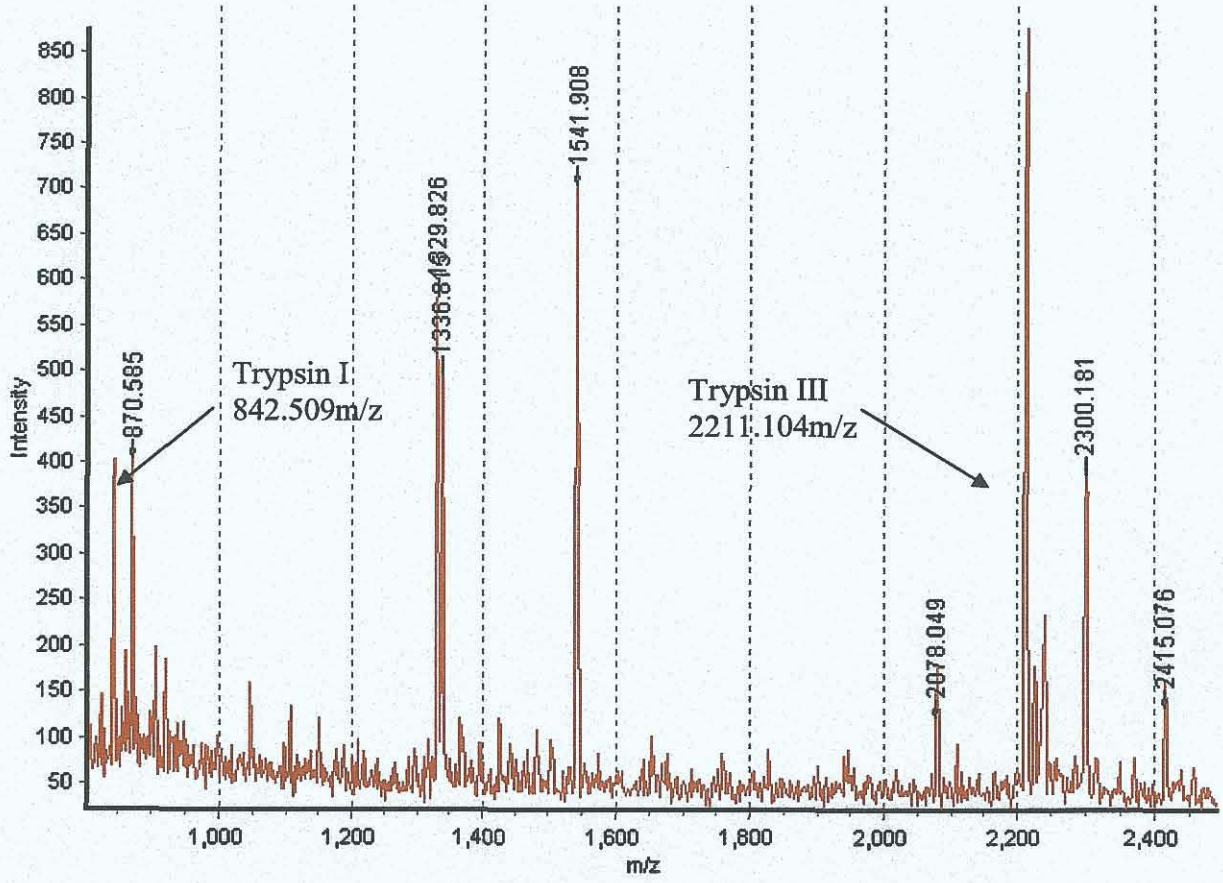
Edit  Delete

**Figure 3.11 A. External calibrants using Pep 4 Mixture (LaserBio Labs) containing Bradykinin fragment 1-5, Angiotensin II, Neurotensin and ACTH fragment 18-39 which is used to calibrate a MALDI-Target Slide**




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.406 | gi\|6114638\|emb\|CAB59385.1 - T cell receptor beta chain variable region [Homo sapiens] | 30.7 | 4.9 | 10.02 |

**Figure 3.11 B. Spot 1678 analysed by MALDI-ToF MS. Peptides were calibrated externally using pep4 from position one on the slide. Spot 1678 was twelve positions away on the slide from the calibrant mix. ProFound search resulted in an ambiguous protein identification result.**




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 * | 0.000 | gi\|29726394\|pdb\|1L8O\|B - Chain B, Molecular Basis For The Local Conformational Rearrangement Of Human Phosphoserine Phosphatase | 34.7 | 5.5 | 25.21 |

**Figure 3.11 C. Spot 1678 analysed by MALDI-ToF MS. Peptides were calibrated internally using tryspin autodigestion peaks (Trypsin I and III). Spot 1678 was identified from ProFound as Molecular Basis For The Local Conformational Rearrangement Of Human Phosphoserine Phosphatase with an excellent expectation score of 0.000**

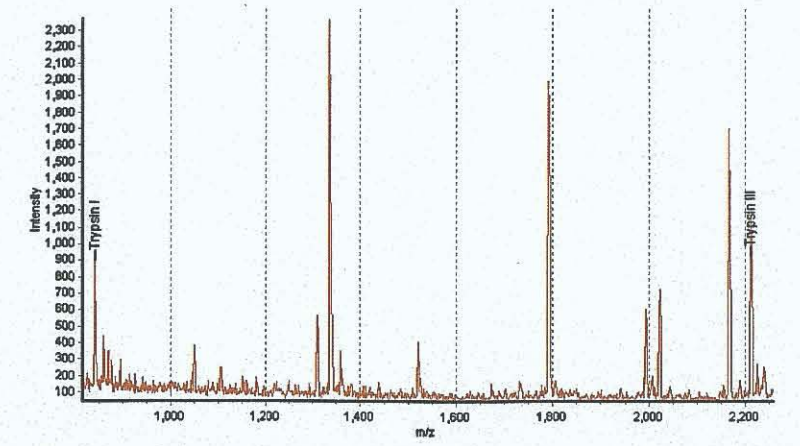
| Mass list of peptides generated from spot 1678 after external calibration m/z [Da] | Mass list of peptides generated from spot 1678 after internal calibration m/z [Da] |
|---|---|
| 870.459 | 870.585 |
| 1330.060 | 1329.826 |
| 1337.050 | 1336.815 |
| 1542.121 | 1541.908 |
| 2077.685 | 2078.049 |
| 2299.363 | 2300.181 |
| 2413.976 | 2415.076 |

**Figure 3.11 D. A comparison of the peptide masses generated using external and internal calibration.**

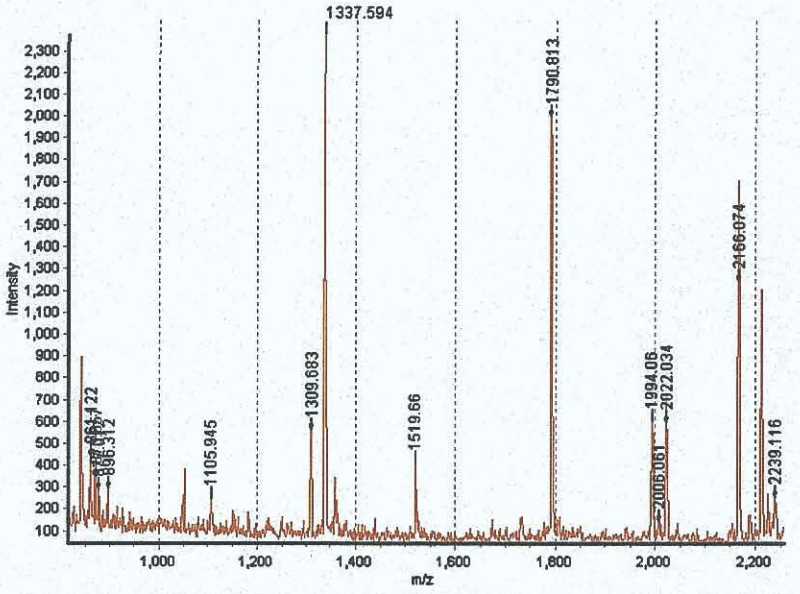
Peptides from spot number 1973 were analysed by MALDI-ToF MS. Mass spectrum acquisition and mass peak assignment was automatically carried out by the Ettan Software (figure 3.12 A).

Peak detection was performed after automatic calibration of the mass spectrum with internal standards (figure 3.12 B). The Ettan Pro Evaluation software (GE Healthcare) can automatically search for the autoproteolytic peptide masses of trypsin 1, 2 and/or 3 (842.509, 1046 and 2211.104 respectively) and if found can automatically internally calibrate each spectra. The calibrated peptide masses were automatically searched with the NCBI database. Spot 1973 was automatically identified as Cofilin 1 (*Homo sapiens*) (figure 3.12 C).

(A)




(B)



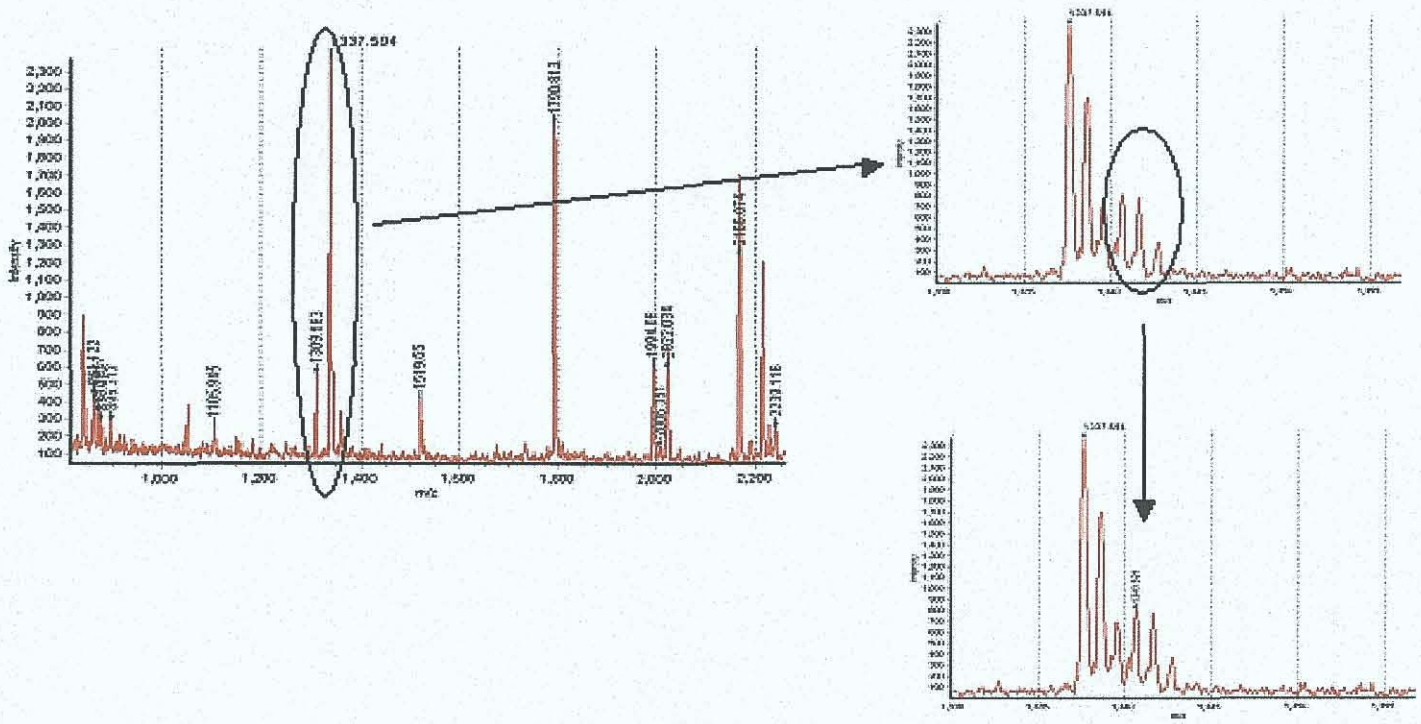

(C)

| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|30582531\|gb\|AAP35492.1 - cofilin 1 (non-muscle) [Homo sapiens] | 32.5 | 8.5 | 18.71 |

**Figure 3.12 (A). MALDI-TOF MS was automatically carried out on spot number 1973. Mass peaks were internally calibrated after the software automatically detected trypsin autodigestion peaks Trypsin I and III (B). The calibrated peptide masses were automatically entered into ProFound against the NCBI database (homo-sapien) with an unambiguous identification of the protein Cofilin 1 (C).**

Incorrect assignment of peak masses, missed assignments of peaks and incorrect assignment of trypsin digestion peaks resulting in inaccurate spectral calibrations are the main disadvantages of this automated MALDI-TOF MS analysis and database searching software. Manual interpretation of MALDI-ToF MS spectra by the addition of peptides that the software fails to identify can result in an increase in the number of peptides assigned to a protein and therefore improve the overall amino acid percentage coverage of protein identification (figure 3.13 A). The masses were automatically associated to peptides from spot 1973 however on manual inspection of the spectra, a peptide mass at 1337.594m/z had not been assigned automatically. When this peak was manually assigned and the peptide masses searched again there was approximately 8% increase in the percentage coverage (figure 3.13 A) when compared to the previous score (figure 3.12 C).

Proteins that were not identified by MALDI-ToF MS with direct spotting were subjected to desalting and concentration using the ZipTip$^{TM}$ (Millipore).



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|30582531\|gb\|AAP35492.1 - cofilin 1 (non-muscle) [Homo sapiens] | 40.4 | 8.5 | 18.71 |

**Fig 3.13 (A). Automatic MALDI-TOF MS on spot number 1973. The peptide mass 1337.594 m/z was not automatically detected by the Ettan Pro software. When this peptide was manually assigned and the peptide masses searched again the percentage coverage improved by approximately 8%.**

## 3.8 Sample Preparation Prior to MALDI-ToF MS using ZipTip$_{C18}$ Tips

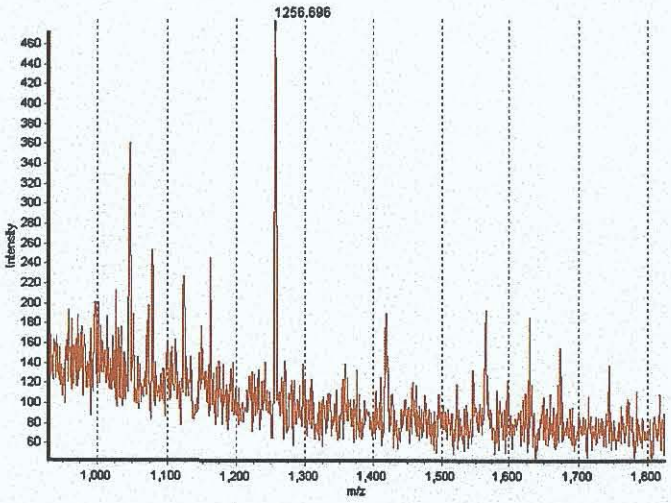
ZipTip$^{TM}$ is a 10µl (P10) pipette tip with a bed of chromatography media fixed at its end such that there is no dead volume. It is intended for purifying and concentrating femtomoles to picomoles of peptide samples prior to analysis providing better data quality. The sample is aspirated and dispensed through ZipTip to bind, wash and elute. Recovered samples are contaminant-free and eluted in a 0.4µl for direct transfer to a MALDI-ToF MS target.

It was not possible to ZipTip all samples as this is an expensive and time consuming procedure so all samples were processed using direct sample spotting with the Ettan Spotter. MALDI-ToF MS was carried out. GE Healthcare state that the low-salt digestion method used by Ettan Spot robots do not require a clean-up step. However all samples that were not identified with high confidence using the robot dried droplet method were then subjected to sample clean up and concentration with Zip Tips.

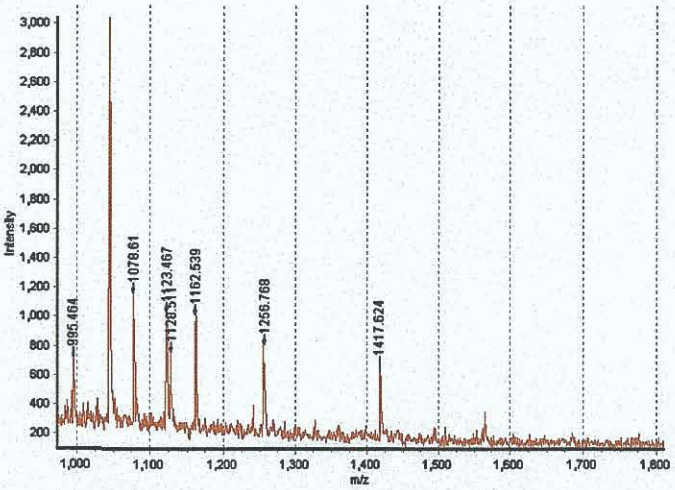
Figure 3.14 shows spectra before and after desalting with ZipTip$_{C18}$. Sample before zip tipping produced a spectra with just one peptide and as a result poor confidence score from ProFound database search. Spectra from the same sample after ZipTip resulted in a concentration and desalting and now 7 peptides were present allowing for a high confidence result from ProFound database search.

Although the benefits of sample clean up and concentration with $C_{18}$ ZipTips is well documented and observed in our laboratory, it has been noted that some high molecular weight peptides sometime remained tightly bound to the column after elution (figure 3.19). This can be problematic as the presence of trypsin III at 2211.104m/z which frequently remained bound to the ZipTip is a very useful as an internal calibrant.

It has also been observed that there is some batch to batch variation in the quality of the ZipTips (results not shown) resulting in poor MS spectra acquired post sample clean up and concentration.




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.822 | gi\|7019976\|dbj\|BAA90948.1 - unnamed protein product [Homo sapiens] | 8.5 | 5.5 | 48.07 |




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|45708738\|gb\|AAH33785.1 - SFRS1 protein [Homo sapiens] | 31.8 | 7.9 | 22.56 |

**Figure 3.14. MALDI-ToF MS spectra of spot number 1425. Spectra taken before and after desalting with ZipTip$_{C18}$. ZipTip sample resulted in a high confidence Identification using ProFound.**

## 3.9 Internal calibration methods for mass accuracy identifications of proteins by MALDI-TOF MS

The first step in obtaining a peptide mass is the in-gel digestion of a protein spot by trypsin, a protease that specifically cuts the carboxyl-terminal of Arg or Lys residue. The trypsin enzyme can result in autodigestion peaks at 842.509, 1045.5642 and 2211.104m/z which are exploited as internal standards for mass calibration and are useful when internal calibrants are not added.

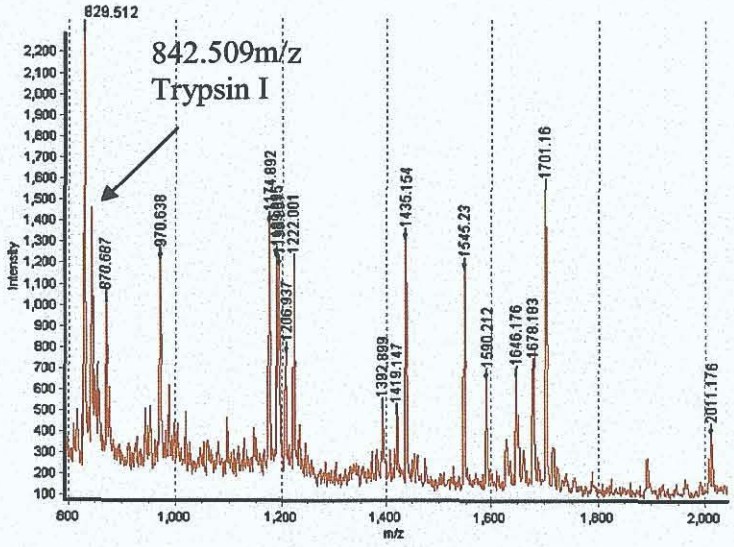
Getting accurate mass values depends on how well the data is calibrated. We compared the MS spectrum from two different calibration protocols. (1) Internal calibration using trypsin autodigestion peaks, and (2) Internal calibration with commercially available peptide 4 standard (LaserBio Labs).

The peptides from spot number 731 were desalted and concentrated using ZipTips. Two elutions from the Zip Tip was carried out. The first elution was directly spotted onto the MALDI target slide and the second elution from the same tip was mixed with a commercially available peptide standard. Both samples were analysed by MALDI-TOF MS.

Figure 3.15 shows the spectrum from the first elution containing 17 peptides. Only one autodigestion trypsin peak is present at 842.509m/z and this is used to internally calibrate the 17 peptides present. After peptide mass searching using ProFound, a low confidence protein identification is made as HSD24.

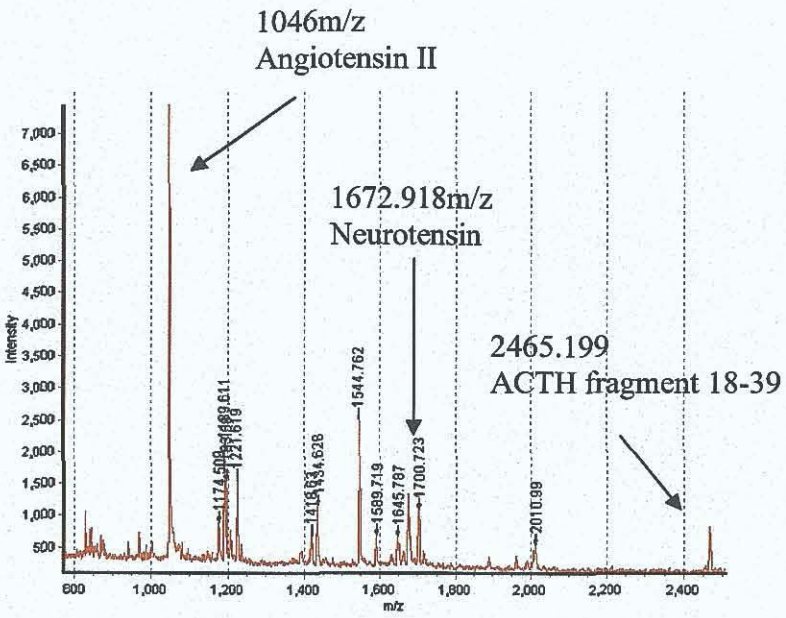
Figure 3.16 shows the spectrum from the second elution of the Zip Tip sample 731 which is mixed with commercially available peptide standards. 10 peptides are present from the sample and three calibrant peptides are also present (Angiotensin II at 1046.5424m/z, Neurotensin at 1672.918m/z and ACTH fragment 18-39 at 2465.199m/z). Using the three calibrant peptides to calibrate the 10 sample peptides a very high confidence score for the identification of spot 731 is given as Aldehyde dehydrogenase 1 following ProFound database search.

Poor mass accuracy of the peptides present in the first elution from the Zip Tip sample 731 as only one internal calibrant was present resulted in a low confidence protein identification of that sample. Using the high mass accuracy peptide measured from the second elution at 1700.723m/z we re-calibrated the peptides measured from the first elution spectra using 842.509m/z (trypsin autodigestion peak) and the sample peptides 1700.723m/z. The new peptide masses were searched using ProFound a very high confidence score resulted matching the identification from the second elution as Aldehyde dehydrogenase 1 (figure 3.17).



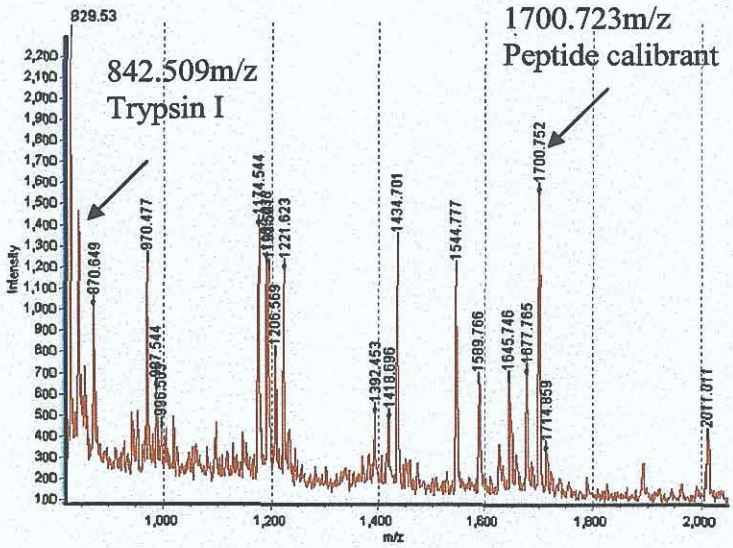

| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.167 | gi\|37955219\|gb\|AAP20058.1 - HSD24 [Homo sapiens] | 19.0 | 10.4 | 10.23 |

**Figure 3.15. MALDI-ToF MS spectra of spot number 731. Sample was desalted with ZipTip$_{C18}$. ZipTip and trypsin autodigestion peak at 842.509m/z was used as internal calibrant. The peptide masses were searched using ProFound and resulted in a low confidence score.**




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.001 | gi\|32815082\|gb\|AAP88039.1 - aldehyde dehydrogenase 1 family, member A1 [Homo sapiens] | 19.2 | 6.3 | 55.47 |

**Figure 3.16. MALDI-ToF MS spectra of spot number 731. Sample was desalted with ZipTip$_{C18}$. ZipTip and mixed with commercially available internal calibrants. The peptide masses were searched using ProFound and the protein was identified with high confidence.**




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|32815082\|gb\|AAP88039.1 - aldehyde dehydrogenase 1 family, member A1 [Homo sapiens] | 30.9 | 6.3 | 55.47 |

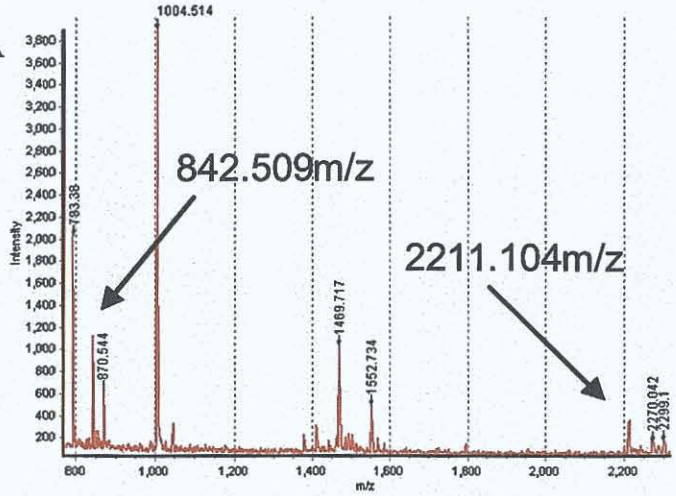
**Figure 3.17. MALDI-ToF MS spectra of spot number 731. The peptide masses were calibrated using the autodigestion trypsin peak 842.509 along with a calibrated peak from the same peptide list that was calibrate using commercially available peptide standards. The searched peptide mass using ProFound resulted in a very high confidence protein identification score.**

### 3.10 Identification of low molecular weight proteins by MALDI ToF MS.

Identification of low molecular weight proteins have proved very difficult by MALDI-ToF MS. A combination of limited peptides available due to the size of the protein and some peptides ionising better that other peptides making accurate mass determination of these peptides difficult.

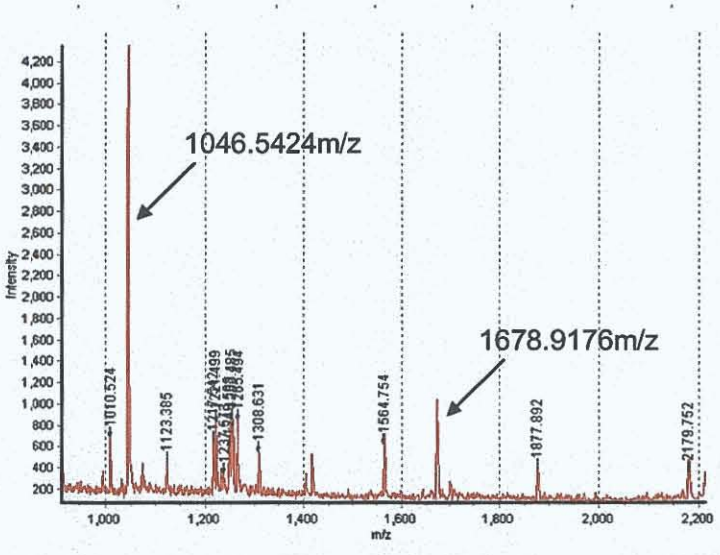
We have discovered that a combination of manually picking the low molecular protein spots to give a large amount of protein available for digestion, increasing the laser intensity to allow some of the less dominant peptides to ionize and internal calibrants with pep4 has proved successful in the identification of some low molecular weight proteins (figure 3.18). Using the Ettan spot picker, master number 2213 was picked, the protein plug was destained, digested and desalted and concentrated before being analysed by MALDI-TOF MS. The resulting peptide list was searched using ProFound. However a low confidence protein identification score was generated (figure 3.18). Using a Pasteur pipette a larger gel plug of the same protein spot was picked. The protein plug was destained, digested and desalted and concentrated prior to analysis by MALDI-TOF MS. A significant number of peptides were generated, the list was searched using ProFound and the resulting protein identification had a very high score (figure 3.18).

A




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.307 | gi\|9844545\|emb\|CAC03973.1 - T-cell receptor beta chain [Homo sapiens] | 19.7 | 6.7 | 8.70 |

B




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.001 | gi\|30582729\|gb\|AAP35591.1 - cytochrome c oxidase subunit VIb [Homo sapiens] | 65.1 | 6.6 | 10.40 |

**Figure 3.18. Comparison of MALDI-ToF MS spectra generated from peptides from Ettan Spot picker sample (A) and manual picked spot (B) (with a sterile pasteur pipette). Both samples were cleaned up with ZipTips. The first sample used autodigestion peaks for internal calibrant and the second used spiked calibration peptides as internal standards.**

## 3.11 Peptide retention on ZipTip Columns

Although the benefits of sample clean up and concentration have been discussed, during the course of this study it has been observed that some high molecular peptides appear to bind permanently to the Zip Tip Columns. This was illustrated by Spot no 1826 (identified as DJ-1) which was analysed by direct spotting and then by Zip Tip. In direct spot (dried droplet spotting method) the autodigestion peptide trypsin III at 2211.1m/z is detected along with a peptides at 2229.2 and 2300.1 m/z (Figure 3.19 A). Following Zip Tip sample clean up of the sample spot, some high molecular weight peptides have remained bound to the Zip Tip are not present in the spectra (figure 3.19 B). A strategy of acquiring MALDI-TOF MS data from both Zip and non-Zip from the same peptide is essential to ensure that the benefits of both techniques utilized for high confidence protein identification.

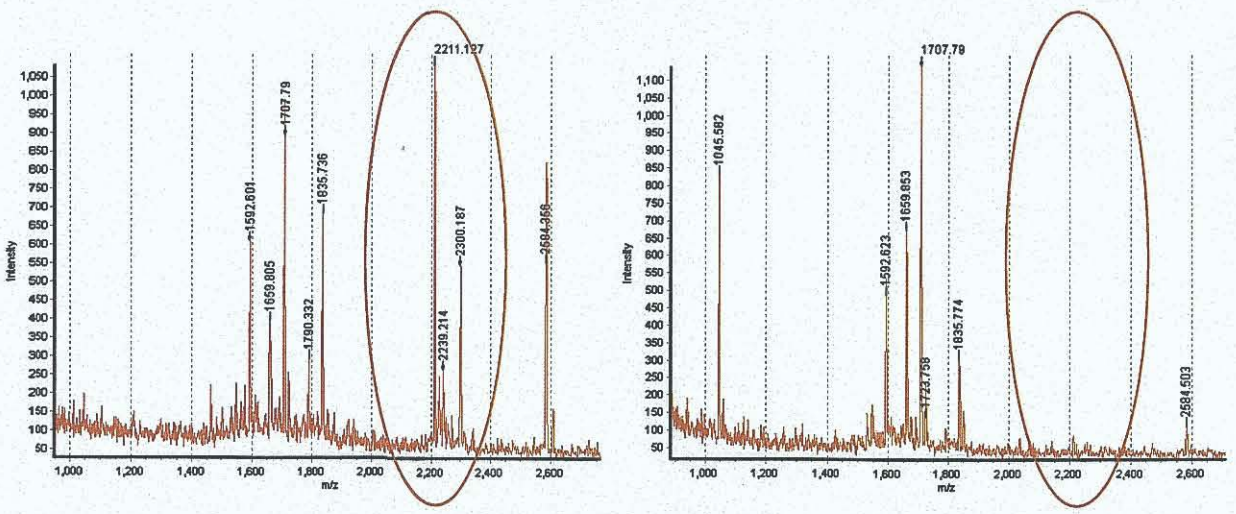



**Figure 3.19. Direct spot of sample 1826. High molecular weight peptides present at 2211.1, 2229.2 and 2300.1 m/z. Zip Tip sample clean up of the same sample resulted in the same high molecular peptides retained on the Zip Tip and not present in the spectra.**

### 3.12 Protein identification results

Using a combination of the various techniques including direct peptide spotting, sample concentration and desalting, internal calibration and spot picking, 23 proteins were identified by MALDI-TOF MS. 10 proteins were differentially up-regulated and 13 were down regulated. One protein (spot 2100) was identified as two separate proteins (table 4).

These proteins identified by MALDI-TOF Mass Spectrometry are list in table 3.4.

## Table 3.4 MALDI-ToF MS Protein Identification results

| Master spot no. | Fold change | T test | ID | Expected | Coverage % | pI | Mw (kDa) | Comments |
|---|---|---|---|---|---|---|---|---|
| 141 | 1.62 | 7.10E-05 | gi\|57162652\|emb\|CAI39687.1 - ribosomal protein S6 kinase, 90kDa, polypeptide 3 [Homo sapiens] | 0.004 | 5.9 | 6.4 | 84.09 | Matched 6 peptides |
| 160 | 1.82 | 2.90E-07 | gi\|24657579\|gb\|AAH39174.1 - Vinculin protein [Homo sapiens] | 0.000 | 23.5 | 5.8 | 117.3 | Matched 23 peptides |
| 234 | -1.65 | 0.0013 | gi\|2274968\|emb\|CAA04006.1 - Glucosidase II [Homo sapiens] | 0.003 | 6.9 | 5.9 | 86.27 | Matched 10 peptides |
| 511 | -1.51 | 0.0028 | gi\|21040386\|gb\|AAH30634.1 - heat shock 70kD protein 9B (mortalin-2) [Homo sapiens] | 0.000 | 17.8 | 6.0 | 74.12 | Matched 9 peptides |
| 552 | -1.57 | 0.00021 | gi\|23307791\|gb\|AAN17824.1 - serum albumin [Bos taurus] | 0.003 | 14.5 | 5.8 | 71.3 | Matched 7 peptides |
| 550 | -1.64 | 7.90E-05 | gi\|23307791\|gb\|AAN17824.1 - serum albumin [Bos taurus] | 0.000 | 21.4 | 5.8 | 71.3 | Matched 11 peptides |
| 551 | -1.73 | 0.00033 | gi\|23307791\|gb\|AAN17824.1 - serum albumin [Bos taurus] | 0.000 | 27 | 5.8 | 71.3 | Matched 16 peptides |
| 674 | 1.54 | 4.8E-005 | gi\|5566500\| UDP-N-Acteyglucosamine pyrophosphorylase | 0.007 | 10.7 | 6.0 | 57.41 | Matched 5 peptides |
| 731 | -1.75 | 3.90E-07 | gi\|2183299\|gb\|AAC51652.1 - aldehyde dehydrogenase 1 [Homo sapiens] | 0.000 | 26.1 | 6.3 | 55.44 | Matched 14 peptides |
| 780 | -1.52 | 0.0046 | gi\|62896667\|dbj\|BAD96274.1 - splicing factor 3a, subunit 3 variant [Homo sapiens] | 0.007 | 12.4 | 5.3 | 59.2 | Matched 8 peptides |
| 810 | 1 62 | 0.037 | gi\|547749\|sp\|P13645\|K1C10_HUMAN - Keratin, type I cytoskeletal 10 (Cytokeratin 10) (K10) (CK 10) | 0.000 | 23.6 | 5.1 | 59.73 | Matched 14 peptides |
| 893 | -1.63 | 0.0012 | gi\|21430983\|gb\|AAM51148.1 - polymerase (DNA directed), delta 2, regulatory subunit (50kD) [Homo sapiens] | 0.000 | 29.2 | 5.3 | 51.9 | Matched 10 peptides |
| 1059 | 1.59 | 0.021 | gi\|18088719\|gb\|AAH20946.1 - Tubulin, beta polypeptide [Homo sapiens] | 0.000 | 51.8 | 4.7 | 50.11 | Matched 20 peptides |
| 1412 | -1.63 | 4.90E-06 | gi\|62088696\|dbj\|BAD92795.1 - splicing factor, arginine/serine-rich 1 (splicing factor 2, alternate splicing factor) variant [Homo sapiens] | 0.000 | 40.8 | 9.3 | 25.78 | Matched 12 peptides |

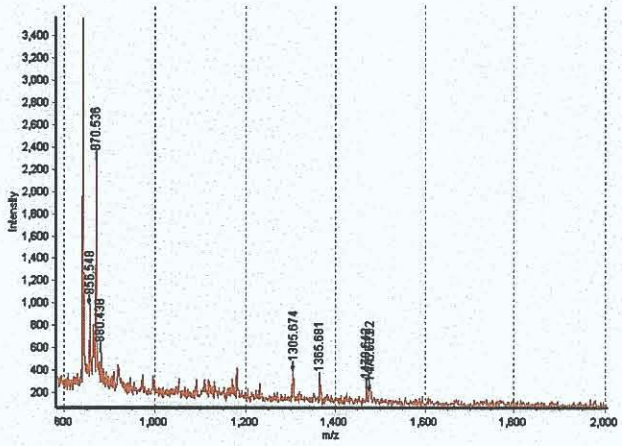
| Master spot no. | Fold change | T test | ID | Expected | Coverage % | pI | Mw (kDa) | Comments |
|---|---|---|---|---|---|---|---|---|
| 1423 | -1.66 | 0.0056 | gi\|57209889\|emb\|CAI4128 4.1 - retinoblastoma binding protein 7 [Homo sapiens] | 0.024 | 13.6 | 4.9 | 47.60 | Matched 7 peptides |
| 1425 | -1.73 | 6.10E-06 | gi\|62088696\|dbj\|BAD92795 .1 - splicing factor, arginine/serine-rich 1 (splicing factor 2, alternate splicing factor) variant [Homo sapiens] | 0.001 | 26.4 | 5.5 | 27.76 | Matched 5 peptides |
| 1432 | -1.56 | 0.00031 | gi\|7546546\|pdb\|1EF7\|B - Chain B, Crystal Structure Of Human Cathepsin X | 0.001 | 37.2 | 5.5 | 27.76 | Matched 8 peptides |
| 1511 | -1.56 | 0.0023 | gi\|5748523\|emb\|CAB53072 .1 - MAPRE1 [Homo sapiens] | 0.000 | 27.2 | 5.0 | 30.15 | Matched 7 peptides |
| 1518 | -1.65 | 0.002 | gi\|21429608\|gb\|AAM50090. 1 - superoxide dismutase copper chaperone [Homo sapiens] | 0.000 | 16.4 | 5.3 | 29.52 | Matched 5 peptides |
| 1678 | 1.58 | 0.00019 | gi\|1890331\|emb\|CAA71318 .1 Molecular Basis For The Local Conformational Rearrangement Of Human Phosphoserine Phosphatase] | 0.003 | 34.7 | 5.5 | 25.17 | Matched 5 peptides |
| 1826 | 1.63 | 2.00E-06 | gi\|33358056\|pdb\|1PE0\|B - Chain B, Crystal Structure Of The K130r Mutant Of Human Dj-1 | 0.008 | 22.3 | 6.5 | 21.14 | Matched 5 peptides |
| 1973 | -2.34 | 0.00014 | gi\|30582531\|gb\|AAP35492. 1 - cofilin 1 (non-muscle) [Homo sapiens] | 0.000 | 36.1 | 8.5 | 18.71 | Matched 7 peptides |
| 2100 | 2.46 | 4.40E-06 | gi\|179318\|gb\|AAA51811.1 - beta-2-microglobulin [Homo sapiens] | 0.003 | 26.9 | 6.1 | 13.83 | Matched 5 peptides |
| 2100 | 2.46 | 4.4E-06 | gi\|229995\|pdb\|1HLA\|M - Chain M, Human Class I Histocompatibility Antigen A2 (HLA-A2, Human Leucocyte Antigen) | 0.003 | 39.2 | 6.5 | 11.58 | Matched 5 peptides |

| Master spot no. | Fold change | T test | ID | Expected | Coverage % | pI | Mw (kDa) | Comments |
|---|---|---|---|---|---|---|---|---|
| 2104 | 2.20 | 7.10E-07 | gi\|30582729\|gb\|AAP35591.1 - cytochrome c oxidase subunit VIb [Homo sapiens] | N/A for LTQ | 18.39 | 6.5 | 10.30 | Matched 2 peptides |
| 2113 | 3.09 | 2.00E-09 | gi\|30582729\|gb\|AAP35591.1 - cytochrome c oxidase subunit VIb [Homo sapiens] | 0.000 | 65.1 | 6.6 | 10.4 | Matched 5 peptides |

**Table 3.4. Identified proteins by Mass Spectrometry. Two protein Identifications from the one gel plug spot 2100.**

## MALDI-TOF MS Spectra results

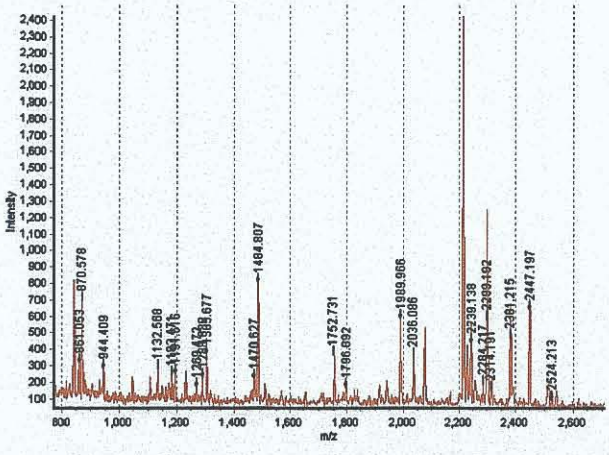
### Master number: 141
### Identified by MALDI-TOF MS (Direct spotting) - ribosomal protein S6 kinase




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.004 | gi\|57162652\|emb\|CAI39687.1 - ribosomal protein S6 kinase, 90kDa, polypeptide 3 [Homo sapiens] | 5.9 | 6.4 | 84.06 |

**Figure 3.20 Spot 141 ribosomal protein S6 kinase**

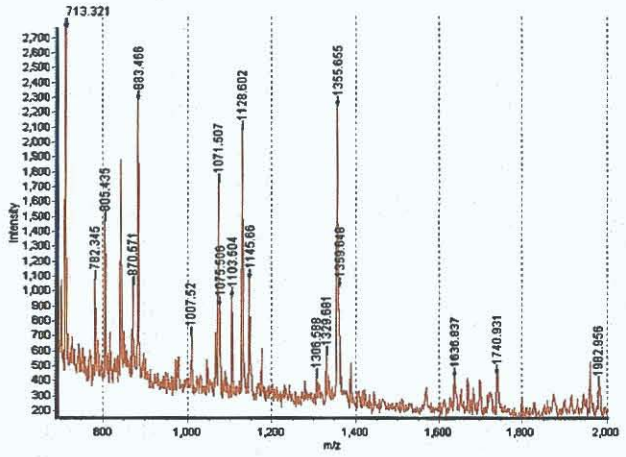
### Master number: 160
### Identified by MALDI-TOF MS (Zip Tip) - Vinculin protein




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.002 | gi\|24657579\|gb\|AAH39174.1 - Vinculin protein [Homo sapiens] | 16.0 | 5.8 | 117.29 |

**Figure 3.21 Spot 160 Vinculin**

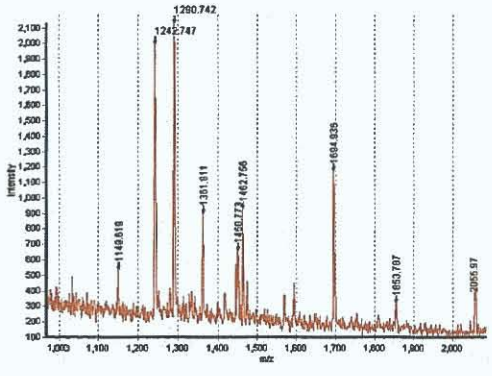
## Master number: 234
## Identified by MALDI-TOF MS (Zip Tip)– Alpha Glucosidase II




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.003 | gi\|38371758\|ref\|NP_055425.3 - alpha glucosidase II alpha subunit isoform 1 [Homo sapiens] | 9.9 | 5.9 | 86.27 |

**Figure 3.22 Spot 234 Vinculin**

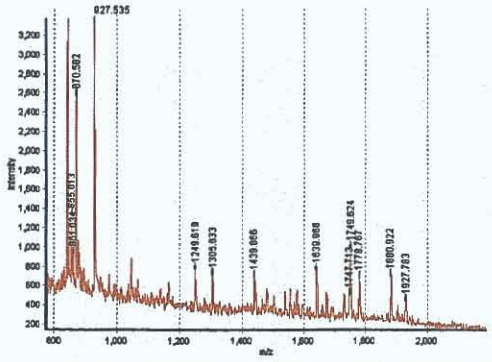
## Master number: 511
## Identified by MALDI-TOF MS (Zip Tip) – Heat shock 70Kd




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|21040386\|gb\|AAH30634.1 - heat shock 70kD protein 9B (mortalin-2) [Homo sapiens] | 17.8 | 6.0 | 74.12 |

**Figure 3.23 Spot 511 Heat shock 70kD**

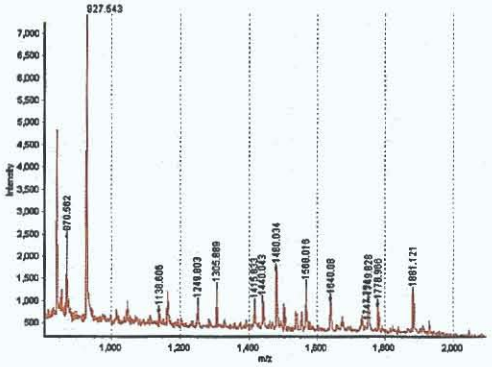
## Master number: 550
## Identified by MALDI-TOF MS (Direct spotting) – Serum Albumin




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|162648\|gb\|AAA51411.1 - albumin [Bos taurus] | 19.1 | 5.8 | 71.27 |

**Figure 3.24 Spot 550 Albumin**

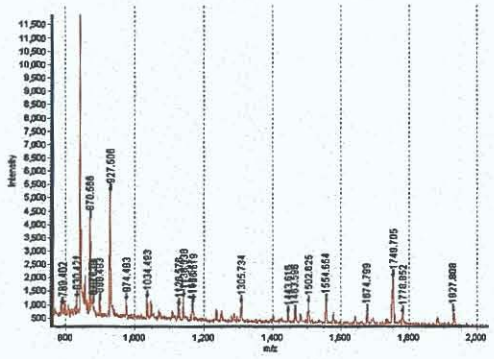
## Master number: 551
## Identified by MALDI-TOF MS (Direct spotting) – Serum Albumin




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|23307791\|gb\|AAN17824.1 - serum albumin [Bos taurus] | 17.8 | 5.8 | 71.30 |

**Figure 3.25 Spot 551 Albumin**

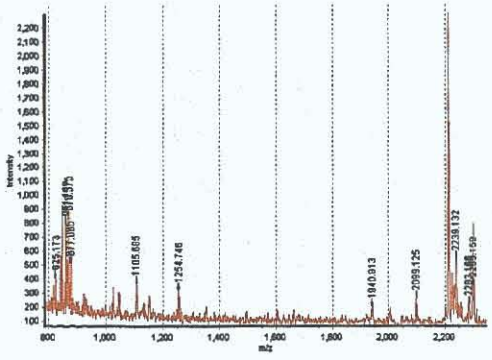
**Master number: 552**
**Identified by MALDI-TOF MS (Zip Tip) – Serum Albumin**




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|23307791\|gb\|AAN17824.1 - serum albumin [Bos taurus] | 29.3 | 5.8 | 71.30 |

**Figure 3.26 Spot 552 Albumin**

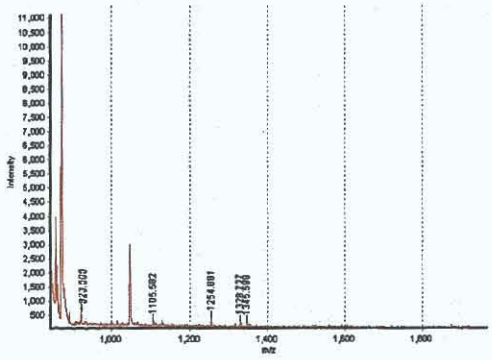
**Master number: 674**
**Direct spot analysis by MALDI-TOF MS–No Identification required Zip Tip**



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.409 | gi\|2145095\|gb\|AAB58419.1 - cis Golgi-localized calcium-binding protein | 6.5 | 4.7 | 81.92 |

**Figure 3.27 Spot 674 Calcium binding protein**

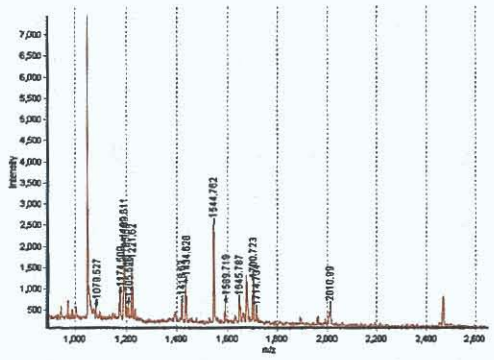
**Master number: 674**
**Zip Tip resulting in identification by MALDI-TOF MS–UDP-N-acteylglucosamine pyrophosphorylase1**



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.007 | gi\|55665560\|emb\|CAH72978.1 - UDP-N-acteylglucosamine pyrophosphorylase 1 [Homo sapiens] | 10.7 | 6.0 | 57.41 |

**Figure3.28 Spot 674 UDP-N-acteylglucosamine pyrophosphorylase 1**

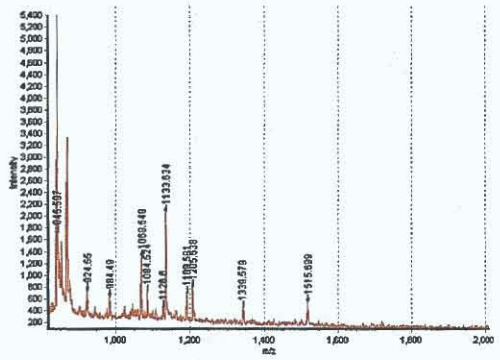
## Master number: 731
## Identified by MALDI-TOF MS (Zip Tip) – Aldehyde dehydrogenase 1



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.001 | gi\|2183299\|gb\|AAC51652.1 - aldehyde dehydrogenase 1 [Homo sapiens] | 19.2 | 6.3 | 55.44 |

**Figure 3.29 Spot 731 Aldehyde dehydrogenase**

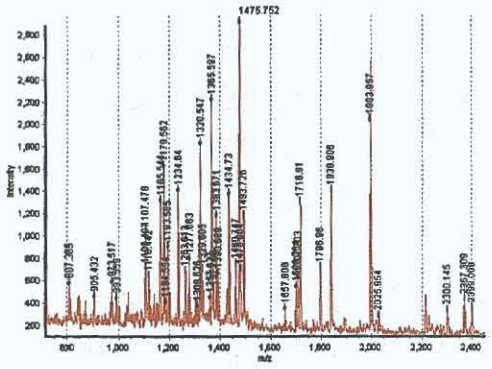
## Master number: 780
## Identified by MALDI-TOF MS (Direct spotting) – Splicing factor 3



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.008 | gi\|55665971\|emb\|CAH69930.1 - splicing factor 3a, subunit 3, 60kDa [Homo sapiens] | 15.8 | 5.3 | 59.17 |

**Figure 3.30 Spot 780 Slicing factor 3a**

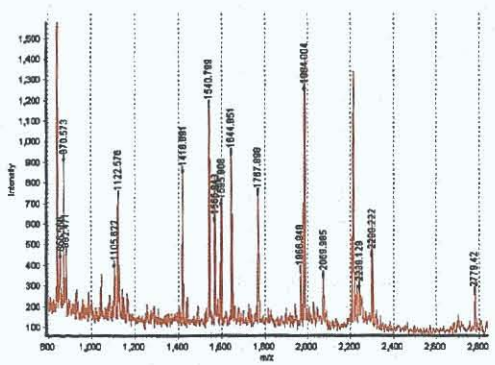
## Master number: 810
## Identified by MALDI-TOF MS (Direct spotting) – Keratin type 1




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.001 | gi\|547749\|sp\|P13645\|K1C10_HUMAN - Keratin, type I cytoskeletal 10 (Cytokeratin 10) (K10) (CK 10) | 22.1 | 5.1 | 59.73 |

**Figure 3.31 Spot 810 Keratin 1**

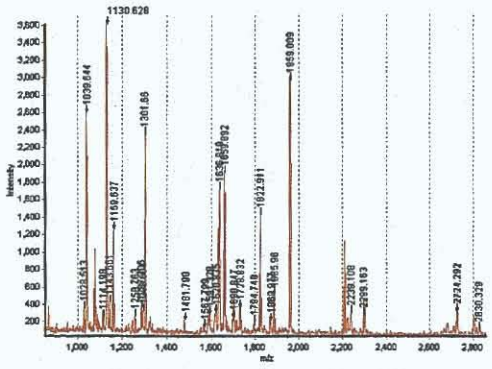
## Master number: 893
## Identified by MALDI-TOF MS (Zip Tip)– Polymerase




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|21430983\|gb\|AAM51148.1 - polymerase (DNA directed), delta 2, regulatory subunit (50kD) [Homo sapiens] | 29.2 | 5.3 | 51.90 |

**Figure 3.32 Spot 893 Polymerase 1**

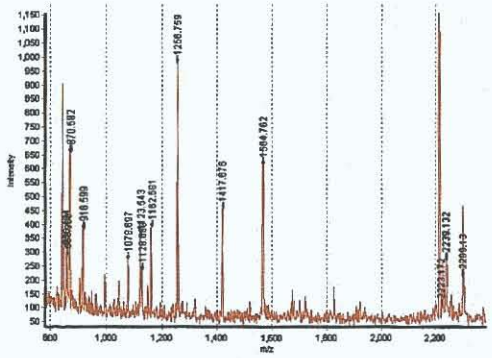
## Master number: 1059
## Identified by MALDI-TOF MS (Direct spotting) – Beta Tubulin



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|18088719\|gb\|AAH20946.1 - Tubulin, beta polypeptide [Homo sapiens] | 35.4 | 4.7 | 50.11 |

**Figure 3.33 Spot 1059 Beta Tubulin**

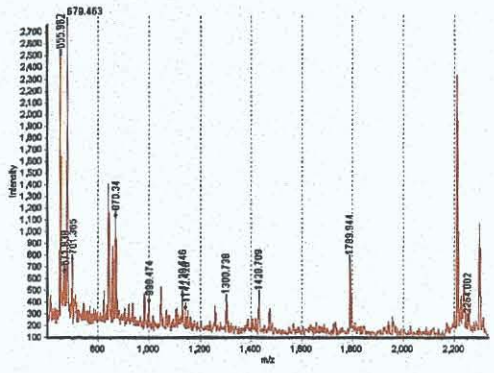
## Master number: 1412
## Identified by MALDI-TOF MS (Direct spotting) – Splicing factor



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|62088696\|dbj\|BAD92795.1 - splicing factor, arginine/serine-rich 1 (splicing factor 2, alternate splicing factor) variant [Homo sapiens] | 32.6 | 9.3 | 25.78 |

**Figure 3.34 Spot 1412 Splicing factor 2**

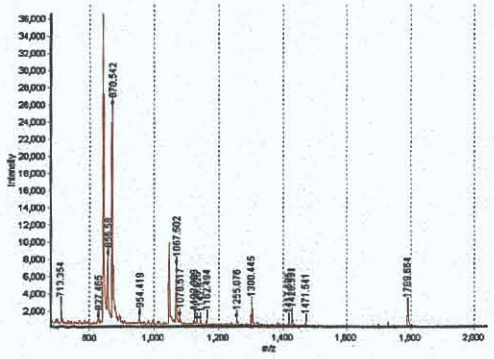
**Master number: 1423**
**Identified by MALDI-TOF MS— Direct spotting - Retinoblastoma binding protein 7 poor confidence**



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.024 | gi\|57209889\|emb\|CAI41284.1 - retinoblastoma binding protein 7 [Homo sapiens] | 13.6 | 4.9 | 47.60 |

**Figure 3.35a Spot 1423 Retinoblastoma binding protein 7**

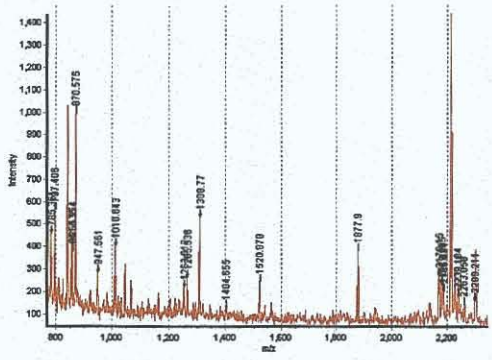
**Master number: 1423**
**Identified by MALDI-TOF MS— Zip Tip - Retinoblastoma binding protein 7**



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 * | 0.000 | gi\|57209888\|emb\|CAI41283.1 - retinoblastoma binding protein 7 [Homo sapiens] | 16.8 | 5.1 | 52.70 |

**Figure 3.35b Spot 1423 Retinoblastoma binding protein 7**

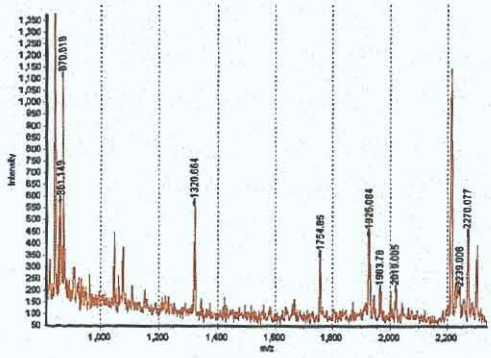
**Master number: 1425**
**Identified by MALDI-TOF MS (Direct spotting)– Cathepsin X**




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.002 | gi\|7546546\|pdb\|1EF7\|B - Chain B, Crystal Structure Of Human Cathepsin X | 33.1 | 5.5 | 27.76 |

**Figure 3.36 Spot 1425 Cathepsin X**

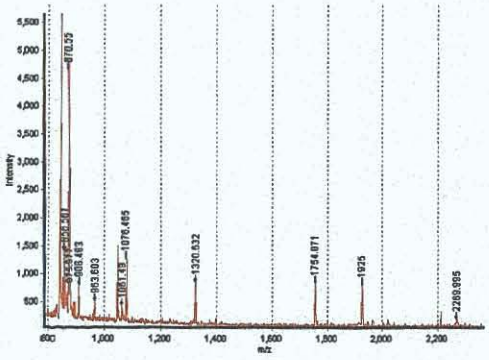
**Master number: 1511**
**Identified by MALDI-TOF MS (Direct spotting) – low confidence protein identification**




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.217 | gi\|26984283\|gb\|AAN85282.1 - J-type co-chaperone HSC20; HSCB [Homo sapiens] | 14.5 | 7.9 | 27.72 |

**Figure 3.37a Spot 1511 Low confidence HSC20**

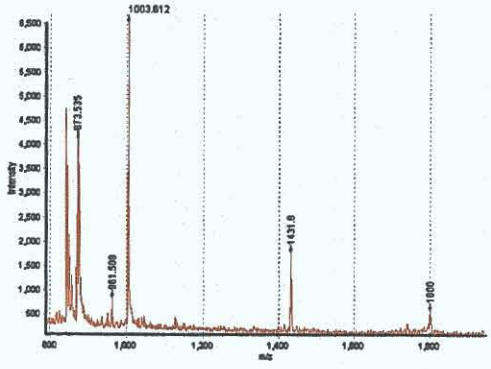
**Master number: 1511**
**Identified by MALDI-TOF MS (Zip Tip)– MAPRE 1**




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|5748523\|emb\|CAB53072.1 - MAPRE1 [Homo sapiens] | 27.2 | 5.0 | 30.15 |

**Figure 3.37b Spot 1511 MAPRE 1**

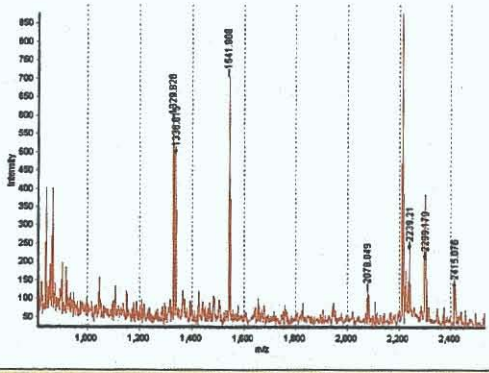
## Master number: 1518
## Identified by MALDI-TOF MS (Zip Tip) – Superoxide dismutase copper chaperone




| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.005 | gi\|21429608\|gb\|AAM50090.1 - superoxide dismutase copper chaperone [Homo sapiens] | 16.4 | 5.3 | 29.52 |

**Figure 3.38 Spot 1518 Superoxide dismutase copper chaperone**

## Master number: 1678
## Identified by MALDI-TOF MS (Zip Tip) – Molecular Basis For The Local Conformational Rearrangement Of Human Phosphoserine Phosphatase

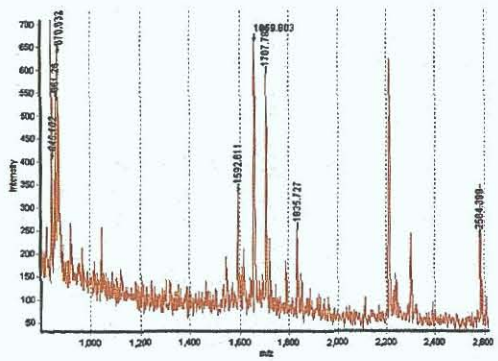



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.003 | gi\|29726394\|pdb\|1L8O\|B - Chain B, Molecular Basis For The Local Conformational Rearrangement Of Human Phosphoserine Phosphatase | 34.7 | 5.5 | 25.21 |

**Figure 3.39 Spot 1678 Molecular Basis For The Local Conformational Rearrangement Of Human Phosphoserine Phosphatase**

## Master number: 1826

## Identified MALDI-TOF MS (Direct spot) – Crystal Structure Of Human Dj-1 with a low confidence score

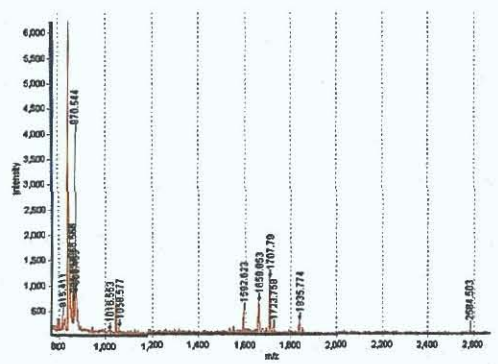


| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.033 | gi\|42543006\|pdb\|1J42\|A - Chain A, Crystal Structure Of Human Dj-1 | 33.3 | 6.3 | 20.06 |

**Figure 3.40a Spot 1826 Dj-1**

## Master number: 1826

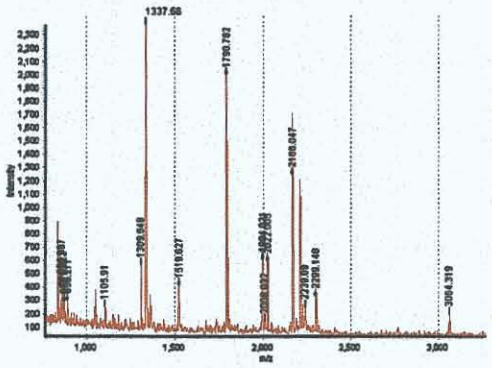
## Identified by MALDI-TOF MS (Zip Tip) – Crystal Structure Of Human Dj-1



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|42543006\|pdb\|1J42\|A - Chain A, Crystal Structure Of Human Dj-1 | 49.2 | 6.3 | 20.06 |

**Figure 3.40b Spot 1826 Dj-1**

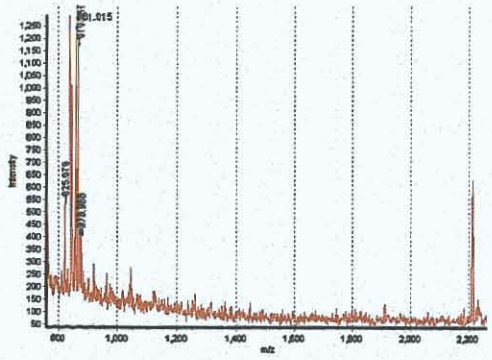
**Master number: 1973**
**Identified by MALDI-TOF MS (Zip Tip) – Cofilin**



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|30582531\|gb\|AAP35492.1 - cofilin 1 (non-muscle) [Homo sapiens] | 36.1 | 8.5 | 18.71 |

**Figure 3.41 Spot 1973 Cofilin 1**

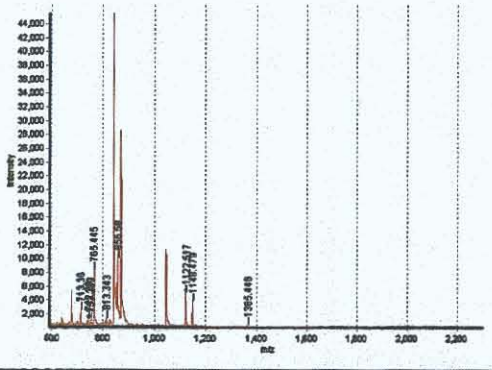
**Master number: 2100**
**Direct spot sample no Identification by MALDI-TOF MS**



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | | | | | |

**Figure 3.42a Spot 2100 No protein identification possible**

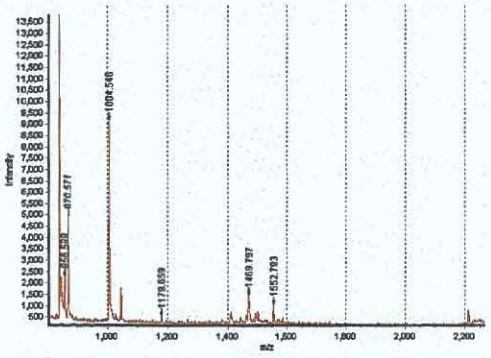
**Master number: 2100**
**Identified by MALDI-TOF MS (Zip Tip) – Human Class I Histocompatibility Antigen A2 and gi|179318|gb|AAA51811.1 - beta-2-microglobulin [Homo sapiens]**



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.000 | gi\|229995\|pdb\|1HLA\|M - Chain M, Human Class I Histocompatibility Antigen A2 (HLA-A2, Human Leucocyte Antigen) | 39.2 | 6.5 | 11.58 |
| 1 * | - | gi\|179318\|gb\|AAA51811.1 - beta-2-microglobulin [Homo sapiens] | 26.9 | 6.1 | 13.83 |

**Figure 3.42b Spot 2100 Two proteins identified Human Class I Histocompatibility Antigen A2 and beta-2-microglobulin**

**Master number: 2113**
**Identified by MALDI-TOF MS (Zip Tip) - Cytochrome c oxidase subunit VIb [Homo sapiens]**



| Rank | Expectation | Protein information | Coverage | pI | Mass [kDa] |
|---|---|---|---|---|---|
| 1 | 0.002 | gi\|30582729\|gb\|AAP35591.1 - cytochrome c oxidase subunit VIb [Homo sapiens] | 53.5 | 6.6 | 10.40 |

**Figure 3.43 Spot 2113 Cytochrome C oxidase**

### 3.13 Protein Identification using nanoLC-MS/MS

NanoLC-MS/MS was carried out using the Ettan MDLC (GE Healthcare) and LTQ-Ion Trap (Thermo). At least two peptides of experimental MS/MS data matched to the theoretical candidate protein internal sequence was considered as a successful identification.

**High-throughput configuration**

Nanobore Reverse Phase Chromatography (RPC) is a method commonly used for peptide separation of less complex mixtures prior to identification by mass spectrometry. In the typical set up, this method requires a set of two columns–an RPC trap and an RPC column–connected in series.

Ettan MDLC increases the duty cycle of the MS instrument by using a combination of two parallel sets of RPC trap/RPC columns in a fully automated and synchronized setup. While the sample is run on the first set of columns, equilibration is performed on the second set, which ensures uninterrupted sample analysis and a major improvement in productivity. The high-throughput configuration is designed for analysis of larger sample sets of less complex mixtures, e.g. digested gel plug. Sample injection is handled by Autosampler A-905, which enables automated injections of up to 384 samples.

The Thermo Finnigan LTQ was configured in the high-throughput mode using two C18 traps and a 10 cm x 75mm PicoFrit column (Agilent). Ten microliters of the trypsin-digested peptides were injected using the MDLC. Peptides were eluted with

a 180-minute gradient from 0.1% formic acid to 65% acetonitrile with 0.1% formic acid at a flow rate of 300 nL/min. The Finnigan LTQ was run in positive ion mode using the nanospray source.

Spray voltage: 1.8 kV

Capillary temperature: 200.00C

Capillary voltage: 44.00

Tube lens (V): 220.00

Full MS: 2000,200 m/z

Data Dependent MS/MS Data Dependent MS/MS. The method used consisted of a continual cycle beginning with one scan of MS (scan one), which recorded all of the *m/z* values of the ions present at that moment in the gradient, followed by three rounds of MS/MS. The initial MS/MS scan was of the first most-intense ion recorded from the MS scan. The second MS/MS scan was of the second most-intense ion recorded from scan one. The third MS/MS scan was the third most-intense ion recorded from scan one. Dynamic exclusion was activated which improves the protein identification capacity during the analysis. Dynamic exclusion prevents an abundant peptide ion, with a broad elution profile, from being continually selected for MS/MS. The intense ions prevent other lower-abundance peptides from being selected. Ions that were selected for MS/MS were placed onto a list. While on this list, they will be excluded for a period of time so that other ions may be selected. The exclusion time set in this study was 90 seconds.

The data was searched using TurboSEQUEST in BioWorks 3.1, selecting modification by carboxymidomethlyation of cysteines (+57.02 Da) and partial modification of methionine (+14.03 Da). One internal trypsin cleavage site was allowed and a tolerance of 1.4Da for precursor ions.

Iodacetamide is used in the 2DE process to prevent formation of disulfide bonds following IEF which reacts with cysteine residues forming S-carboxyamidomethylcysteine which adds +57.02Da to each modified cysteine (Lapko *et al,* 2000). Partial oxidation of methionines causes an addition of 14.03Da

After acquisition of MS/MS spectra, the software algorithm SEQUEST (Yates *et al,* 1998) was employed for peptide/protein identification. SEQUEST matches an MS/MS spectrum with a peptide sequence using the following steps. (1) All theoretical peptides with molecular masses isobaric to that of the precursor ion for an MS/MS spectrum are extracted from the database (SWISS-PROT). (2) Each peptide was given a preliminary score by examining the number of predicted fragment ions from the database peptide that match the acquired fragment ions in the MS/MS spectrum. (3) The 500 best-matching peptides then undergo a more rigorous ion matching algorithm that generates a SEQUEST cross-correlation score (Xcorr). A list of the best matching peptides are returned with the top-scoring peptide being considered the best candidate (Peng *et al,* 2003).

Cross-correlation scores (XCorrs) are generated using the SEQUEST algorithm which takes into account fragment ions derived from precursor ions matching them

to model spectra in the database. The Xcorr of a peptide is based on the "fit" of the MS/MS data to the theoretical distribution of ions produced for the peptide.

A SEQUEST identification was deemed correct if it met the following criteria: XCorr > 1.5 for singly charged ions, XCorr > 2.0 for doubly charged ions, and XCorr > 2.5 for triply charged ions, peptide length was at least seven residues, at least 30% of predicted b/y ions were accounted for within the spectrum, the peptide was fully tryptic (containing a Lys or Arg residue on the C-terminus), and at least one other peptide from the same protein was identified (Frewen *et al,* 2006). Using these parameters, BioWorks identified seven proteins. Coverage was typically excellent, e.g., 36% coverage for bovine serum albumin from 13 peptides distributed throughout the protein.

Peptide samples were trapped and concentrated on C18 trap columns before being separated using an acetonitrile gradient (0.1-65% over 180mins) through a C18 reverse phase column. Peptides were analysed by MS and MS/MS. Fragmented peptides were searched using SEQUEST™ against the SWISS PROT database. Spot 160-vinculin (figure 3.13.1), 810-Keratin 1 (figure 3.13.2), 1059-Beta Tubulin (figure 3.13.3), 2213 Cytochrome C (figure 3.13.4), 1432 Cathepsin X (figure 3.13.5) and 1678 Molecular Basis For The Local Confomational Rearrangement Of Human (figure 3.13.6) and 2104 Cytochrome C (figure 3.13.7) confirmed the protein identifications from the MALDI-TOF MS results.

Spot 2104 was not identified by MALDI-TOF MS however using nanoLC-Ion Trap Mass Spectrometery it was identified as Cytochrome C (figure 3.50).

Results below show the full MS spectrum of all the peptide peaks eluted from the RPC column during the 180 minute Acetonitrile gradient. Relative abundance of the peaks is plotted against time. An individual MS spectrum shows the parent ions at any one time. The three most abundant ions are automatically selected for MSMS fragmentation and are then excluded from fragmentation for 90 seconds. The next three abundant peptides are then selected for fragmentation and they are then excluded for 90 seconds. Once a parent ion (peptide) selected it is fragmented with collisional energy to produce a fragmentation pattern of daughter ions (MS/MS spectrum).

There are three different types of bonds that can fragment along the amino acid backbone: the NH-CH, CH-CO and CO-NH bonds. Each bond breakage gives rise to two species, one neutral and the other one charged, and only the charged species is monitored by the mass spectrometer. The charge can stay on either of the two fragments depending on the chemistry and relative proton affinity of the two species. Hence there are six possible fragment ions for each amino acid residue and these are either a, b and c ions having the charge retained on the N-terminal fragment and x, y and z ions having the charge retained on the C-terminal fragment. The most common cleavage sites are at the CO-NH bonds which give rise to the b and/or the y ions. The mass difference between two adjacent b ions or y ions is indicative of a particular amino acid residue.

In total 24 proteins were identified using a combination of MALDI-TOF MS and nanoLC-Ion trap MS.

### 3.13.1 Master number: 160

**Identified by nanoLC-MSMS - Vinculin protein**

Having identified Vinculin by MALDI-TOF MS the remaining peptides were resuspended in 20μL of 0.1%FA. The peptides were separated at 300nL/min over a long gradient 0-65% Acetonitrile in 180 minutes directly into an Ion Trap Mass Spectrometers where MS and MSMS analysis was carried out. Figure 3.44a shows a base peak chromatogram plotting Relative abundance vs Time of ions detected by the Mass Spectrometer. Two prominent peaks are the double charged species of trypsin 1 and trypsin 2 eluting at 55.87 and 80.71 minutes respectively. Peptides related to Vinculin are low abundant and not visually detected in this Chromatographic view.

Figure 3.44b shows the full MS spectrum generated at 49 minutes into the gradient. Ions at 313, 341, 371 and 445m/z are baseline peaks present at all times and are mixture of solvent peaks, plasticer peaks and silica peaks. These ions are on an exclusion list and no MSMS data is generated from them. The parent ion at 523m/z is also on an exclusion list. This ion is the double charge species of tryspin 2 (1046m/z) and no MSMS data is generated from this ion. The parent ion at 736.0m/z was automatically selected for MSMS fragmentation and figure 3.44c shows the fragmentation pattern generated. Figure 3.44d shows the SEQUEST result of the fragmentation pattern generated in figure 3.44c. SEQUEST used crossed-correlation analysis between the experimental MSMS spectrum and theoretically generated

mass spectra from peptide sequences with a similar molecular weight to the experimental parent ion. In this case the fragmentation ions generated from the parent ion at 736.0m/z matches the majority of b and y ions associated with the amino acid sequence DPSASPGDAGEQAIR of the protein vinculin with a cross correlation (XCorr) score of 5.13. Twenty other peptide sequences relating to vinculin were identified in that experiment resulting in 26% amino acid coverage of the protein by MSMS analysis (figure 3.44e).

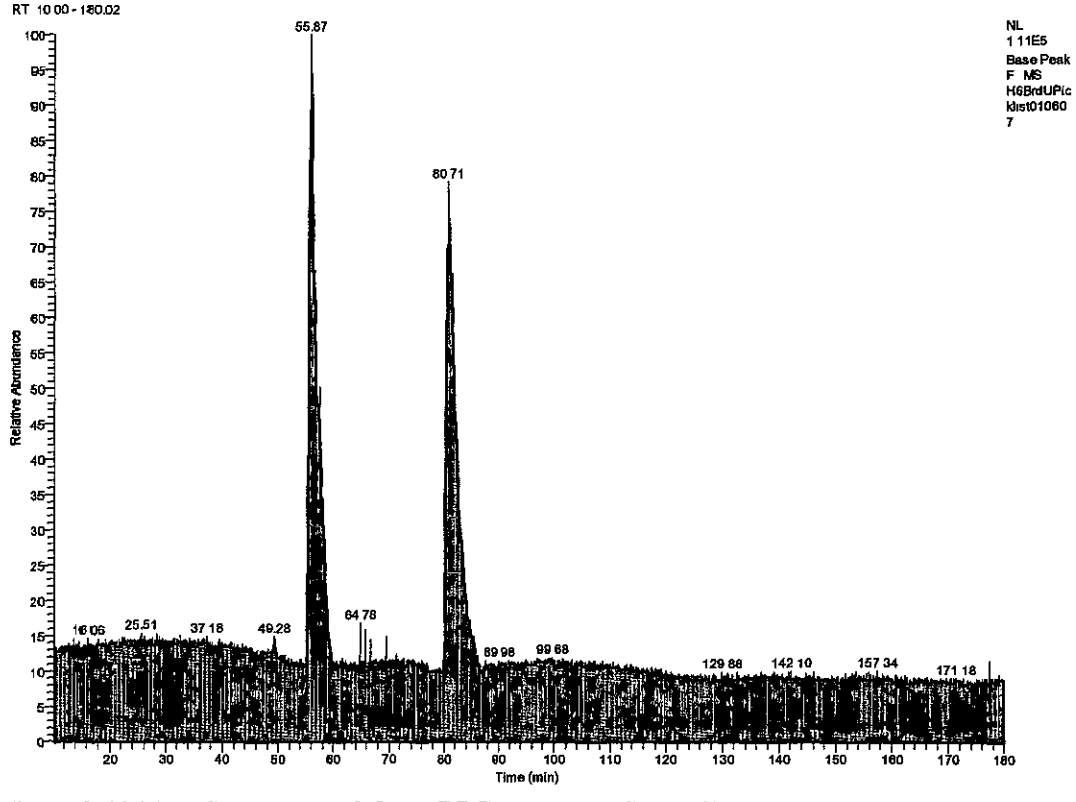



**Figure 3.44 (a). MS spectrum of eluate RPC separation Spot 160. The digested peptide sample was redissolved in RPC solvent A (0.1% formic acid), and injected onto Ettan MDLC equipped with a Zorbax™ SB-300 C18 3.5 mm column. Peptides were eluted with RPC solvent B (0.1% formic acid, 86% acetonitrile) at a flow rate of approximately 300 nl/min and analyzed with a Finnigan LTQ (Thermo Electron Corp.) fitted with a nanospray interface.**

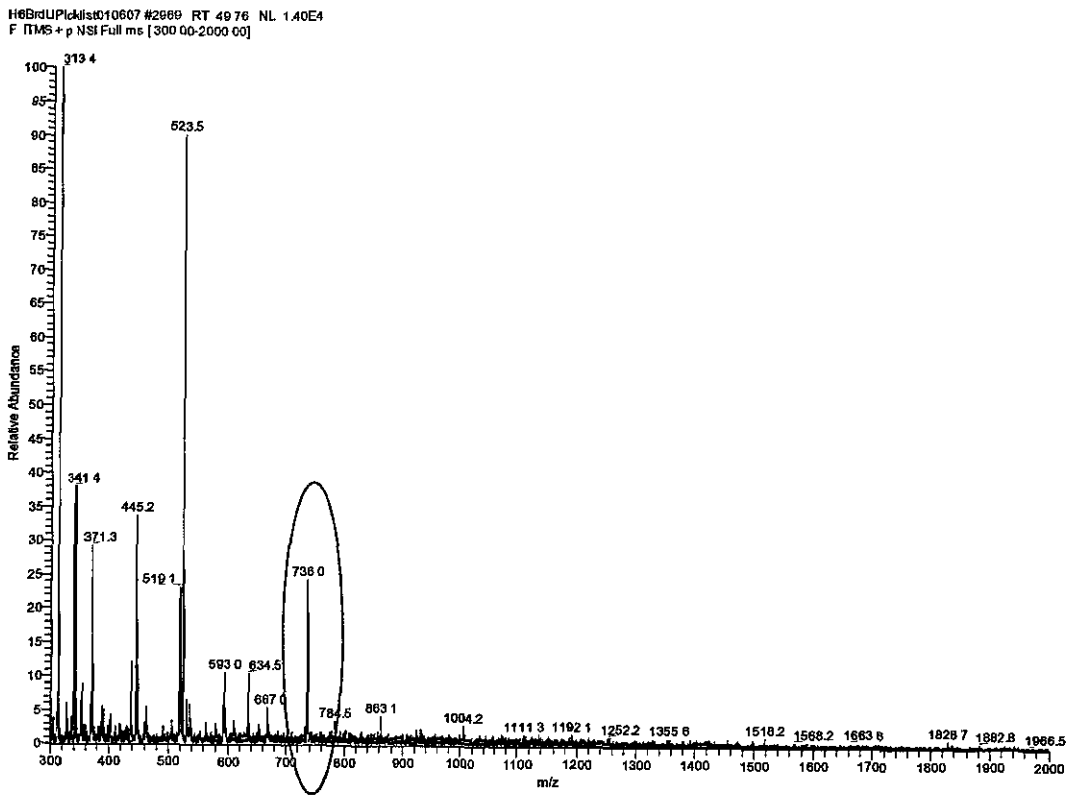



**Figure 3.44 (b) Mass spectrum for retention time 49.76. Parent Ion 736.0 m/z was chosen for MS/MS**

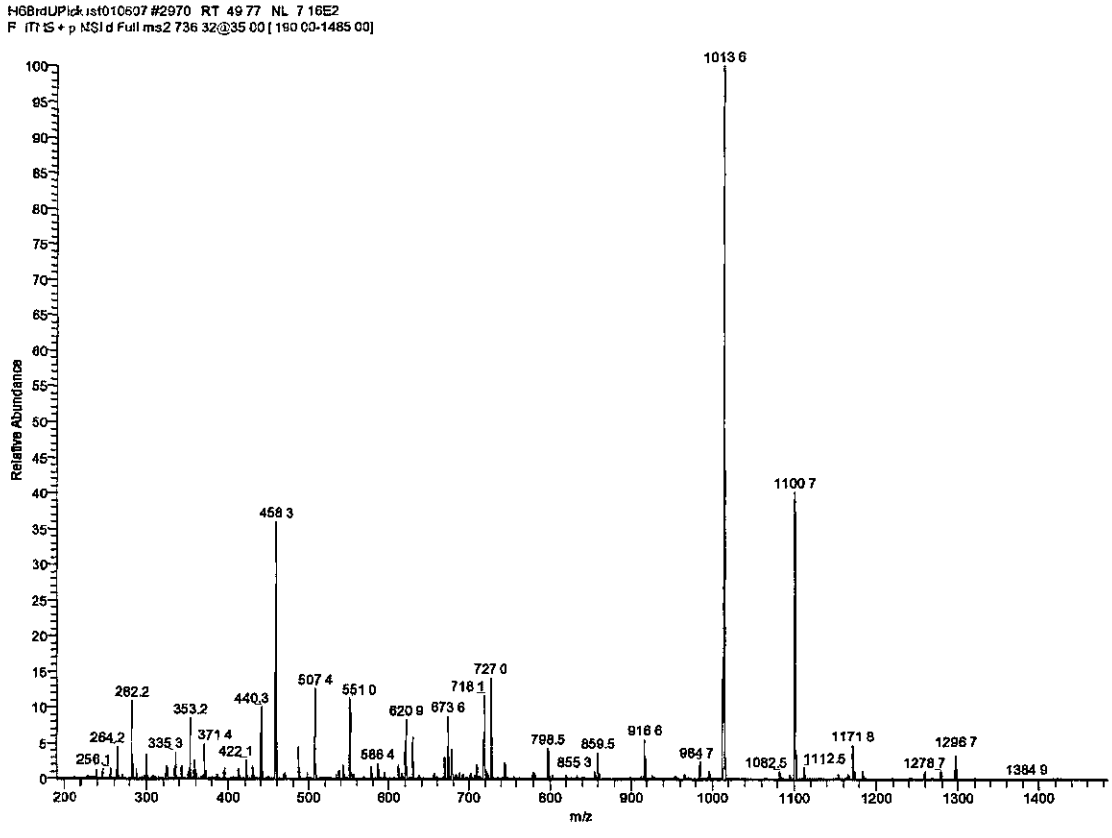



**Figure 3.44 (c) MS/MS spectrum of the generated product ions from doubly charged ion 736.3m/z**

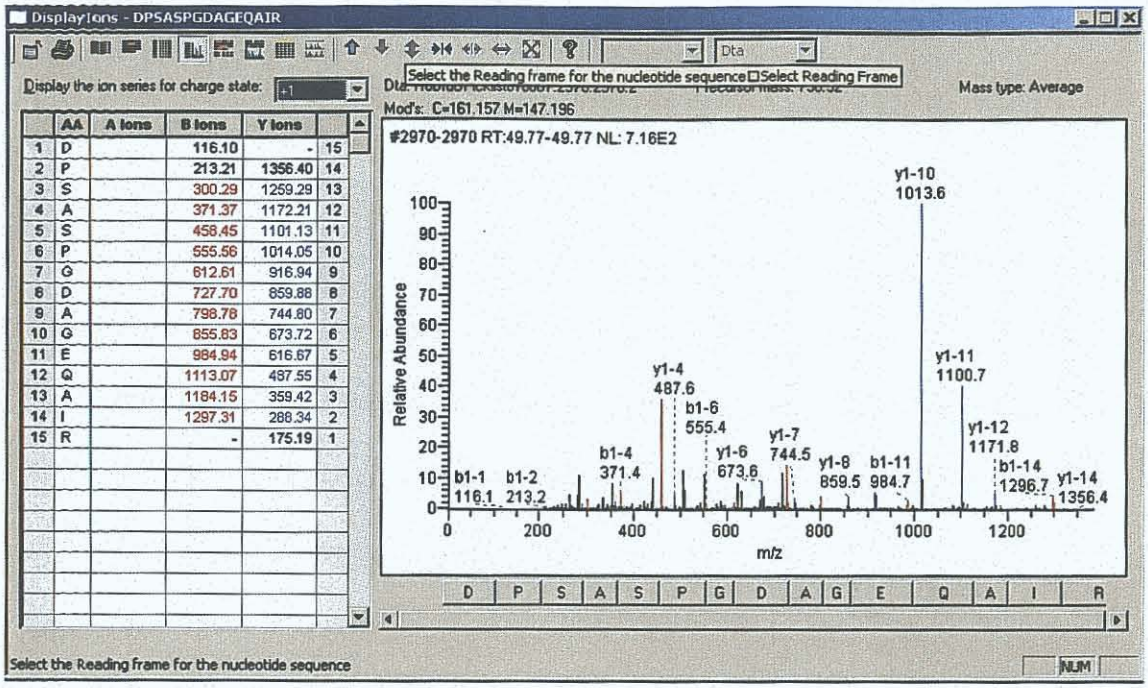



**Figure 3.44 (d) Based on this analysis, the peptide DPSASPGDAGEQAIR of the protein vinculin was identified. TurboSEQUEST™ cross correlation (XCorr) = 5.13**

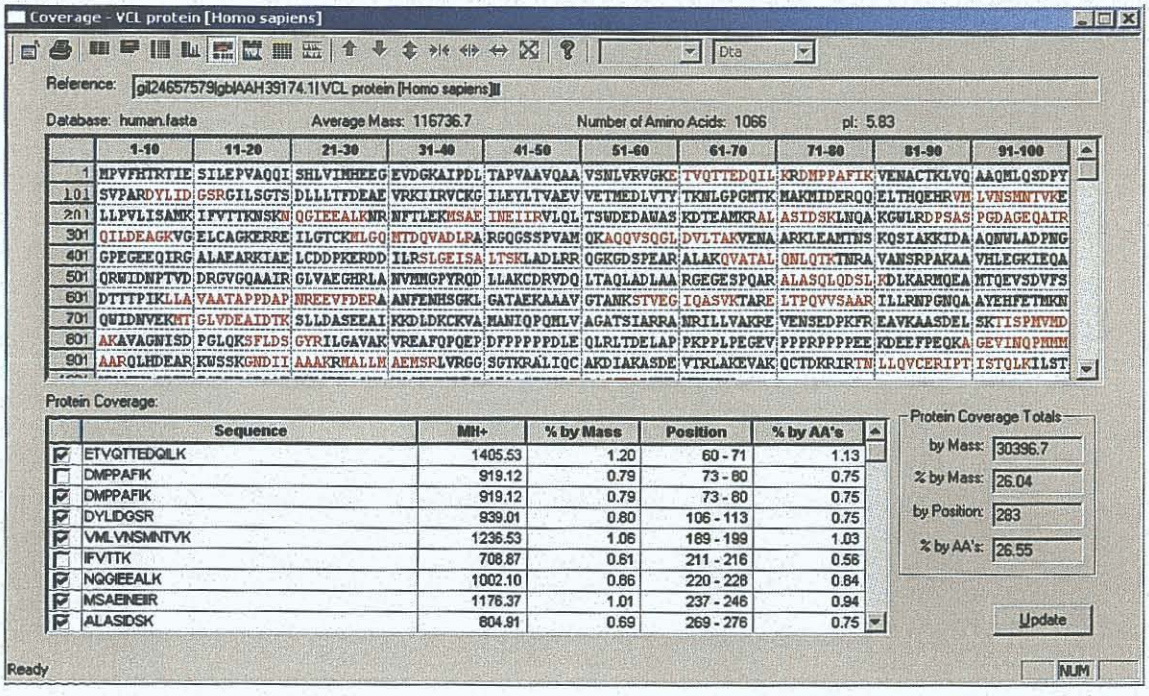



**Figure 3.44 (e) BioWorks identified Vinculin (VCL protein-Homo sapiens) with a % amino acid coverage of 26.55% from 25 peptides distributed throughout the protein.**

### 3.13.2 Master number: 810

**Identified by nanoLC-MSMS - Keratin 1 protein**

Having identified Keratin by MALDI-TOF MS the remaining peptides were resuspended in 20µL of 0.1%FA. These peptides were separated at 300nL/min over a long gradient 0-65% Acetonitrile in 180 minutes directly into an Ion Trap Mass Spectrometers where MS and MSMS analysis was carried out. Figure 3.45a shows a base peak chromatogram plotting Relative abundance vs Time of ions detected by the Mass Spectrometer.

Figure 3.45b shows the full MS spectrum generated at 37.51 minutes into the gradient. Ions at 313, 341, 371 and 445m/z are contamination peaks present at all times and are mixture of solvent peaks, plasticer peaks and silica peaks. These ions are on an exclusion list so no MSMS data is generated from them. Four parent ions are present at 639.5m/z, 767.5m/z, 939.5m/z and 1022.5m/z are present in scan 4293 at 58.1 minutes into the run (figure 3.45b). The parent ion at 639.5m/z was automatically selected for MSMS fragmentation and figure 3.45c shows the fragmentation pattern generated. Figure 3.45d shows the SEQUEST result of the fragmentation pattern generated in figure 3.45c. SEQUEST used crossed-correlation analysis between the experimental MSMS spectrum and theoretically generated mass spectra from peptide sequences with a similar molecular weight to the experimental parent ion. In this case the fragmentation ions generated from the parent ion at 639.5m/z matches the majority of b and y ions associated with the

amino acid sequence LALDLEIATYR of the protein Keratin with a cross correlation (XCorr) score of 4.44. Sixteen other peptide sequences relating to Keratin 1 were identified in that experiment resulting in 31% amino acid coverage of the protein by MSMS analysis (figure 3.45e).

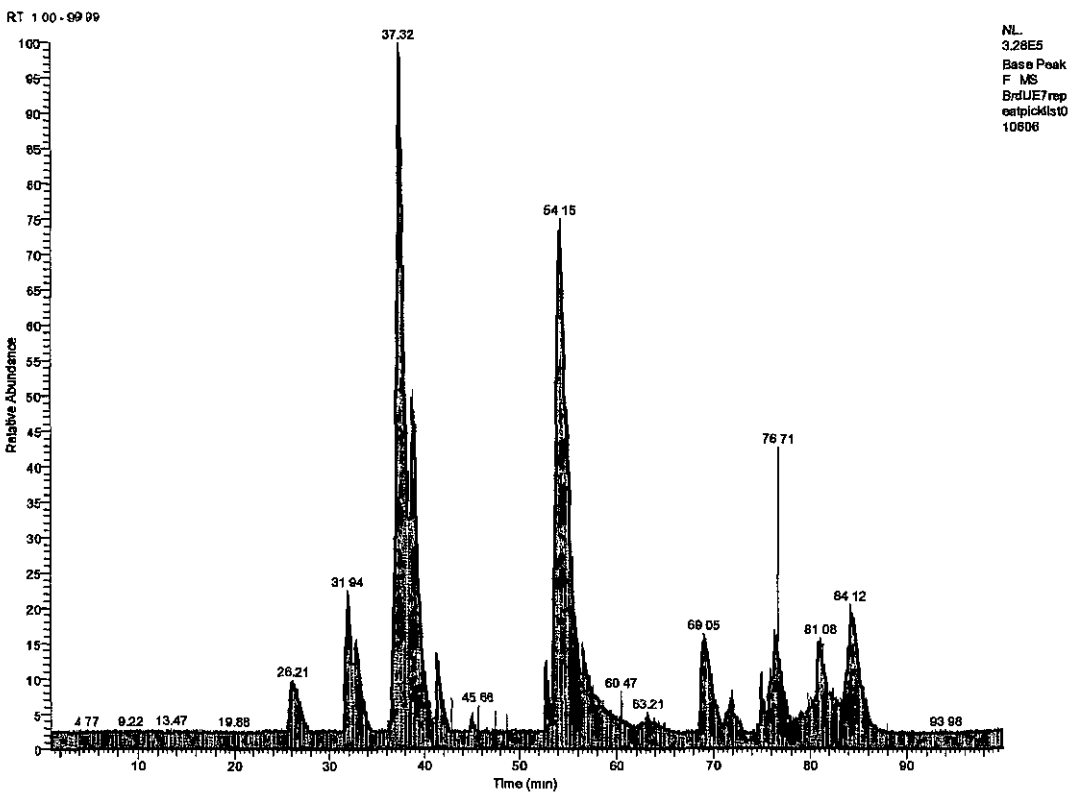



**Figure 3.45 (a) MS spectrum of eluate from RPC separation of spot 810. The digested peptide sample was injected onto Ettan MDLC equipped with a Zorbax™ SB-300 C18 3.5 mm column. Peptides were eluted with RPC solvent B (0.1% formic acid, 86% acetonitrile) at a flow rate of approximately 300 nl/min and analyzed with a Finnigan LTQ (Thermo Electron Corp.) fitted with a nanospray interface.**

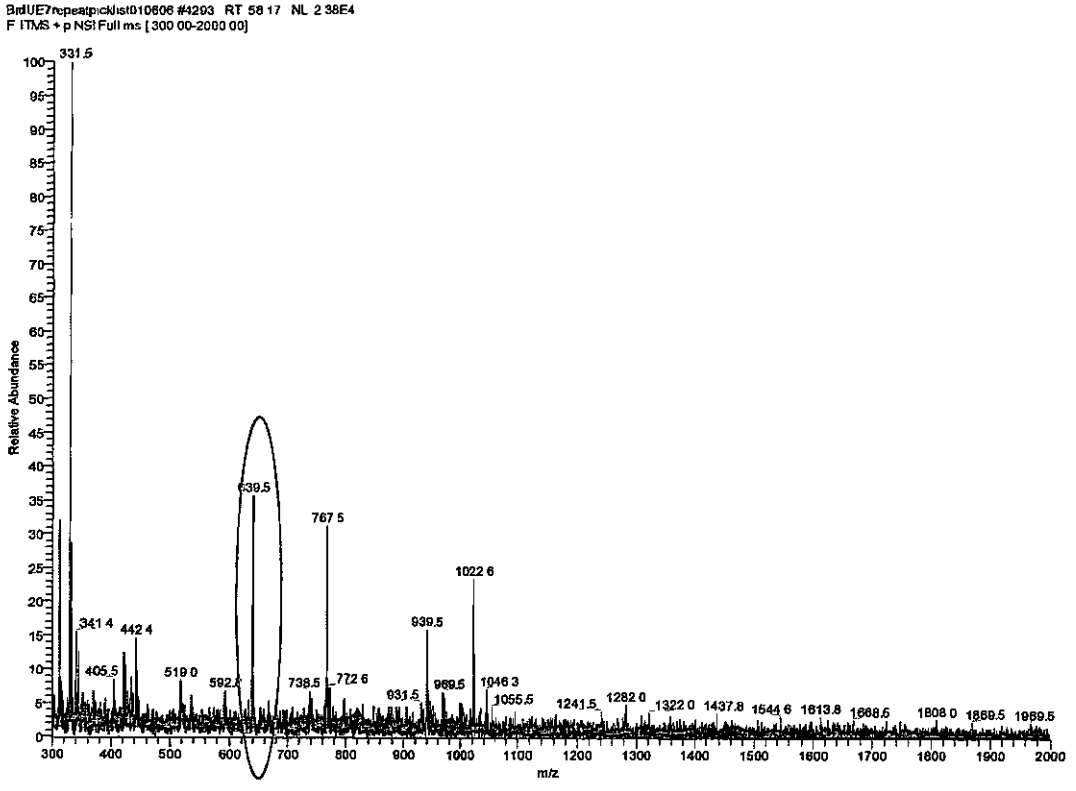



**Figure 3.45 (b) (A) Mass spectrum for retention time 37.51 min. Parent Ion 639.5m/z was chosen for MS/MS**

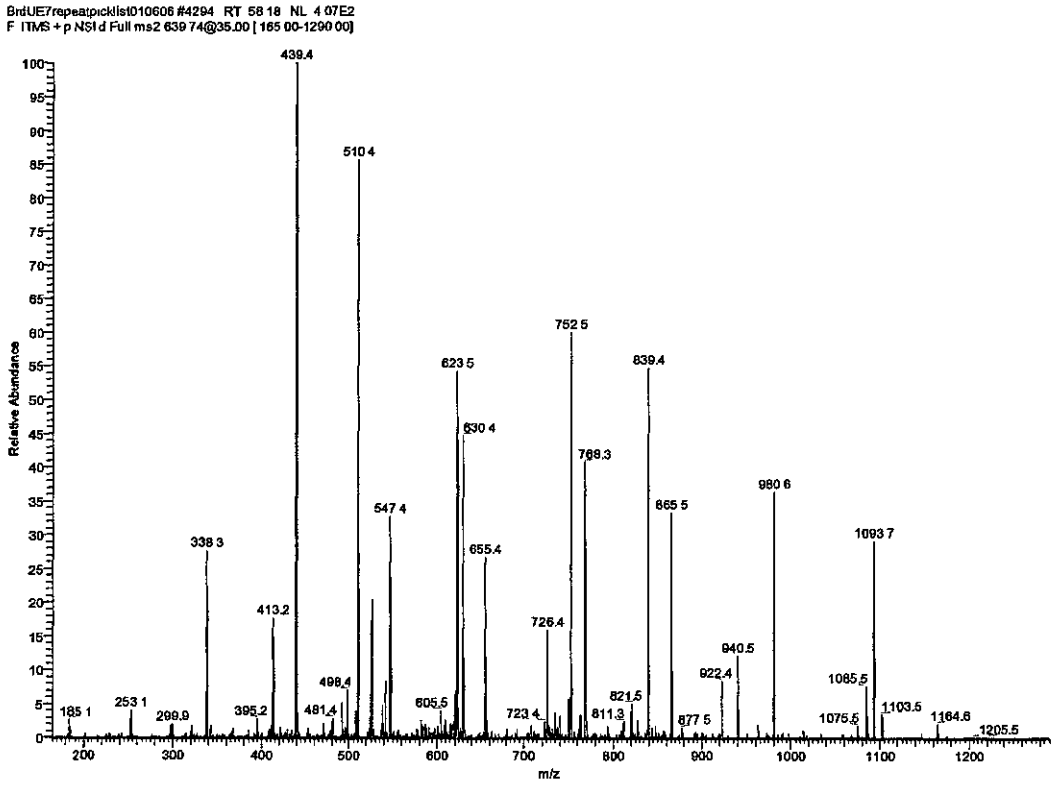



**Figure 3.45 (c) MS/MS spectrum of the generated product ions from doubly charged ion 639.7**

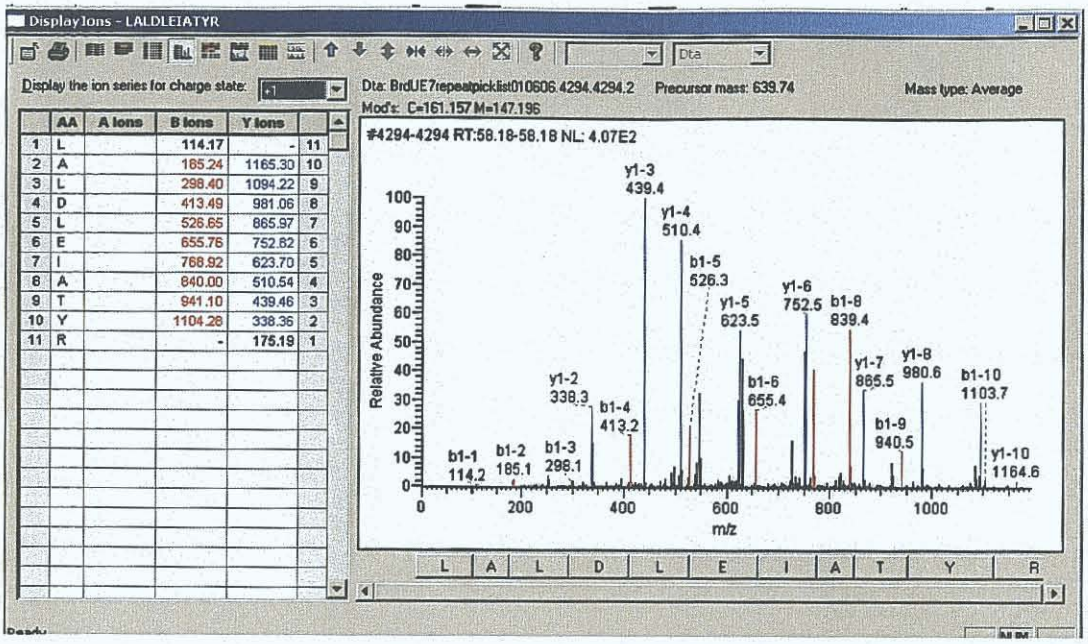



**Figure 3.45 (d) Based on this analysis, the peptide LALDLEIATYR of the protein Keratin 1 was identified (Figure 3.45e). TurboSEQUEST™ cross correlation (XCorr) = 4.44**

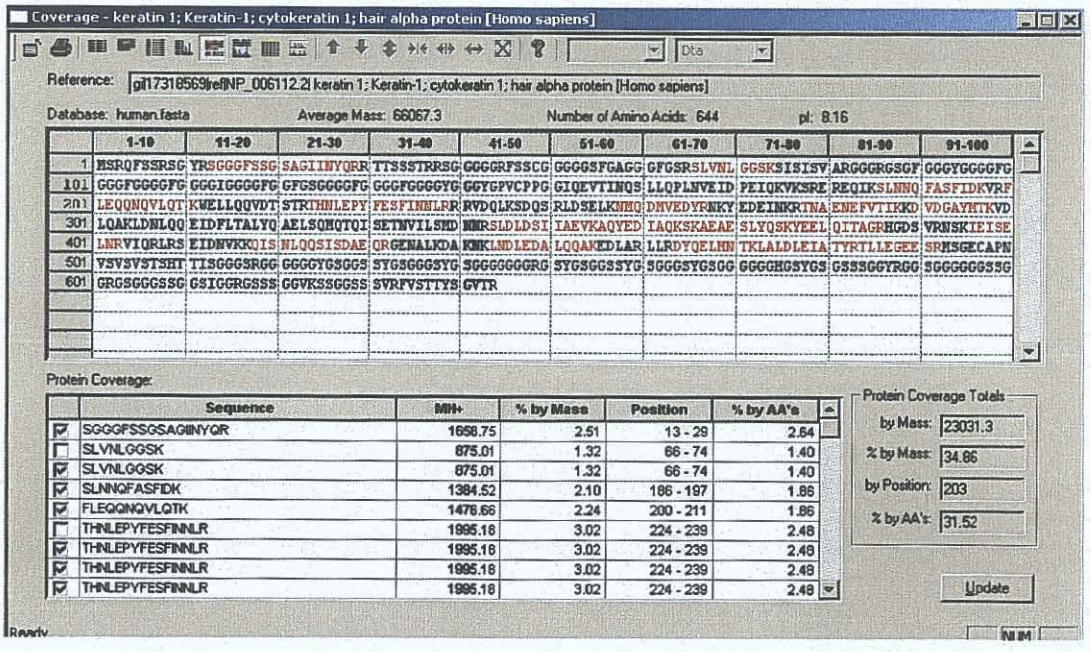



**Figure 3.45 (e) BioWorks identified Keratin 1(Homo sapiens) with a % amino acid coverage of 31.52% from 16 peptides distributed throughout the protein.**

### 3.13.3 Master number: 1059

### Identified by nanoLC-MSMS – Beta Tubulin

Having identified Beta-Tublin by MALDI-TOF MS the remaining peptides were resuspended in 20μL of 0.1%FA. These peptides were separated at 300nL/min over a long gradient 0-65% Acetonitrile in 180 minutes directly into an Ion Trap Mass Spectrometers where MS and MSMS analysis was carried out. Figure 3.46a shows a base peak chromatogram plotting Relative abundance vs Time of ions detected by the Mass Spectrometer.

Figure 3.46b shows the full MS spectrum generated at 67.53 minutes into the gradient. The parent ion at 668.6 m/z was automatically selected for MSMS fragmentation and figure 3.46c shows the fragmentation pattern generated. Figure 3.46d shows the SEQUEST result of the fragmentation pattern generated in figure 3.46c. SEQUEST used crossed-correlation analysis between the experimental MSMS spectrum and theoretically generated mass spectra from peptide sequences with a similar molecular weight to the experimental parent ion. In this case the fragmentation ions generated from the parent ion at 668.6 m/z matches the majority of b and y ions associated with the amino acid sequence IMTFSVVPSPK of the protein Beta-tubulin with a cross correlation (XCorr) score of 4.69. Ten other peptide sequences relating to beta-tubulin were identified in that experiment resulting in 32.27% amino acid coverage of the protein by MSMS analysis (see figure 3.46e).

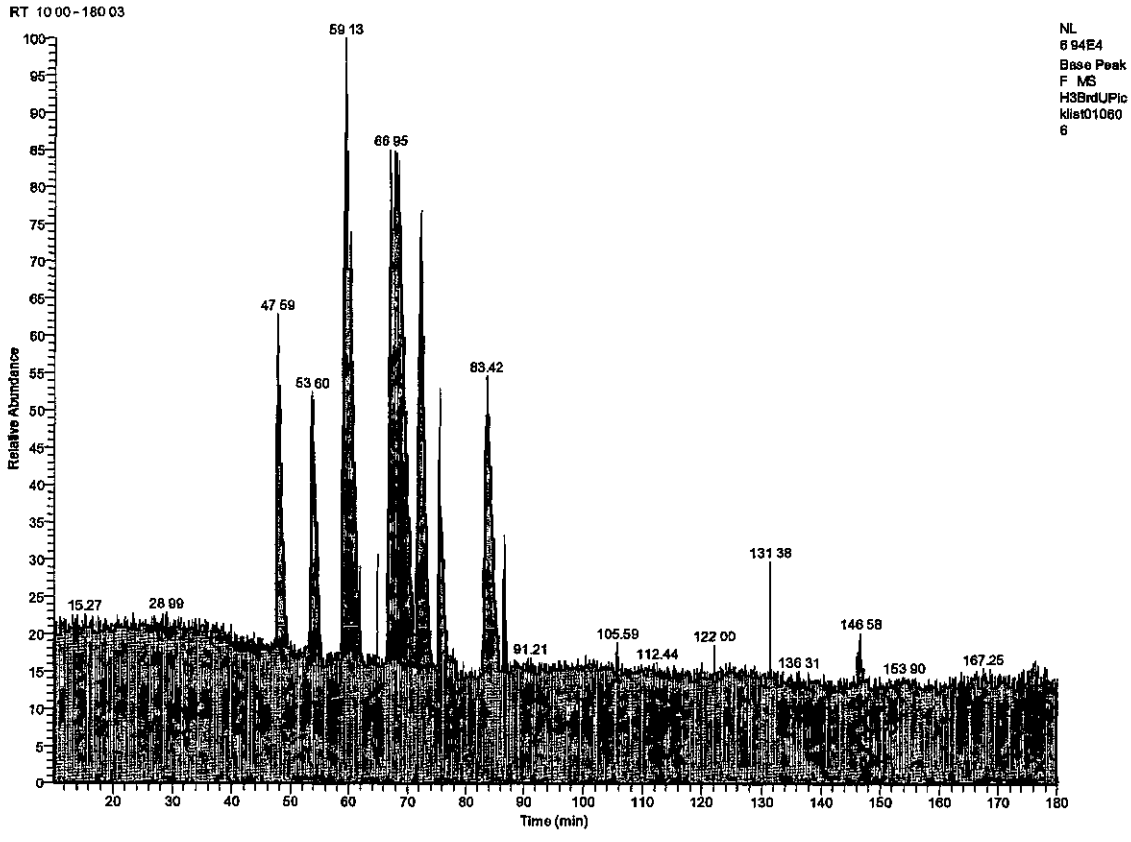



**Figure 3.46 (a) MS spectrum of eluate from RPC separation of spot 1059. The digested peptide sample was injected onto Ettan MDLC equipped with a Zorbax™ SB-300 C18 3.5 mm column. Peptides were eluted with RPC solvent B (0.1% formic acid, 86% acetonitrile) at a flow rate of approximately 300 nl/min and analyzed with a Finnigan LTQ (Thermo Electron Corp.) fitted with a nanospray interface.**

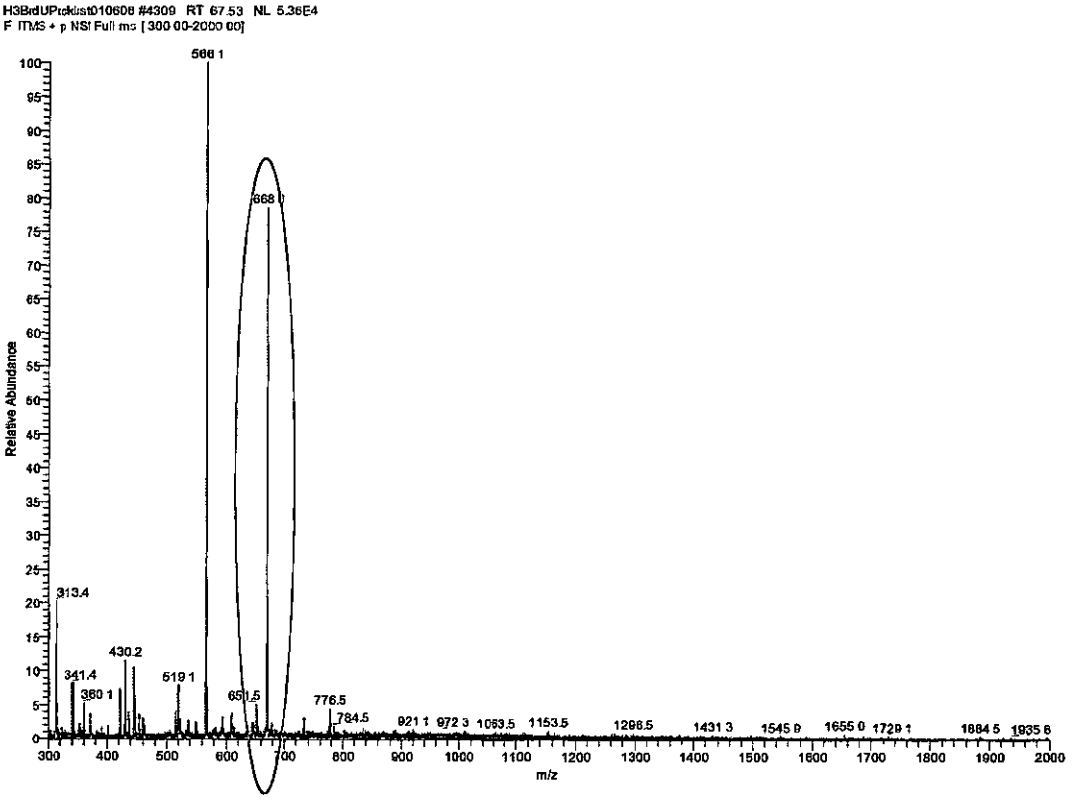



**Figure 3.46 (b) Mass spectrum for retention time 67.53 min as shown in Figure 2bB. Ion 668.6 was chosen for MS/MS**

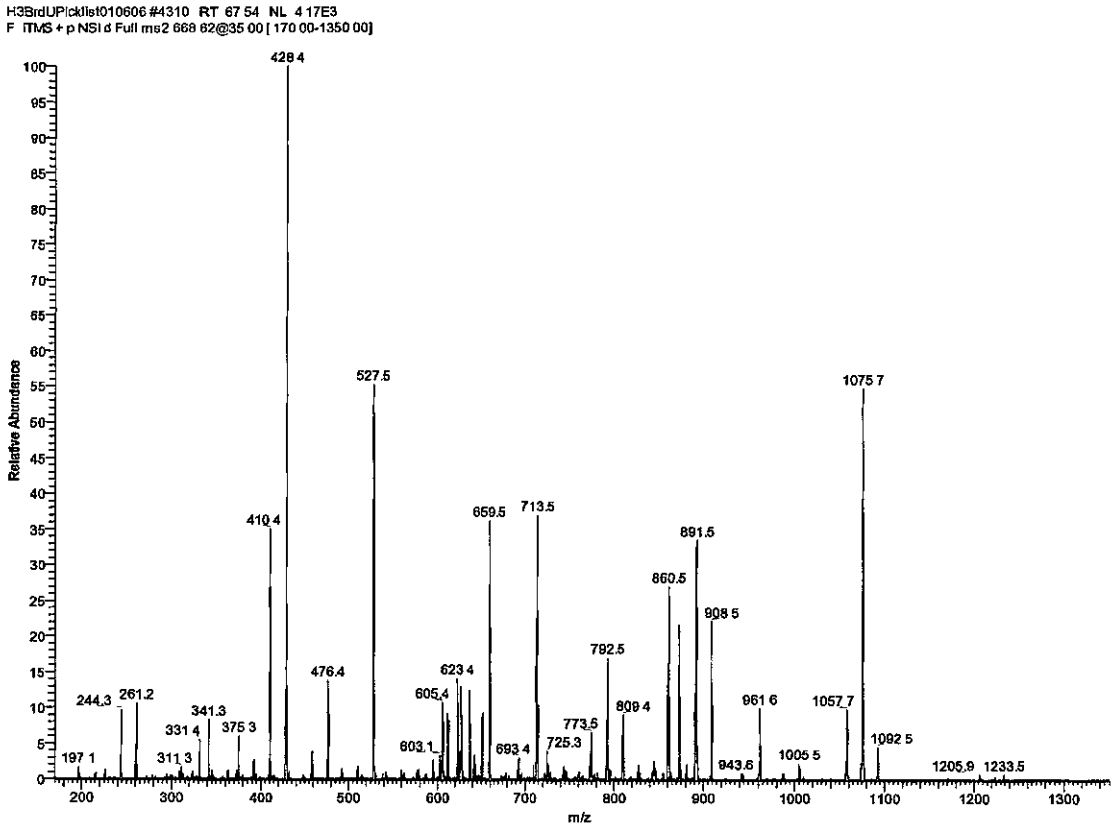



**Figure 3.46 (c) MS/MS spectrum of the generated product ions from doubly charged ion 668.6**

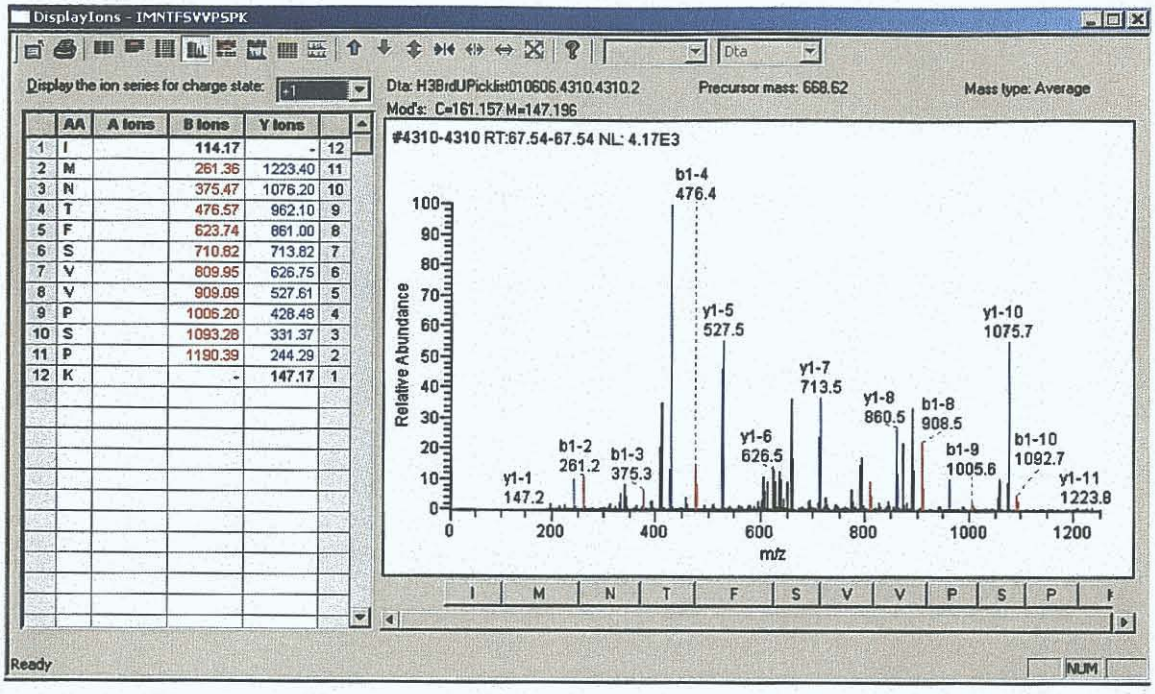
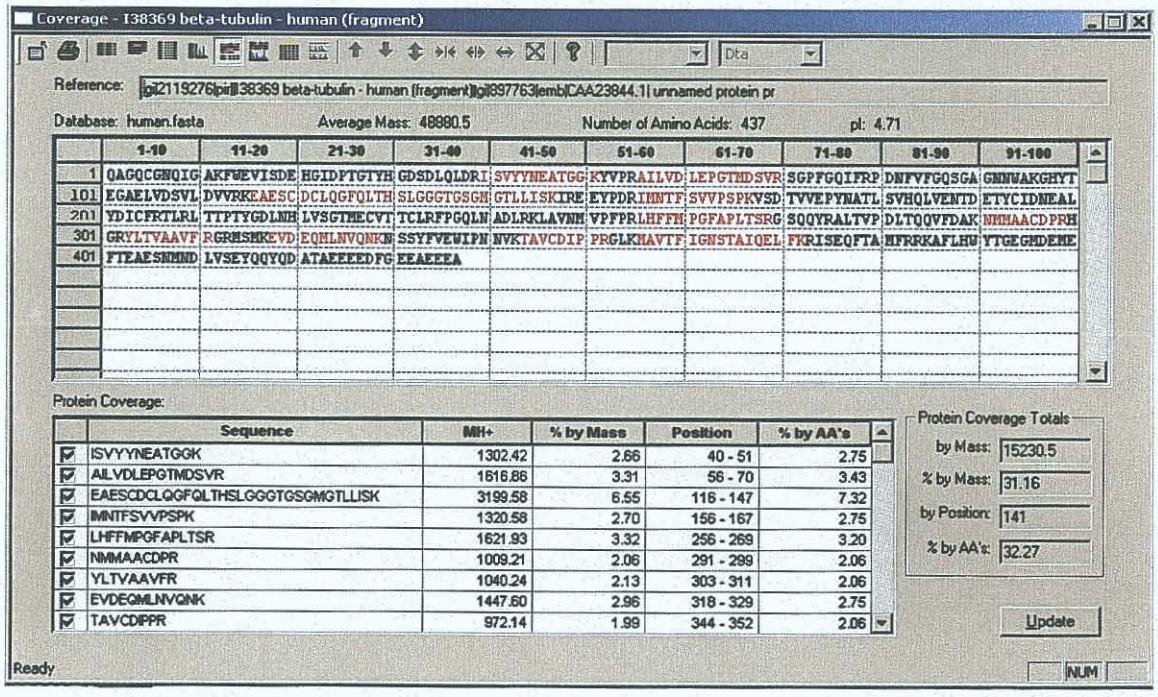



**Figure 3.46 (d) Based on this analysis, the peptide IMNTFSVVPSPK of the protein BetaTubulin was identified. TurboSEQUEST™ cross correlation (XCorr) = 4.69**




| Sequence | MH+ | % by Mass | Position | % by AA's |
|---|---|---|---|---|
| ISVYYNEATGGK | 1302.42 | 2.66 | 40 - 51 | 2.75 |
| AILVDLEPGTMDSVR | 1616.86 | 3.31 | 56 - 70 | 3.43 |
| EAESCDCLQGFQLTHSLGGGTGSGMGTLLISK | 3199.58 | 6.55 | 116 - 147 | 7.32 |
| IMNTFSVVPSPK | 1320.58 | 2.70 | 156 - 167 | 2.75 |
| LHFFMPGFAPLTSR | 1621.93 | 3.32 | 256 - 269 | 3.20 |
| NMMAACDPR | 1009.21 | 2.06 | 291 - 299 | 2.06 |
| YLTVAAVFR | 1040.24 | 2.13 | 303 - 311 | 2.06 |
| EVDEQMLNVQNK | 1447.60 | 2.96 | 318 - 329 | 2.75 |
| TAVCDIPPR | 972.14 | 1.99 | 344 - 352 | 2.06 |

**Figure 3.46 (e) BioWorks identified Beta Tubulin-human with a amino acid % coverage of 32.27% coverage for Beta Tubulin from 11 peptides distributed throughout the protein**

### 3.13.4 Master number: 2213

**Identified by nanoLC-MSMS – Cytochrome c oxidase**

Having identified Cytochrome C by MALDI-TOF MS the remaining peptides were resuspended in 20μL of 0.1%FA. These peptides were separated at 300nL/min over a long gradient 0-65% Acetonitrile in 180 minutes directly into an Ion Trap Mass Spectrometers where MS and MSMS analysis was carried out. Figure 3.47 shows a base peak chromatogram plotting Relative abundance vs Time of ions detected by the Mass Spectrometer..

Figure 3.47b shows the full MS spectrum generated at 22.83 minutes into the gradient. Ions at 331.5 and 502.9m/z were not selected for MS/MS fragmentation as they are contamination peaks and are on an exclusion list. The parent ion at 403.9 m/z was automatically selected for MSMS fragmentation and figure 3.47c shows the fragmentation pattern generated. Figure 3.47d shows the SEQUEST result of the fragmentation pattern generated in figure 3.47c. SEQUEST used crossed-correlation analysis between the experimental MSMS spectrum and theoretically generated mass spectra from peptide sequences with a similar molecular weight to the experimental parent ion. In this case the fragmentation ions generated from the parent ion at 404.09 m/z matches the majority of b and y ions associated with the amino acid sequence AEGTFPGK of the protein Cytochrome C oxodase with a cross correlation (XCorr) score of 3.07. One other peptide sequence relating to Cytochrome C was identified in that experiment resulting in 18.39% amino acid coverage of the protein by MSMS analysis (figure 3.47e).

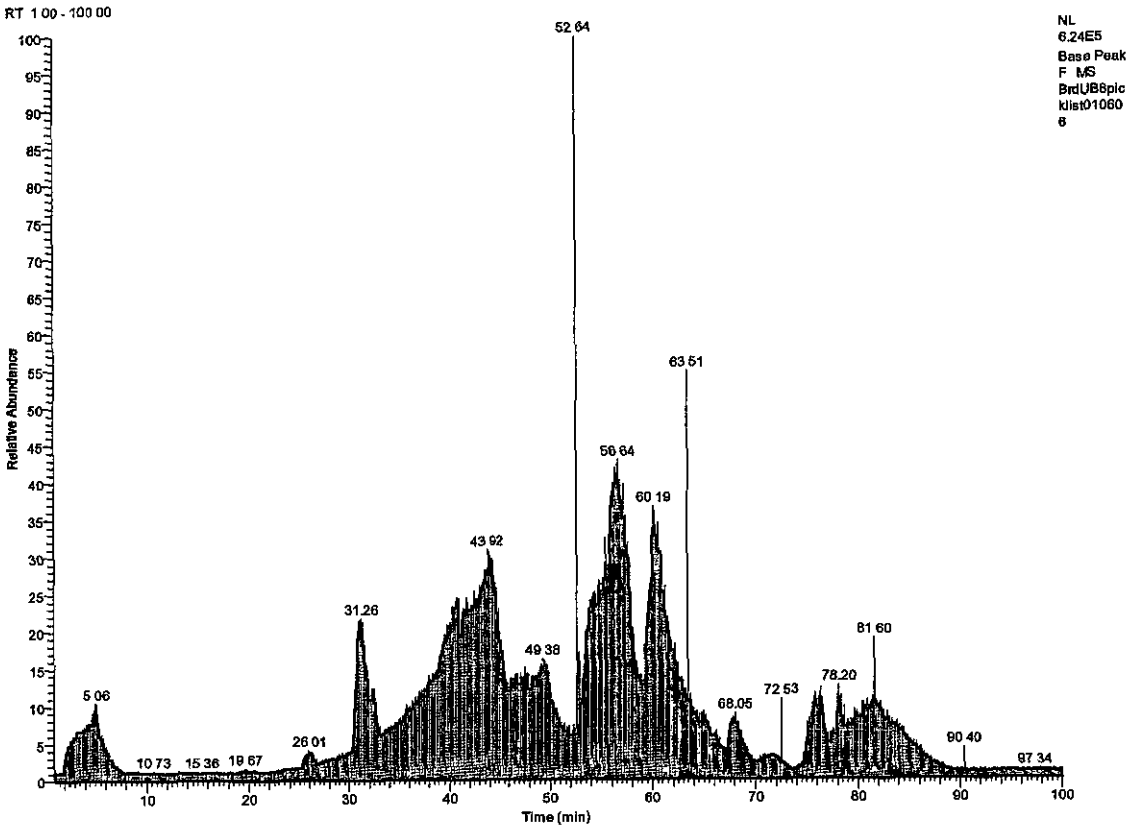



**Figure 3.47 (a) MS spectrum of eluate from RPC separation of spot 2113. The digested peptide sample was injected onto Ettan MDLC equipped with a Zorbax™ SB-300 C18 3.5 mm column. Peptides were eluted with RPC solvent B (0.1% formic acid, 86% acetonitrile) at a flow rate of approximately 300 nl/min and analyzed with a Finnigan LTQ (Thermo Electron Corp.) fitted with a nanospray interface**

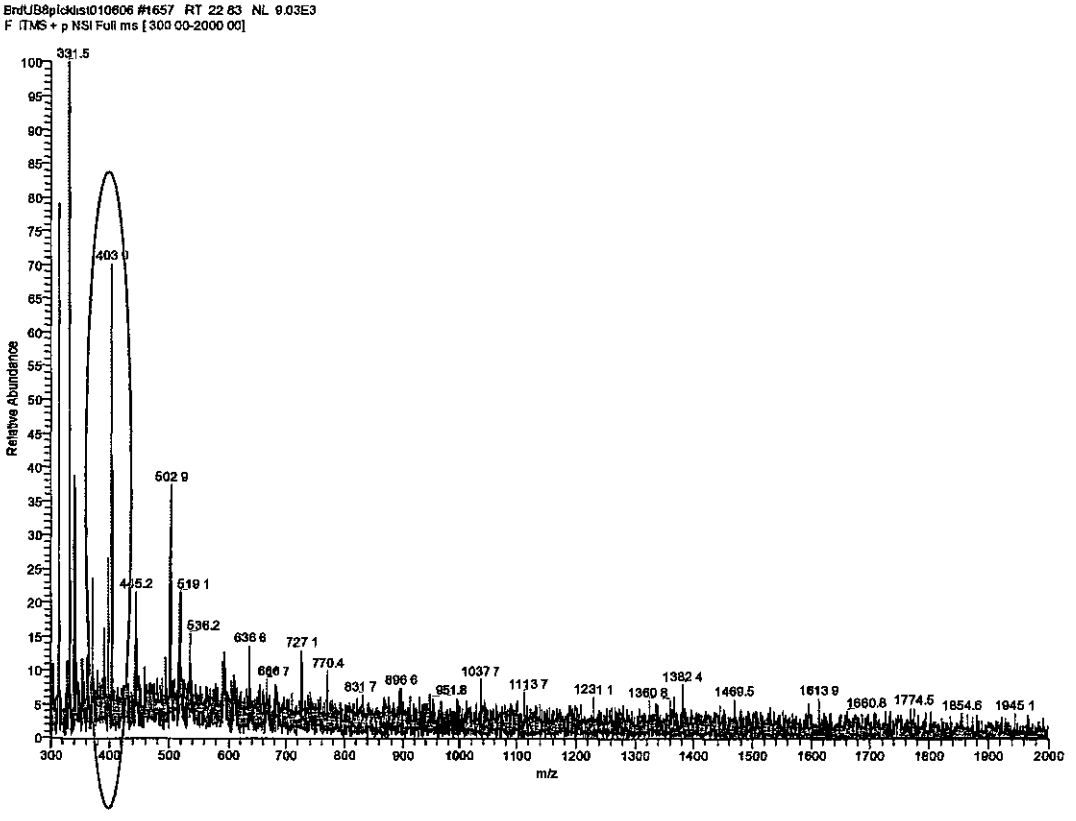



**Figure 3.47 (b) Mass spectrum for retention time 22.83 min. Parent Ion 403.9m/z was chosen for MS/MS**

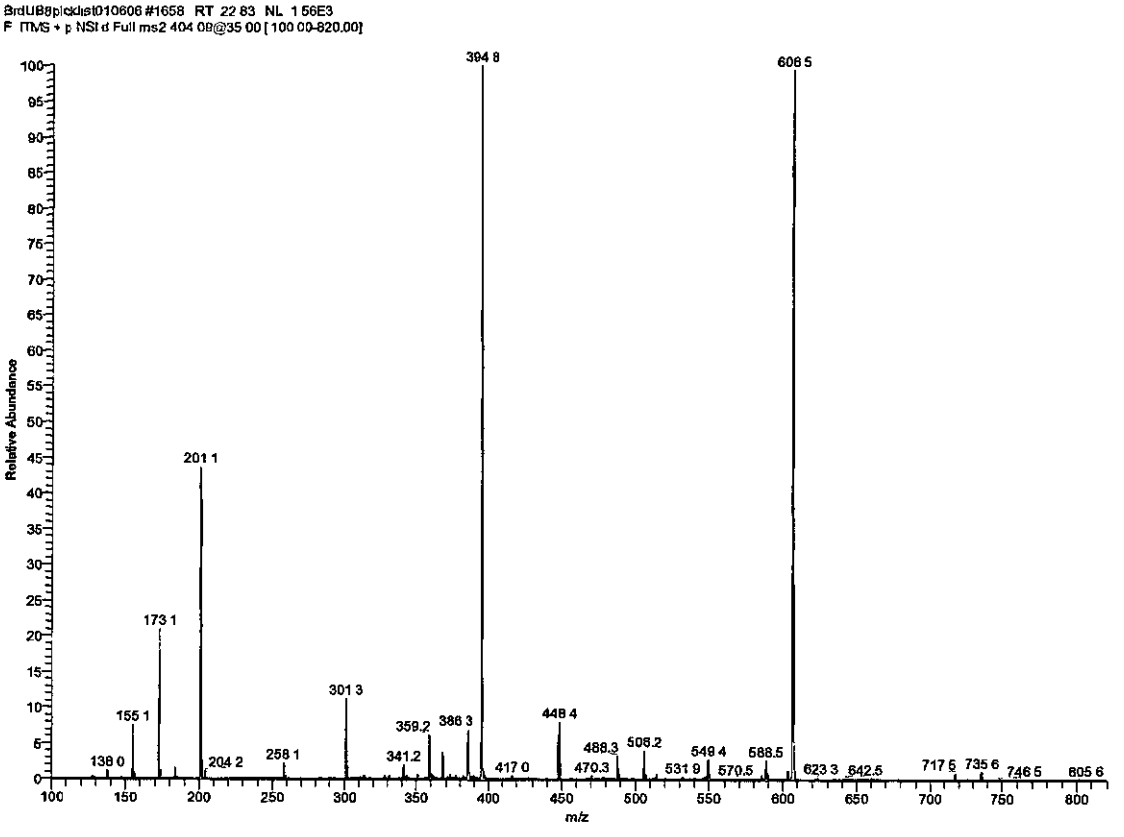



**Figure 3.47 (c) MS/MS spectrum of the generated product ions from doubly charged ion 404.0**

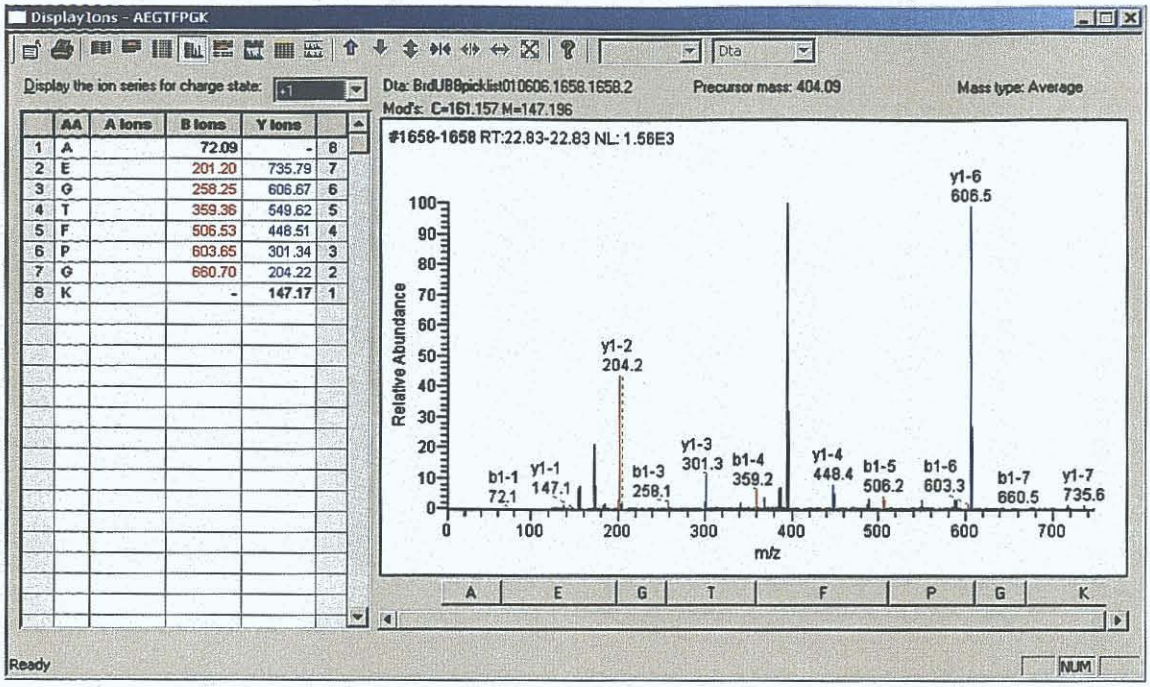



**Figure 3.47 (d) Based on this analysis, the peptide AEGTFPGK of the protein Cytochrome C was identified (Figure 3.47e). TurboSEQUEST™ cross correlation (XCorr) = 3.07**

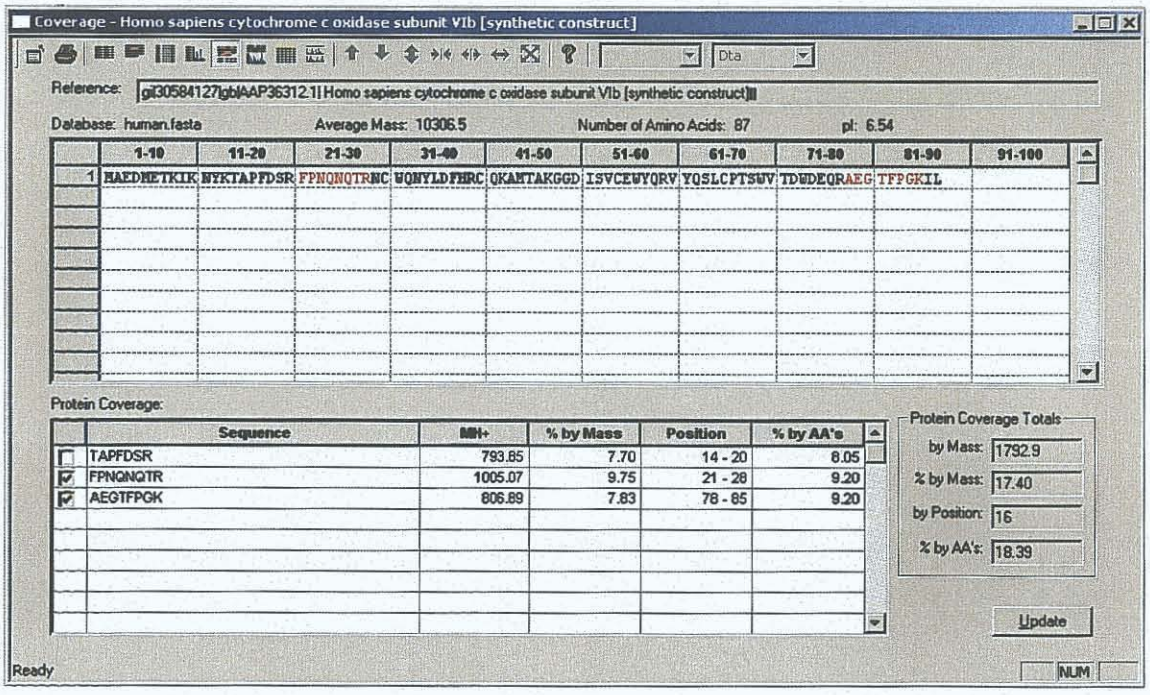



**Figure 3.47 (e) BioWorks identified Cytochrome C (Homo sapiens) with an amino acid coverage % of 18.39% coverage from 2 peptides distributed throughout the protein**

### 3.14.5 Master number: 1432

**Identified by nanoLC-MSMS – Cathepsin X**

Having identified Cathespin X by MALDI-TOF MS the remaining peptides were resuspended in 20μL of 0.1%FA. These peptides were separated at 300nL/min over a long gradient 0-65% Acetonitrile in 180 minutes directly into an Ion Trap Mass Spectrometers where MS and MSMS analysis was carried out. Figure 3.48a shows a base peak chromatogram plotting Relative abundance vs Time of ions detected by the Mass Spectrometer.

Figure 3.48b shows the full MS spectrum generated at 54.38 minutes into the gradient. Ions at 313.4, 371.3, 445.2 and 519.1m/z were not selected for MS/MS fragmentation as they are contamination peaks and are on an exclusion list. Parent ion at 422.0m/z was also on an exclusion list as the ion is the double charged species of Trypsin I (842m/z). The parent ion at 655.1 m/z was automatically selected for MSMS fragmentation and figure 3.48c shows the fragmentation pattern generated. Figure 3.48d shows the SEQUEST result of the fragmentation pattern generated in figure 3.48c. SEQUEST used crossed-correlation analysis between the experimental MSMS spectrum and theoretically generated mass spectra from peptide sequences with a similar molecular weight to the experimental parent ion. In this case the fragmentation ions generated from the parent ion at 655.1 m/z matches the majority of b and y ions associated with the amino acid sequence NVDGVNYASITR of the protein Cathepsin X with a cross correlation (XCorr) score of 4.69. Three other peptide sequence relating to Cathepsin X were identified in that experiment resulting in 12.87% amino acid coverage of the protein by MSMS analysis (figure 3.48e).

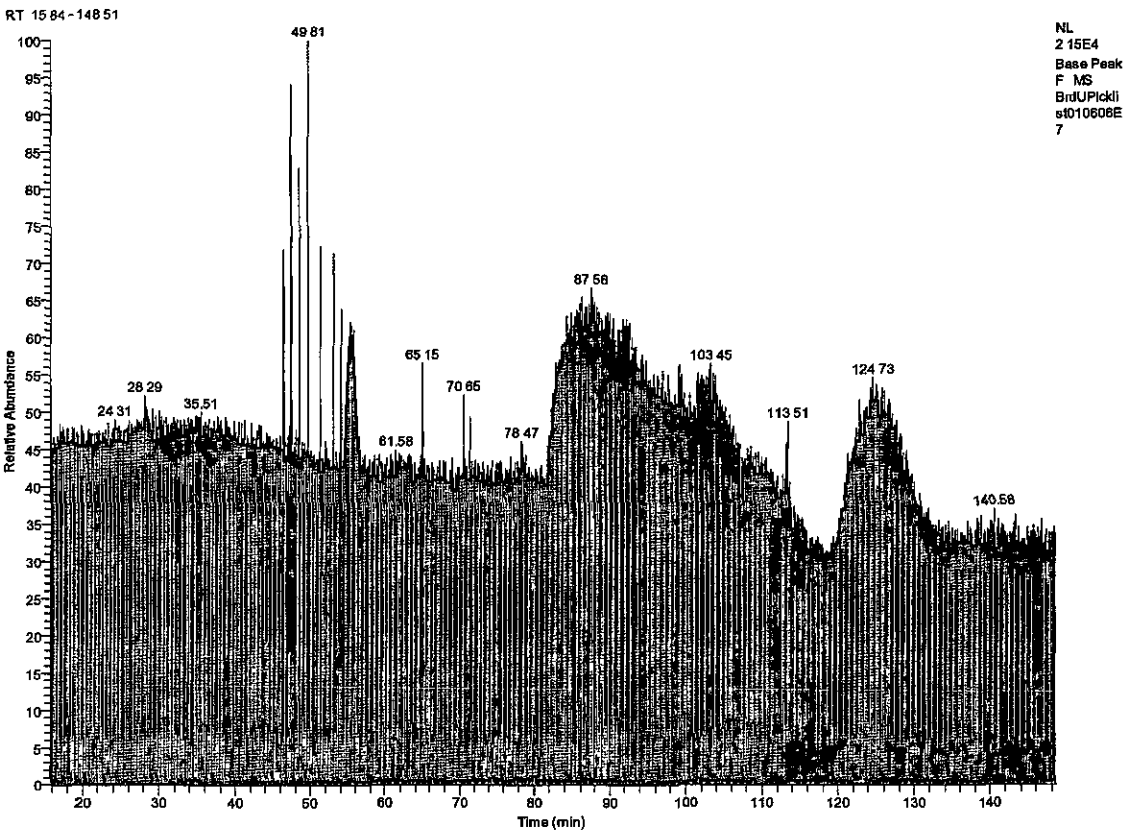



**Figure 3.48 (a) MS spectrum of eluate from RPC separation of spot 1432. The digested peptide sample was injected onto Ettan MDLC equipped with a Zorbax™ SB-300 C18 3.5 mm column. Peptides were eluted with RPC solvent B (0.1% formic acid, 86% acetonitrile) at a flow rate of approximately 300 nl/min and analyzed with a Finnigan LTQ (Thermo Electron Corp.) fitted with a nanospray interface**

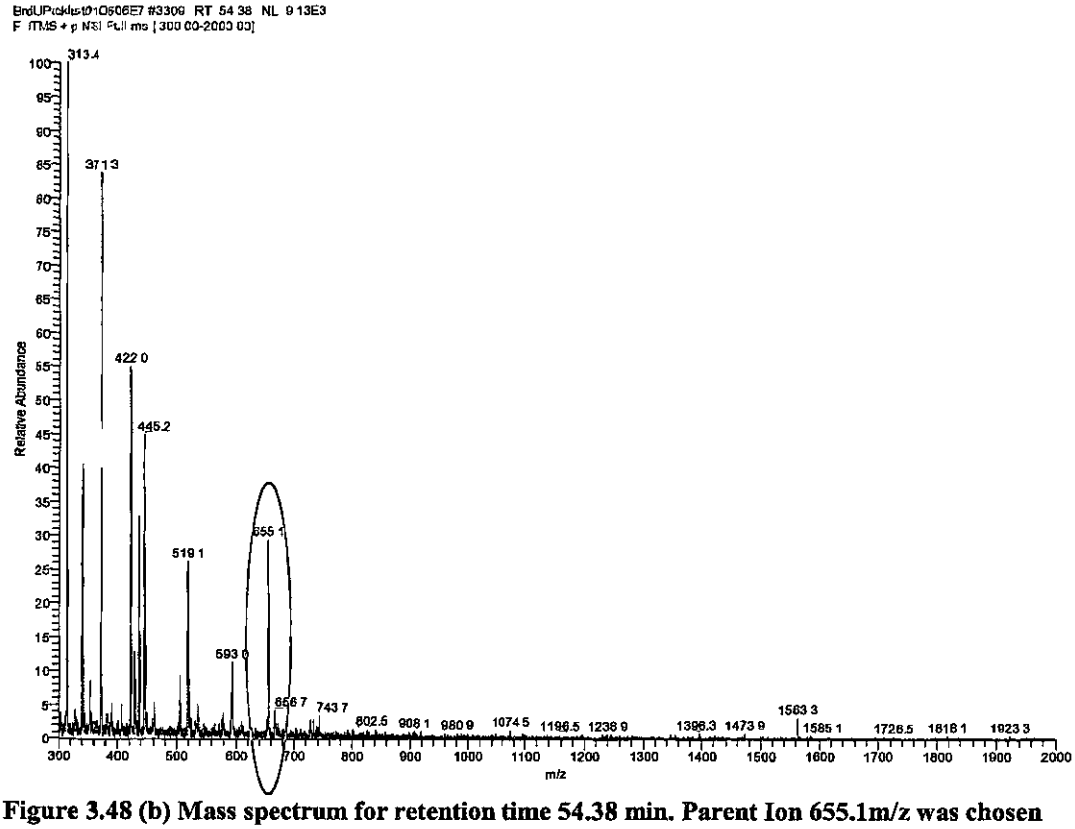



**Figure 3.48 (b) Mass spectrum for retention time 54.38 min. Parent Ion 655.1m/z was chosen for MS/MS**

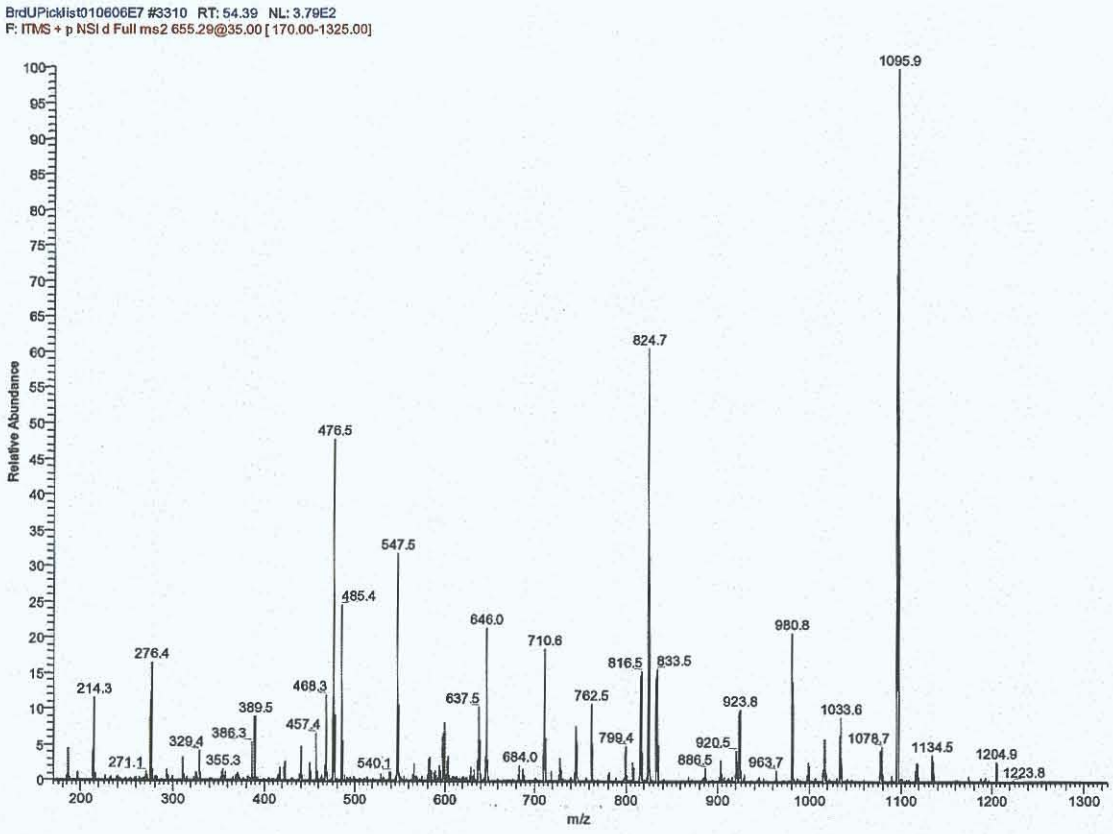
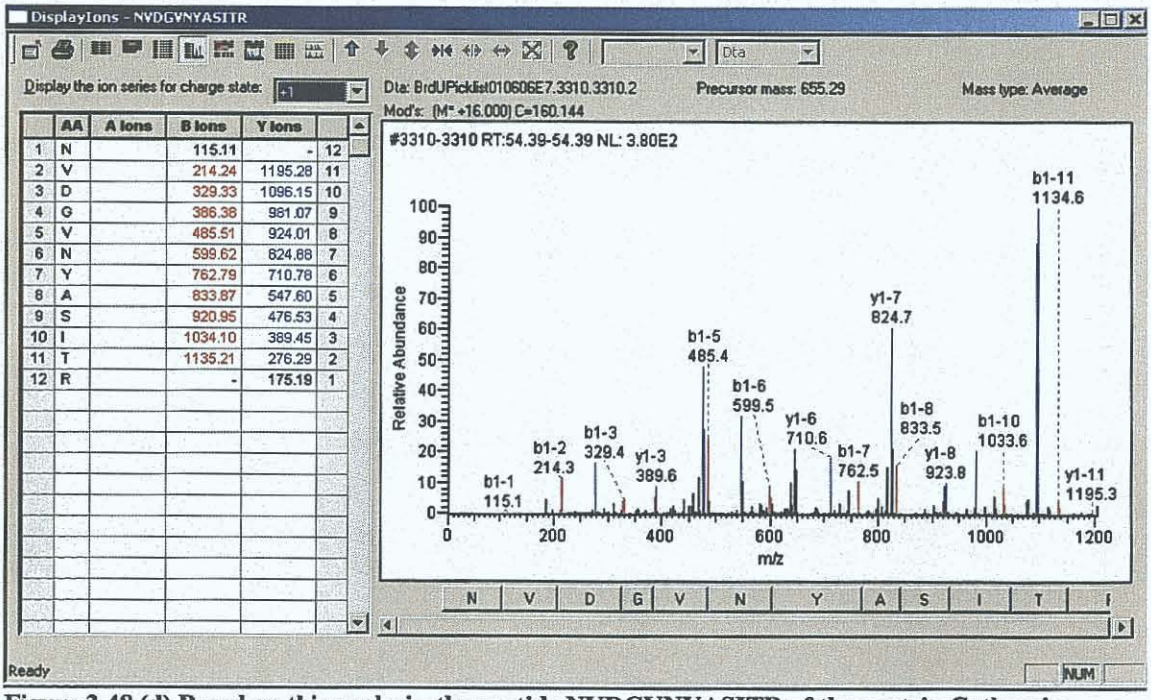



**Figure 3.48 (c) MS/MS spectrum of the generated product ions from doubly charged ion 655.29**




| | AA | A Ions | B Ions | Y Ions | |
|---|---|---|---|---|---|
| 1 | N | | 115.11 | - | 12 |
| 2 | V | | 214.24 | 1195.28 | 11 |
| 3 | D | | 329.33 | 1096.15 | 10 |
| 4 | G | | 386.36 | 981.07 | 9 |
| 5 | V | | 485.51 | 924.01 | 8 |
| 6 | N | | 599.62 | 824.88 | 7 |
| 7 | Y | | 762.79 | 710.78 | 6 |
| 8 | A | | 833.87 | 547.60 | 5 |
| 9 | S | | 920.95 | 476.53 | 4 |
| 10 | I | | 1034.10 | 389.45 | 3 |
| 11 | T | | 1135.21 | 276.29 | 2 |
| 12 | R | | - | 175.19 | 1 |

**Figure 3.48 (d) Based on this analysis, the peptide NVDGVNYASITR of the protein Cathepsin X was identified (Figure 30e). TurboSEQUEST™ cross correlation (XCorr) = 4.69**

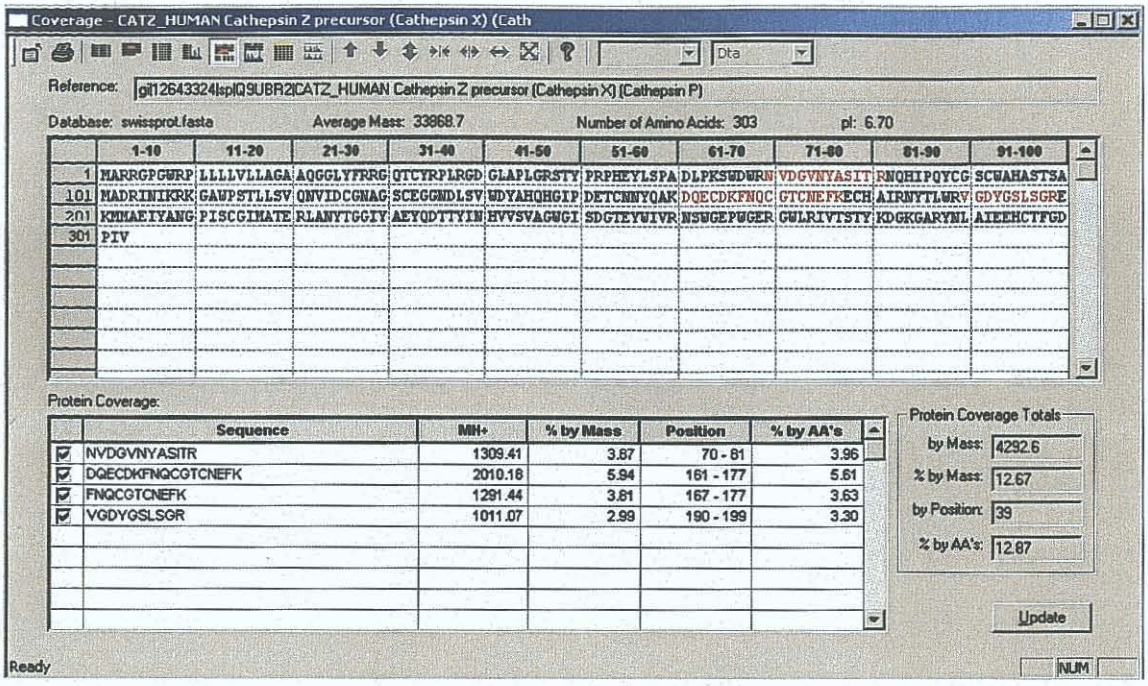



**Figure 3.48 (e) BioWorks identified Cathepsin X. Coverage was 12.87% coverage for Cathepsin X from 4 peptides distributed throughout the protein**

### 3.13.6 Master number: 1678

**Identified by nanoLC-MSMS – Protein Molecular Basis for the local conformational rearrangement of human Phosphoserine Phosphatase**

Having identified Protein Molecular Basis for the local conformational rearrangement of human Phosphoserine Phosphatase by MALDI-TOF MS the remaining peptides were resuspended in 20μL of 0.1%FA. The peptides were separated at 300nL/min over a long gradient 0-65% Acetonitrile in 180 minutes directly into an Ion Trap Mass Spectrometers where MS and MSMS analysis was carried out. Figure 3.49a shows a base peak chromatogram plotting Relative abundance vs Time of ions detected by the Mass Spectrometer.

Figure 3.49b shows the full MS spectrum generated at 88.96 minutes into the gradient. Ions at 313.4, 341.5, 371.3, 445.2 and 519.1m/z were not selected for MS/MS fragmentation as they are contamination peaks and are on an exclusion list. The parent ion at 669.0 m/z was automatically selected for MSMS fragmentation and figure 3.49c shows the fragmentation pattern generated. Figure 3.49d shows the SEQUEST result of the fragmentation pattern generated in figure 3.49c. SEQUEST used crossed-correlation analysis between the experimental MSMS spectrum and theoretically generated mass spectra from peptide sequences with a similar molecular weight to the experimental parent ion. In this case the fragmentation ions generated from the parent ion at 669.24 m/z matches the majority of b and y ions associated with the amino acid sequence NVQVFLISGGFR of the Protein Molecular Basis for the local conformational rearrangement of human Phosphoserine Phosphatase with a cross correlation (XCorr) score of 4.69. Eight

other peptide sequences relating to Protein Molecular Basis for the local conformational rearrangement of human Phosphoserine Phosphatase was identified in that experiment resulting in 56.89% amino acid coverage of the protein by MSMS analysis (figure 3.49e).

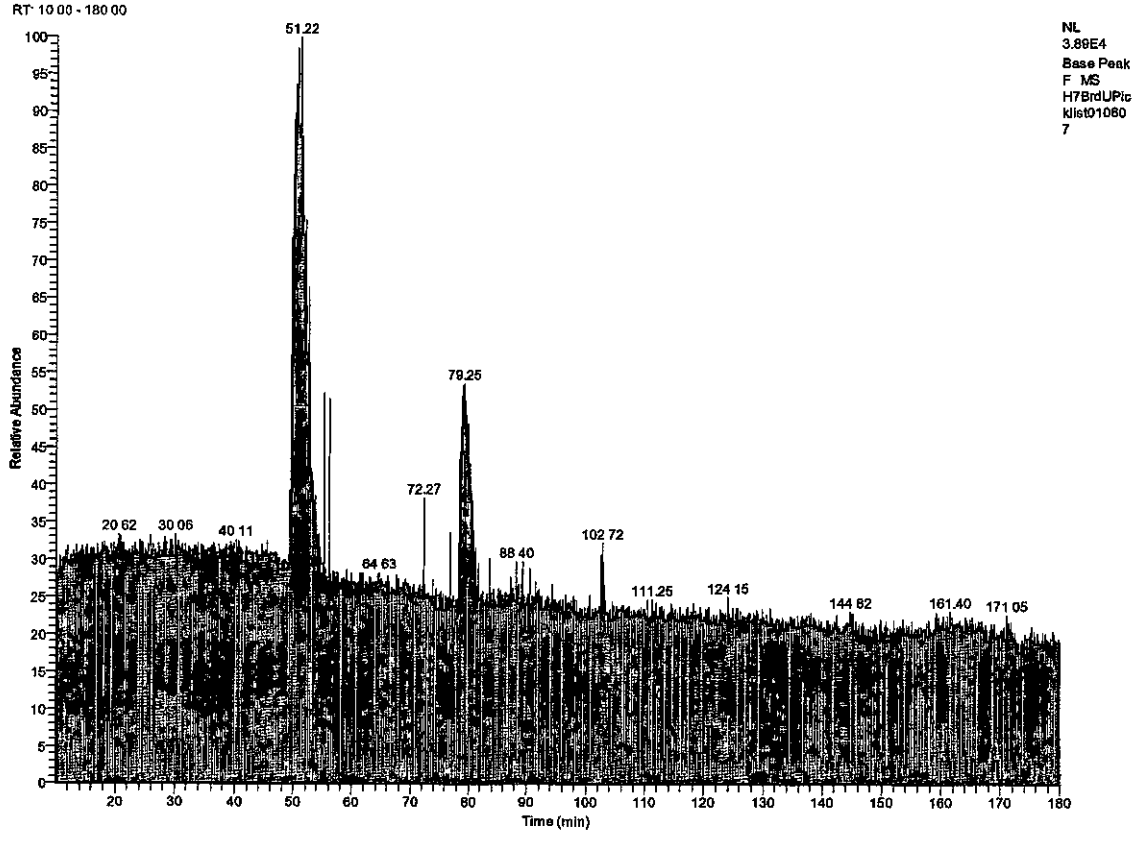



**Figure 3.49 (a) MS spectrum of eluate from RPC separation of spot 1678. The digested peptide sample was injected onto Ettan MDLC equipped with a Zorbax™ SB-300 C18 3.5 mm column. Peptides were eluted with RPC solvent B (0.1% formic acid, 86% acetonitrile) at a flow rate of approximately 300 nl/min and analyzed with a Finnigan LTQ (Thermo Electron Corp.) fitted with a nanospray interface**

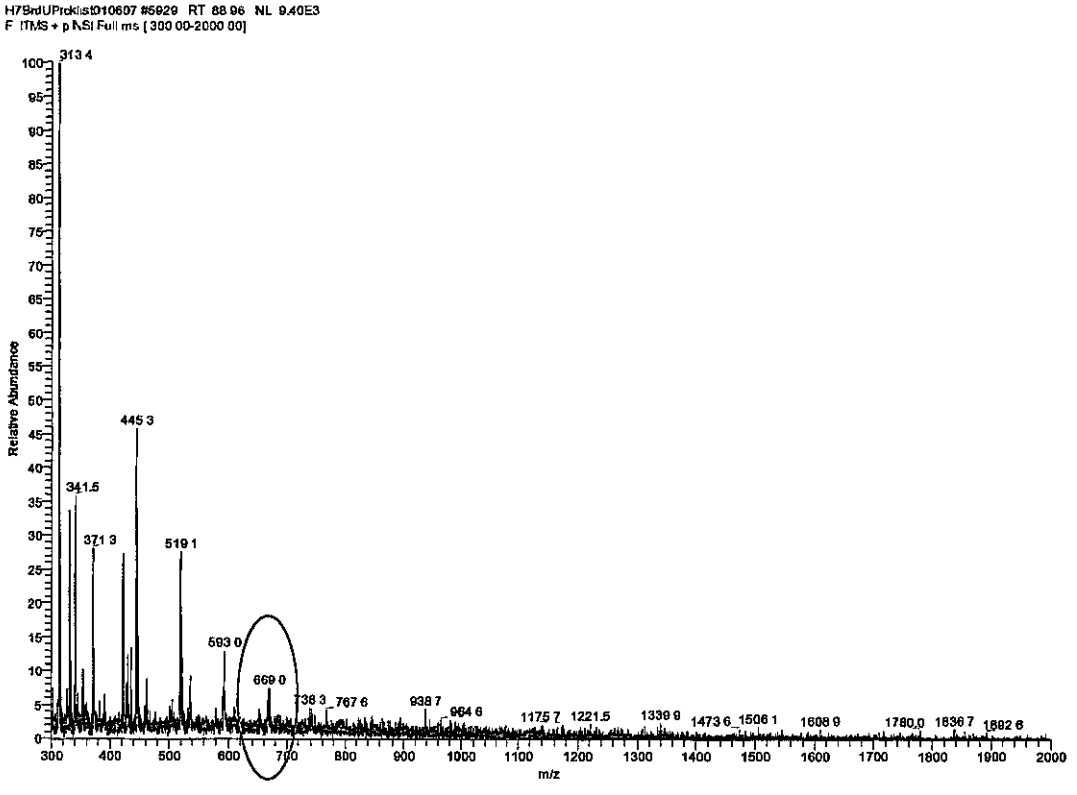



**Figure 3.49 (b) Mass spectrum for retention time 88.96 min. Parent Ion 669.24 was chosen for MS/M**

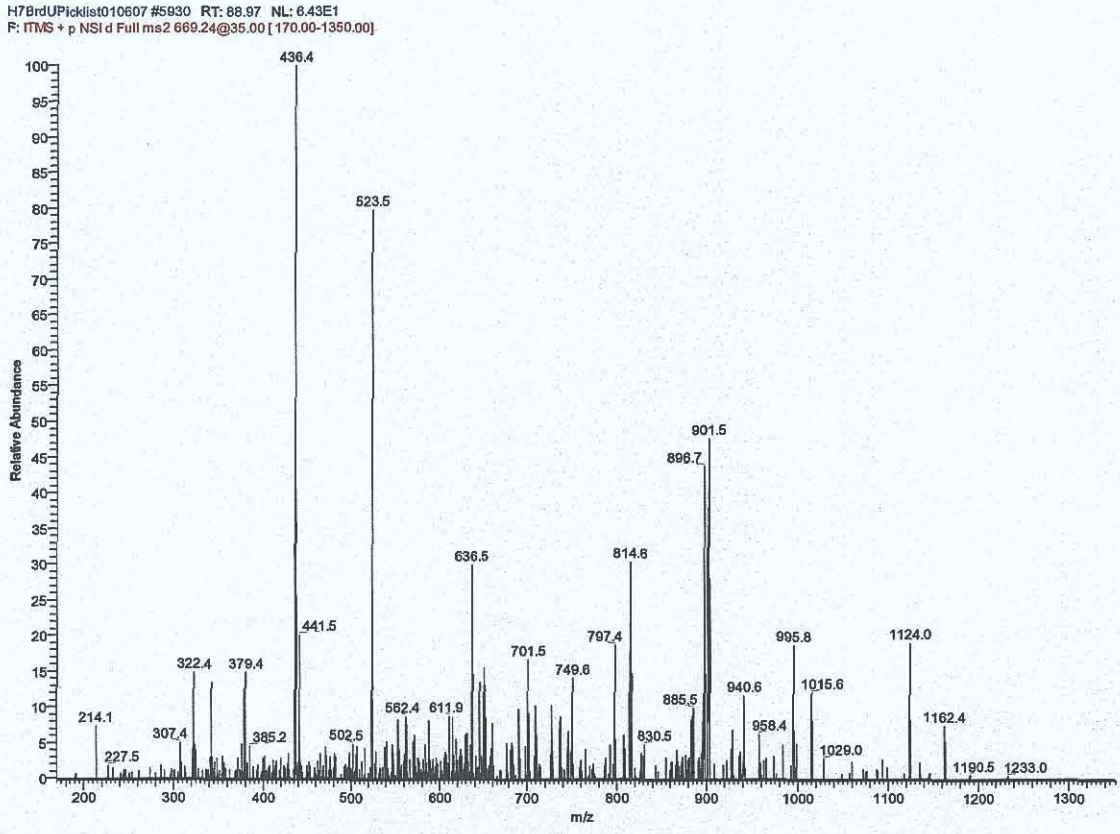
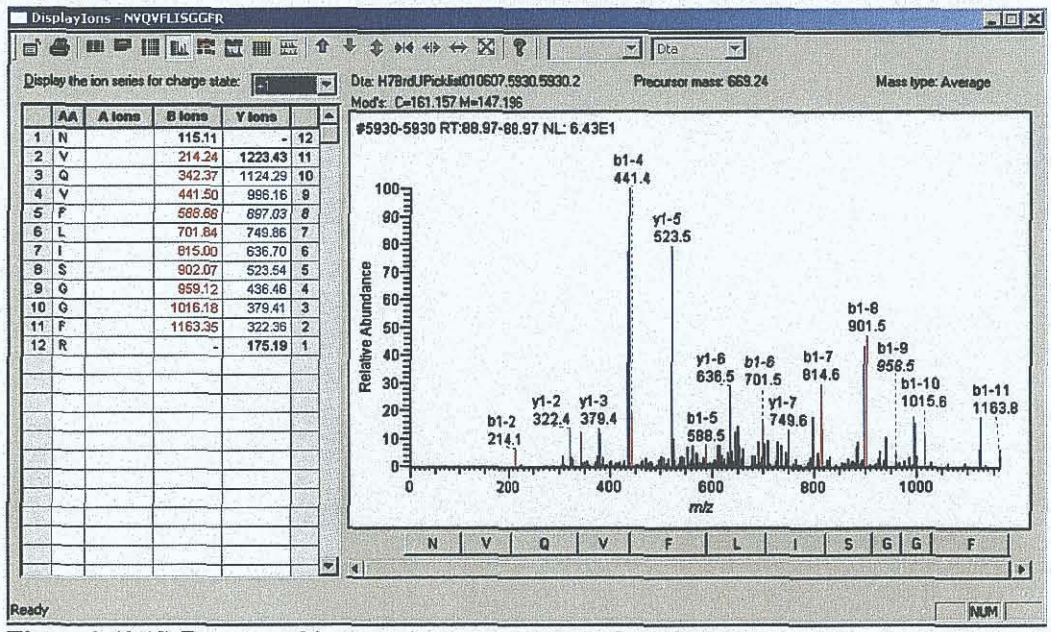



**Figure 3.49 (c) MS/MS spectrum of the generated product ions from doubly charged ion 669.24**




| | AA | A Ions | B Ions | Y Ions | |
|---|---|---|---|---|---|
| 1 | N | | 115.11 | - | 12 |
| 2 | V | | 214.24 | 1223.43 | 11 |
| 3 | Q | | 342.37 | 1124.29 | 10 |
| 4 | V | | 441.50 | 996.16 | 9 |
| 5 | F | | 588.66 | 897.03 | 8 |
| 6 | L | | 701.84 | 749.86 | 7 |
| 7 | I | | 815.00 | 636.70 | 6 |
| 8 | S | | 902.07 | 523.54 | 5 |
| 9 | G | | 959.12 | 436.46 | 4 |
| 10 | G | | 1016.18 | 379.41 | 3 |
| 11 | F | | 1163.35 | 322.36 | 2 |
| 12 | R | | - | 175.19 | 1 |

**Figure 3.49 (d) Based on this analysis, the peptide NVQVFLISGGFR of the protein Molecular Basis for the local conformational rearrangement of human Phosphoserine Phosphatase was identified (Figure 31e). TurboSEQUEST™ cross correlation (XCorr) = 4.69**

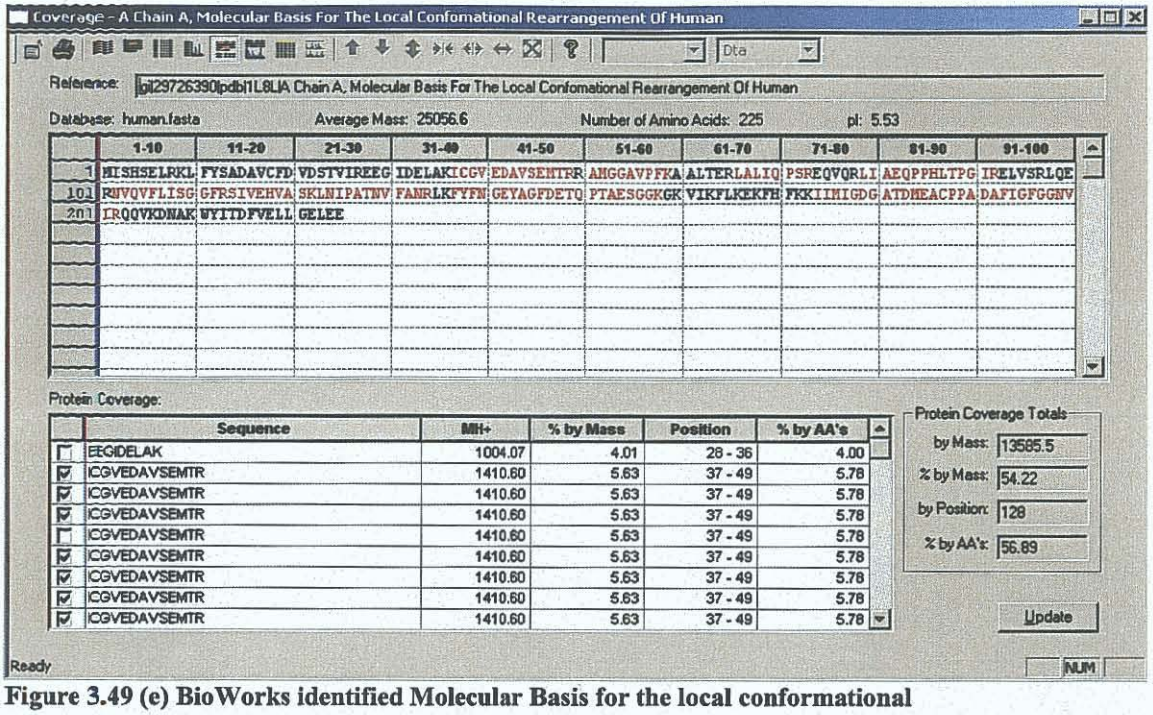



**Figure 3.49 (e) BioWorks identified Molecular Basis for the local conformational rearrangement of human Phosphoserine Phosphatase. Coverage was 56.89% coverage for Molecular Basis for the local conformational rearrangement of human Phosphoserine Phosphatase from 9 peptides distributed throughout the protein**

### 3.13.7 Master number: 2104

**Identified by nanoLC-MSMS – Cytochrome C**

MALDI-TOF MS was unable to identify spot number 2014. The peptides from spot 2104 were analysed after dried droplet method and sample clean up with Zip Tip by MALDI-ToF MS analysis without success. The remaining peptides were resuspended in 20μL of 0.1%FA. The peptides were separated at 300nL/min over a long gradient 0-65% Acetonitrile in 180 minutes directly into an Ion Trap Mass Spectrometers where MS and MSMS analysis was carried out. Figure 3.50a shows a base peak chromatogram plotting Relative abundance vs Time of ions detected by the Mass Spectrometer.

Figure 3.50b shows the full MS spectrum generated at 40.40 minutes into the gradient. Ions at 313.4, 341.5, 371.3, 445.2 and 519.1m/z were not selected for MS/MS fragmentation as they are contamination peaks and are on an exclusion list. The very low abundant parent ion at 404.1 m/z was automatically selected for MSMS fragmentation and figure 3.50c shows the fragmentation pattern generated. Figure 3.50d shows the SEQUEST result of the fragmentation pattern generated in figure 3.50c. SEQUEST used crossed-correlation analysis between the experimental MSMS spectrum and theoretically generated mass spectra from peptide sequences with a similar molecular weight to the experimental parent ion. In this case the fragmentation ions generated from the parent ion at 404.32 m/z matches the majority of b and y ions associated with the amino acid sequence AEGTFPGK of

Cytochrome C Oxidase with a cross correlation (XCorr) score of 4.69. Eight other peptide sequences relating to Cytochrome C oxidase was identified in that experiment resulting in 18.39% amino acid coverage of the protein by MSMS analysis (figure 3.50e).

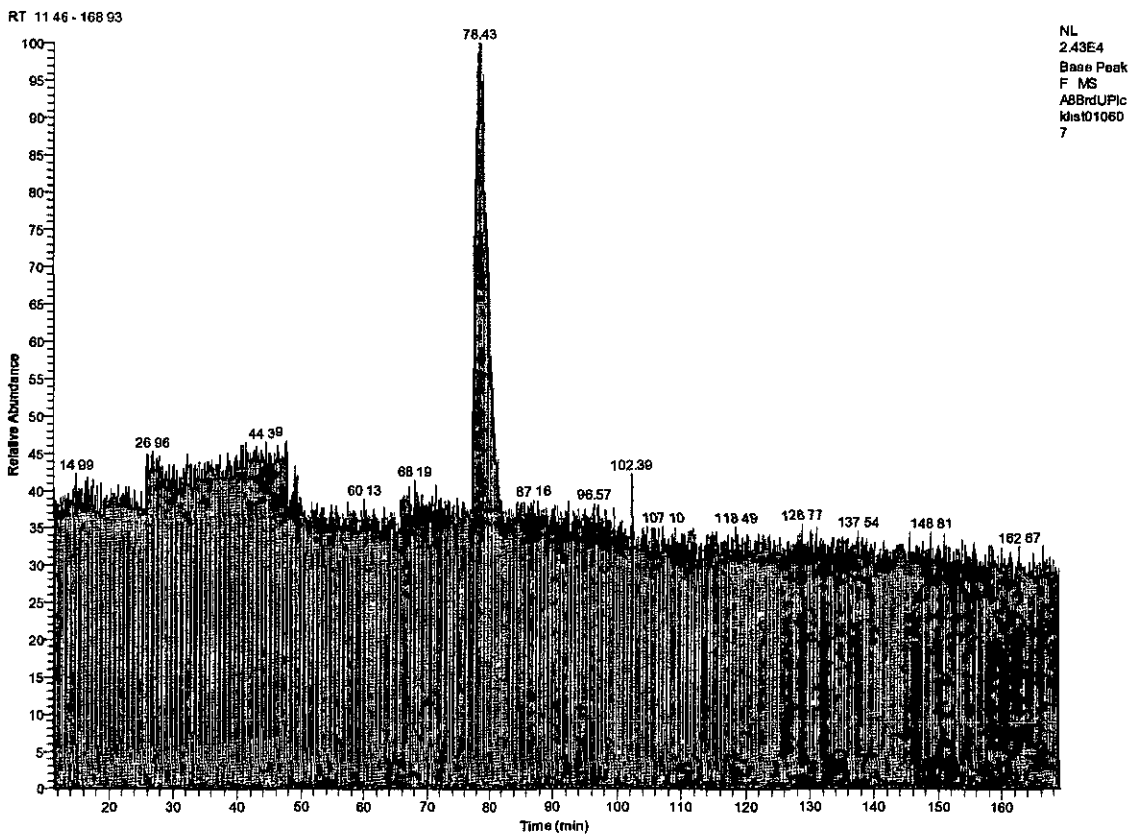



**Figure 3.50 (a) MS spectrum of eluate from RPC separation of spot 2213. The digested peptide sample was injected onto Ettan MDLC equipped with a Zorbax™ SB-300 C18 3.5 mm column. Peptides were eluted with RPC solvent B (0.1% formic acid, 86% acetonitrile) at a flow rate of approximately 300 nl/min and analyzed with a Finnigan LTQ (Thermo Electron Corp.) fitted with a nanospray interface.**

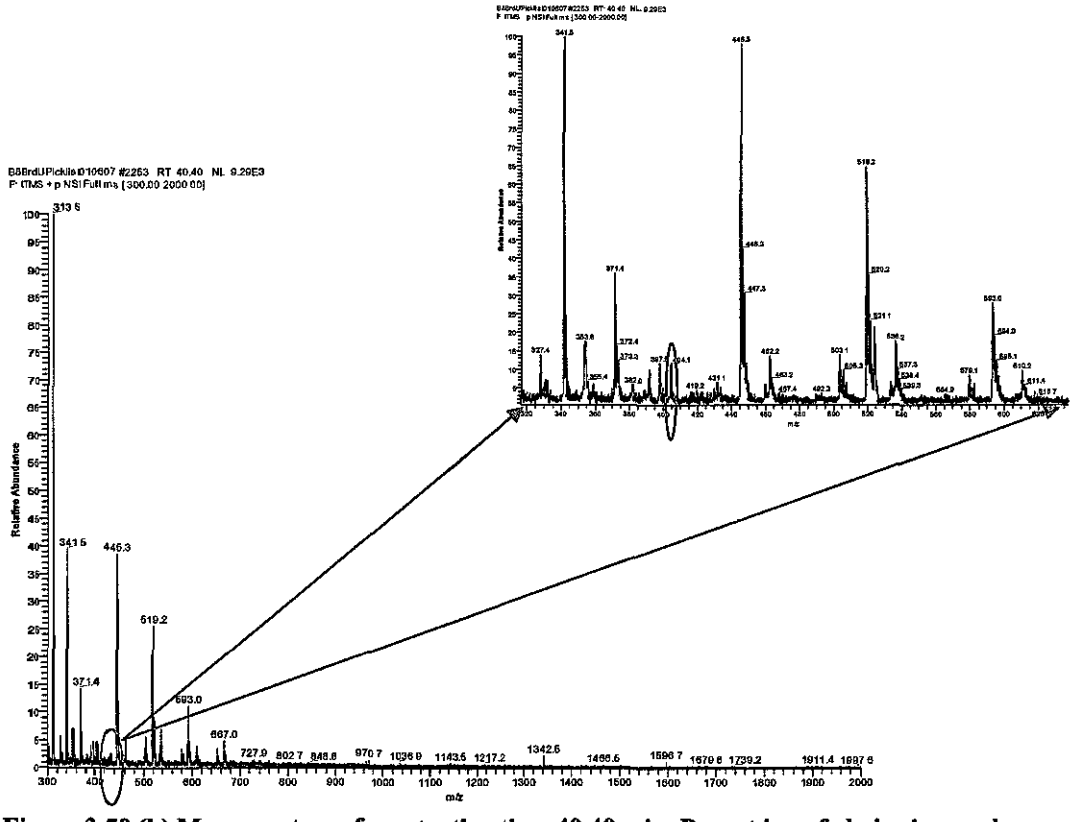



**Figure 3.50 (b) Mass spectrum for retention time 40.40 min. Parent ion of choice is very low abundant. Spectrum is expanded out to show peak of interest. Parent Ion 404.32m/z was chosen for MS/MS**

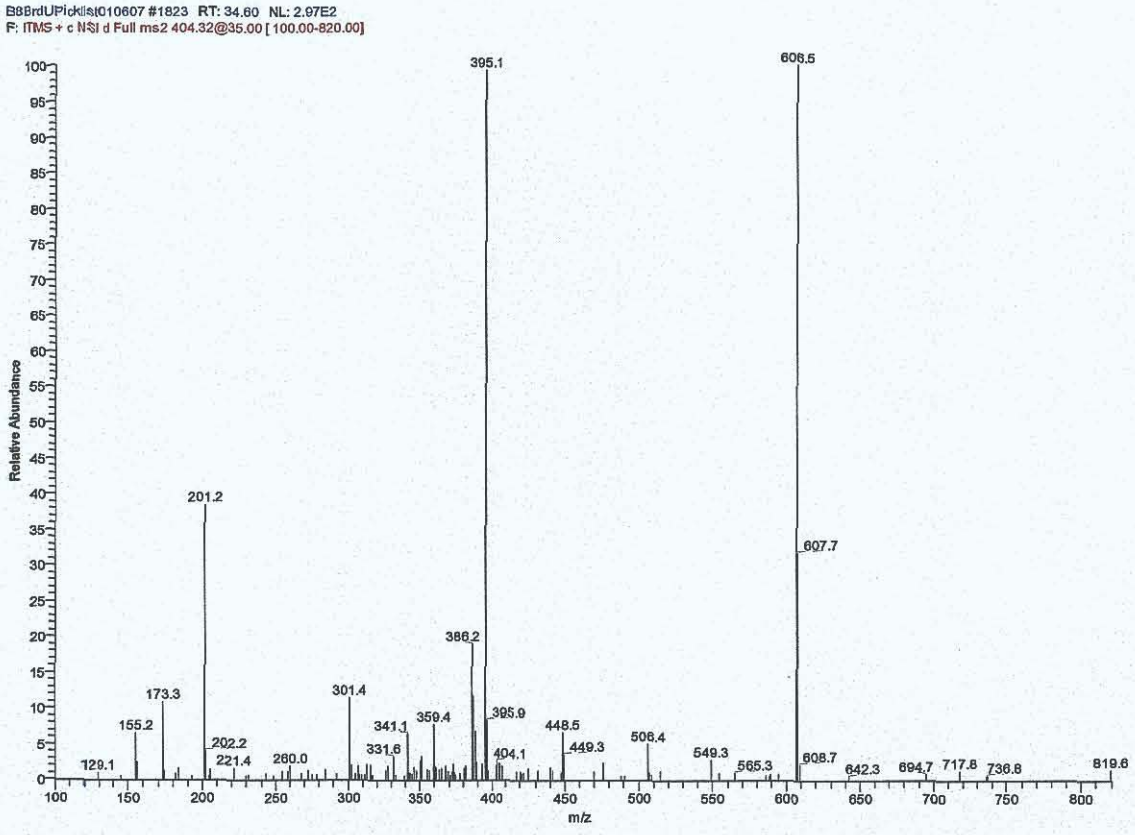
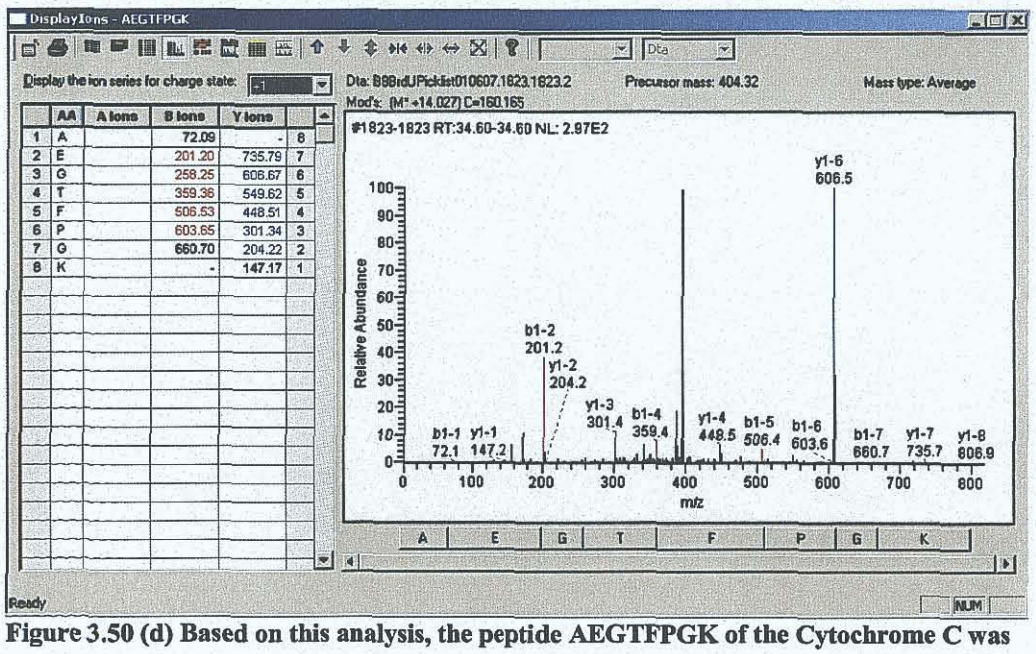



**Figure 3.50 (c) MS/MS spectrum of the generated product ions from doubly charged ion 404.32**




| | AA | A Ions | B Ions | Y Ions | |
|---|---|---|---|---|---|
| 1 | A | | 72.09 | - | 8 |
| 2 | E | | 201.20 | 735.79 | 7 |
| 3 | G | | 258.25 | 606.67 | 6 |
| 4 | T | | 359.36 | 549.62 | 5 |
| 5 | F | | 506.53 | 448.51 | 4 |
| 6 | P | | 603.65 | 301.34 | 3 |
| 7 | G | | 660.70 | 204.22 | 2 |
| 8 | K | | - | 147.17 | 1 |

**Figure 3.50 (d) Based on this analysis, the peptide AEGTFPGK of the Cytochrome C was identified (Figure 32e). TurboSEQUEST™ cross correlation (XCorr) = 4.69**

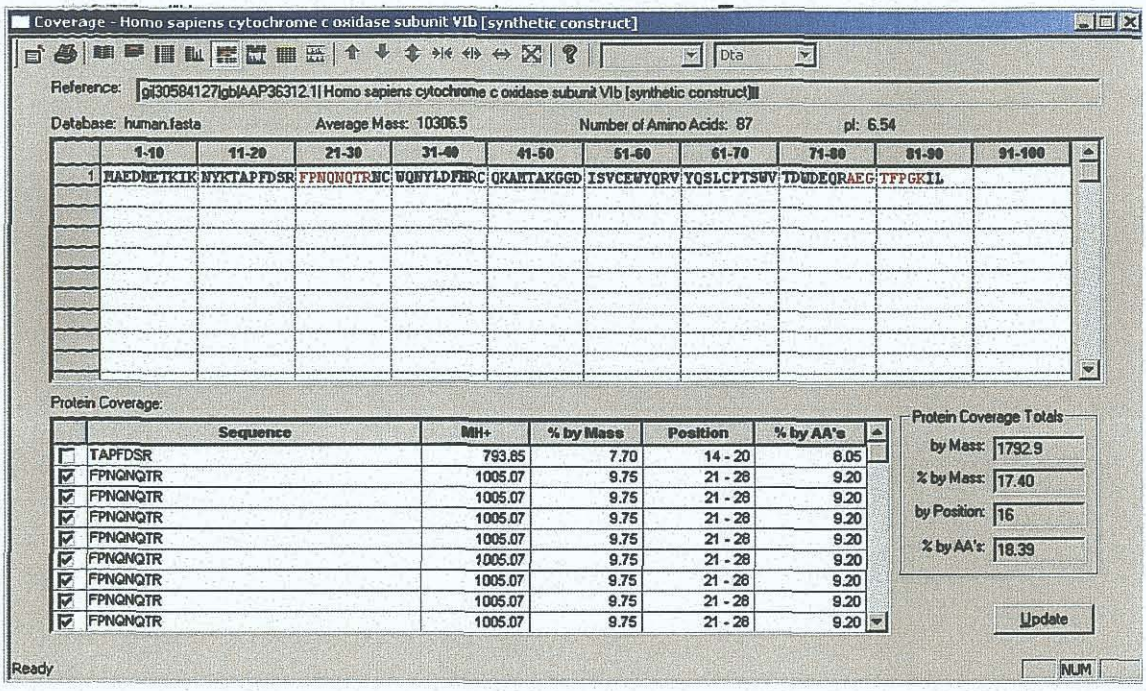



**Figure 3.50 (e) BioWorks identified Cytochrome C. Coverage was 18.39% coverage for Cytochrome C from 2 peptides distributed throughout the protein**

## 3.14 Protein identification results from Tandem Mass Spectrometry

Using nanoLC and Tandem Mass Spectrometry seven proteins were identified. Master spot numbers 160, 810, 1059, 1432, 1678 and 2104 confirmed with the protein identifications obtained by MALDI-MS. Master number 2213 was only identified by Tandem Mass Spectrometry

The proteins identified by nanoLC-Tandem Mass Spectrometry can be seen below (table 3.5).

| Master spot no. | Fold change | T test | ID | Coverage % | pI | Mw (kDa) |
|---|---|---|---|---|---|---|
| 160 | 1.82 | 2.90E-07 | gi\|24657579\|gb\|AAH39174.1 - Vinculin protein [Homo sapiens] | 26.55 | 5.8 | 116.7 |
| 810 | 1.62 | 0.037 | gi\|1731856945\|NP_006112.2- Keratin, type I | 31.52 | 8.16 | 66.06 |
| 1059 | 1.59 | 0.021 | gi\|2119276\|pir\|38369 - Tubulin, beta polypeptide [Homo sapiens] | 32.27 | 4 71 | 48.88 |
| 1432 | -1.56 | 0.00031 | gi\|1264332\|so\|Q9UBR3\|CATZ - Human Cathepsin X | 12.87 | 6.70 | 33.86 |
| 1678 | 1.58 | 0.00019 | gi\|29726390\|pdb\|1L8\|A Molecular Basis For The Local Conformational Rearrangement Of Human Phosphoserine Phosphatase] | 56.89 | 5.53 | 25.06 |
| 2104 | 2.20 | 7.10E-07 | gi\|30582729\|gb\|AAP35591.1 - cytochrome c oxidase subunit VIb [Homo sapiens] | 18.39 | 6.54 | 10.30 |
| 2213 | 3.09 | 2.00E-09 | gi\|305584\|gb\|AAP36312.1 - cytochrome c oxidase subunit VIb [Homo sapiens] | 18.39 | 6.54 | 10.30 |

**Table 3.5. Summary of results of Tandem Mass Spectrometry results**

## 3.15 Optimisation of MLDC nano-LC and LTQ MS/MS methods using 100fmol Bovine Serum Albumin (BSA) digest

Using a commercially available BSA digest (Waters) at a concentration of 100fmol we looked at three different acetonitrile (ACN) gradients using the MDLC at flow rates of approximately 300nl/min to separate the peptides using Zorbax™ SB-300 C18 3.5 mm column (Agilent) directly in the Finnigan LTQ (Thermo Electron Corp.) fitted with a nanospray interface. The data generated was searched using SEQUEST.

The three gradients were

1. Long Gradient 0.1%FA – 65% ACN over 180mins
2. Medium Gradient 0.1%FA – 65%ACN over 120mins
3. Short Gradient 0.1%FA – 65% ACN over 60mins

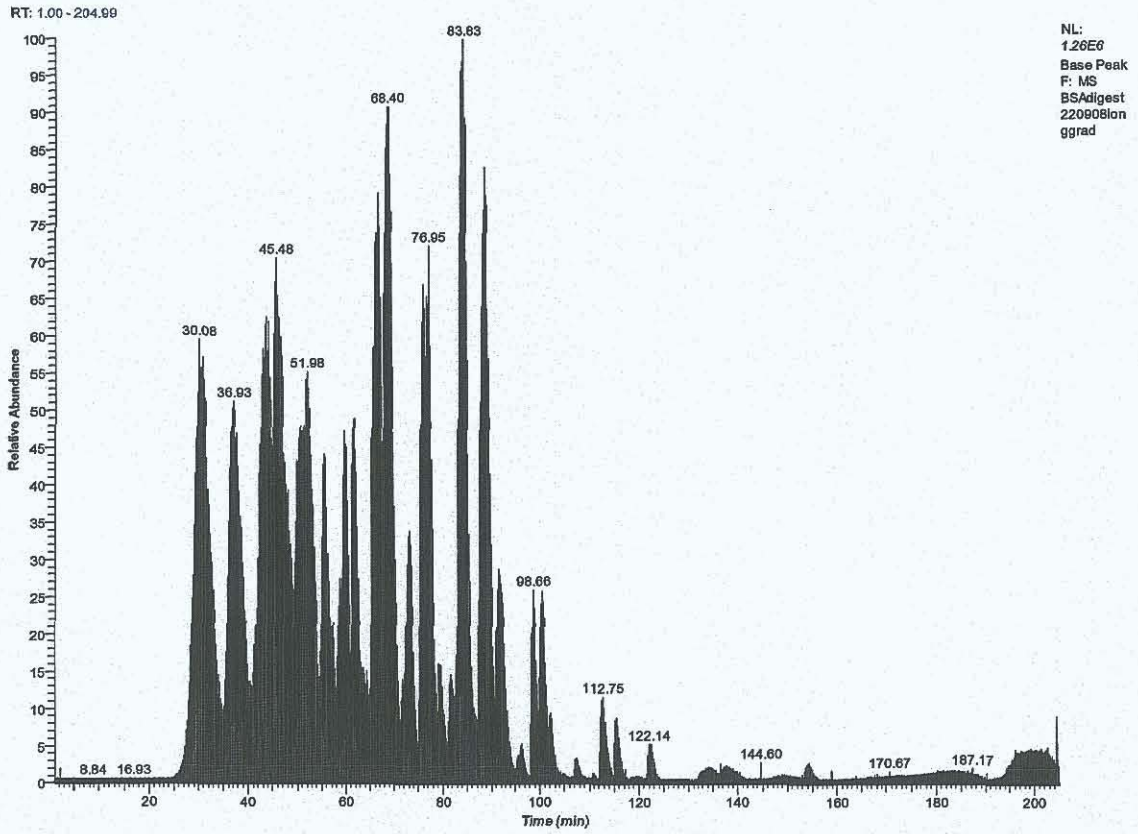



**Figure 3.51 (a) MS spectrum of BSA digest using the long Gradient 0.1%FA – 65% ACN over 180mins**

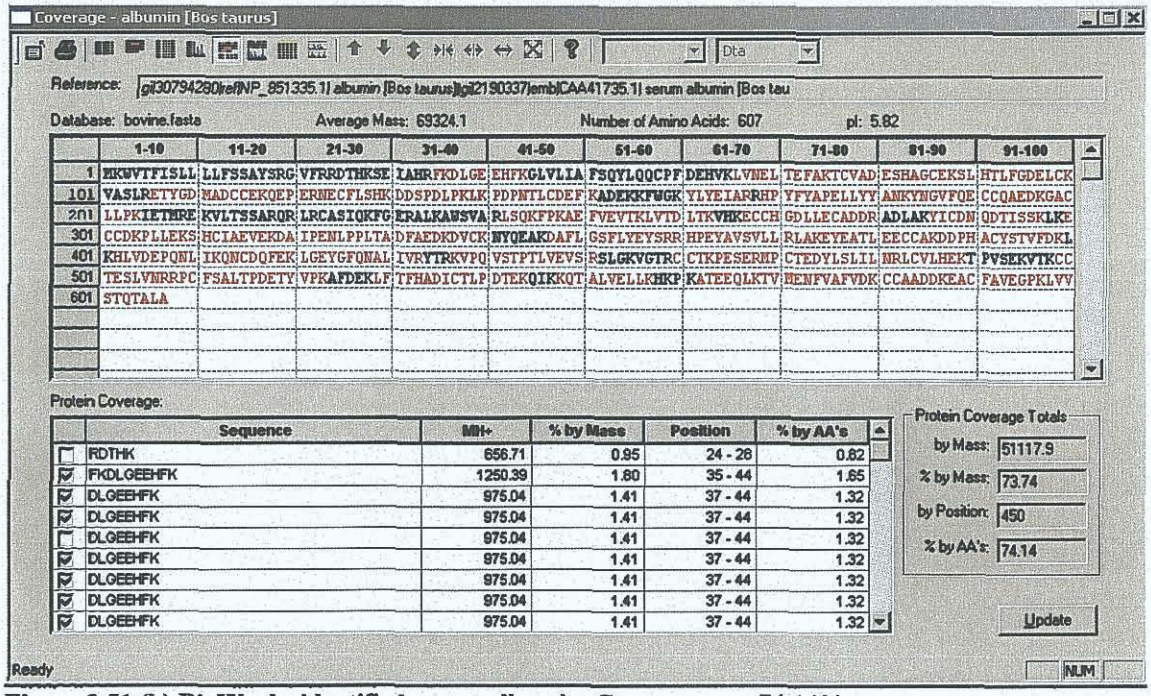



**Figure 3.51 (b) BioWorks identified serum albumin. Coverage was 74.14%**

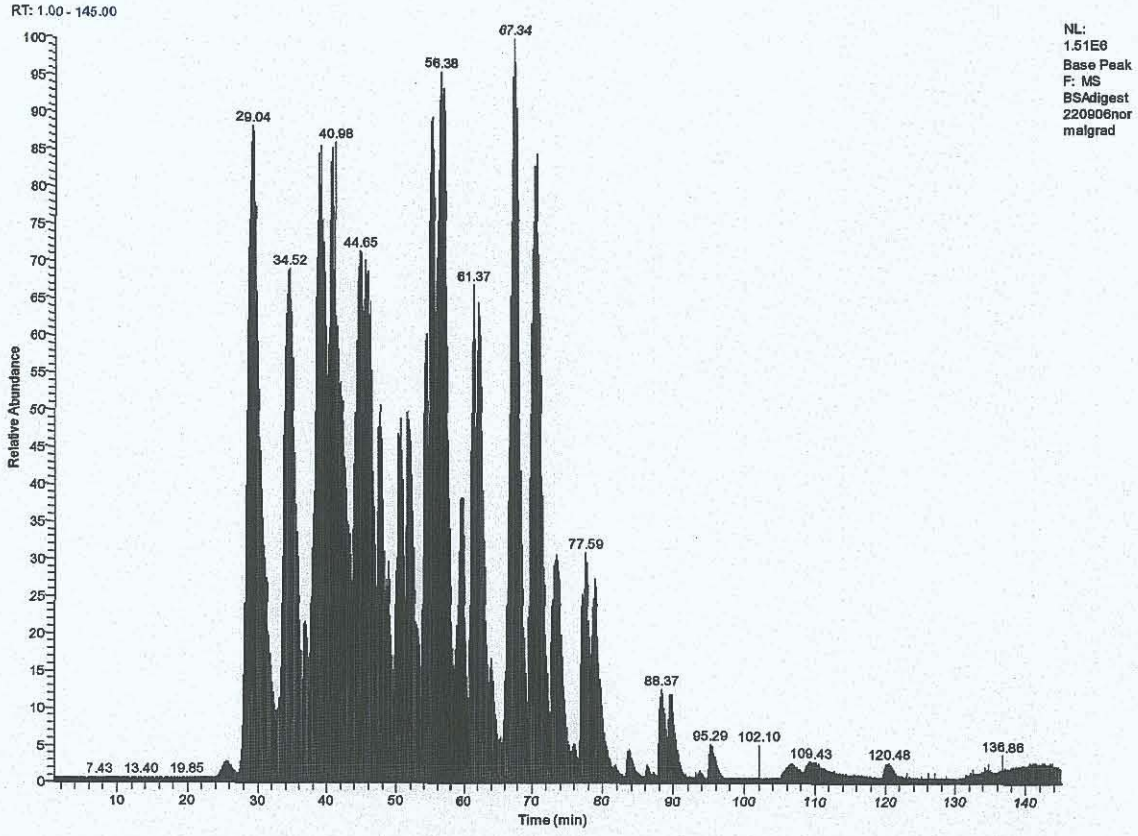



**Figure 3.52(a) MS spectrum of BSA digest using the long Gradient 0.1%FA – 65% ACN over 120mins**

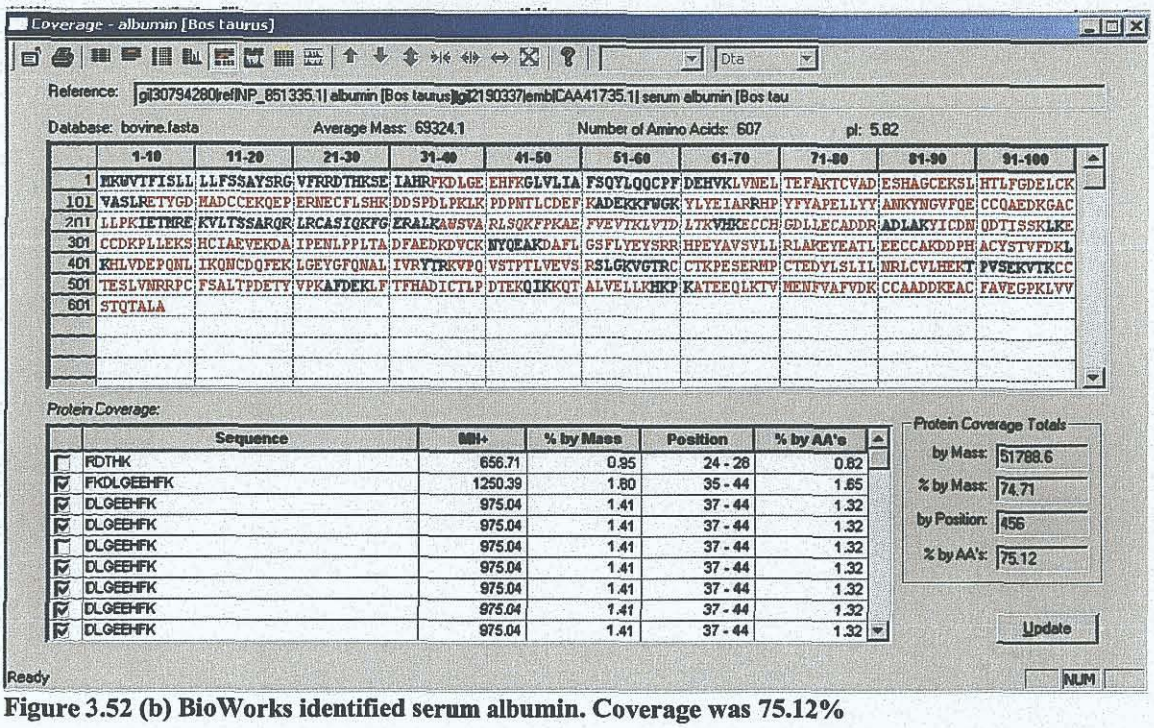



**Figure 3.52 (b) BioWorks identified serum albumin. Coverage was 75.12%**

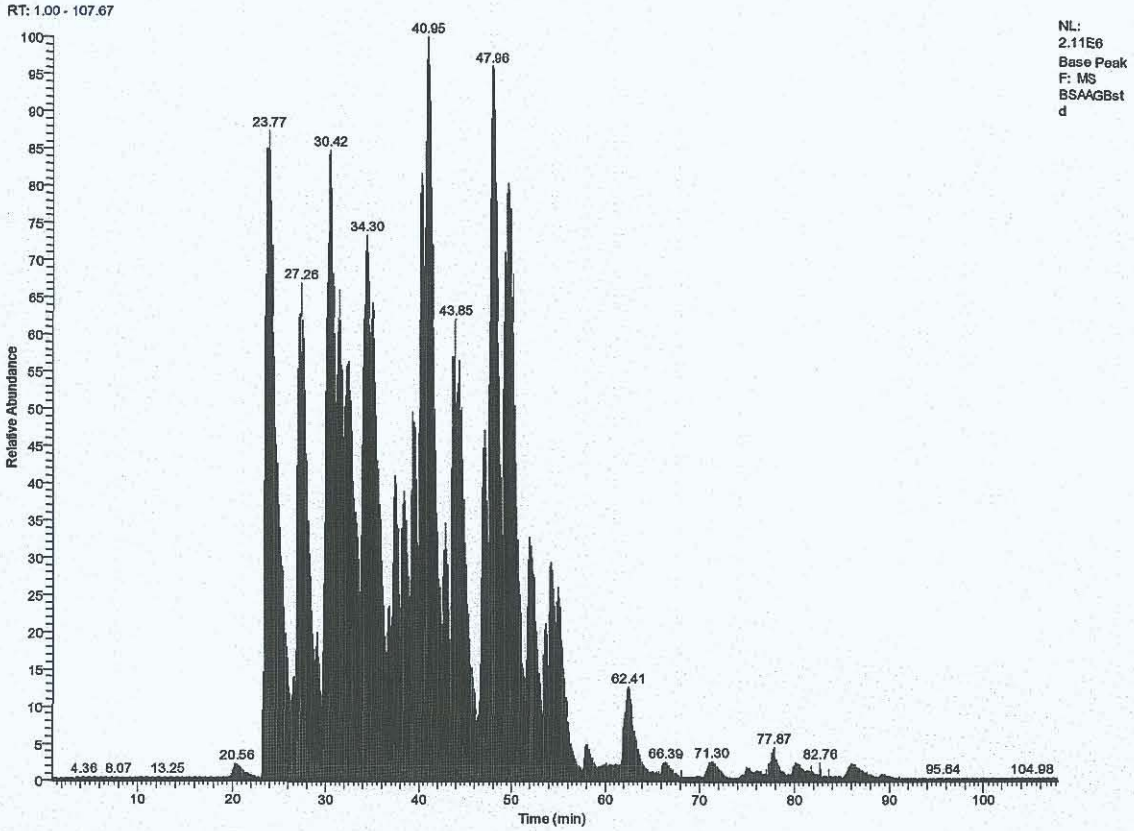



**Figure 3.53 (a) MS spectrum of BSA digest using the long Gradient 0.1%FA – 65% ACN over 60mins**

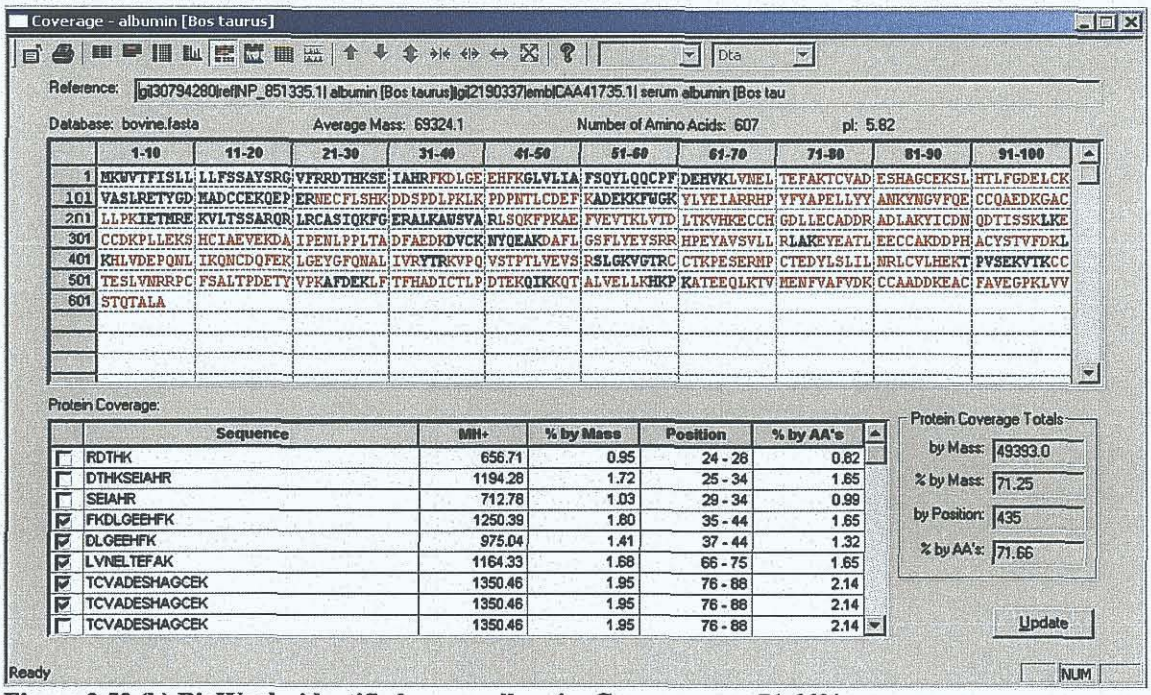



**Figure 3.53 (b) BioWorks identified serum albumin. Coverage was 71.66%**

Using three different gradients to separate a single protein tryptic digest (in this case 100fmol of BSA) indicated that all three methods were capable of separating the complex mix of 50 peptides. The long gradient 180mins resolved the peptides allowing a SEQUEST result for BSA giving coverage of 74% by amino acids and the medium gradient of 120mins gave slightly higher coverage of 75% for BSA by amino acids (figure 3.51+3.52). The short gradient at 60mins resolved the peptides enough to allow coverage of 71% for BSA by amino acid coverage (figure 3.53).

## 5.0 Conclusions

The main focus of this thesis was to optimise methods of Mass Spectrometry for protein identification. Proteins separated by 2D-gel electrophesis were stained with Coomasie blue and gel-plugs were subjected to in-gel digestion.

Optimisation procedures included combination of MALDI-TOF MS data from Zip Tip and non-Zip Tip samples. The use of calibration strategies with trypsin autodigestion peaks, commercially available calibrations peaks and protein peptide peaks to obtain high confidence protein identifications. The use of larger gel plugs to enhance mass spectral information and obtain protein identifications from low molecular weight proteins.

Setting up of liquid chromatography tandem mass spectrometry methods to allow for peptide sequence information

Using these optimised protein identification methods we looked at the effects of 5-bromo-2'-deoxyuridine (BrdU) in the poorly differentiated lung carcinoma cell line DLKP using 2-DE gel electrophoresis and mass spectrometry.

The following conclusions were made as a result of this study:

Following treatment of DLKP with 10μM BrdU, DIGE technology allowed the comparison of BrdU treated DLKP versus control DLKP. Using three biological replicates, 43 statistically significant abundance changes between control and treated DLKP were identified.

Using a combination of peptide clean up, calibration methods and increased volume of spot we increased the number proteins identified by MALDI-ToF MS possible compared to the high-throughput robotic methods currently in place in our laboratory. We identified 22 differentially regulated proteins using MALDI-ToF MS.

Using nano-LC and tandem Mass spectrometry methods it was possible to confirm 6 protein identifications that had been obtained by MALDI-ToF MS. An additional protein was identified by this method that was too low abundant for a MALDI-ToF MS identification.

In total, 23 proteins were identified by Mass Spectrometry. These proteins were involved in cell growth, maintenance, metabolism and immune response.

Through optimisation of the LC methods run times for sample identification were reduced from 180 minutes to 60 minutes with a minimal reduction of % amino acid coverage.

## 6.0 Future work

1. Time course study of the effects of BrdU on DLKP. During the course of this study DLKP was exposed to BrdU on day four and day seven. A global proteomic study using DIGE to examine the effects of BrdU on DLKP at a number of time points would be interesting.

2. Fraction of DLKP-BrdU treated proteome. A limitation of the total proteomic experiment is the bias toward highly abundant proteins such as proteins involved in metabolism and those found in the cytoskeleton. Fractionation of the BrdU-treated DLKP proteome into nuclear, cytoplasmic, mitochondrial, membrane etc would reveal specific regulation of proteins involved in the various processes investigated.

3. During the course we looked at a narrow range of protein from the total proteome (pH 4-7). Investigation of other ranges (e.g pH 3-5, 6-11 etc.) could reveal other differentially regulated proteins.

4. Western blotting will be used to confirm the identity of selected proteins obtained by Mass Spectrometry and to also confirm changes in expression level afte BrdU treatment. Immunocytochemistry and confocal microscopy will also be used to investigate the changes in protein expression, as protein function is often related to its location. Confocal microscopy will also allow for studies on protein interaction and co-localisation to be conducted.

5. Use quantitative Mass Spectrometry to compare differentially regulated proteins obtained by ITRAQ labelling versus 2D-DIGE.

## 7.0 Bibliography

1. Adams, L. D. & Weaver, K. M. Detection and recovery of proteins from gels following zinc chloride staining. *Appl. Theor. Electrophor.* **1**, 279-282 (1990).

2. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422**, 198-207 (2003).

3. Aebersold, R., Rist, B. & Gygi, S. P. Quantitative proteome analysis: methods and applications. *Ann. N. Y. Acad. Sci.* **919**, 33-47 (2000).

4. Alban, A. *et al.* A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics* **3**, 36-44 (2003).

5. Anderson, N. L. & Anderson, N. G. Analytical techniques for cell fractions. XXII. Two-dimensional analysis of serum and tissue proteins: multiple gradient-slab gel electrophoresis. *Anal. Biochem.* **85**, 341-354 (1978).

6. Arthur, J. M. Proteomics. *Curr. Opin. Nephrol. Hypertens.* **12**, 423-430 (2003).

7. Audebert, S. *et al.* The carboxy-terminal sequence Asp427-Glu432 of beta-tubulin plays an important function in axonemal motility. *Eur. J. Biochem.* **261**, 48-56 (1999).

8. Bamburg, J. R., McGough, A. & Ono, S. Putting a new twist on actin: ADF/cofilins modulate actin dynamics. *Trends Cell Biol.* **9**, 364-370 (1999).

9. Beavis, R. C. & Chait, B. T. Matrix-assisted laser desorption ionization mass-spectrometry of proteins. *Methods Enzymol.* **270**, 519-551 (1996).

10. Beere, H. M. *et al.* Heat-shock protein 70 inhibits apoptosis by preventing recruitment of procaspase-9 to the Apaf-1 apoptosome. *Nat. Cell Biol.* **2**, 469-475 (2000).

11. Benjamin, I. J. & McMillan, D. R. Stress (heat shock) proteins: molecular chaperones in cardiovascular biology and disease. *Circ. Res.* **83**, 117-132 (1998).

12. Berggren, K. N. *et al.* An improved formulation of SYPRO Ruby protein gel stain: comparison with the original formulation and with a ruthenium II tris (bathophenanthroline disulfonate) formulation. *Proteomics* **2**, 486-498 (2002).

13. Beussman, D. J., Vlasak, P. R., McLane, R. D., Seeterlin, M. A. & Enke, C. G. Tandem reflectron time-of-flight mass spectrometer utilizing photodissociation. *Anal. Chem.* **67**, 3952-3957 (1995).

14. Biemann, K. & Scoble, H. A. Characterization by tandem mass spectrometry of structural modifications in proteins. *Science* **237**, 992-998 (1987).

15. Bjellqvist, B. *et al.* Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J. Biochem. Biophys. Methods* **6**, 317-339 (1982).

16. Blackstock, W. P. & Weir, M. P. Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol.* **17**, 121-127 (1999).

17. Bonfils, C., Bec, N., Lacroix, B., Harricane, M. C. & Larroque, C. Kinetic analysis of tubulin assembly in the presence of the microtubule-associated protein TOGp. *J. Biol. Chem.* (2006).

18. Bonfils, C., Bec, N., Lacroix, B., Harricane, M. C. & Larroque, C. Kinetic analysis of tubulin assembly in the presence of the microtubule-associated protein TOGp. *J. Biol. Chem.* (2006).

19. Boveris, A. & Cadenas, E. Mitochondrial production of superoxide anions and its relationship to the antimycin insensitive respiration. *FEBS Lett.* **54**, 311-314 (1975).

20. Brinkley, W. Microtubules: a brief historical perspective. *J. Struct. Biol.* **118**, 84-86 (1997).

21. Burridge, K. & Chrzanowska-Wodnicka, M. Focal adhesions, contractility, and signaling. *Annu. Rev. Cell Dev. Biol.* **12**, 463-518 (1996).

22. Cadenas, E. & Davies, K. J. Mitochondrial free radical generation, oxidative stress, and aging. *Free Radic. Biol. Med.* **29**, 222-230 (2000).

23. Canas, B., Lopez-Ferrer, D., Ramos-Fernandez, A., Camafeita, E. & Calvo, E. Mass spectrometry technologies for proteomics. *Brief Funct. Genomic Proteomic* **4**, 295-320 (2006).

24. Carlier, M. F. Control of actin dynamics. *Curr. Opin. Cell Biol.* **10**, 45-51 (1998).

25. Cheng, Y. F. & Kramer, R. H. Human microvascular endothelial cells express integrin-related complexes that mediate adhesion to the extracellular matrix. *J. Cell. Physiol.* **139**, 275-286 (1989).

26. Chronowski, G. M. *et al.* An elevated serum beta-2-microglobulin level is an adverse prognostic factor for overall survival in patients with early-stage Hodgkin disease. *Cancer* **95**, 2534-2538 (2002).

27. Clauser, E. Comparative study of the structure and molecular functions of angiotensin II and vasopressin receptors. *C. R. Seances Soc. Biol. Fil.* **189**, 179-190 (1995).

28. Cohen, S. L. Domain elucidation by mass spectrometry. *Structure* **4**, 1013-1016 (1996).

29. Colinge, J., Masselot, A., Giron, M., Dessingy, T. & Magnin, J. OLAV: towards high-throughput tandem mass spectrometry data identification. *Proteomics* **3**, 1454-1463 (2003).

30. Corthals, G. L., Wasinger, V. C., Hochstrasser, D. F. & Sanchez, J. C. The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* **21**, 1104-1115 (2000).

31. Davis, B. J. Disc Electrophoresis II. Method and application to Human Serum Proteins. *Ann. N. Y. Acad. Sci.* **121**, 404-427 (1964).

32. Dolbeare, F. Bromodeoxyuridine: a diagnostic tool in biology and medicine, Part III. Proliferation in normal, injured and diseased tissue, growth factors, differentiation, DNA replication sites and in situ hybridization. *Histochem. J.* **28**, 531-575 (1996).

33. Dowsey, A. W., Dunn, M. J. & Yang, G. Z. The role of bioinformatics in two-dimensional gel electrophoresis. *Proteomics* **3**, 1567-1596 (2003).

34. Duester, G. Involvement of alcohol dehydrogenase, short-chain dehydrogenase/reductase, aldehyde dehydrogenase, and cytochrome P450 in the control of retinoid signaling by activation of retinoic acid synthesis. *Biochemistry* **35**, 12221-12227 (1996).

35. Dufner, A. & Thomas, G. Ribosomal S6 kinase signaling and the control of translation. *Exp. Cell Res.* **253**, 100-109 (1999).

36. Dunn, M. J. Electroelution of proteins from polyacrylamide gels. *Methods Mol. Biol.* **244**, 339-343 (2004).

37. Dunn, M. J. & Bradd, S. J. Separation and analysis of membrane proteins by SDS-polyacrylamide gel electrophoresis. *Methods Mol. Biol.* **19**, 203-210 (1993).

38. Durr, E. *et al.* Direct proteomic mapping of the lung microvascular endothelial cell surface in vivo and in cell culture. *Nat. Biotechnol.* **22**, 985-992 (2004).

39. Edman, P. & Begg, G. A protein sequenator. *Eur. J. Biochem.* **1**, 80-91 (1967).

40. Esumi, N. *et al.* Procoagulant activity of human neuroblastoma cell lines, in relation to cell growth, differentiation and cytogenetic abnormalities. *Jpn. J. Cancer Res.* **80**, 438-443 (1989).

41. Ezzell, R. M., Goldmann, W. H., Wang, N., Parasharama, N. & Ingber, D. E. Vinculin promotes cell spreading by mechanically coupling integrins to the cytoskeleton. *Exp. Cell Res.* **231**, 14-26 (1997).

42. Feng, J., Romaniouk, A. V., Samal, S. K. & Vijay, I. K. Processing enzyme glucosidase II: proposed catalytic residues and developmental regulation during the ontogeny of the mouse mammary gland. *Glycobiology* **14**, 909-921 (2004).

43. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64-71 (1989).

44. Feyles, V., Sikora, L. K., McGarry, R. C. & Jerry, L. M. Effects of retinoic acid and bromodeoxyuridine on human melanoma-associated antigen expression in small cell lung carcinoma cells. *Oncology* **48**, 58-64 (1991).

45. Frewen, B. E., Merrihew, G. E., Wu, C. C., Noble, W. S. & MacCoss, M. J. Analysis of peptide MS/MS spectra from large-scale proteomics experiments using spectrum libraries. *Anal. Chem.* **78**, 5678-5684 (2006).

46. Friedman, D. B. *et al.* Proteome analysis of human colon cancer by two-dimensional difference gel electrophoresis and mass spectrometry. *Proteomics* **4**, 793-811 (2004).

47. Galbraith, C. G., Yamada, K. M. & Sheetz, M. P. The relationship between force and focal complex development. *J. Cell Biol.* **159**, 695-705 (2002).

48. Gattiker, A., Bienvenut, W. V., Bairoch, A. & Gasteiger, E. FindPept, a tool to identify unmatched masses in peptide mass fingerprinting protein identification. *Proteomics* **2**, 1435-1444 (2002).

49. Geiger, B., Tokuyasu, K. T., Dutton, A. H. & Singer, S. J. Vinculin, an intracellular protein localized at specialized sites where microfilament bundles terminate at cell membranes. *Proc. Natl. Acad. Sci. U. S. A.* **77**, 4127-4131 (1980).

50. Gobom, J. *et al.* A calibration method that simplifies and improves accurate determination of peptide molecular masses by MALDI-TOF MS. *Anal. Chem.* **74**, 3915-3923 (2002).

51. Goldmann, W. H. & Ingber, D. E. Intact vinculin protein is required for control of cell shape, cell mechanics, and rac-dependent lamellipodia formation. *Biochem. Biophys. Res. Commun.* **290**, 749-755 (2002).

52. Gorg, A., Boguth, G., Obermaier, C., Posch, A. & Weiss, W. Two-dimensional polyacrylamide gel electrophoresis with immobilized pH gradients in the first dimension (IPG-Dalt): the state of the art and the controversy of vertical versus horizontal systems. *Electrophoresis* **16**, 1079-1086 (1995).

53. Gorg, A. *et al.* The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* **21**, 1037-1053 (2000).

54. Gorg, A., Postel, W. & Gunther, S. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* **9**, 531-546 (1988).

55. Graves, P., Pritsker, A. & Davies, T. F. Post-translational processing of the natural human thyrotropin receptor: demonstration of more than two cleavage sites. *J. Clin. Endocrinol. Metab.* **84**, 2177-2181 (1999).

56. Grey, H. M. *et al.* The small subunit of HL-A antigens is beta 2-microglobulin. *J. Exp. Med.* **138**, 1608-1612 (1973).

57. Griffin, T. J. & Aebersold, R. Advances in proteome analysis by mass spectrometry. *J. Biol. Chem.* **276**, 45497-45500 (2001).

58. Haebel, S. *et al.* Electrophoresis-related protein modification: alkylation of carboxy residues revealed by mass spectrometry. *Electrophoresis* **19**, 679-686 (1998).

59. Hager, J. W. & Le Blanc, J. C. High-performance liquid chromatography-tandem mass spectrometry with a new quadrupole/linear ion trap instrument. *J. Chromatogr. A* **1020**, 3-9 (2003).

60. Han, D., Antunes, F., Canali, R., Rettori, D. & Cadenas, E. Voltage-dependent anion channels control the release of the superoxide anion from mitochondria to cytosol. *J. Biol. Chem.* **278**, 5557-5563 (2003).

61. Hansen, J. B., Petersen, R. K., Jorgensen, C. & Kristiansen, K. Deregulated MAPK activity prevents adipocyte differentiation of fibroblasts lacking the retinoblastoma protein. *J. Biol. Chem.* **277**, 26335-26339 (2002).

62. Harding, J. D., Przybyla, A. E., MacDonald, R. J., Pictet, R. L. & Rutter, W. J. Effects of dexamethasone and 5-bromodeoxyuridine on the synthesis of amylase mRNA during pancreatic development in vitro. *J. Biol. Chem.* **253**, 7531-7537 (1978).

63. Hardman, M. & Makarov, A. A. Interfacing the orbitrap mass analyzer to an electrospray ion source. *Anal. Chem.* **75**, 1699-1705 (2003).

64. Hart, C., Schulenberg, B., Steinberg, T. H., Leung, W. Y. & Patton, W. F. Detection of glycoproteins in polyacrylamide gels and on electroblots using Pro-Q Emerald 488 dye, a fluorescent periodate Schiff-base stain. *Electrophoresis* **24**, 588-598 (2003).

65. Hart, G. W., Haltiwanger, R. S., Holt, G. D. & Kelly, W. G. Glycosylation in the nucleus and cytoplasm. *Annu. Rev. Biochem.* **58**, 841-874 (1989).

66. Haynes, P. A. & Yates, J. R.,3rd. Proteome profiling-pitfalls and progress. *Yeast* **17**, 81-87 (2000).

67. Hayward, I. P., Bridle, K. R., Campbell, G. R., Underwood, P. A. & Campbell, J. H. Effect of extracellular matrix proteins on vascular smooth muscle cell phenotype. *Cell Biol. Int.* **19**, 839-846 (1995).

68. He, L., Wu, J. T., Parus, S. & Lubman, D. M. Development of a capillary high-performance liquid chromatography tandem mass spectrometry system using SWIFT technology in an ion trap/reflectron time-of-flight mass spectrometer. *Rapid Commun. Mass Spectrom.* **11**, 1739-1748 (1997).

69. Hehlgans, S., Haase, M. & Cordes, N. Signalling via integrins: implications for cell survival and anticancer strategies. *Biochim. Biophys. Acta* **1775**, 163-180 (2007).

70. Hensbergen, P. *et al.* Proteomic profiling identifies an UV-induced activation of cofilin-1 and destrin in human epidermis. *J. Invest. Dermatol.* **124**, 818-824 (2005).

71. Henzel, W. J. *et al.* Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 5011-5015 (1993).

72. Herceg, Z. & Wang, Z. Q. Failure of poly(ADP-ribose) polymerase cleavage by caspases leads to induction of necrosis and enhanced apoptosis. *Mol. Cell. Biol.* **19**, 5124-5133 (1999).

73. Honnappa, S., John, C. M., Kostrewa, D., Winkler, F. K. & Steinmetz, M. O. Structural insights into the EB1-APC interaction. *EMBO J.* **24**, 261-269 (2005).

74. Hoorn, E. J., Hoffert, J. D. & Knepper, M. A. Combined proteomics and pathways analysis of collecting duct reveals a protein regulatory network activated in vasopressin escape. *J. Am. Soc. Nephrol.* **16**, 2852-2863 (2005).

75. Hynes, R. Phosphorylation of vinculin by pp60src: what might it mean? *Cell* **28**, 437-438 (1982).

76. Ishola, D. A.,Jr *et al.* Albumin-bound fatty acids induce mitochondrial oxidant stress and impair antioxidant responses in proximal tubular cells. *Kidney Int.* **70**, 724-731 (2006).

77. Jang, W. G. *et al.* Analysis of proteome and transcriptome of tumor necrosis factor alpha stimulated vascular smooth muscle cells with or without alpha lipoic acid. *Proteomics* **4**, 3383-3393 (2004).

78. Jensen, O. N., Podtelejnikov, A. & Mann, M. Delayed extraction improves specificity in database searches by matrix-assisted laser desorption/ionization peptide maps. *Rapid Commun. Mass Spectrom.* **10**, 1371-1378 (1996).

79. Jiang, L., He, L. & Fountoulakis, M. Comparison of protein precipitation methods for sample preparation prior to proteomic analysis. *J. Chromatogr. A* **1023**, 317-320 (2004).

80. Johansson, A. C., Steen, H., Ollinger, K. & Roberg, K. Cathepsin D mediates cytochrome c release and caspase activation in human fibroblast apoptosis induced by staurosporine. *Cell Death Differ.* **10**, 1253-1259 (2003).

81. Josic, D. *et al.* Use of selective extraction and fast chromatographic separation combined with electrophoretic methods for mapping of membrane proteins. *Electrophoresis* **26**, 2809-2822 (2005).

82. Jovinge, S., Ares, M. P., Kallin, B. & Nilsson, J. Human monocytes/macrophages release TNF-alpha in response to Ox-LDL. *Arterioscler. Thromb. Vasc. Biol.* **16**, 1573-1579 (1996).

83. Jungblut, P. *et al.* Resolution power of two-dimensional electrophoresis and identification of proteins from gels. *Electrophoresis* **17**, 839-847 (1996).

84. Jungblut, P. R. *et al.* Comparative proteome analysis of Mycobacterium tuberculosis and Mycobacterium bovis BCG strains: towards functional genomics of microbial pathogens. *Mol. Microbiol.* **33**, 1103-1117 (1999).

85. Junn, E. *et al.* Interaction of DJ-1 with Daxx inhibits apoptosis signal-regulating kinase 1 activity and cell death. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 9691-9696 (2005).

86. Kaprielian, Z. & Patterson, P. H. The molecular basis of retinotectal topography. *Bioessays* **16**, 1-11 (1994).

87. Karas M, Bachmann D, Bahr U and Hillenkamp F. Matrix-assisted ultraviolet Laser desorption of non-volatile compounds. *Int J Mass Spectrom & Ion Proc* **78**, 53 (1987).

88. Karas, M. & Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **60**, 2299-2301 (1988).

89. Katsumi, A., Orr, A. W., Tzima, E. & Schwartz, M. A. Integrins in mechanotransduction. *J. Biol. Chem.* **279**, 12001-12004 (2004).

90. Khalsa-Moyers, G. & McDonald, W. H. Developments in mass spectrometry for the analysis of complex protein mixtures. *Brief Funct. Genomic Proteomic* **5**, 98-111 (2006).

91. Khurana, T., Khurana, B. & Noegel, A. A. LIM proteins: association with the actin cytoskeleton. *Protoplasma* **219**, 1-12 (2002).

92. Kim, H. J., Hida, H., Jung, C. G., Miura, Y. & Nishino, H. Treatment with deferoxamine increases neurons from neural stem/progenitor cells. *Brain Res.* **1092**, 1-15 (2006).

93. 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* **26**, 231-243 (1975).

94. Knepper, M. A. Proteomics and the kidney. *J. Am. Soc. Nephrol.* **13**, 1398-1408 (2002).

95. Koller, A. *et al.* Proteomic survey of metabolic pathways in rice. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11969-11974 (2002).

96. Kratchmarova, I., Blagoev, B., Haack-Sorensen, M., Kassem, M. & Mann, M. Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. *Science* **308**, 1472-1477 (2005).

97. Krause, E., Wenschuh, H. & Jungblut, P. R. The dominance of arginine-containing peptides in MALDI-derived tryptic mass fingerprints of proteins. *Anal. Chem.* **71**, 4160-4165 (1999).

98. Krutchinsky, A. N., Kalkum, M. & Chait, B. T. Automatic identification of proteins with a MALDI-quadrupole ion trap mass spectrometer. *Anal. Chem.* **73**, 5066-5077 (2001).

99. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685 (1970).

100. Lane, H. A., Fernandez, A., Lamb, N. J. & Thomas, G. p70s6k function is essential for G1 progression. *Nature* **363**, 170-172 (1993).

101. Lapko, V. N., Smith, D. L. & Smith, J. B. Identification of an artifact in the mass spectrometry of proteins derivatized with iodoacetamide. *J. Mass Spectrom.* **35**, 572-575 (2000).

102. Lappalainen, P. & Drubin, D. G. Cofilin promotes rapid actin filament turnover in vivo. *Nature* **388**, 78-82 (1997).

103. Laugesen, S. *et al.* Phosphoproteins analysis in plants: a proteomic approach. *Phytochemistry* **67**, 2208-2214 (2006).

104. Le Moguen, K. *et al.* Comparative proteomic analysis of cisplatin sensitive IGROV1 ovarian carcinoma cell line and its resistant counterpart IGROV1-R10. *Proteomics* **6**, 5183-5192 (2006).

105. Lee, K. Y., Ladha, M. H., McMahon, C. & Ewen, M. E. The retinoblastoma protein is linked to the activation of Ras. *Mol. Cell. Biol.* **19**, 7724-7732 (1999).

106. Lee, W. H. *et al.* Upregulation of class II beta-tubulin expression in differentiating keratinocytes. *J. Invest. Dermatol.* **124**, 291-297 (2005).

107. Levander, F., Rognvaldsson, T., Samuelsson, J. & James, P. Automated methods for improved protein identification by peptide mass fingerprinting. *Proteomics* **4**, 2594-2601 (2004).

108. Li, B., Zhang, D. M., Luo, Y. M. & Chen, X. G. Three new and antitumor anthraquinone glycosides from Lasianthus acuminatissimus MERR. *Chem. Pharm. Bull. (Tokyo)* **54**, 297-300 (2006).

109. Li, W., Wang, G. J., Cui, J. W., Chen, X. F. & Zhang, X. M. The proteomics study of apoptotic NB4 cells induced by sodium butyrate. *Zhonghua Xue Ye Xue Za Zhi* **27**, 436-440 (2006).

110. Liao, J., Lowthert, L. A., Ghori, N. & Omary, M. B. The 70-kDa heat shock proteins associate with glandular intermediate filaments in an ATP-dependent manner. *J. Biol. Chem.* **270**, 915-922 (1995).

111. Lilley, K. S. & Friedman, D. B. All about DIGE: quantification technology for differential-display 2D-gel proteomics. *Expert Rev. Proteomics* **1**, 401-409 (2004).

112. Lin, J. D., Huang, C. C., Weng, H. F., Chen, S. C. & Jeng, L. B. Comparison of membrane proteins from benign and malignant human thyroid tissues by two-dimensional polyacrylamide gel electrophoresis. *J. Chromatogr. B. Biomed. Appl.* **667**, 153-160 (1995).

113. Lin, S. Y. & Riggs, A. D. Lac operator analogues: bromodeoxyuridine substitution in the lac operator affects the rate of dissociation of the lac repressor. *Proc. Natl. Acad. Sci. U. S. A.* **69**, 2574-2576 (1972).

114. Lindahl, R. Aldehyde dehydrogenases and their role in carcinogenesis. *Crit. Rev. Biochem. Mol. Biol.* **27**, 283-335 (1992).

115. Lippens, S., Denecker, G., Ovaere, P., Vandenabeele, P. & Declercq, W. Death penalty for keratinocytes: apoptosis versus cornification. *Cell Death Differ.* **12 Suppl 2**, 1497-1508 (2005).

116. Loboda, A. Novel ion mobility setup combined with collision cell and time-of-flight mass spectrometer. *J. Am. Soc. Mass Spectrom.* **17**, 691-699 (2006).

117. Loo, J. A. *et al.* High sensitivity mass spectrometric methods for obtaining intact molecular weights from gel-separated proteins. *Electrophoresis* **20**, 743-748 (1999).

118. Luche, S., Santoni, V. & Rabilloud, T. Evaluation of nonionic and zwitterionic detergents as membrane protein solubilizers in two-dimensional electrophoresis. *Proteomics* **3**, 249-253 (2003).

119. Mann, M. A shortcut to interesting human genes: peptide sequence tags, expressed-sequence tags and computers. *Trends Biochem. Sci.* **21**, 494-495 (1996).

120. Mann, M. & Wilm, M. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal. Chem.* **66**, 4390-4399 (1994).

121. Martin, J. A., Solla, A., Woodward, S. & Gil, L. Fourier transform-infrared spectroscopy as a new method for evaluating host resistance in the Dutch elm disease complex. *Tree Physiol.* **25**, 1331-1338 (2005).

122. Martinat, C. *et al.* Sensitivity to oxidative stress in DJ-1-deficient dopamine neurons: an ES- derived cell model of primary Parkinsonism. *PLoS Biol.* **2**, e327 (2004).

123. Mathioudaki, K. *et al.* SR-A1, a member of the human pre-mRNA splicing factor family, and its expression in colon cancer progression. *Biol. Chem.* **385**, 785-790 (2004).

124. Mattow, J. *et al.* Identification of proteins from Mycobacterium tuberculosis missing in attenuated Mycobacterium bovis BCG strains. *Electrophoresis* **22**, 2936-2946 (2001).

125. McBride, S., Meleady, P., Baird, A., Dinsdale, D. & Cylnes, M. Human lung carcinoma cell line DLKP contains 3 distinct subpopulations with different growth and attachment properties. *Tumor Biol.* **19**(2), 88-103 (1998).

126. McBride, S., Walsh, D., Meleady, P., Daly, N. & Clynes, M. Bromodeoxyuridine induces keratin protein synthesis at a posttranscriptional level in human lung tumour cell lines. *Differentiation* **64**, 185-193 (1999).

127. McCormack, A. L. *et al.* Direct analysis and identification of proteins in mixtures by LC/MS/MS and database searching at the low-femtomole level. *Anal. Chem.* **69**, 767-776 (1997).

128. Meleady, P. & Clynes, M. Bromodeoxyuridine induces integrin expression at transcriptional (alpha2 subunit) and post-transcriptional (beta1 subunit) levels, and alters the adhesive properties of two human lung tumour cell lines. *Cell. Commun. Adhes.* **8**, 45-59 (2001).

129. Merril, C. R., Goldman, D., Sedman, S. A. & Ebert, M. H. Ultrasensitive stain for proteins in polyacrylamide gels shows regional variation in cerebrospinal fluid proteins. *Science* **211**, 1437-1438 (1981).

130. Molloy, M. P. Two-dimensional electrophoresis of membrane proteins using immobilized pH gradients. *Anal. Biochem.* **280**, 1-10 (2000).

131. Molloy, M. P., Brzezinski, E. E., Hang, J., McDowell, M. T. & VanBogelen, R. A. Overcoming technical variation and biological variation in quantitative proteomics. *Proteomics* **3**, 1912-1919 (2003).

132. Nikitin, A. Y., Shan, B., Flesken-Nikitin, A., Chang, K. H. & Lee, W. H. The retinoblastoma gene regulates somatic growth during mouse development. *Cancer Res.* **61**, 3110-3118 (2001).

133. O'Farrell, P. H. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250**, 4007-4021 (1975).

134. Ohnuma, S. & Harris, W. A. Neurogenesis and the cell cycle. *Neuron* **40**, 199-208 (2003).

135. Ohsawa, I., Nishimaki, K., Yasuda, C., Kamino, K. & Ohta, S. Deficiency in a mitochondrial aldehyde dehydrogenase increases vulnerability to oxidative stress in PC12 cells. *J. Neurochem.* **84**, 1110-1117 (2003).

136. Opiteck, G. J., Ramirez, S. M., Jorgenson, J. W. & Moseley, M. A.,3rd. Comprehensive two-dimensional high-performance liquid chromatography for the isolation of overexpressed proteins and proteome mapping. *Anal. Biochem.* **258**, 349-361 (1998).

137. ORNSTEIN, L. DISC ELECTROPHORESIS. I. BACKGROUND AND THEORY. *Ann. N. Y. Acad. Sci.* **121**, 321-349 (1964).

138. Papayannopoulos, I. A. & Biemann, K. Fast atom bombardment and tandem mass spectrometry of synthetic peptides and byproducts. *Pept. Res.* **5**, 83-90 (1992).

139. Pappin, D. J. Peptide mass fingerprinting using MALDI-TOF mass spectrometry. *Methods Mol. Biol.* **211**, 211-219 (2003).

140. Pappin, D. J., Hojrup, P. & Bleasby, A. J. Rapid identification of proteins by peptide-mass fingerprinting. *Curr. Biol.* **3**, 327-332 (1993).

141. Patterson, S. D. & Latter, G. I. Evaluation of storage phosphor imaging for quantitative analysis of 2-D gels using the Quest II system. *BioTechniques* **15**, 1076-1083 (1993).

142. Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J. & Gygi, S. P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J. Proteome Res.* **2**, 43-50 (2003).

143. Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551-3567 (1999).

144. Pieper, R. *et al.* Multi-component immunoaffinity subtraction chromatography: an innovative step towards a comprehensive survey of the human plasma proteome. *Proteomics* **3**, 422-432 (2003).

145. Pivazyan, A. D., Birks, E. M., Wood, T. G., Lin, T. S. & Prusoff, W. H. Inhibition of poly(ADP-ribose)polymerase activity by nucleoside analogs of thymidine. *Biochem. Pharmacol.* **44**, 947-953 (1992).

146. Pullen, N. *et al.* Phosphorylation and activation of p70s6k by PDK1. *Science* **279**, 707-710 (1998).

147. Quadroni, M. & James, P. Proteomics and automation. *Electrophoresis* **20**, 664-677 (1999).

148. Rabilloud, T. Use of thiourea to increase the solubility of membrane proteins in two-dimensional electrophoresis. *Electrophoresis* **19**, 758-760 (1998).

149. Rantalainen, M. *et al.* Statistically integrated metabonomic-proteomic studies on a human prostate cancer xenograft model in mice. *J. Proteome Res.* **5**, 2642-2655 (2006).

150. Ravagnan, L. *et al.* Heat-shock protein 70 antagonizes apoptosis-inducing factor. *Nat. Cell Biol.* **3**, 839-843 (2001).

151. Reinders, J., Lewandrowski, U., Moebius, J., Wagner, Y. & Sickmann, A. Challenges in mass spectrometry-based proteomics. *Proteomics* **4**, 3686-3703 (2004).

152. Richardson, A. & Parsons, T. A mechanism for regulation of the adhesion-associated proteintyrosine kinase pp125FAK. *Nature* **380**, 538-540 (1996).

153. Rodriguez Fernandez, J. L., Geiger, B., Salomon, D. & Ben-Ze'ev, A. Suppression of vinculin expression by antisense transfection confers changes in cell morphology, motility, and anchorage-dependent growth of 3T3 cells. *J. Cell Biol.* **122**, 1285-1294 (1993).

154. Ruth, M. C. *et al.* Analysis of membrane proteins from human chronic myelogenous leukemia cells: comparison of extraction methods for multidimensional LC-MS/MS. *J. Proteome Res.* **5**, 709-719 (2006).

155. Salplachta, J., Rehulka, P. & Chmelik, J. Identification of proteins by combination of size-exclusion chromatography with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and comparison of some desalting procedures for both intact proteins and their tryptic digests. *J. Mass Spectrom.* **39**, 1395-1401 (2004).

156. Santamaria, I. *et al.* Cathepsin L2, a novel human cysteine proteinase produced by breast and colorectal carcinomas. *Cancer Res.* **58**, 1624-1630 (1998).

157. Schafer, D. A. & Cooper, J. A. Control of actin assembly at filament ends. *Annu. Rev. Cell Dev. Biol.* **11**, 497-518 (1995).

158. Schmidt, A. & Hall, M. N. Signaling to the actin cytoskeleton. *Annu. Rev. Cell Dev. Biol.* **14**, 305-338 (1998).

159. Schuchard, M. D. *et al.* Artifactual isoform profile modification following treatment of human plasma or serum with protease inhibitor, monitored by 2-dimensional electrophoresis and mass spectrometry. *BioTechniques* **39**, 239-247 (2005).

160. Schwartz, J. C., Senko, M. W. & Syka, J. E. A two-dimensional quadrupole ion trap mass spectrometer. *J. Am. Soc. Mass Spectrom.* **13**, 659-669 (2002).

161. Seike, M. *et al.* Proteomic signature of human cancer cells. *Proteomics* **4**, 2776-2788 (2004).

162. Shadforth, I. & Bessant, C. Genome annotating proteomics pipelines: available tools. *Expert Rev. Proteomics* **3**, 621-629 (2006).

163. Shaw, J. *et al.* Evaluation of saturation labelling two-dimensional difference gel electrophoresis fluorescent dyes. *Proteomics* **3**, 1181-1195 (2003).

164. Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68**, 850-858 (1996).

165. Shinjyo, N. & Kita, K. Up-regulation of heme biosynthesis during differentiation of Neuro2a cells. *J. Biochem. (Tokyo)* **139**, 373-381 (2006).

166. Simpkins, F., Czechowicz, J. A., Liotta, L. & Kohn, E. C. SELDI-TOF mass spectrometry for cancer biomarker discovery and serum proteomic diagnostics. *Pharmacogenomics* **6**, 647-653 (2005).

167. Simpson, R. J. & Morgan, F. J. Isolation and complete amino acid sequence of a basic low molecular weight protein from black swan egg white. *Int. J. Pept. Protein Res.* **22**, 476-481 (1983).

168. Simpson, R. J. & Morgan, F. J. Isolation and complete amino acid sequence of a basic low molecular weight protein from black swan egg white. *Int. J. Pept. Protein Res.* **22**, 476-481 (1983).

169. Sinha, P., Poland, J., Schnolzer, M. & Rabilloud, T. A new silver staining apparatus and procedure for matrix-assisted laser desorption/ionization-time of flight analysis of proteins after two-dimensional electrophoresis. *Proteomics* **1**, 835-840 (2001).

170. Sladek, N. E., Kollander, R., Sreerama, L. & Kiang, D. T. Cellular levels of aldehyde dehydrogenases (ALDH1A1 and ALDH3A1) as predictors of therapeutic responses to cyclophosphamide-based chemotherapy of breast cancer: a retrospective study. Rational individualization of oxazaphosphorine-based cancer chemotherapeutic regimens. *Cancer Chemother. Pharmacol.* **49**, 309-321 (2002).

171. Smith, C. *et al.* Purification and partial characterization of a human hematopoietic precursor population. *Blood* **77**, 2122-2128 (1991).

172. Smolka, M. B., Zhou, H., Purkayastha, S. & Aebersold, R. Optimization of the isotope-coded affinity tag-labeling procedure for quantitative proteome analysis. *Anal. Biochem.* **297**, 25-31 (2001).

173. Srisomsap, C. *et al.* Detection of cathepsin B up-regulation in neoplastic thyroid tissues by proteomic analysis. *Proteomics* **2**, 706-712 (2002).

174. Steen, H. & Pandey, A. Proteomics goes quantitative: measuring protein abundance. *Trends Biotechnol.* **20**, 361-364 (2002).

175. Steinberg, T. H. *et al.* Global quantitative phosphoprotein analysis using Multiplexed Proteomics technology. *Proteomics* **3**, 1128-1144 (2003).

176. Strunck, E., Frank, K., Tan, M. I. & Vollmer, G. Expression of l-3-phosphoserine phosphatase is regulated by reconstituted basement membrane. *Biochem. Biophys. Res. Commun.* **281**, 747-753 (2001).

177. Tapscott, S. J., Lassar, A. B., Davis, R. L. & Weintraub, H. 5-bromo-2'-deoxyuridine blocks myogenesis by extinguishing expression of MyoD1. *Science* **245**, 532-536 (1989).

178. Taupin, P. BrdU immunohistochemistry for studying adult neurogenesis: paradigms, pitfalls, limitations, and validation. *Brain Res. Brain Res. Rev.* **53**, 198-214 (2007).

179. Terskikh, A. V. *et al.* From hematopoiesis to neuropoiesis: evidence of overlapping genetic programs. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 7934-7939 (2001).

180. Thiede, B. *et al.* Analysis of missed cleavage sites, tryptophan oxidation and N-terminal pyroglutamylation after in-gel tryptic digestion. *Rapid Commun. Mass Spectrom.* **14**, 496-502 (2000).

181. Tomer, K. B. Separations combined with mass spectrometry. *Chem. Rev.* **101**, 297-328 (2001).

182. Unlu, M., Morgan, M. E. & Minden, J. S. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* **18**, 2071-2077 (1997).

183. Urwin, V. E. & Jackson, P. Two-dimensional polyacrylamide gel electrophoresis of proteins labeled with the fluorophore monobromobimane prior to first-dimensional isoelectric focusing: imaging of the fluorescent protein spot patterns using a cooled charge-coupled device. *Anal. Biochem.* **209**, 57-62 (1993).

184. Van den Bergh, G. & Arckens, L. Fluorescent two-dimensional difference gel electrophoresis unveils the potential of gel-based proteomics. *Curr. Opin. Biotechnol.* **15**, 38-43 (2004).

185. VanBogelen, R. A. & Molloy, M. P. Exploring the proteome: reviving emphasis on quantitative protein profiling. *Proteomics* **3**, 1833-1834 (2003).

186. Vartiainen, M. K. *et al.* The three mouse actin-depolymerizing factor/cofilins evolved to fulfill cell-type-specific requirements for actin dynamics. *Mol. Biol. Cell* **13**, 183-194 (2002).

187. Virag, L. Structure and function of poly(ADP-ribose) polymerase-1: role in oxidative stress-related pathologies. *Curr. Vasc. Pharmacol.* **3**, 209-214 (2005).

188. Vorm, O. & Roepstorff, P. Peptide sequence information derived by partial acid hydrolysis and matrix-assisted laser desorption/ionization mass spectrometry. *Biol. Mass Spectrom.* **23**, 734-740 (1994).

189. Walsh, D., Meleady, P., Power, B., Morley, S. J. & Clynes, M. Increased levels of the translation initiation factor eIF4E in differentiating epithelial lung tumor cell lines. *Differentiation* **71**, 126-134 (2003).

190. Wang, D. & Gao, L. Proteomic analysis of neural differentiation of mouse embryonic stem cells. *Proteomics* **5**, 4414-4426 (2005).

191. Washburn, M. P., Ulaszek, R. R. & Yates, J. R.,3rd. Reproducibility of quantitative proteomic analyses of complex biological mixtures by multidimensional protein identification technology. *Anal. Chem.* **75**, 5054-5061 (2003).

192. Weinberg, R. A. The retinoblastoma protein and cell cycle control. *Cell* **81**, 323-330 (1995).

193. Werth, D. K., Niedel, J. E. & Pastan, I. Vinculin, a cytoskeletal substrate of protein kinase C. *J. Biol. Chem.* **258**, 11423-11426 (1983).

194. Westermeier, R. Isoelectric focusing. *Methods Mol. Biol.* **244**, 225-232 (2004).

195. Wiese, S., Reidegeld, K. A., Meyer, H. E. & Warscheid, B. Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. *Proteomics* **7**, 340-350 (2007).

196. Wilkins, D. G., Valdez, A. S., Krueger, G. G. & Rollins, D. E. Quantitative analysis of l-alpha-acetylmethadol, l-alpha-acetyl-N-normethadol, and l-alpha-acetyl-N,N-dinormethadol in human hair by positive ion chemical ionization mass spectrometry. *J. Anal. Toxicol.* **21**, 420-426 (1997).

197. Williams, J. G. & Tomer, K. B. Disposable chromatography for a high-throughput nano-ESI/MS and nano-ESI/MS-MS platform. *J. Am. Soc. Mass Spectrom.* **15**, 1333-1340 (2004).

198. Williams, J. N. *et al.* Proteomic analysis of outer membranes and vesicles from wild-type serogroup B Neisseria meningitidis and a lipopolysaccharide deficient mutant. *Infect. Immun.* **75**, 1364-1372 (2007).

199. Wilson, C. M. Studies and critique of Amido Black 10B, Coomassie Blue R, and Fast Green FCF as stains for proteins after polyacrylamide gel electrophoresis. *Anal. Biochem.* **96**, 263-278 (1979).

200. Xu, J. *et al.* The Parkinson's disease-associated DJ-1 protein is a transcriptional co-activator that protects against neuronal apoptosis. *Hum. Mol. Genet.* **14**, 1231-1241 (2005).

201. Xu, W., Baribault, H. & Adamson, E. D. Vinculin knockout results in heart and brain defects during embryonic development. *Development* **125**, 327-337 (1998).

202. Yan, J. X., Harry, R. A., Spibey, C. & Dunn, M. J. Postelectrophoretic staining of proteins separated by two-dimensional gel electrophoresis using SYPRO dyes. *Electrophoresis* **21**, 3657-3665 (2000).

203. Yang, Y. *et al.* Inactivation of Drosophila DJ-1 leads to impairments of oxidative stress response and phosphatidylinositol 3-kinase/Akt signaling. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 13670-13675 (2005).

204. Yates, J. R.,3rd. Mass spectrometry. From genomics to proteomics. *Trends Genet.* **16**, 5-8 (2000).

205. Yates, J. R.,3rd, Carmack, E., Hays, L., Link, A. J. & Eng, J. K. Automated protein identification using microcolumn liquid chromatography-tandem mass spectrometry. *Methods Mol. Biol.* **112**, 553-569 (1999).

206. Yates, J. R.,3rd, Eng, J. K., McCormack, A. L. & Schieltz, D. Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal. Chem.* **67**, 1426-1436 (1995).

207. Yates, J. R.,3rd, Morgan, S. F., Gatlin, C. L., Griffin, P. R. & Eng, J. K. Method to compare collision-induced dissociation spectra of peptides: potential for library searching and subtractive analysis. *Anal. Chem.* **70**, 3557-3565 (1998).

208. Yen, A., Forbes, M., DeGala, G. & Fishbaugh, J. Control of HL-60 cell differentiation lineage specificity, a late event occurring after precommitment. *Cancer Res.* **47**, 129-134 (1987).

209. Yen, A. & Forbes, M. E. c-myc down regulation and precommitment in HL-60 cells due to bromodeoxyuridine. *Cancer Res.* **50**, 1411-1420 (1990).

210. Yokota, T. *et al.* Down regulation of DJ-1 enhances cell death by oxidative stress, ER stress, and proteasome inhibition. *Biochem. Biophys. Res. Commun.* **312**, 1342-1348 (2003).

211. Yoon, K. A., Nakamura, Y. & Arakawa, H. Identification of ALDH4 as a p53-inducible gene and its protective role in cellular stresses. *J. Hum. Genet.* **49**, 134-140 (2004).

212. Yoshikawa, K. Cell cycle regulators in neural stem cells and postmitotic neurons. *Neurosci. Res.* **37**, 1-14 (2000).

213. Yu, S. W. *et al.* Mediation of poly(ADP-ribose) polymerase-1-dependent cell death by apoptosis-inducing factor. *Science* **297**, 259-263 (2002).

214. Zacharius, R. M., Zell, T. E., Morrison, J. H. & Woodlock, J. J. Glycoprotein staining following electrophoresis on acrylamide gels. *Anal. Biochem.* **30**, 148-152 (1969).

215. Zamir, E. *et al.* Dynamics and segregation of cell-matrix adhesions in cultured fibroblasts. *Nat. Cell Biol.* **2**, 191-196 (2000).

216. Zargham, R. & Thibault, G. Alpha 8 integrin expression is required for maintenance of the smooth muscle cell differentiated phenotype. *Cardiovasc. Res.* **71**, 170-178 (2006).

217. Zhang, L. J. *et al.* Proteomic analysis of low-abundant integral plasma membrane proteins based on gels. *Cell Mol. Life Sci.* **63**, 1790-1804 (2006).

218. Zhang, S. *et al.* c-Jun N-terminal kinase mediates hydrogen peroxide-induced cell death via sustained poly(ADP-ribose) polymerase-1 activation. *Cell Death Differ.* (2007).

219. Zhang, W. & Chait, B. T. ProFound: an expert system for protein identification using mass spectrometric peptide mapping information. *Anal. Chem.* **72**, 2482-2489 (2000).

220. Zhou, G. *et al.* 2D differential in-gel electrophoresis for the identification of esophageal scans cell cancer-specific protein markers. *Mol. Cell. Proteomics* **1**, 117-124 (2002).

221. Zhou, M. *et al.* An investigation into the human serum "interactome". *Electrophoresis* **25**, 1289-1298 (2004).

222. Zhukov, N. A., Shirinskaia, N. V., Dolgikh, T. I. & Akhmedov, V. A. Dynamics of expression of cytokines and lactoferrin in patients with chronic alcohol pancreatitis and chronic relapsing pancreatitis. *Eksp. Klin. Gastroenterol.* **(5)**, 67-71, 194 (2003).