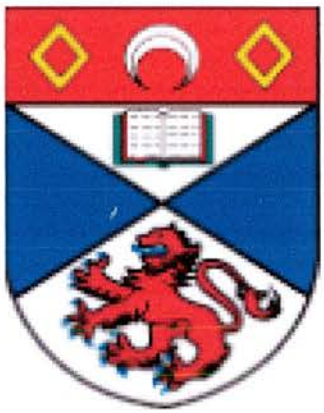
The Role of Acid Labile Surfactant-I in Protein Extraction from Sodium Dodecyl Sulfate Polyacrylamide Gels



University of St. Andrews

A thesis presented for the degree of
Master of Philosophy
to the
University of St. Andrews
submitted 13th September 2005

By
Carolyn Rice Buckner

I, Carolyn Buckner, hereby certify that this thesis, which is approximately 23,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Date 16.Nov.2005 Signature of candidate

I was admitted as a research student in September 2004 and as well as a candidate for the degree of Master of Philosophy; the higher study for which this is a record was carried out in the University of St. Andrews between September 2004 and September 2005.

Date 16Nov2005 Signature of candidate

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Master of Philosophy in the University of St. Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 16th Nov 2005 Signature of supervisor ....

In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being force, subject to any copyright vested in the work and being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to a bona fide library or research worker.

Date 16Nov2005 Signature of candidate

Acknowledgements

I would like to thank Dr. Catherine Botting for her help, support and guidance throughout this research endeavour.

I would like to thank Robin Antrobus for his help and patience in the initial stages of this research, Alex Houston for his help and encouragement and Sally Shirran for her knowledge and support.

Many thanks to Professor D. O' Hagan and Professor R. T. Hay for their kind gifts of the proteins fluorinase, SUMO-1, NFKB p50, and p65. Thanks also to Dr. N. P. Botting for the kind gift of the protein QPRTase and the use of his lab facilities for the synthesis of the compounds involved in this research.

I would also like to express my gratitude for the help provided to me by the technical staff of the School of Chemistry and the Biomedical Sciences Building of the University of St. Andrews. Thanks to Melanja Smith (NMR), Marjory Parker (Stores), Colin Millar (Stores), Ellis Jaffray (BMS technician), and ??? (BMS stores).

Finally I would like to thank my family for providing me with encouragement and support and Daniel, without whom this would never have been possible.

Abstract

The surfactant known as acid labile surfactant I (ALS-I) (RapiGest) was investigated to probe its effectiveness in enhancing protein digestion and subsequent protein identification and to determine its suitability for use in whole protein extraction methods. Tryptic solution and in-gel digests of several standard proteins showed a distinct increase in the number of peptides matched in ALS-I aided digests versus digests performed without ALS-I. Several experiments gave no identification unless ALS-I was present.

The ability to analyse whole protein samples with detergent or extracted from gels has eluded scientists and slowed certain areas of research such as the investigation and analysis of membrane proteins. ALS-I was ascertained to be compatible with and enhance whole protein extractions from sodium dodecyl sulfate-polyacrylamide gels (SDS-PA). Extraction was optimized by testing a series of extraction solutions containing ALS-I, DTT and tris/glycine. Extraction efficiency was shown to vary with incubation time and larger proteins proved harder to extract from a gel than smaller proteins. Subsequent analysis of protein extracts was performed by matrix-assisted laser desorption/ionisation mass spectrometry and liquid chromatography-electrospray ionisation mass spectrometry.

In order to analyse ALS-I extracted protein by MALDI-TOF MS it was necessary to develop a novel slow crystallisation using neutral matrix systems and sample preparation techniques. The neutral matrices 2-amino-4-methyl-5-nitropyridine, 4-nitroaniline and 6-aza-2-thiothymine gave good whole protein signals at M+H and M+2H.

Whole protein extracts from SDS-polyacrylamide gels were also analysed by LC-ESI MS, giving good spectra for a range of proteins from superoxide dismutase (SOD) to bovine serum albumin (BSA). Despite small mass inaccuracies, the novel possibility of analysing whole protein extracts from SDS-polyacrylamide gels containing detergent and sufficient protein to give clear protein spectra by LC-ESI MS is very promising, and progress in this technique may prove invaluable to further research into the proteome.

Contents

4. Conclusions and Further Work

Contents of Figures, Tables, Schemes and Spectra

1. Introduction

2. Materials and Methods

3. Results and Discussion

List of Abbreviations

ALS-I – Acid Labile Surfactant I

SDS-PAGE – Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (1D and 2D)

DTT – Dithiothreitol

MALDI-TOF MS – Matrix Assisted Laser/Desorption Ionisation-Time of Flight Mass Spectrometry

ESI MS – Electrospray Ionisation Mass Spectrometry

AMNP – 2-amino-4-methyl-5-nitropyridine

NA – 4-nitroaniline

ATT – 6-aza-2-thiothymine

SOD – Superoxide Dismutase

BSA – Bovine Serum Albumin

CBB – Coomassie Brilliant Blue

RuBPS – Ruthenium II Bathophenanthroline Disulfonate Chelate

HPLC – High Performance Liquid Chromatography

MudPIT – Multidimensional Protein Identification Technology

cIEF-MS – Capillary Isoelectric Focusing Mass Spectrometry

PMF – Peptide Mass Fingerprinting

MS/MS – Tandem Mass Spectrometry

PSD – Post Source Decay

HCCA - α-cyano-4-hydroxycinnamic acid

SA – Sinapinic Acid

DHB – 2,5-dihydroxybenzoic acid

CID – Collision Induced Dissociation

FT-ICR MS – Fourier Transform – Ion Cyclotron Resonance Mass Spectrometry

CMC – Critical Micelle Concentration

ALS-PAGE – Acid Labile Surfactant – Polyacrylamide Gel Electrophoresis

FWI – Formic acid/Water/Isopropanol (1:3:2)

FAPW – Formic acid/Acetonitrile/Isopropanol/Water

LG-B - β-lactoglobulin B

CCD – Charge Coupled Device

IAA – Iodoacetamide

NEIAA – *N*-Ethyl Iodoacetamide

N-d5-EIAA – *N*-[$^2H_5$]-ethyl iodoacetamide

CA – Carbonic Anhydrase

ADH – Alcohol Dehydrogenase

β-gal - β-galactosidase

MOWSE – Molecular Weight Search

1. Introduction

1.1 Proteomics

The proteome can be defined as all the proteins that are expressed in a cell at any one time, including isoforms and modifications. The identification of proteins and their relative quantitation between two cell states are the two main objectives of proteomics. However one genome can code for several different proteomes as the protein profile changes over time or with conditions. It is therefore the proteome that dictates phenotype. Proteomics can better be described as the study of all proteins expressed by a given cellular state (1).

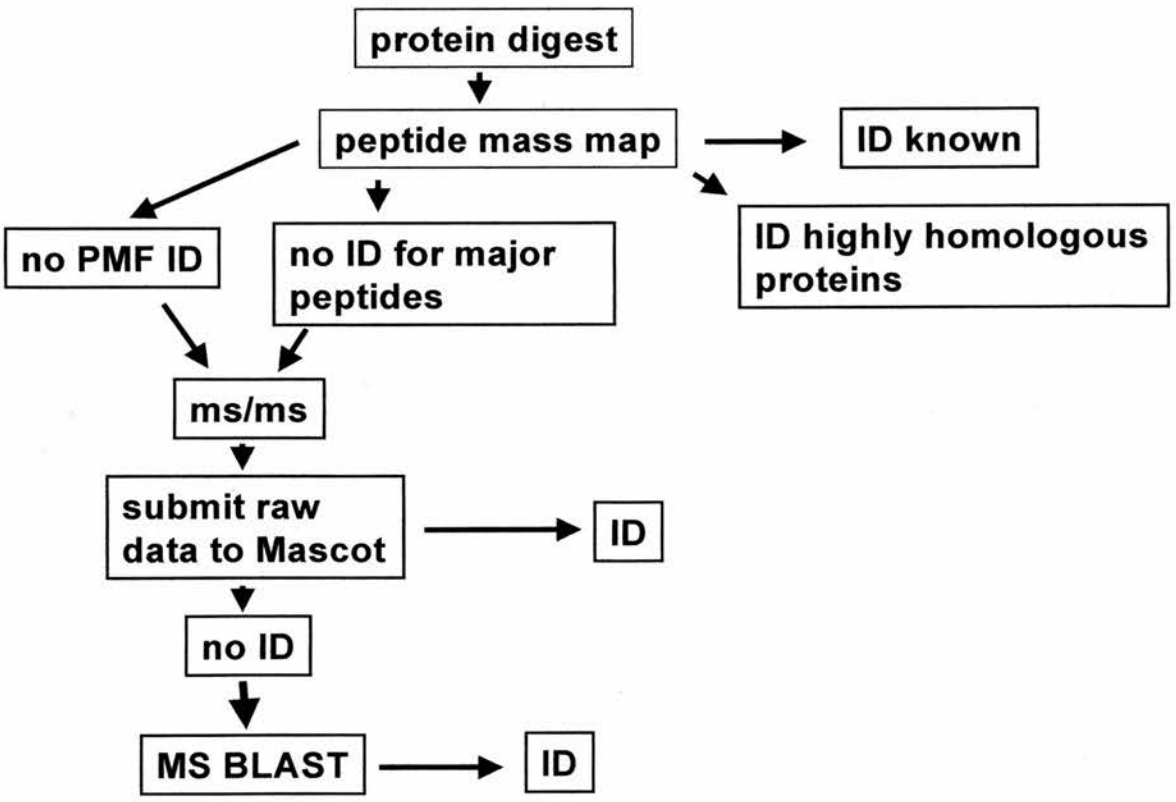
Protein identifications are normally made using mass spectrometry as described later. Mass spectrometry is therefore a vital tool in proteomics research. This method of identification of proteins is crucial to research in many areas of biochemistry. Proteomics focuses on determining which proteins are characteristic of a certain state and determining the differences between states. For example studies have been carried out on the differences in proteomes from normal and cancerous cells in search of a specific protein or set of proteins that are phenotypic of the cancerous cells (2).

A comparative study like this will involve both qualitative and quantitative analyses. Quantitation may be either absolute or relative. Absolute quantitation is performed by introducing a labelled peptide analogue of the peptide being measured into the sample mixture (3). If the only difference between the two peptides is their isotopic composition they can be expected to ionize equally in the mass spectrometer and therefore peak intensity can be compared. Relative quantitation is very important in determining the relative abundance of proteins in different states of a similar

biological system (2). This can either be performed by spot intensity measurements on a 2D gel or by isotopic labelling methods (4, 5). Information derived from proteomic analyses can be used in research to identify the specific expression patterns of proteins from pathological cells such as cancer cells or in biochemical studies to aid in identifying drug-binding proteins (6).

Many areas of science must cooperate very closely and make use of several different techniques to quickly and accurately deduce the identity of the proteins in question. One and two-dimensional gel electrophoresis and other methods of separation together with mass spectrometry and the advances made everyday in bioinformatics have improved the accuracy of identification and time it takes for proteins to be identified.

Figure 1: *A method map of the natural progression of techniques and technologies used in the Proteomics field, specifically depicting the steps taken to identify a protein using techniques including peptide mass fingerprinting, tandem mass spectrometry and protein databases such as Mascot.*



1.2 Protein Separation

1.2.1 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

One-dimensional Sodium Dodecyl Sulphate polyacrylamide gel electrophoresis (1D SDS-PAGE) is a widely used method of protein separation. This method separates proteins by mass, utilising the anionic detergent SDS (7). The protein solutions are loaded on to a polyacrylamide gel under denaturing conditions in a solution of SDS. The SDS binds stoichiometrically to the protein, forming a complex of around one SDS molecule to two amino acids. This gives the protein linearity and a uniform net negative charge directly proportional to the mass of the protein. Electric current is passed through an outer bath of running buffer giving it a positive charge. This buffer contacts the bottom of the gels. The inner bath, which contacts the tops of the gels, has a negative charge running through it. Proteins are electrophoresed from top to bottom with the negatively charged proteins attracted to the positively charged bottom of the gels.

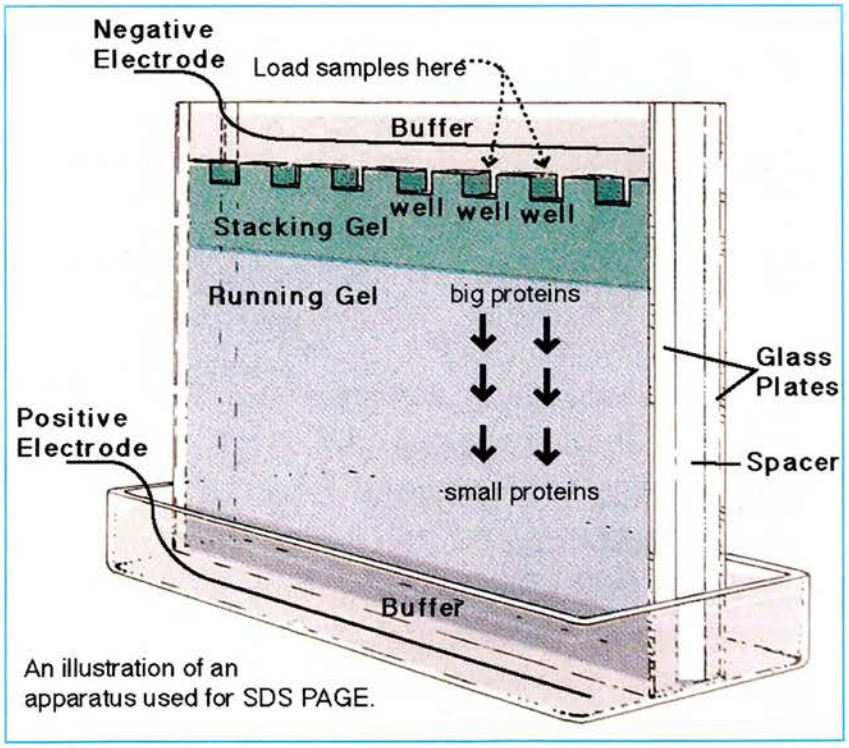


Figure 2: *Diagram of 1D Sodium Dodecylsulfate-Polyacrylamide Gel Electrophoresis*

Migration of the protein through the gel is subject to the viscosity of the gel and thus there is an inverse relationship between the molecular weight of the protein and its movement through the gel, with longer proteins unable to travel as fast through the gel. Gels are stained in order to visualize the protein bands. Stains including Coomassie Brilliant Blue R-250 and G-250, Silver stain, Zinc-Imidazole reverse stain, Sypro Ruby and other fluorescent stains are among the stains commonly used to visualize proteins. In a proteomics experiment the protein-containing bands are then excised for further analysis such as tryptic digestion and identification. One-dimensional PAGE separates proteins solely by mass and it does not have a high resolving power thus cannot separate a mixture of proteins if they are of approximately the same molecular weight (8, 9).

1.2.1.1 Visualization of SDS-Polyacrylamide Protein Bands

The most commonly used stain in SDS-PAGE is Coomassie Brilliant Blue (CBB), CBB R-250, which has a reddish hue, and CBB G-250, which has a greenish hue. CBB stains by binding to basic amino acids through an ionic interaction (10). An acidic environment is necessary to enhance this interaction. The more popular of the two types is CBB R-250. However, the binding interaction of CBB R-250 causes difficulties in mass spectrometric analysis due to adduct formation. CBB G-250 may be used on a colloidal form. Colloidal particles are unable to penetrate the gel, but protein in the gel extracts the dye particles from the colloid. This gives less background staining and is more sensitive, but longer staining times are required. It is, however more mass spectrometry compatible.

Another popular staining method is silver staining. This staining method is the most sensitive and as such is used for detecting small amounts of proteins. Silver staining is based on the affinity of silver ions for nucleophilic and aromatic groups (11). The gel is exposed to silver ions, which bind to the proteinaceous areas of the gel. Excess silver ions are washed away and the silver ions are then reduced to metallic silver and the protein bands become visible. This stain must be used with polyacrylamide gels as it produces too much background in other gels. While silver stain is the most sensitive staining method available, revealing protein bands containing as little protein as 1-3 ng, it is work intensive and can be highly incompatible with mass spectrometric analyses (12). Most silver staining techniques involve fixing the protein within the gel, which calls for the use of gluteraldehyde. Aldehydes are known to modify lysines, which will prevent subsequent trypsin digestion. This reduces the efficiency of peptide extraction. Schevchenko *et al.* published the main protocol for identifying proteins from silver-stained polyacrylamide gels by mass spectrometry (13). However not all the protein bands identified by silver staining will contain enough protein for mass spectrometric analyses.

Negative staining, or reverse staining, techniques are very useful for SDS-PAGE, particularly if the gel band is to be subsequently analysed by procedures such as biological activity tests, C-terminal analysis with exopeptidases and N-terminal deblocking. In negative staining protein bands remain unchanged in colour and the background of the gel becomes opaque and/or coloured allowing protein bands to be distinguished from the rest of the gel. Two examples of negative stain are Zinc and Copper imidazole complexes (14, 15). They work by depositing a metal precipitate in the gel, which is unable to bind to proteins due to their SDS coating. These stains are

useful as they do not fix the protein in the gel and thus are ideal for protein bands, which will be subjected to further analysis by mass spectrometry. However reverse stains show a lesser degree of sensitivity than silver staining meaning a larger amount of protein is required in order to visualize the protein band. Reverse stains are at least as sensitive as Coomassie Brilliant Blue.

A newer class of stain is comprised of fluorescent stains such as Sypro Orange, Red, Ruby and Tangerine, Nile Red, RuBPS and Deep Purple. Fluorescent stains have a high level of sensitivity and are known not to modify proteins covalently (16). Thus they can be used in conjunction with in-gel digestion and further analysis by mass spectrometry. Fluorescent dyes also produce a more linear gradient of band colouring over a broad range of protein concentrations than silver staining. Nile Red is a non-covalent dye that is useful as its staining times are as low as a few minutes and can be used for microsequencing or immunodetection. However it is less sensitive than reverse stains such as copper and zinc. Sypro Orange and Red were designed as quick staining, highly sensitive alternatives to silver staining with detection limits between 2.5 ng and 5 ng of protein (17). Staining times are less than one hour but they show reduced sensitivity when used to stain 2D gels (18). Sypro Ruby is known to be the most sensitive of the Sypro series. It effectively stains lipo- and glycoproteins as well and does not stain nucleic acids. One drawback to the Sypro Ruby stain is it requires twelve hours staining time versus the hour of Sypro Red and Orange. RuBPS is a metal chelate that can be easily excited by UV and visible light. It has strong similarities to Sypro Ruby but has been shown to have slightly less sensitivity than Sypro Ruby (19). Deep Purple has been identified as the most sensitive of the

fluorescent stains (20, 21). Deep Purple staining of polyacrylamide gels results in dark bands and low-level background staining.

1.2.2 Two Dimensional Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

Another commonly practised method of separation is two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) (22). This technique is a two-step process involving first the separation of a protein mixture using isoelectric focusing to electrophorese the proteins along an immobilised pH gradient until they reach their pI, thus separating them in one dimension. This is followed by a separation on a SDS polyacrylamide gel as described above. The protein is separated in the first dimension by charge and then in the second dimension by molecular weight; hence this is a two-dimensional analysis (5). As a result, proteins appear on the gel as spots. As with one-dimensional SDS-PAGE these spots can then be excised for further analysis.

Currently the resolving powers of 2D gels are in the region of 1000 proteins (23). However, these separation methods can be time consuming, highly expensive to automate and have a low throughput in comparison with other methods. One and 2D SDS-PAGE also allow for some quantitation by measuring the stain density of bands or spots on a gel. However, one spot can contain several proteins thus making their quantitation unreliable (24, 25). 2D-PAGE systems are also known to function poorly with high molecular weight proteins, low molecular weight proteins, basic proteins and hydrophobic proteins leaving a relevant mass range of 20kDa-100kDa (22).

1.2.3 High Performance Liquid Chromatography

Liquid chromatography is a very important analytical technique and is an alternative method of separating peptides and proteins (26). It involves bonding and non-bonding interactions between analyte molecules and a mobile phase and a stationary phase. Liquid chromatography utilizes a liquid mobile phase to separate the analytes.

The analytes are dissolved in a solvent and passed under high pressure through a chromatographic column. The mixture of analytes is resolved into separate components by the column. The extent of the resolution depends on the binding interactions between the analytes and the stationary phase, which is packed tightly within the column. HPLC is widely used due to its overall versatility. Many types of compound can be analysed due to the assortment of solvent systems and stationary phases.

1.2.3.1 Ion-Exchange Chromatography

Ion-Exchange Chromatography is a method of HPLC commonly used for purifying proteins (7). In IEC proteins are separated by their net charge. Therefore a protein with a positive net charge will bind to a column containing negatively charged functional groups such as carboxylate groups. Once a mixture of proteins is bound, by slowly increasing the salt concentration of the solvent, proteins will be competed off the stationary phase depending on the strength of their net charge. Thus a protein with fewer positively charged groups will be eluted off the negatively charged column before a protein containing a high number of positive charges. While this is a very good method for purifying and separating proteins, it is incompatible with mass spectrometry due to the salt content of the solvent systems.

1.2.3.2 Reverse Phase Chromatography

One common form of liquid chromatography is reverse phase chromatography (26). This type of chromatography employs a stationary phase that has strong hydrophobic binding interactions. Typically the mixture of analytes is introduced to the column in a high aqueous and low organic solvent system. Once the analytes have bound to the column the organic solvent content is gradually increased to give a gradient elution. Each component elutes off the column as the interaction with the stationary phase is disrupted by the organic solvent.

1.2.3.3 Multidimensional Protein Identification Technology

An alternative separation method to SDS-PAGE is multidimensional protein identification technology (MudPIT) originally described by Washburn *et al.* (27). In MudPIT the separations are carried out using two-dimensional liquid chromatography. The protein mixture is first digested with trypsin and then the separation is performed at the peptide level. The peptides are separated first by cation exchange chromatography and then by reverse phase chromatography. Cation exchange chromatography consists of binding a positively charged sample to a negatively charged immobile phase, washing away any impurities and subsequently eluting the charged peptides of the immobile phase with a salt containing solvent. The sample collection is interfaced with the ion source of a mass spectrometer, utilizing tandem, or MS/MS, mass spectrometry techniques.

1.2.4 Capillary Isoelectric Focusing

Another type of separation technique, which can be linked to mass spectrometry, is Capillary Isoelectric Focusing or cIEF-MS (28). This technique is based on the same principal as isoelectric focusing in 2D gels. Here different zwitterions, which contain both acidic and basic groups, have different isoelectric points. The isoelectric point is the point at which the charges of the different groups are balanced. Thus in cIEF the capillary is filled with a solution of ampholytes with a broad range of pIs. An electric current is applied to the capillary creating a pH gradient. Proteins in this environment will migrate to a region where they are uncharged. This is known as focusing. Proteins that migrate out of that region become charged again and return to the region where they are uncharged. The analytes are observed by passing the focused zones through a detector such as a mass spectrometer (29).

1.3 Protein Digestion

Up until now mass spectrometry has struggled to identify a protein through the analysis of a whole protein. Recently a method called 'top down' proteomics has been demonstrated which attempts to identify proteins beginning with analysis of the whole protein (30, 31). Below is a diagram of the process involved in a typical top down characterization of a protein.

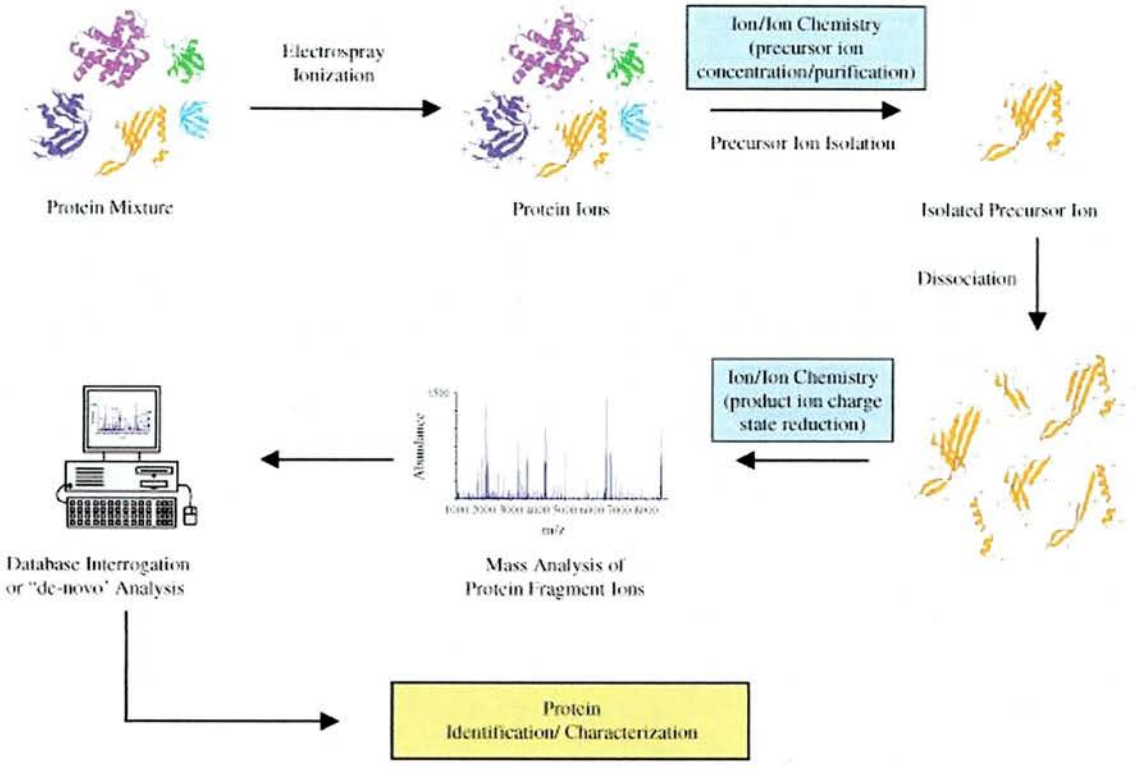


Figure 3: *Schematic diagram of the top down characterization of a protein*

Here the whole protein is fragmented by the mass spectrometer. This has been successful with smaller proteins ranging in size from 7 to 36 kDa. However, it is more common to digest whole proteins into peptide fragments (32). This is conversely known as 'bottom up' protein identification. The peptides are then analysed by mass spectrometry. Current technologies in mass spectrometers allow for molecular weight determination with an accuracy of better than 0.1 mass units on peptides that have a molecular weight of approximately 800 to 2500 mass units.

Protein digestion protocols involve the reduction and alkylation of cysteines to prevent disulphide bond cross-linked peptides prior to subsequent digestion of a protein or mixture of proteins. The digestion is typically performed with the protease, trypsin, which cleaves the protein at the carboxy terminal side of arginine and lysine

residues, unless either is followed by proline, yielding a set of peptide masses that is exclusive to that protein and therefore acts as a fingerprint for that protein (32). Other proteases used in protein digestion include Lys-C, Glu-C, and Arg-C, which cleave at the C-terminus of their respective amino acids and Asp-N, which cleaves at the N-terminus of aspartic acid. Once analysed by mass spectrometry, the peptide mass list or peptide mass fingerprint (PMF) can be compared to a database of proteins for which predictive digests have been created and the identity of the protein determined based on the number of matching peptide molecular weights (33). Peptide sequencing by tandem mass spectrometry (MS/MS) (see later) can also be performed. Having some sequence data is even more specific to an individual protein and is particularly useful when analysing a mixture of several proteins.

1.4 Mass Spectrometry

With the upsurge in biochemical research that has been seen in recent years, the area of mass spectrometry has become an important tool in the fast and accurate determination of molecular weights of peptides and proteins up to approximately 300kDa (34). All mass spectrometers follow a similar pattern in their procedures. First a sample is ionized by a variety of methods depending on the type and physical state of the sample. During ionisation the sample is transferred to the gas phase. The ions are then analysed by several different mass analyzers including time of flight analyzers.

The techniques that have become increasingly important and allowed proteins and peptides to be studied by mass spectrometry are matrix assisted laser desorption/ionisation-time of flight mass spectrometry (MALDI-TOF MS) and

Electrospray Ionisation (ESI) mass spectrometry (32). These are soft ionisation methods. ESI-MS can be easily coupled with HPLC. Both techniques are suitable for analysing simple mixtures of proteins, however ESI is more suited to analysing digests of mixtures of proteins than MALDI-TOF MS. MALDI-TOF MS is a fast method of analysing digests of single proteins. Tandem mass spectrometry is a powerful mass spectrometric technique, which, due to the sequencing data acquired, allows for much greater specificity in determining the identity of proteins and mapping modifications. These mass spectrometry methods, in combination with current technologies, that include growing databases of proteins from genome sequencing projects, result in mass spectrometry being a very powerful tool in protein identification.

1.4.1 MALDI-TOF Mass Spectrometry

Matrix assisted laser desorption/ionisation-time of flight mass spectrometry is a powerful tool in the analysis of proteins and identification of a protein from peptides derived from a protease-digested protein. Developed in 1988 by Karas and Hillenkamp, it involves the use of a powerfully UV absorbent solid matrix with which the sample is co-crystallised (35).

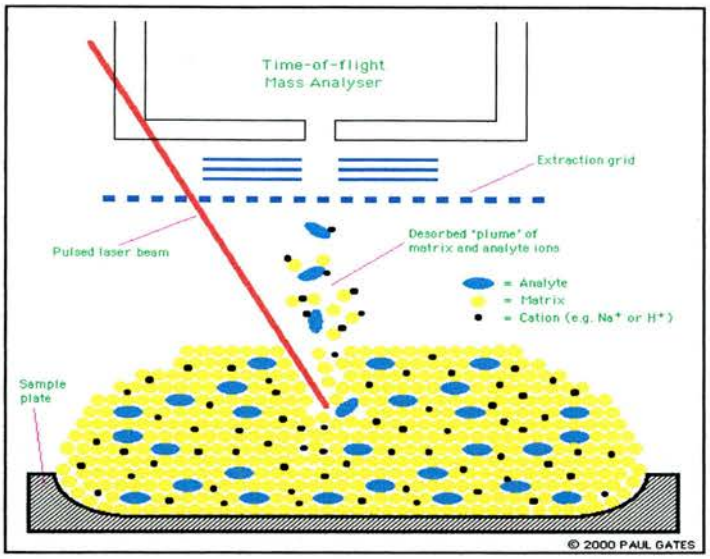


Figure 4: *Pictorial depiction of MALDI-TOF ionization and analysis*

The matrix readily vapourises on impact with UV laser radiation, carrying some of the sample with it. MALDI is considered a soft ionisation technique with little or no fragmentation of sample observed.

This ionization technique, along with the use of TOF mass spectrometry, has the sensitivity to produce molecular ions for low femtomole to low picomole concentrations as well as having a practical mass range of up to 300 kDa (34). The number of matrix molecules is far greater than the number of sample molecules thus ensuring sample separation in the co-crystal and preventing the formation of sample clusters. MALDI-TOF MS can also tolerate millimolar salt concentrations and detergent.

There are many ways of creating this matrix/analyte co-crystal mixture, including the commonly used dried-drop method and slow crystallisation (36). Samples, for example, of digested protein are co-crystallised with a matrix and applied to the sample probe, typically a 100 spot stainless steel plate. The sample plate is inserted

into the mass spectrometer and subjected to high voltage (up to 30 kV) under vacuum. The matrix/analyte co-crystals absorb radiation from a pulsed laser beam, typically a N_2 337 nm laser, causing the matrix material to desorb carrying some of the digested protein sample with it. The pulsed laser beam causes the analyte/matrix co-crystals to be excited leading to the formation of cations by protonation or cation attachment in the gas phase. The matrix also absorbs much of the resulting energy causing less damage to the sample.

Despite low amounts of fragmentation due to the matrix absorbing the majority of the laser energy, the analyte does absorb a small amount of laser energy; enough to cause a degree of fragmentation. There are three types of fragmentation seen in MALDI-TOF MS (29). They are prompt fragmentation, which occurs on the sample surface, fast fragmentation, which occurs in the source but before acceleration, and post-source decay (PSD), which occurs after acceleration. Prompt and fast fragments will always be seen in spectra but it is more difficult to detect PSD fragments. All fragmentation leads to the broadening of peaks and subsequent decrease in mass accuracy, as the fragment ion generally cannot be separated from the parent ion.

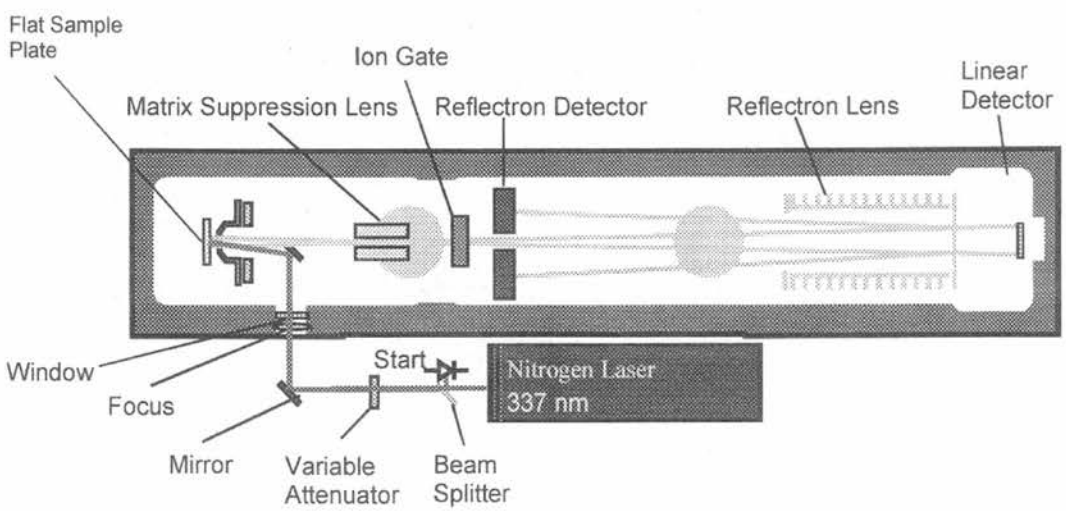
Once the analyte/matrix co-crystals are in the gas phase the matrix may play a role in ionising the analyte molecules. The charged molecules are then accelerated towards the mass analyser, directed by electrostatic lenses away from the ionisation source. The ions are directed in to a field-free drift region, separated from matrix ions that are prevented from entering this region by a matrix suppression lens. Linear Time-of-flight analysis is used to determine the m/z ratio from the ion's time of arrival at the detector. The arrival time of an ion is dependent on its mass, charge and kinetic

energy. As the voltage applied to each ion is the same, the velocity of each ion is proportional to its mass. Therefore ions with the same mass to charge ratio have the same velocity. The distance each ion travels to the detector is also the same, thus the time each ion takes to travel to the detector is directly related to the mass of the ion. It can be seen from the equation below that as the distance and voltage are the constant in any given experiment, the lower the mass of the ion the shorter the amount of time it will take to reach the detector.

$$t^2 = m/z\ (d^2/2V_s e) \tag{29}$$

Where t is the time travelled to the detector, m/z is the mass to charge ratio of a specific ion, d is the distance travelled to the detector, V_s is the accelerated potential of the ion and e is the charge of an electron. This type of mass analyser has low resolving power, but no practical upper mass limitation. Mass resolution is poorer when there is a wide distribution of flight times for ions with the same mass to charge ratio due to variations in their kinetic energy.

Figure 5: *Schematic diagram of a MALDI – TOF mass spectrometer, TofSpec 2E (Micromass, Manchester, UK)*



In MALDI-TOF MS analysis of peptides a reflectron assembly is used to improve mass resolution as shown in the diagram above. The reflectron lens is located at the end of the drift-free region and acts to slow the generated molecular ions to a stop and accelerates them back towards the reflectron detector. Molecular ions with high kinetic energy penetrate the reflectron more deeply than those with low kinetic energy. The reflectron therefore corrects any difference in the arrival times of ions. This allows ions to be focused, in both time and space, on to the reflectron detector to improve mass resolution. While the reflectron is a good addition to the MALDI-TOF mass spectrometer it does however have a decreased sensitivity and a limited mass range (29, 34).

Due to the limited mass range of the reflectron assembly, MALDI-TOF MS data of whole proteins is acquired in linear mode, allowing a mass range of up to 300 kDa.

1.4.1.1 Sample Preparation for MALDI-TOF Analysis

One major difference between other mass spectrometric techniques and MALDI-TOF MS is the method of sample introduction. As described above, MALDI data is acquired from solid matrix/analyte co-crystals. Optimisation of sample preparation and choice of matrix are the most important steps in MALDI-TOF MS analysis. There are several methods of crystallisation used in conjunction with several different matrices. The sample being analysed and the type of analysis being performed determines the matrix that is used.

There are many criteria to consider before choosing an appropriate matrix (29). Firstly, the wavelength of absorption for the matrix must overlap the wavelength of

the laser. This can be difficult to determine, as it is difficult to derive the absorption coefficient of a solid. Secondly, the matrix cannot be innately reactive in structure. Proteins are very susceptible to covalent bonding, oxidation, and other such reactions. Thus compounds that might promote these reactions are unsuitable for use as matrices. Matrices also have to be able to withstand a considerable amount of laser energy without becoming photochemically unstable. Any instability might result in reactive chemical fragments, which could result in the formation of adducts. Most matrices in use still produce matrix adducts however the best matrices give low intensity adduct signals. Thirdly, and most importantly, the matrix must allow for the formation of matrix/analyte co-crystals. Thus when the co-crystals are fired upon both the matrix and the analyte are desorbed, allowing the analyte to be ionised. This is fundamental to the principal of MALDI MS analysis. There are other factors that must be assessed in order to fine tune matrix selection, but often these are experiment specific.

The matrix plays an important role by absorbing most of the secondary laser energy thus protecting the integrity of the sample. It also increases the effectiveness of the transfer of energy from the laser to the analyte; another reason why MALDI-TOF MS is a highly sensitive technique (29, 34, 37). MALDI-TOF MS can be used to analyse most samples at the same laser intensity due to the laser energy being absorbed by the matrix and not the analyte.

The most common method of crystallisation is the dried drop method (36). For this method a sample is prepared by mixing an aliquot of the analyte with an appropriate

amount of saturated matrix solution. A small amount of this analyte/matrix solution is loaded on to the target plate and allowed to dry for several minutes.

A variation of this method involves the direct application of the matrix solution to the target plate followed by the application of acid to promote crystallisation. The sample is loaded on top of this mixture and the plates are left to dry, allowing any solvent to evaporate (36).

Some samples can also be prepared for MALDI-TOF MS using slow crystallisation. This method, utilized by C. Botting (38) and first reported by Xiang and Beavis (39), has been shown to enhance MALDI-MS sensitivity of proteins presented in buffer constituents normally detrimental to mass spectrometric analysis, as is the case when analysing whole proteins eluted from gel slices (36). Slow crystallisation occurs when a saturated matrix solution is added to a solution of protein in an open tube and the mixture is left to stand at room temperature allowing some of the solvent to evaporate. Large matrix/analyte co-crystals are formed. The co-crystals are then washed in water or formic acid depending on the nature of the matrix to wash away salt contaminants. The crystals are re-suspended in a small amount of water or formic acid, again depending on the matrix. These crystals are loaded directly onto the target plate and analysed by MALDI-TOF MS in linear mode.

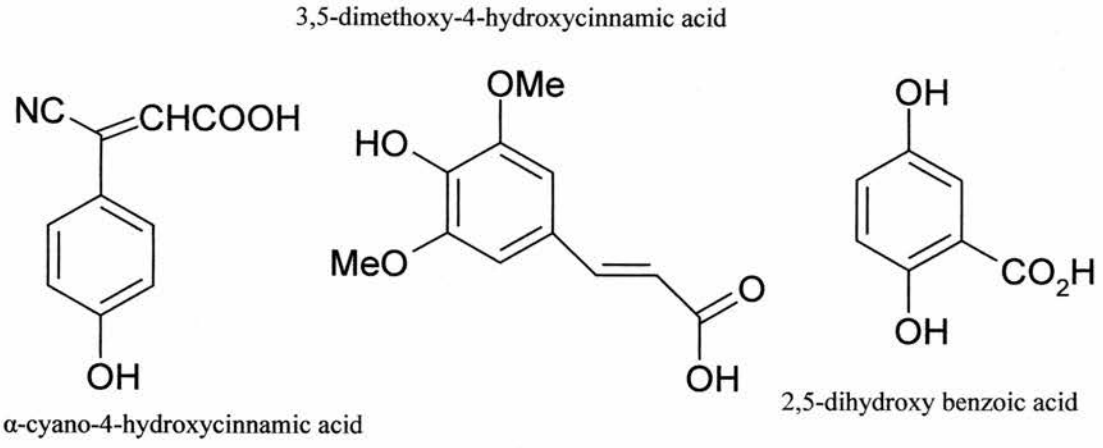


Figure 6: *Structures of Common Matrices*

There are several matrices currently in standard use. Each matrix fulfils criteria that make it suitable for certain experiments (40). The matrix α-cyano-4-hydroxycinnamic acid (HCCA) is commonly used to analyse peptides as it gives very intense signals in the 500-2500 Da mass range. Thus it is commonly utilised in the dried drop method and is the matrix of choice for peptide mass fingerprinting (37). It is also used to analyse proteins and is generally applicable in slow crystallisation. Another widely used matrix is trans-3,5-dimethoxy-4-hydroxycinnamic acid or Sinapinic acid (SA). It is useful for analysing proteins (10-150 kDa), including glycoproteins and membrane proteins. Sinapinic acid also has the advantage of being very tolerant of salts and other contaminants including lipids and carbohydrates. The third matrix that is commonly used in MALDI-TOF MS analysis is 2,5-dihydroxy benzoic acid (DHB), also known as gentisic acid. This matrix is useful for a wide range of peptides and proteins as well as polymers and organic compounds. It is particularly useful for peptides and small molecules in the 200-1000 Da mass range.

The above matrices are not applicable for samples that are acid sensitive, as they all require the addition of acid to promote crystallisation. This has led to research

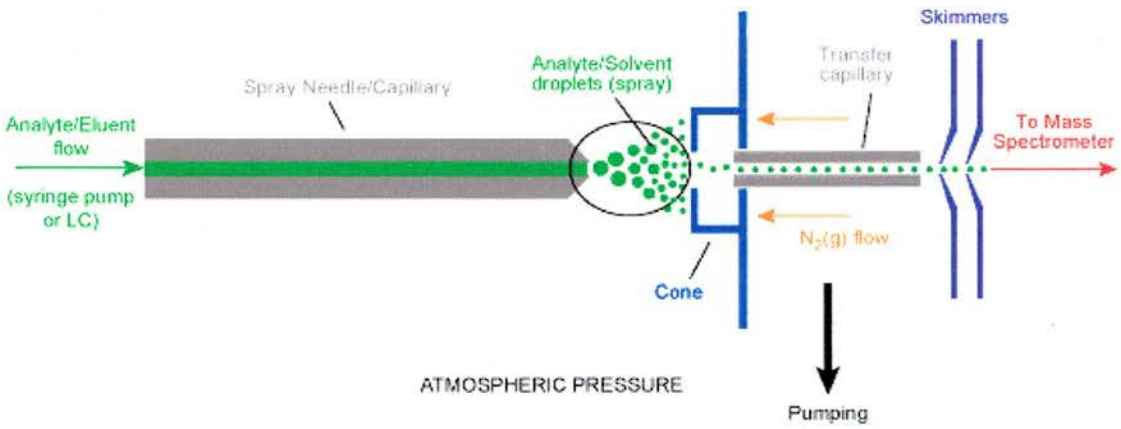
involving basic and neutral matrices, which allow for crystallisation at a neutral or basic pH. A comprehensive study of 37 different basic amino group-containing matrices was undertaken by Fitzgerald *et al.* (41). These matrices were tested for their compatibility with the analysis of whole proteins and oligonucleotides. Samples were prepared for analysis using the dried drop method. In this study the matrix 2-amino-4-methyl-5-nitropyridine gave good results for proteins up to 12,000 Da. The best matrix for analysing oligonucleotides smaller than 20 nucleotides in length was identified as 2-amino-5-nitropyridine. Another study performed on protein-protein interactions used the matrix 6-aza-2-thiothymine (ATT) to determine the effect of matrix concentration and pH on protein complexes (42). It was shown that dimers of leucine zipper polypeptides (m/z 6533) could be seen in MALDI-TOF MS spectra when analysed with ATT, but the dimers were not observed when analysed with HCCA. Up to now, data has not been acquired for the analysis of larger proteins in conjunction with neutral matrices.

1.4.2 ESI Mass Spectrometry

Electrospray Ionisation is a method of mass spectrometry that produces ions from a liquid solution (29, 32). J.B. Fenn *et al.* further developed this technique in order to ionise high mass biological molecules and subsequently gather meaningful mass spectrometric data (43). A flow of solvent and analyte passes through a narrow capillary to produce a fine spray at the capillary end. A voltage of 3000-4000 V is applied to the capillary. Due to the high charge on the droplets they are attracted to the mass spectrometer inlet. The droplets are subjected to dry heat or gas that evaporates the solvent from the surface before they enter the vacuum of the mass spectrometer. As the size of the ion decreases there is a concurrent increase in the charge density on

its surface. Like charges repel and when this repulsion overcomes the forces of surface tension, smaller droplets are formed through what is known as a Coulombic explosion. These smaller droplets undergo the same procedure until the ions are completely desolvated.

***Figure 7**: Schematic Representation of Electrospray Ionisation Mass Spectrometer (44)*



Molecules with nominal mass over 1000 generally have several sites available for ionization. Multiple ionizations result in multiply charged molecules and as all mass spectrometry measures the mass to charge ratio (m/z) it enables analysis of heavier molecules with a spectrometer capable only of analysing low mass to charge ratios. The consumption of material in this method of mass spectrometry necessary to obtain data is approximately 20 femtomoles (20×10^{-15} moles), which for a protein of molecular weight 10,000 Da is only 200 picograms of material (29), however ESI-MS is far less salt tolerant than MALDI-TOF MS.

1.4.3 MS/MS

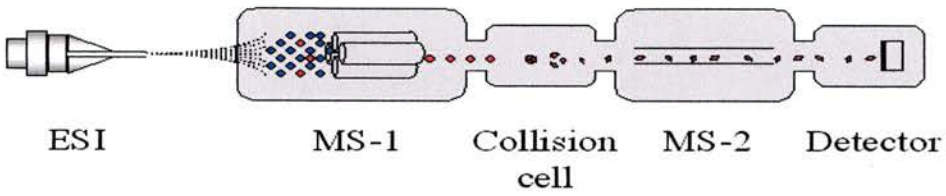


***Figure 8**: Schematic Representation of an LC-ESI MS/MS Mass Spectrometer*

Tandem mass spectrometry involves the further fragmentation of ions produced by MALDI or ESI. In peptide MS/MS fragmentation occurs by collision-induced dissociation (CID) (29, 32). Alternative methods of fragmentation exist in FT-ICR MS but they will not be discussed here. These fragmentations tend to occur at the peptide bonds thus giving sequence information on the peptide. This information is more specific than a PMF. Therefore MS/MS data for one peptide may be enough to allow an unequivocal identification. MS/MS data is very powerful, particularly when identifying proteins in mixtures.

Figure 9 shows the fragmentation of a peptide at the peptide bond giving rise to y and b ions. Figure 10 is a stylized spectrum depicting only y ions showing how a peptide sequence can be obtained by MS/MS analysis. The peptide sequence reads GTFAR.

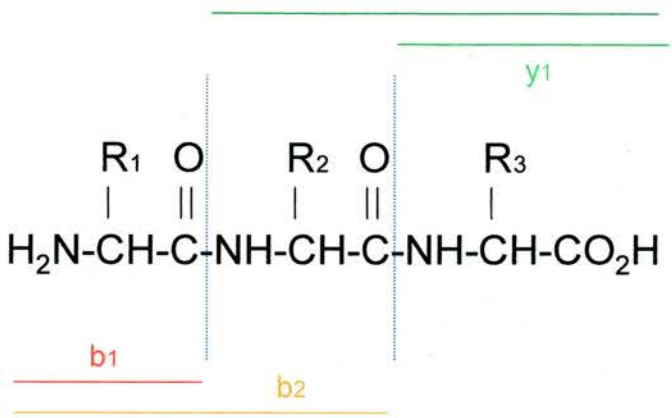
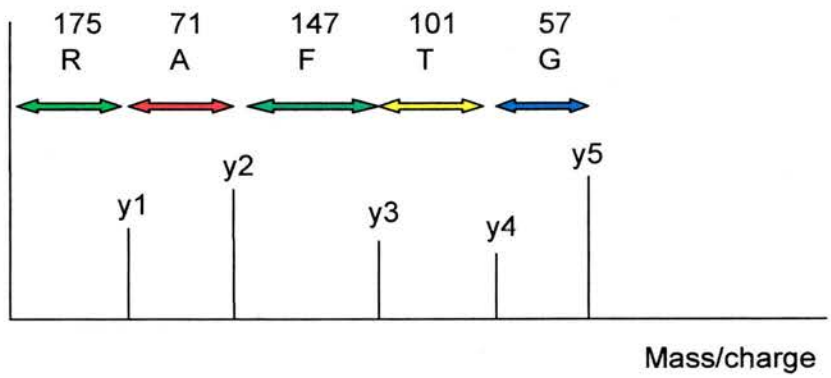


***Figure 9**: Peptide Backbone Fragmentation*

***Figure 10**: MS/MS Spectrum*



1.4.4 Peptide Mass Spectrometry

Bottom up protein characterization requires the analysis of peptides. Peptide analysis is performed by MALDI-TOF MS and latterly LC-ESI MS/MS. The peptides are formed by digesting the protein, either in solution or in-gel with a protease, typically trypsin. This produces peptides for which, if the protein sequence is known, masses can be predicted (32). This list of peptide masses is referred to as the peptide mass

fingerprint. MALDI-TOF MS is very useful for analysing peptide solutions and is also capable of analysing the mixtures of peptides obtained from a protein digest (1). MALDI analysis of peptides is limited to a mass range of 800-5000 Da. Matrix interference occurs below about 800 Da, and larger peptides tend not to ionize well in the MALDI-TOF MS instruments (34). MALDI-TOF MS instruments only provide peptide mass data, however recent developments have led to MALDI-TOF/TOF instruments, which also perform MS/MS experiments (45). An alternative is LC-ESI MS/MS analysis where the peptides are separated first by nano flow HPLC and then subjected to ESI-MS/MS. This provides detailed sequence information of peptides allowing for the correct identification of several proteins in a mixture. LC-ESI MS/MS also has a limited mass range of approximately 600-3000 Da as only doubly and triply charged peptides give easily interpretable MS/MS spectra. Molecules tend to pick up an extra charge for every 1000 Da under ESI conditions.

1.4.4.1 Data Analysis

A popular search engine for tryptic digestion analysis is Mascot (Matrix Science). For peptide mass fingerprinting a peptide peak list is submitted and this is compared to a predicted peak list for all the proteins in the database being searched e.g. the NCBI database (46). The probability of matches occurring randomly is calculated and hence the probability of an identification being valid is determined. The algorithm used in this procedure is based on molecular weight search (MOWSE). (47).

Similarly MS/MS spectra obtained from MALDI or ESI experiments are submitted to Mascot or similar software packages where they are fitted to predicted spectra and the probability of a match calculated.

1.4.5 Whole Protein Mass Spectrometry

It is possible to analyse whole proteins by both MALDI-TOF MS and LC-ESI MS. Whole protein samples can be prepared for MALDI-TOF MS using the same methods as protein digests, such as the dried drop method, but many of the methods are matrix specific. For example the dried drop method works well with whole proteins when the matrix sinapinic acid is used but not so well when performed with α-cyano-4-hydroxycinammic acid (37). However the dried drop method can be amended such that the HCCA matrix is heated to give a supersaturated solution before being added to the sample. The mixture of sample and matrix is added to the target plate and the spot scratched to encourage co-crystals to form. This method produces whole protein signals comparable to or better than those given by SA.

Analysis of whole proteins above approximately 7000 Da by MALDI-TOF MS requires data to be acquired in linear mode due to the mass limitations of the reflectron assembly (34). MALDI-TOF MS data for whole proteins generally displays broad peaks. This broadness is the result of the spread in kinetic energies displayed between ions with the same mass to charge ratio as described above (29). Peak broadness reduces the accuracy with which the mass can be determined. Despite this, MALDI-TOF MS analysis is ideal for analysing proteins that have been extracted from gels as it has a much higher tolerance than ESI for detergents and other contaminants, particularly when combined with slow crystallisation.

ESI analysis of whole proteins often requires considerable sample clean up such as reverse phase chromatography, to allow signals to be obtained. This is due to the poor salt tolerance of ESI-MS. ESI has low tolerance for detergents and other compounds

used in protein science (29). However, ESI analysis is an invaluable tool in determining the accurate mass of proteins. ESI-MS produces spectra with highly resolved peaks allowing for masses to be determined to within 1 or 2 Da. Matrix adducts are not a problem, as they can be in MALDI-TOF MS. ESI-MS is also useful for analysing whole proteins as it produces multiply charged analytes. Whole proteins become multiply charged under ESI conditions, thus presenting with mass to charge ratios in approximately the 600-2000 range. This means that an instrument with low mass capabilities can analyse samples with masses up to 70,000 Da and above. Furthermore these low mass to charge species can be focused by a reflectron thus giving improved resolution.

The mass of a multiply charged ion is calculated by first determining the charge series of the peaks. These charges can then be applied to an equation, which relates peak values (p_n) to the charge (z_1) and the average mass of the protein (M_r). Thus,

$$p = m/z \tag{29}$$

$$p_1 = (M_r + z_1)/z_1$$

$$p_2 = \{M_r + (z_1\text{-}1)\}/(z_1\text{-}1)$$

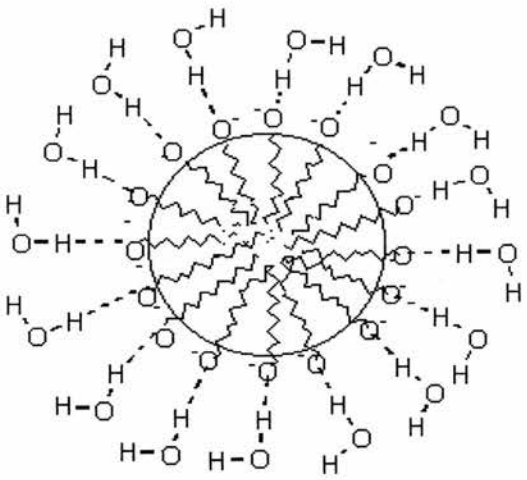
$$M_r = [(P_1\text{-}1)/(P_2\text{-}P_1)] * P_2\text{-}1$$

where m is the total mass of an ion, z is the total charge, $p_1 = m/z$ is the value for peak one, $p_2 = m/z$ is the value for the adjacent peak in the series of higher mass to charge ratios and z_1 is the charge on p_1.

1.5 Detergents

Detergents are used in many biochemical applications, primarily to facilitate the solubilization of hydrophobic compounds in water and disrupt hydrophobic interactions between molecules. As stated by Meng *et al.*, "the use of a strong, ionic detergent with polyacrylamide gel electrophoresis (i.e., SDS-PAGE) is critical for efficient solubilization and fractionation of proteins based primarily on molecular weight." (31) Detergents are amphipathic molecules meaning they have both a polar head and a non-polar tail, generally a hydrocarbon chain of varying length (48). The polar groups in detergents form hydrogen bonds with water molecules and the hydrophobic groups interact with more hydrophobic species. For these reasons, detergents are soluble in water, forming micelles in aqueous solution (49). A micelle is a cluster of several detergent molecules with the hydrophobic tails protected from the aqueous environment by a shield of polar head groups.

***Figure 11**: Diagram of a Soap Micelle*



A Soap Micelle

The detergent has to reach a certain concentration for this to occur, known as its critical micelle concentration (50). The average size of the micelle is defined by the

aggregation number, i.e. the number of detergent monomers that group together to form an individual micelle. This number and the critical micelle concentration are typically inversely related meaning the lower the critical micelle concentration the larger the average micelle size (51). The larger the micellar size the harder it is to remove by dialysis or other sample clean up methods. There are three standard types of detergents; ionic, non-ionic and zwitterionic, based on the nature of the hydrophilic head group (48).

1.5.1 Ionic Detergents

Ionic detergents are classed as such due to the net charge on the head group, either anionic or cationic (51). The common detergent sodium dodecyl sulfate (SDS) is an anionic detergent due to the negatively charged sulfate group. Most detergents are anionic with a negatively charged head group and the majority of these are alkyl sulfates. The detergent Acid Labile Surfactant-I is an anionic surfactant. Ionic detergents have either a hydrocarbon straight chain or a more complicated inflexible steroidal structure such as in sodium deoxycholate.

Cationic detergents contain a positively charged polar head group (51). Most are derivatives of ammonia and many have germicidal properties, which makes them useful cleansing agents in hospitals.

Bile acid salts are compounds that contain a rigid hydrophobic structure (52). Bile acid salts are ionic detergents with a negatively charged head group and a steroidal hydrophobic group such as the sodium salts of deoxycholic and cholic acid. They are characterised by a negatively charged carboxyl group at the end of the short alkyl

chain and hydroxyl groups on the steroid structure. This results in an undefined polar head group and instead the molecule has a polar face and an apolar face. Unlike their larger alkyl ionic detergent counterparts, bile acid salts form small kidney shaped micelles.

1.5.2 Non-Ionic Detergents

Non-ionic detergents have an uncharged hydrophilic head group and are thus better at breaking lipid-lipid and lipid-protein interactions (51). They are non-denaturant and their main use in biology includes isolating membrane proteins in their biologically active form. They are also commonly found in household dish detergents since they do not react with hard water ions.

Due to the lack of charge on the head group it follows that salts have very little effect on the micellar size of non-ionic detergents. One common non-ionic detergent is Triton X-100 that has a polyoxyethylene moiety as a head group. The other major group of non-ionic detergents have a glycosidic head group such as in octyl glucoside and dodecyl maltoside. The polyoxyethylene chains are further removed from the hydrophobic centre of the micelles due to the formation of random coils and those with short chains form aggregates and viscous solutions in water. However detergents with slightly longer polyoxyethylene chains don't aggregate.

1.5.3 Zwitterionic Detergents

Zwitterionic detergents have properties of both ionic and non-ionic detergents (53). They have no net charge as in non-ionic detergents and as such lack conductivity and do not bind to ion-exchange resins. Zwittergents are however efficient at breaking

protein-protein interactions. They vary in their denaturing properties. CHAPS for example, which has a rigid steroid ring structure like that in a bile acid salt, is very non-denaturing.

1.5.4 Detergent Effects on Mass Spectrometry

Depending on which detergent is present the mass spectrometry detection limit for a given sample will vary. As shown in the table below, MALDI-TOF MS and ESI-MS are capable of withstanding different levels of detergent. Electrospray ionisation is more sensitive to detergent. This may be because it affects the droplet formation. Several studies have been performed looking at detergent effects (54, 55). The tables below contain information used as a guide for submitting protein samples to mass spectrometry services.

The numbers one through four are a numerical representation of the detergent effect on the spectrum and signal intensity, with the numeral one representing the least effect and the numeral four representing the loss of protein or peptide signal altogether.

Table 1: *Detergent Effects on MALDI-TOF mass spectra*

Detergent	Class
n-octyl-glucoside	1
n-dodecyl-glucoside	1
octanoyl-N-methylglucamide	1
decanoyl-N-methylglucamide	1
n-dodecyl-beta-D-maltoside	2

octylphenolpoly(ethyleneglycolether) (Triton X-100)	3
octylphenolpoly(ethyleneglycolether) (Triton X-114)	3
polyethylene glycol (PEG 2000)	3
dodecylpoly(ethyleneglycolether) (Thesit)	4
isotridecylpoly(ethyleneglycolether)	4
CHAPS	4
CHAPSO	4
n-dodecyl-N,N-dimethyl-3-ammonio-1-propanesulfonate	4
sodium dodecylsulfate (SDS)	4

The Ogorzalek Loo study used three concentrations of detergent, 1%, 0.1% and 0.01%. Detergents at a level of 1% suppressed most protein signals. At 0.01% detergent concentration there was little spectral interference except for with SDS and sodium taurocholate which reduced protein signal by a factor of twenty. The table below reflects the effect on protein signals by a 0.1% detergent concentration and gives the percentage signal intensity with detergent compared to 100% signal intensity without detergent.

Table 2 :*Detergent Effects on ESI Mass Spectra*

Detergent	**Signal**
sodium dodecylsulfate (SDS)	< 10%
sodium taurocholate	< 10%
sodium cholate	21-30%

CTAB	< 10%
LDAO	< 10%
CHAPS	31-60%
Tween 20	10-20%
Thesit	< 10%
Triton X-100	< 10%
NP40	< 10%
n-octyl sucrose	10-20%
n-dodecyl sucrose	10-20%
n-dodecyl maltoside	21-30%
octyl glucoside	21-30%
octyl thioglucoside	21-30%
n-hexyl glucoside	30-60%
n-dodecyl glucoside	> 60%

Most mass spectrometers can accommodate small concentrations of SDS (29, 56). However SDS specifically interacts with proteins to give a stoichiometrically bound outer layer of SDS (7, 9). This creates greater problems for mass spectrometry, as there are few clean up processes efficient enough to allow signals to be observed. One common method for effective sample clean up is reverse phase chromatography.

Research showed that ALS-I was compatible with mass spectrometric techniques due to degradation of the detergent on addition of acid (56, 57, 58, 59). It was also

determined that ALS-I was compatible with MALDI-TOF MS and LC-MS/MS analysis prior to degradation of the detergent.

1.5.5 Detergents in Mass Spectrometry

Detergents have several applications in biochemistry as demonstrated above. However, with the increasing use of mass spectrometry in biochemistry the use of detergents upstream of mass spectrometric protein analysis and identification is a real problem (60). Detergents are known to have several adverse effects on mass spectra, the three main effects being: background ions of high intensity that can 'drown out' protein signals, analyte signal suppression and the formation of adducts on the protein. Detergents can also cause the shift of the charge envelope in electrospray ionization mass spectrometry, although the specific relationship between the detergent and the charge envelope shift is yet unknown (54, 61).

One commonly used detergent, sodium dodecyl sulphate or SDS, is highly incompatible with mass spectrometry as it suppresses signals (54). Stoichiometric binding between SDS and proteins means that SDS is also very difficult to wash away.

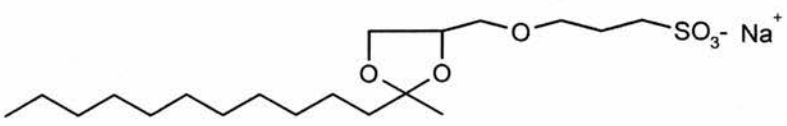
Triton X-100 and NP-40 are non-ionic detergents with common applications in biochemistry. They have a similar effect on mass spectrometric signals to SDS but for different reasons. Triton X-100 has a spread of molecular weights centred around approximately 537 Da and 631 Da and NP-40 has a molecular weight of 603 Da, and although they have been found to be compatible with MALDI-TOF MS analysis up to a concentration of 1%, they are known to form adducts with proteins thus causing

severe signal suppression (54, 60). Both detergents ionize well showing large detergent peaks in spectra, which also suppresses protein signals in ESI. In MALDI-TOF MS analysis these detergents prevent co-crystallization of the matrix and analyte. Triton X-100 has a relatively low CMC at 0.2-0.9 mM but it also has a large average micellar weight. NP-40 has a CMC of 0.1-0.3 mM that makes it difficult to remove by dialysis. Both detergents are complex mixtures of polymers that are easily fragmented by reverse phase chromatography, which results in background signals that can obscure the protein signal (61).

Despite the difficulties of using detergents in mass spectrometry, they are often necessary in order to solubilize proteins that are difficult to handle such as membrane proteins (49). Due to the distinct incompatibility of these common detergents with mass spectrometry there has been further research in order to synthesise and identify detergents that are compatible with mass spectrometric techniques. Research performed by Gharahdaghi *et al.* in 1996 determined that MALDI-TOF analysis was most compatible with gluco- and maltopyranoside detergents such as octyl-β-glucoside (62).

Acid labile surfactants (ALS) are an important step in the development of mass spectrometry-compatible detergents. As stated before ALS-I is an anionic detergent and is a long chain derivative of 1,3 dioxolane sodium propyloxy sulfate.

***Figure 12**: Structure of Acid Labile Surfactant I*



It was first synthesised in two steps by Yamamura *et al.* in a study of destructible surfactants with varying lengths of carbon chains at the 2 position (63). A more recent synthesis achieved by Ross *et al.* derived the first step from Yamamura *et al.* (64).

Konig *et al.* performed experiments substituting ALS-I for SDS in PAGE protein separation techniques (57, 58). Protein samples were separated both by SDS-PAGE and ALS-PAGE and then stained with either Coomassie Brilliant Blue or Silver stain. Bands were excised and subjected to in-gel digestion. The resultant peptides were analysed by MALDI-TOF mass spectrometry. This showed that samples digested from ALS polyacrylamide gels contained more peptides at higher abundance than those from SDS polyacrylamide gels. However it was still necessary to clean up the digests with a ZipTip, a small reverse phase column packed in a pipette tip. Konig *et al.* concluded that separating proteins by ALS-PAGE was more sensitive than SDS-PAGE (57). For example the Mascot score obtained for γ-Neuronal enolase 2 (47 kDa) separated by ALS-PAGE was 266 compared with the score of 194 for the same protein separated by SDS-PAGE.

Another study was performed by Ross *et al.* to determine the benefits of ALS-PAGE over SDS-PAGE (64). Protein bands were subjected to in-gel digestion, extracted and desalted. The most efficient extraction of peptides occurred when eluting them directly into 2 μL of matrix solution. They were analysed by MALDI-TOF MS. Proteins were found to migrate more slowly in ALS polyacrylamide gels than in SDS-polyacrylamide gels indicating that ALS-I does not hold as high a charge on proteins as SDS. Despite this ALS-PAGE gave better results than SDS-PAGE. Peptide signals were more intense for ALS samples thus identifying more peptides than in SDS. A

samples from an ALS-PA gel, for example an ALS sample of yeast enolase, matched thirteen peptides, whereas only six peptides were identified in the same protein from an SDS gel.

Likewise, Nomura *et al.* found that MALDI-MS detection of tryptic peptides was improved by 1D and 2D ALS-PAGE. They also found that pretreating SDS polyacrylamide gel pieces with ALS prior to tryptic digest improved elution of peptides from the gel. They also showed that ALS-I does not inhibit trypsin activity during in-gel digestion unlike SDS (56).

ALS-I is commercially available from Waters, who market the compound as RapiGest™ SF. Product information for RapiGest™ SF states that it is "a reagent used to enhance in-solution enzymatic digestions of proteins." (59) Digestion is reported to occur within sixty minutes. Upon addition of HCl to a concentration of 50 mM and heating of the mixture the ALS-I readily degrades allowing the samples to be analysed by MALDI-TOF MS or HPLC/MS.

Another recent study has investigated the use of a novel cleavable detergent (65). The detergent 3-[3-(1,1-bisalkyloxyethyl)pyridine-1-yl] propane-1-sulfonate (PPS) was used to extract proteins from within a cell by disrupting the cell membrane. The detergent also effectively solubilized the protein once extracted from the cell. The benefit of this detergent is that at low pH it is degraded. Norris *et al.* found that PPS improved MALDI sensitivity of whole proteins without relinquishing spectral quality.

1.6 Whole Protein Extraction From SDS-Polyacrylamide Gels

There are many instances when it would be advantageous to obtain a molecular weight of the accuracy obtainable by mass spectrometry for a protein band on a SDS-polyacrylamide gel. As SDS-polyacrylamide gels are run under denaturing conditions, they are efficient at separating the components of tightly bound complexes for example an antibody from an antigen. They also allow the separation of species, which may be intractable by other purification methods. This could be a result of N or C-terminal clipping, giving two closely related species. Determination of their masses with good mass accuracy may allow the site of the clipping to be determined (30). A further example of this would be a mixture of a post translationally modified and unmodified form of a protein. It may be possible to determine the type of modification occurring simply by measuring the mass change between the two species with the accuracy that mass spectrometry can give. Furthermore, if it was possible to obtain whole protein mass spectrometric data from samples separated by SDS-PAGE, running samples out on a polyacrylamide gel first could be seen as a good route for removing mass spectrometrically incompatible contaminants, e.g. polymeric detergents, urea etc. from the sample.

Over the years a number of groups have worked towards obtaining mass spectral data from proteins separated by SDS-PAGE. There are three major ways in which this has been achieved: 1) direct analysis of polyacrylamide gels, 2) electroelution and 3) passive elution. Scientists also showed that signals could be achieved by directly analysing a polyacrylamide gel slice at the expense of mass accuracy. Extraction efficiency was shown to be dependent on time, the size of the proteins and the

addition of a detergent to aid in the solubilization of the proteins. However they did not attempt to obtain mass spectrometric signals from the whole protein extract.

Ogorzalek Loo *et al.* performed a study on the direct analysis of ultrathin polyacrylamide gels by MALDI-TOF MS (66). Their experiments included analysis of isoelectric focusing, native and SDS gels. Once dry, ultrathin gels are no more than 10 μm thick. The isoelectric focusing gels utilized were 5% polyacrylamide gels; the native and SDS gels were 12.5% polyacrylamide gels. However the SDS precast gels only contained SDS in the agarose buffer strips and not in the separating gel. Separation was performed and gels were then left to dry, sometimes augmented by a vacuum desiccator or a stream of air. The matrix, sinapinic acid, was spotted on dried gels prior to analysis by MALDI-TOF MS. Spectra were acquired in linear mode on a Vestec LaserTec Research instrument. Signals were detected for samples of as little as 600 fmol of α and β chain haemoglobin and 1 pmol of horse heart myoglobin. Despite obtaining good protein signals for these smaller proteins, this study recognised the challenge of determining the accurate mass of a protein based on the broader peaks of MALDI-TOF mass spectra. Their remedy for this was to acquire several calibrations. Re-analysis of the band also proved difficult as protein signals were degraded.

In 1996 Castellanos-Serra *et al.* investigated a reversible negative staining method in order to avoid fixing the protein with acid as occurs with both Coomassie and silver stains (67). SDS-polyacrylamide gels were stained with an imidazole-SDS-zinc stain, which causes the protein to be immobilized but is fully reversible. Whole protein was eluted from the gel by crushing the gel into fine pieces or cutting it into pieces no bigger than 1 mm^3 and incubating this crushed material in a tris, glycine and EDTA

buffer. The EDTA chelates the zinc ions. No detergent was necessary to solubilize the proteins. However as SDS is a major component of the gels and stain, the residual SDS might be expected to cause suppression of mass spectrometry signals. The group showed protein extracts could be N-terminal sequenced and also that the extracts retained their biological activity. However, no mass spectrometric data was reported with this investigation.

In 2000 another study was undertaken to obtain MALDI-TOF MS signals from passively eluted and electroeluted whole proteins (68). Passive elution was performed by placing destained crushed gel pieces in a solution of formic acid, water and 2-propanol (FWI) and vortexing for an extended period of time. Another set of experiments was performed in the same way but involved a ten-minute to two-hour incubation in an ultrasonic bath at 26 °C using an extraction solution of formic acid, acetonitrile, 2-propanol and water (FAPW). Protein was also electroeluted in a Centrilutor using electrophoresis buffer. It was found that ultrasonic passive elution gave the best MALDI mass spectra. Samples were prepared for analysis using the matrices sinapinic acid and s-DHB and signals were acquired in linear mode. Using the dried drop techniques, this resulted in only a small portion of the extraction being analysed. Two stains, Coomassie and Zn-imidazole, were also tested in conjunction with the protein elution methods. No significant difference was found in the quality of the signals. However by increasing the extraction time from ten minutes to two hours a broadening of peaks was observed. A mass shift attributed to acrylamide adducts on cysteines and formylation adducts on serines and threonines were also observed. Formylation was also found to be affected by extraction time. No ESI-MS data was

acquired. Extraction efficiency was a problem in acquiring spectra, the researchers obtained signals only if 400 pmoles of protein were loaded on the gel.

Several other studies have also performed whole protein extraction utilizing the passive elution technique and the extraction solutions FWI and FAPW. A study performed by Ehring et al. used ultrasonic passive elution and incubated samples at 35 °C for ten to thirty minutes (69). Extraction efficiency was not improved by longer incubation and destaining times. MALDI-TOF mass spectra show good signals for extracted proteins mixed with the matrix s-DHB. However, again it was noted that formylation [M+28] and acrylamide [M+71] adducts were the cause of signal degradation observed. Formylation adducts split the signal into several peaks thus signals are not as intense as a single peak.

Several studies have been performed on the effects of non-polymerized acrylamide on analytes during electrophoresis (66, 68, 69, 70). Galvani *et al.* passively extracted β-lactoglobulin B (LG-B) from SDS-polyacrylamide gels into a solution containing formic acid, acetonitrile, isopropanol and water (FAPW) and subsequently analysed the extracts by MALDI-TOF MS (68). The extracts were co-crystallised with the matrices s-DHB and sinapinic acid by loading them on to the target plate using the dried-drop method and a method called the three-layer deposition. The latter involved applying 0.6 µL of sample extract between two layers of matrix solution. They found that extracted LG-B was subject to acrylamide adducts which were attributed to unpolymerised acrylamide ineracting with cysteine residues. It was reported that acrylamide adducts as well as formylation were responsible for the broadening of MALDI-TOF MS peaks thus reducing mass accuracy.

Another study was performed by Ehring *et al.*, which demonstrated a link between extraction times of whole proteins into FAPW and the broadness of peaks acquired by MALDI-TOF MS. They determined that longer extraction times of protein into FAPW led to an increase in formylation adducts. They also observed acrylamide adducts which they attributed to a reaction of cysteine residues with unpolymerised acrylamide.

Another extraction method developed by Cohen and Chait involves extracting the protein directly into the matrix solution (71). Proteins were separated by 1D SDS-PAGE and stained with copper-reverse stain. Protein bands were excised, destained and then crushed before the protein was extracted. Protein extraction was carried out by adding enough matrix solution, in this case HCCA, to cover the gel pieces. Samples were vortexed for one to two hours and then sample vials were left open to induce co-crystallization of the matrix and analyte, thus combining passive elution and slow crystallization. Cohen and Chait determined that this method of extraction was more effective than passive elution alone for samples of less than 25 pmoles as it allows the sample to be eluted in a relatively large sample volume and then analysing a concentrated portion of the whole. They showed that slow crystallisation greatly enhances the sensitivity of MALDI-MS and promotes visualization of high-mass components.

Previous work in the Botting laboratory has built upon these methods and developed extraction methods, which gave MALDI signals for proteins separated by SDS-PAGE. Initial experiments used zinc-imidazole reverse stained protein bands. This method utilizing an extraction buffer containing 190 mM glycine, 25 mM Tris, 100

mM DTT and 0.1% Triton X-100 achieved a 50% extraction efficiency. Samples were incubated at 4 °C overnight. The extract was then slow crystallised for 3-5 hours and the detergent and contaminant were subsequently washed away with water. The crystals were resuspended in 2 μL of water and analysed by MALDI-TOF MS. The Coomassie stained bands could be analysed utilizing the following destaining method. The Coomassie stained protein band is excised and destained via a series of washes with water, acetonitrile, methanol and a mixture of formic acid, water and isopropanol followed by the addition of the extraction buffer. An extraction buffer containing the detergent Triton X-100 to keep the protein soluble, dithiothreitol to maintain the reduced disulfide bonds and tris/glycine to maintain the protein in the same buffer a it had been in the gel. The extract was then subjected to slow crystallisation in a separate step. The detergent is therefore washed away from the co-crystals and good MALDI signals were obtained for concentrations down to 2.5 pmoles (72). However, there was still considerable signal suppression, possibly due to the presence of SDS. Furthermore, although accurate masses were obtained for smaller proteins, larger proteins, e.g. BSA showed a considerable mass shift. The MALDI-TOF MS mass accuracy was not good enough to be able to determine the nature of any modifications that be causing this shift. It was felt that the properties of ALS-I might lend it to improving such an extraction procedure replacing the Triton X-100 with a mass spectrometry compatible detergent and potentially allow ESI signals of whole proteins extracted from polyacrylamide gels to be obtained.

2. Materials and Methods

2.1 Materials

The proteins bovine serum albumin, myoglobin, superoxide dismutase, aldolase, alcohol dehydrogenase, carbonic anhydrase and ubiquitin were obtained from Sigma-Aldrich (Dorset, UK). The proteins SUMO-1, NFKB p50 and p65 were kind gifts from Professor R. T. Hay (University of St. Andrews, UK). The protein fluorinase was a kind gift from Pofessor D. O' Hagan (University of St. Andrews, UK). The protein QPRTase was a kind gift from Dr. N. Botting (University of St. Andrews, UK). Trypsin was obtained from Promega (Wisconsin, USA). Potassium pentachloro aquo ruthenate was obtained from Alfa Aesar, Johnson Matthey (Karlsruhe, Germany) and anhydrous Bathophenanthrolinedisulfonic acid disodium salt and sodium ascorbate were obtained from Lancaster Synthesis (Lancashire, UK). In the preparation of ALS-I, 2-Tridecanone, *p*-toluenesulfonic acid monohydrate and 1,3-Propane Sultone were obtained from Sigma-Aldrich (Dorset, UK) and 1-(2-pyridylazo)-2-naphthol used in determining the critical micelle concentration of ALS-I was obtained from Lancaster Synthesis (Lancashire, UK). Acrylamide solution, TEMED, Coomassie and other laboratory reagents were purchased from Sigma-Aldrich (Dorset, UK).

Instruments used include the fluoro imaging system FLA-5000 (Fujifilm, Japan) for visualizing gels stained with the fluorescent stains RuBPS and Sypro Ruby and the UVIKON 930 Spectrophotometer (KONTRON Instruments) for CMC determination. The mass spectrometers utilized in the analyses of these experiments were the TofSpec 2E (Micromass, Manchester, UK), the UltiMate nano LC system (Dionex, California, USA) coupled to the QStar Pulsar XL Quadrupole TOF instrument

(Applied Biosystems, Warrington, UK) and the benchtop LCT (Micromass) ESI mass spectrometer coupled to an Alliance HT 2795 HPLC system (Waters). NMR data was collected the Bruker Avance 300 instrument (Bruker, Germany).

2.2 SDS-PAGE

Ten and twelve percent gels were prepared. The 10% resolving gel comprised of 3 mL H_2O, 2.5 mL 30% acrylamide/bisacrylamide 29:1 solution, 1.9 mL separating gel buffer (1.5 M Tris, pH 8.8, 0.1% SDS), 112 µL 10% Ammonium Persulfate and 5 µL TEMED. The stacking gel consisted of 1.0 mL H_2O, 300 µL 30% acrylamide/bisacrylamide 29:1 solution, 444 µL stacking gel buffer (0.5 M Tris, pH 6.8, 0.1% SDS), 28 µL 10% Ammonium Persulfate and 5 µL TEMED. Disruption buffer was added to protein samples which consists of 2.0 mL glycerol, 2.0 mL 10% SDS, 0.25 mg bromophenol blue, 2.5 mL of 4 times concentrated stacking gel buffer and 0.5 mL β-mercaptoethanol. Protein samples were boiled to ensure complete denaturation before being loaded onto the gel. Gels were electrophoresed at 150 V (Biorad) for between 40-90 minutes.

The gels were stained with Coomassie Blue R-250 in 10% acetic acid/40% methanol/50% H_2O for 1-2 hours and then destained in several changes of 20% methanol/10% acetic acid/70% H_2O overnight.

2.3 Investigation of Ruthenium II Bathophenanthroline Disulfonate Chelate

2.3.1 Synthesis of Ruthenium II Bathophenanthroline Disulfonate Chelate

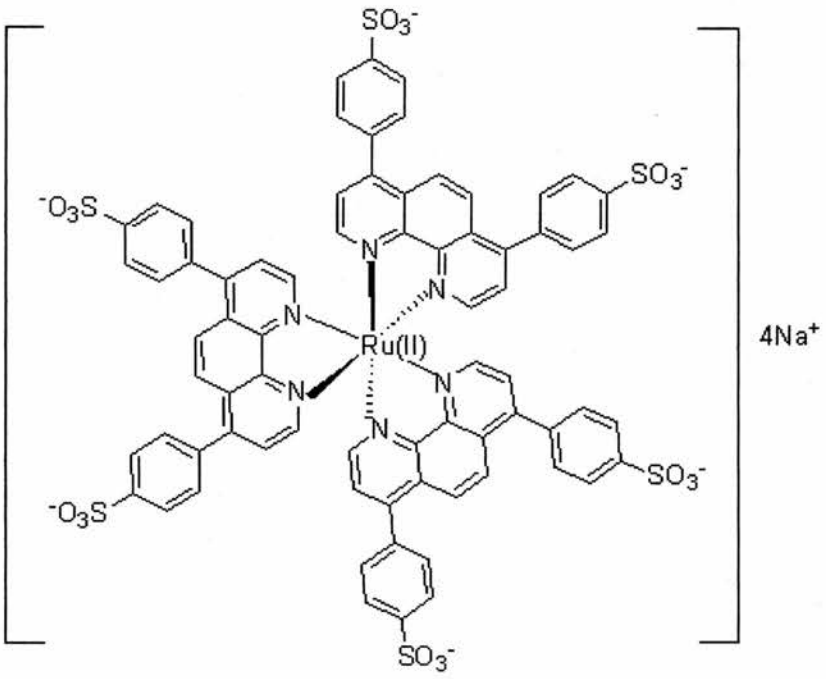


***Figure 13**:Structure of Ruthenium II Bathophenanthroline Disulfonate Chelate*

Potassium pentachloro aquo ruthenate (0.2 g, 0.53 mmoles) was dissolved in 20 mL boiling water to give a deep red solution (19). Anhydrous Bathophenanthrolinedisulfonic acid disodium salt was added at a three molar equivalence (0.9 g, 1.7 mmoles) to the refluxing mixture and refluxing continued for twenty minutes. The addition of the bathophenanthrolinedisulfonic acid disodium salt can lead to considerable foaming. The result was a green/brown solution. Sodium ascorbate (5 mL, 500 mM) was added to the green/brown solution. This caused foaming of the solution. Refluxing was resumed for another 20 minutes after which the reaction mixture was allowed to cool and the pH was adjusted to 8.4 using concentrated sodium hydroxide (4 M) solution.

2.3.2 Staining and analysis with RuBPS

Polyacrylamide gels to be stained with RuBPS were fixed overnight in 30% v/v ethanol/ 10% acetic acid (19). Gels were then rinsed in 20% ethanol (4x30 min) to remove all the acetic acid (acids quench the fluorescence of the chelate). Gels were stained in the dark for 3-6 hours in 20% ethanol containing 150 nM of RuBPS for optimum staining. Once staining was complete, the gels were rinsed in water prior to analysis. Gels were visualised using the fluoro image analyser FLA-5000 (Fujifilm, Japan) in conjunction with the software package Image Reader 1.1 for the FLA-5000 (Fujifilm, Japan). The gel images were accessible through the software Image Gauge V. 3.45 alias (Fujifilm, Japan).

2.4 Synthesis of ALS-I

2.4.1 Synthesis of 4-hydroxymethyl-2-methyl-2-undecyl-1,3-dioxolane

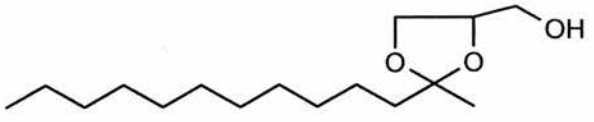


***Figure 14**: Structure of 4-hydroxymethyl-2-methyl-2-undecyl-1,3-dioxolane*

4-hydroxymethyl-2-methyl-2-undecyl-1,3-dioxolane was synthesised according to the method of Yamamura *et al.* (63). 2-tridecanone (4 g, 0.02 mol), glycerol (2.25 g, 0.024 mol), toluene (10 mL) and *p*-toluenesulfonic acid monohydrate (0.02 g, 0.1 mM) were heated under Dean and Stark conditions until no further separation of water was seen. Approximately 1.1 mL of water was collected. The mixture was allowed to cool and was subsequently washed with 5% (w/v) sodium carbonate (10 mL) and water (3 x 10 mL). The toluene was removed at reduced pressure to give 3.1

g of yellow oil. Yield, (3.1 g, 57%), bp.142 °C; ^{1}H-NMR δ_H (300 MHz; $CDCl_3$) 0.88 (t, 3 H, CH_3) 1.2-2.0 (m, 23 H, CH_2 and $C\underline{H_3}COO$) 3.5-4.3 [m, 5 H, OCH_2CH and $(CH_2O)O$]; ^{13}C-NMR δ_C (75.46 MHz; $CDCl_3$) 23.75, 24.99 (CH_3) 23.02-32.25 (9 CH_2) 44.11 (CH_2OH) 209.67 ($\underline{C}OH$); m/z 273.14 ($M+H^+$) 271.17 ($M-H^+$).

2.4.2 Synthesis of sodium 4-[(2-methyl-2-undecyl-1,3-dioxolan-4-yl) methoxyl]-1-propane sulfonate (ALS-I)

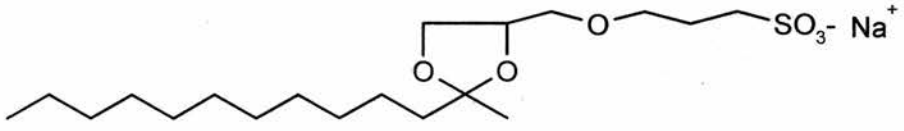


***Figure 15**: Structure of Sodium 4-[(2-methyl-2-undecyl-1,3-dioxolan-4-yl) methoxyl-1-propane sulfonate (ALS-I)*

Sodium 4-[(2-methyl-2-undecul-1,3-dioxolan-4-yl) methoxyl]-1-propane sulfonate (ALS-I) was synthesised according to the method of Ross *et al.* (64). 4-hydroxymethyl-2-methyl-2-undecyl-1, 3-dioxolane (3 g, 0.011 mol), powdered NaOH (0.52 g, 0.013 mol), and toluene (15 mL) were stirred at a constant 50 °C while 1,3-propane sultone (1.59 g, 0.013 mol) was added slowly over 30 minutes. The temperature was then raised to 70-75 °C and the solution was left to stir for 6 hours. The solution was then allowed to cool to room temperature before being poured into boiling ethanol (40 mL). The solvent was then removed at reduced pressure. The resulting solid was subsequently dissolved in boiling ethanol and filtered while hot. The solid residue was extracted once more with boiling ethanol and this was combined with the mother liquor. The solvent was removed under reduced pressure and the resulting white solid was recrystallised in ethanol. The product was found to be hygroscopic at room temperature and thus was filtered under nitrogen to give a

white solid. Yield: (2.0 g, 43%),^{1}H-NMR δ_H (300 MHz; $CDCl_3$) 0.9 (t, 3 H, CH_3) 1.2-1.7 (m, 27 H, O$\underline{CH_2}$, $\underline{CH_2}$ and $\underline{CH_3}$COO) 2.1-2.5 (m, 2H, $\underline{CH_2}SO_3$) 3.5-4.4 (m, 5 H, O$\underline{CH}$ and $CH_2O\underline{CH_2}$); ^{13}C-NMR δ_C (75.46 MHz; $CDCl_3$) 14.51 ($\underline{C}H_3CH_2$) 23.08-32.32 (9 $\underline{C}H_2$) 23.81, 25.06 (-$CH_2\underline{C}H_3$) 39.31, 40.23 ($CH_3\underline{C}H_2$- and $\underline{C}H_2CO$) 63.26, 63.46 ($\underline{C}H_2CH$ and $\underline{C}H_2O$) 66.17-66.19 (O$\underline{C}H_2$), 76.25, 76.91 ($\underline{C}H$) 111.42-111.65 ($\underline{C}H_2O\underline{C}H_2$) 138 ($\underline{C}$).

2.5 Protein Solution Digestions

Standard protein (2.5 µL, 25 pmoles) was placed in a microcentrifuge tube along with ammonium bicarbonate (0.5 µL, 1 M), dithiothreitol (DTT) (1 µL, 0.2 M), and water (3.5-4.5 µL). ALS-I (1 µL, 10 mg/mL in water) was added to half of the solution digestion trials. This solution was incubated for 30 minutes at room temperature. Following incubation, iodoacetamide (1 µL, 0.5 M) was added to the microcentrifuge tube and the solution was allowed to incubate for a further 30 minutes at room temperature. Trypsin (0.5 µL) was then added and the solution was incubated overnight at 37 °C.

2.6 Protein In-gel Digestions and Extractions

In-gel digestion was performed using a method based on that of Schevchenko *et al.* (13). Proteins were separated on a 10% polyacrylamide gel and then stained with Coomassie Blue R-250 stain in 10% acetic acid/40% methanol/50% H_2O for 1-2 hours and then destained in several changes of 20% methanol/10%acetic acid/70% H_2O overnight. The bands were then excised, cut into 1 mm^3 pieces and washed twice with 100 µL of water for 10 minutes each. These washes were followed by two washes of 50 µL MeCN/H_2O (1:1 v/v) for 15 minutes each and one wash of 25 µL

MeCN for 2 minutes. The acetonitrile was then removed and 50 µL of 0.1 M ammonium bicarbonate was added for five minutes. To this 50 µL acetonitrile was added and the gel slices were incubated for a further 15 minutes. All the excess liquid was removed and the gel slices were dried down under reduced pressure. The dried gel slices were incubated at 56 °C in 50 µL 10 mM dithiothreitol (DTT) in 0.1 M ammonium bicarbonate for 45 minutes, reducing the disulfide bonds. The excess liquid was removed and 50 µL of 55 mM iodoacetamide in 0.1 M ammonium bicarbonate was added and the gel slices were incubated at room temperature for 30 minutes in the dark. Following this the gel slices were washed with 25 µL 0.1 M ammonium bicarbonate for 5 minutes. To this was added 25 µL acetonitrile and the gel slices were incubated for a further 15 minutes. The supernatant was removed and the gel slices were then dried down under reduced pressure and each slice was rehydrated in 10-25 µL 50 mM ammonium bicarbonate containing 0.1 pmoles/µL trypsin. Half the digestion trials included ALS-I at a concentration of 0.1% (1 mg/mL) in the digestion solution. Samples were then incubated overnight at 37 °C.

The supernatant was removed and placed in a new microcentrifuge tube. The gel slices were then incubated in 10 µL of 25 mM ammonium bicarbonate for 5 minutes, 10 µL acetonitrile was added and the solutions incubated for a further 15 minutes. The liquid was removed and added to the liquid collected before. This wash pattern was repeated two more times, substituting 5% formic acid for ammonium bicarbonate. All extraction washes were collected together and dried down under reduced pressure. The extracted peptides were then resuspended in 10% formic acid. To ensure degradation of the acid labile surfactant the solutions were incubated at 37 °C for 30 minutes. All the samples were heated in order to ensure trials were comparable.

2.7 Whole Protein Extraction

Whole protein extraction was performed based on a method reported by Mirza *et al.* (73). Myoglobin and superoxide dismutase (25-300 pmoles) were separated by 1D SDS-PAGE on a 10% polyacrylamide gel. The gels were stained and destained as described above. The protein bands were excised, cut into 1 mm^3 pieces or crushed finely with the end of a plastic rod and washed thoroughly with water (2x 150 µL, each 10 minutes). The bands were then washed as described below. Each of the following washes was accompanied by vortexing and rotating. The bands were washed with 10% acetic acid (150 µL, 10 minutes) followed by another water wash (150 µL, 10 minutes). The bands were then washed with methanol (100-150 µL, 20 minutes) followed by a water wash (150 µL, 10 minutes). This was followed by a wash with acetonitrile (100-150 µL, 20 minutes) and a subsequent water wash (150 µL, 10 minutes). The bands were destained in a solution of formic acid: water: isopropanol (FWI) (1:3:2, v/v/v) (100-200 µL, 1-4 hrs, until gel pieces were colourless). A change of FWI was sometimes necessary. Once destained, the bands were washed twice with water. The destained gel slices were crushed into pieces no bigger than 1 mm^3. The protein was then extracted into a solution containing tris/glycine, 100 mM DTT, and detergent. A range of buffer concentrations was tested from, 0.05 M to 1M. Each concentration was tested using either 0.1% Triton X-100 or ALS-I. Several concentrations of ALS-I were tested from 0.05 mg/mL to 1 mg/mL final concentration. Extraction was performed with and without sonication. Samples were incubated for 1 to 8 hours in the extraction solution depending on the size of the protein. Protein samples were analysed by LC-ESI MS or slow crystallised and analysed by MALDI-TOF MS.

2.8 MALDI Sample Preparation Methods

2.8.1 Dried Drop Method

0.5 µL of matrix solution, α-cyano-4-hydroxycinnamic acid (HCCA) (10 mg/mL) in 50% acetonitrile and 50% ethanol, followed by 0.5 µL of sample was applied to the target plate. Using the matrix α-cyano-4-hydroxycinnamic acid the spots had to be acidified to encourage matrix crystallisation. If the sample was in basic buffer 0.5 µL of 0.1% TFA was added to adjust the pH to below 2.4 to allow matrix crystallization. The ideal pH lies between 1.3 and 2.4. The spots were then left to dry in a flow of air, and as the solvent evaporates crystals are formed.

2.8.2 Slow Crystallisation using α-cyano-4-hydroxycinnamic acid

Slow crystallisation was performed as described by Botting (38). A saturated solution of α-cyano-4-hydroxycinnamic acid (HCCA) was prepared in formic acid/water/isopropanol (1:3:2 v/v/v) (FWI). Protein samples (generally 20 µL) were mixed with the saturated matrix solution (25 µL) in microcentrifuge tubes and the side of the tube was scratched to promote crystallization. The tubes were left open at room temperature for 3-4 hours, or until matrix/protein co-crystals had formed. The supernatant was then removed and the crystals were collected by centrifugation or scraping and then washed twice with H_2O. Crystals were resuspended in 2 µL of water and 0.5 µL of this suspension was loaded on to a spot on a 100-spot target plate.

2.8.3 Slow Crystallisation using Neutral Matrices

The three matrices tested 6-aza-2-thiothymine (pH 7), 2-amino-4-methyl-5-nitropyridine (pH 7.1) and 4-nitroaniline (pH 6.4) were each dissolved in solutions of isopropanol/water 2:1 v/v, formic acid/MeCN/isopropanol/water 50:25:15:10 v/v/v/v,

formic acid/MeCN/isopropanol/water 20:15:25:40 v/v/v/v, ethanol/1M ammonium citrate 1:1 v/v, ethanol/water 1:1 v/v, ethanol/0.1% TFA 1:1 v/v, and methanol/water 1:1 v/v at saturating concentrations and at 10 mg/mL. These neutral matrices were set up for slow crystallisation trials by adding 20 μL of saturated matrix solution to 10 μL of protein solution in a microcentrifuge tube. Half of the trials were sonicated for 1 hour and then the tubes were placed on the bench with their lids open until crystals formed, 6-12 hours. The other half did not undergo sonication and were left on the bench with their lids open until crystals formed, 6-12 hours. The neutral matrices required washing with 1-10% formic acid due to the solubility of the neutral matrices at neutral pH. The crystals were then suspended in 2 μL 1% formic acid.

2.8.4 Whole Protein Analysis Using a Supersaturated Solution of HCCA

Whole protein signals were obtained using a supersaturated solution of HCCA in 35% acetonitrile/65% 0.1% TFA. This solution was incubated at 40 °C for 30 minutes prior to spotting sample then matrix on to the target plate. A line was scratched through the spot to encourage crystal formation concentrated in a small area (74).

2.9 Instrumentation, Acquisition and Analysis Methods

2.9.1 MALDI-TOF MS of Peptides

Spectra of peptide extracts were acquired using a TofSpec 2E (Micromass, Manchester, UK) supplied with a 337 nm laser and operated in reflectron, positive ionization mode. Peptide extracts were spotted on the target plate using the dried drop method and the matrix α-cyano-4-hydroxycinammic acid as described in section 2.8.1. They required a laser coarse setting of 20% and a laser fine setting of 40. Matrix suppression was set at 500 Da. The pulse voltage was set at 3350 V.

Calibration for protein digests was performed by spotting a β-galactosidase (β-gal) tryptically digeseted spot at the centre of every eight samples to be analysed. Sample spectra were aligned to the calibration curve that was created for the β-gal digestion. The peptide Glu-fibrinogen (Glu-Fib) was also added to the sample spots in order to allow internal lockmass adjustment of the samples. The Glu-Fib peak was identified in a spectrum and the calibration modified to give exactly 1570.6774 Da for its mass. Protein identification was performed by submitting the list of peptide monoisotopic masses selected using BioLynx ProteinProbe (Micromass, Manchester, UK) to the Mascot Peptide Mass Fingerprint search engine (Matrix science) for comparison against a protein database (NCBI or MSDB). This software compares the peak list to predicted peak lists of peptides created by the specific protein digestion criteria for the proteins on the database. Peptide tolerance was set at 0.15 Da. The fixed modification of Carbamidomethyl modification of cysteines and the variable modification of oxidation on methionine were also selected.

2.9.2 MALDI-TOF MS of Whole Proteins

Whole protein spectra were acquired in linear mode. HCCA requires laser coarse 20% and laser fine 40%. Neutral matrices required a laser coarse setting of 50% energy, a laser fine setting of 40-80. Suppression was set at the maximum value of 5000 in both instances. The pulse voltage was tuned to the protein mass being analysed. The optimum pulse voltage for myoglobin and SOD was 2000 V. Calibration was performed by analysing standard spots of trypsinogen slow crystallised in HCCA in FWI, creating a calibration curve based on the M+H and M+2H signals and adjusting the spectra acquired for experimental samples accordingly.

2.9.3 LC-ESI MS/MS of Peptides

MS/MS spectra for tryptic digests of proteins were acquired using the UltiMate nano LC system (Dionex, California, USA) coupled to the QStar Pulsar XL Quadrupole TOF instrument (Applied Biosystems, Warrington, UK). The upper limit for quadrupole mass selection is 6000 amu and the TOF has a mass range of up to 40,000 amu. This instrument induced fragmentation in the pulsed LINAC collision cell. Data was acquired in information dependent acquisition (IDA) mode. A one second MS scan to assess peptides eluting and two further MS/MS scans of three seconds are performed, selecting the two most intense ions for fragmentation, before the cycle is repeated. Masses are then excluded for one minute to prevent reanalysis of the more intense peaks. Collision energies were automatically adjusted depending on the size of the peptide being fragmented. The UltiMate nano LC system utilizes an LC Packings PepMap C18 reverse phase column, which is 15 cm long by 75 μm ID. A 10 μL sample of a 40 kDa or lighter protein was analysed over a 55 minute gradient, whereas proteins above 40 kDa were analysed over an 85 minute gradient. Both gradients shown below were run at a flow rate of 0.25 μL/min. The gradient was a mixture of solvent A, 5% acetonitrile and 1% formic acid, and solvent B, 95% acetonitrile and 1% formic acid. Data was submitted to Mascot MS/MS Ion Search for analysis. The raw data was automatically converted to files compatible with Mascot using the Applied Biosystems Mascot.dll script or Mascot Daemon.

***Table 3**: LC-ESI MS/MS 55-Minute Gradient*

Gradient	Time (minutes)
3% B	0

3% B	5
35% B	23
50% B	30
99% B	31
99% B	40
3% B	41
3% B	55

***Table 4**: LC-ESI MS/MS 85-Minute Gradient*

Gradient	Time (minutes)
3% B	0
3% B	5
35% B	40
50% B	60
99% B	61
99% B	70
3% B	71
3% B	85

2.9.4 LC-ESI MS of Whole Proteins

Whole protein extracts from polyacrylamide gels were analysed on a bench top LCT (Micromass) ESI mass spectrometer coupled to an Alliance HT 2795 HPLC system

(Waters). Samples were introduced to the setup for analysis via an automated sample racking system. They were injected on to a MassPREP (Waters) On-Line Desalting Cartridge (2.1 x 10 mm) containing a polymer-based, reversed-phase packing material. Protein samples were run at a flow rate of 0.05 mL per minute with a flow restrictor to increase the backpressure in the system and a twenty-two minute gradient shown in the table below. Solvent A consists of acetonitrile and 1% formic acid and solvent C consists of H_2O and 1% formic acid. The Desolvation temperature was set at 300 °C and the source temperature at 100 °C. The capillary voltage was set to 3500 V and the RF Lens to 500. Spectra acquired for whole protein extracts were deconvoluted using the MaxEnt function in MassLynx (75). This algorithm iterates to determine the best fit to the data and hence determine the mass of the whole protein from the charge envelope of the peaks in a spectrum.

***Table 5**: LC-ESI MS 22-Minute Gradient*

Gradient A	Gradient C	Time (minutes)
2% A	98% C	0
2% A	98% C	1
98% A	2% C	6
98% A	2% C	8
2% A	98% C	8.1
2% A	98% C	12
2% A	98% C	22

2.10 Critical Micelle Concentrations of Aqueous Surfactants

A saturated solution of 1-(2-pyridylazo)-2-naphthol (PAN) in pentane (10 mL) was prepared (76). A series of dilutions of ALS-I from 0.06 mM to 10 mM were prepared. To each 5 mL of ALS-I solution was added 500 μL of saturated PAN solution. On evaporation of the pentane, the solutions turned varying intensities of orange. These intensities were measured on the UVIKON 930 Spectrophotometer (KONTRON Instruments) at 470 nm.

2.11 Bradford Assays

Coomassie Brilliant Blue G (0.5 mg/mL) was dissolved in 25% methanol, 42.5% H_3PO_4 and 32.5% H_2O (77). This was then diluted 1 in 4 with H_2O to give the Bradford Reagent with a pH of 1.1. A standard 1 mg/mL BSA solution was prepared in the same buffer as the proteins being assayed. To 1 mL portions of Bradford Reagent were added 1, 2, 4, 6, and 8 μL of BSA standard solution in plastic cuvettes. Likewise a volume of assay protein expected to give a concentration within the range of the standard curve was added to a further cuvette containing Bradford reagent. After mixing, the samples were left for five minutes to allow the colour to develop. The absorbance at 595 nm was measured for the standard BSA and the sample protein solutions. The concentration of the assayed protein was determined by plotting its absorbance on a standard BSA curve. Over a small range of concentrations the plot is linear.

3. Results and Discussion

3.1 Synthesis and use of Ruthenium II Bathophenanthroline Disulfonate Chelate (RuBPS)

Stains are important tools in visualizing proteins by SDS-PAGE. There are several kinds of stains with different properties, including binding properties, sensitivity and methods of visualization.

RuBPS was synthesized as a laboratory alternative to Sypro Ruby and because it was a non-covalent alternative for the protein extraction work that follows. The synthesis of RuBPS is a simple one-pot reaction that produces a deep red chelate (19). It involved the addition of anhydrous bathophenanthrolinedisulfonic acid disodium salt and sodium ascorbate to potassium pentachloro aquo ruthenate dissolved in water. It was reported that there could be considerable foaming on addition of the anhydrous bathophenanthrolinedisulfonic acid disodium salt, however in this case little foaming was observed (19). Once the chelate was formed the pH was adjusted to 8.4. The adjustment of pH required care as once pH 7.0 had been reached the pH increased rapidly on addition of small amounts of sodium hydroxide.

Various dilutions of RuBPS stock in 20% ethanol were tested on 1D SDS polyacrylamide gels in order to find the optimal staining concentration. This was found to be 150 nM. Protein bands on polyacrylamide gels stained with a concentration of stain ten times less (15mM), were not visible on the gel due to the high intensity background staining of the gel. Polyacrylamide gels stained at 75 nM had visible protein bands but showed smearing above and below the main protein band. Background staining was also a problem for higher concentrations of stain. Gels

stained with four times and two times the optimal concentration of RuBPS had large amounts of smearing and protein bands appeared non-distinct. The protein BSA stained better than the protein carbonic anyhdrase (CA) however this could be due to inaccuracies in weighing out protein where a portion of the sample is salt. Gel bands of as little as 10-20 ng of BSA were visible at a stain concentration of 150 nM however 20 ng was the lowest amount of CA visible at this concentration of stain. Occasionally gel bands containing higher amounts of protein appeared less intense or were completely obscured compared to gel bands containing less protein. This was due to uneven staining intensities on different parts of the gel, which appeared to be a major disadvantage with this staining method.

An optimal staining time of between two and six hours was also identified. Longer staining times gave rise to too much background making weaker, 2.5 pmol (ng) total protein, bands harder to identify. Staining times less than two hours caused difficulty in visualizing protein bands containing as much as 50 ng of protein as well as weaker protein bands. The RuBPS stain had similar sensitivity to the Sypro Ruby stain and thus gave similar results, both being more sensitive than Coomassie Brilliant Blue. Sypro Ruby had fewer steps overall in its staining protocol however they both required far more washes than Coomassie Brilliant Blue.

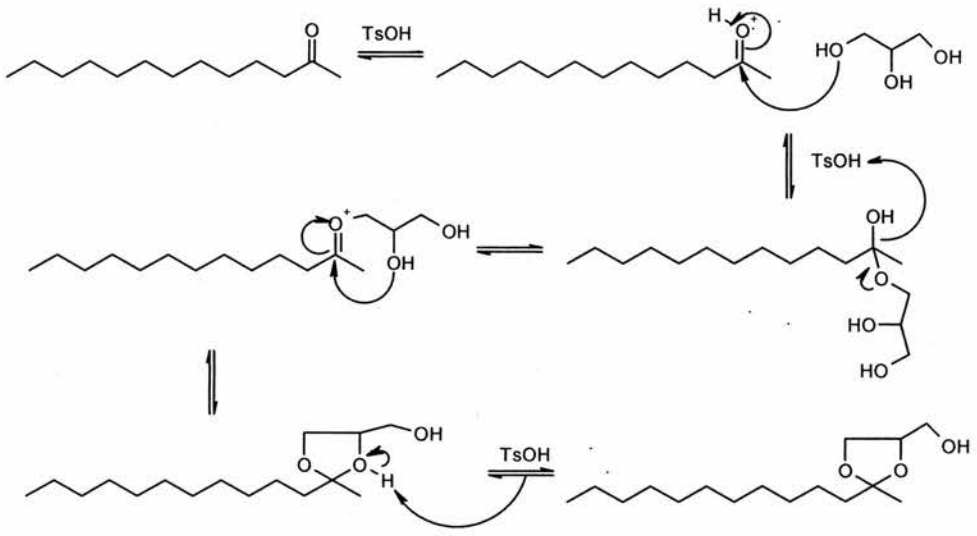
Both RuBPS and Sypro Ruby had previously been tested for mass spectrometric compatibility (19). This was performed by excising protein bands for in-gel digestion and MALDI-TOF analysis. This analysis showed that RuBPS gave better protein coverage overall and thus more peptides were identified from spots stained with RuBPS than spots stained with Sypro Ruby. Neither stain was found to be more

sensitive than silver stain however both were more compatible with mass spectrometric analyses than silver stain with RuBPS having slightly less interference with MALDI-TOF analysis than Sypro Ruby.

The synthesis and optimisation of RuBPS was proposed as a necessary step towards investigating and enhancing whole protein extraction. However RuBPS staining wasn't tested in conjunction with whole protein extraction for several reasons. In trying to develop a simple and generally accessible method for whole protein extraction, it was decided that a method that was applicable to Coomassie stained bands was preferable, and that RuBPS would only be explored if other methods had failed. Furthermore, although protein bands must be visualized by a scanner or charge-coupled device (CCD) camera, in order to excise protein bands from gels stained with fluorescent stain a UV transilluminator is required and the precautions required to limit UV exposure make it awkward to excise bands precisely.

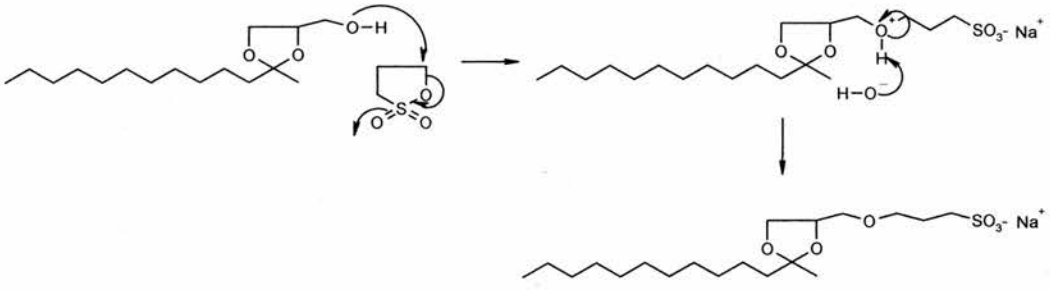
3.2 Synthesis of Acid Labile Surfactant I

The synthesis of ALS-I is a two-step process (63). The first step involves the synthesis of 4-hydroxymethyl-2-methyl-2-undecyl-1,3-dioxolane from 2-tridecanone and glycerol in a *p*-toluenesulfonic acid monohydrate catalysed reaction. The predicted mechanism for the first step is as follows.



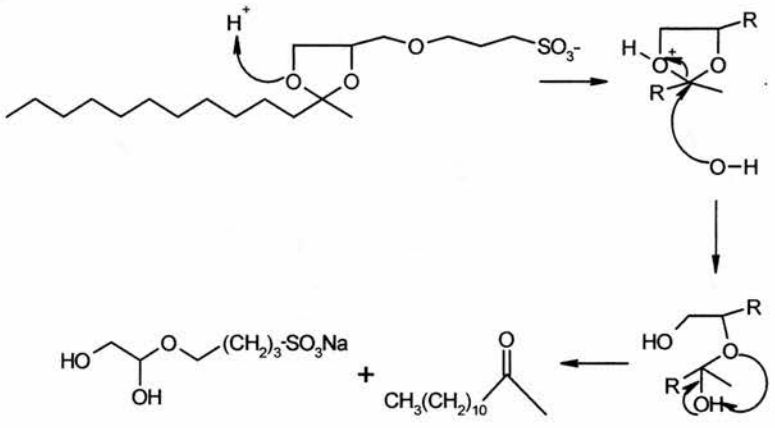
***Scheme 1**: Reaction Mechanism for the Synthesis of 4-hydroxymethyl-2-methyl-2-undecyl-1,3-dioxolane*

The second step involves the addition of 1,3-propane sultone in base to the product from the first step. The predicted mechanism is as follows.



***Scheme 2**: Reaction Mechanism for the synthesis of 4-[(2-methyl-2-undecyl-1,3-dioxolan-4-yl) methoxyl]-1-propane sulfonate*

The product of this reaction, sodium 4-[(2-methyl-2-undecyl-1,3-dioxolan-4-yl) methoxyl]-1-propane sulfonate, degrades when subjected to low pH. Decomposition of ALS-I gives dodeca-2-one and sodium 3-(2,3-dihydroxypropoxy) propanesulfonate and occurs as follows.



***Scheme 3**: Reaction Mechanism for the Acid Hydrolysis of ALS-I*

3.2.1 Synthesis of 4-hydroxymethyl-2-methyl-2-undecyl-1,3-dioxolane

Yamamura et al. showed the acid catalyzed condensation of a range of long chain ketones with glycerol gave only the 4-hydroxymethyl-2,2-disubstituted-1,3-dioxolane product (63). Thus 2-tridecanone was reacted with a slight excess of glycerol in toluene with a catalytic amount of *p*-toluenesulfonic acid monohydrate under Dean and Stark conditions to give 4-hydroxymethyl-2-methyl-2-undecyl-1,3-dioxolane as a yellow oil. NMR analysis shows that there was a slight impurity in this product and different methods of purification were tested in order to purify the product. However, both distillation and column chromatography gave rise to added impurities in the NMR analyses of this product. Therefore subsequent experiments were performed on the product without further purification.

3.2.2 Synthesis of sodium 4-[(2-methyl-2-undecyl-1,3-dioxolan-4-yl) methoxyl]-1-propane sulfonate (ALS-I)

The synthesis of ALS-I was carried out by reaction of the product of the previous synthesis, 4-hydroxymethyl-2-methyl-2-undecyl-1, 3-dioxolane, with a slight excess of powdered sodium hydroxide and 1, 3-propane sultone at 75 °C to yield the appropriate sulfonate compound. The product, a white solid, was recrystallised in ethanol and the product was filtered under nitrogen due to the hygroscopic nature of the product. Yamamura *et al.* achieved a better yield, 86%, than the yield achieved in these experiments, 50%, however this could be due to the larger scale of their reaction (63). Overall the synthesis of ALS-I is a straightforward two-step process with a good yield.

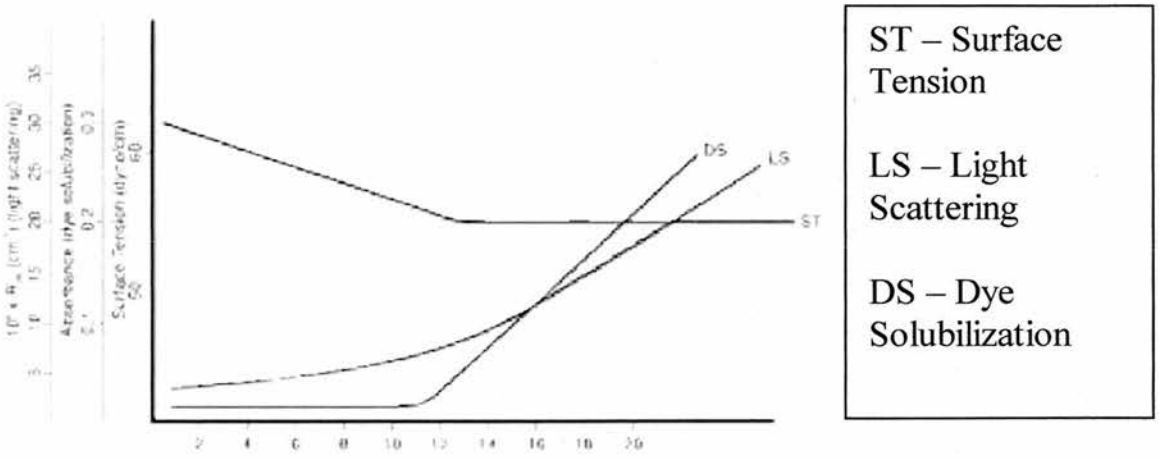
For the purposes of these experiments the ALS-I was needed in very small amounts and low concentrations, no more than 1mg/mL and as little as 0.05 mg/mL. Previous experiments in the literature used ALS-I at a concentration of 1 mg/mL, 0.1% (w/v), as a direct substitute for SDS, which is also used at 0.1% (57, 58).

3.2.3 Determination of Critical Micelle Concentration

The critical micelle concentration (CMC) of a detergent is defined as the lowest concentration at which detergent monomers aggregate to form micelles (51). It is normally given as a range of concentrations. Micelles are formed when detergents are added to water. The non-polar hydrophobic tail of the detergent clumps together and the hydrophilic polar head of the detergent molecule interacts with the water. The CMC is dependent on the length of the alkyl hydrophobic chain, such that the shorter the alkyl chains, the lower the CMC (76). The CMC of a detergent can be altered for

example by compounds such as urea, which breaks up the water structure thus increasing the CMC. The CMC of an ionic detergent can be reduced by the addition of counter ions that also increase micellar size.

Critical micelle concentrations can be determined in several different ways (51). These include measuring changes in surface tension, light scattering and dye solubilization. Surface tension can be used to measure the CMC because the surface tension decreases as the concentration of the detergent decreases and reaches a minimum around the CMC.



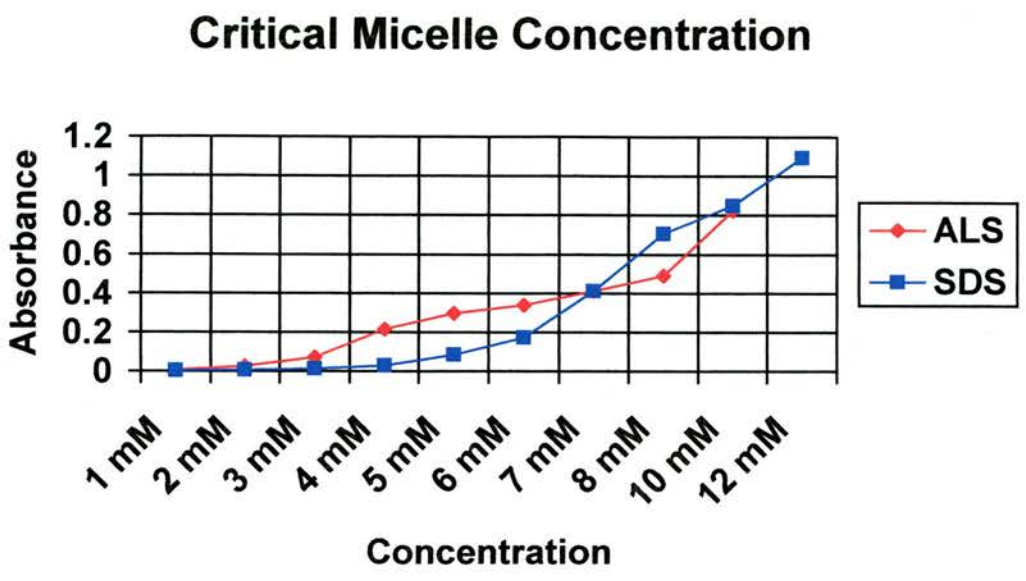
***Graph 1**: Determination of CMC by Three Different Methods*

Light scattering remains approximately constant until the CMC is reached. It then increases exponentially as the concentration of the detergent increases past the CMC. As can be seen from the graph above, the inflection point for light scattering is much less apparent than those given by dye solubilization and surface tension.

Dye solubilization, utilizing a water insoluble dye, is one method of determining the CMC that can be perfomed with simple resources. A specified amount of dye

dissolved in a volatile solvent is added to aqueous solutions of detergent at various concentrations. On evaporation of the solvent, the solutions will either absorb the dye, changing the colour of the solution, or the solution will remain colourless, based on whether or not the CMC has been reached. Spectrophotometric analysis of the solutions allows the determination of the CMC. The point of inflection on a graph of the concentration vs. the absorbance shows the point above which the dye solubilizes and micelles are formed.

In order for a detergent to be effective it must be used at a concentration above that at which it forms micelles. Thus it is important to determine the critical micelle concentration of the detergent. A study in 2003 by Konig *et al.* stated they were unable to determine the CMC of ALS-I (57). However in the original synthesis paper published in 1996 the CMC of ALS-I was determined by surface tension measurements to be 0.6 mM. For the purpose of the experiments to be performed it was important to determine the minimum concentration at which ALS-I would be useful. This was determined by the dye-solubilization method. The dye, 1-(2-pyridylazo)-2-naphthol (PAN), was water insoluble and only dissolved on addition of a suitable amount of ALS-I to the solution. Critical Micelle Concentrations are most often found to be within a small range. Here the CMC of ALS-I was determined to be in the range of 2.5 mM to 4 mM. The CMC of SDS, which is known to be 7 mM to 10 mM, was determined as a control and gave a CMC of 8 mM to 11 mM validating the method (51). Yamamura *et al.* determined the CMC for ALS-I to be 0.6 mM by measuring the surface tension of ALS-I containing solutions with a tensiometer. This differs by approximately a factor of four from the value obtained above.



***Graph 2**: Graph Depicting the Critical Micelle Concentration Determined by Dye Solubilization*

The difference in the CMC given in the literature and the CMC determined by dye solubilization could be due to several factors. Firstly, the method of dye solubilization relies on accurate detergent concentrations. The ALS-I was hygroscopic so the detergent concentration may have been overestimated. The material was also salty which would again cause overestimation. Secondly the point of inflection seen here is not very distinct. An alternative method such as tensiometer measurements might give an improved point of inflection. However the data showed that ALS-I had a CMC lower than that of SDS whilst still in the same order of magnitude range. It was therefore valid for previous researchers to use ALS-I at an equivalent concentration to SDS, although they perhaps should also have tried it at between one tenth and one half the concentration. It should also be noted that the CMC does not dictate the extent to which the mass spectrometric analysis will be affected by detergent ionization or suppression of signals.

3.3 Acid Labile Surfactants in Protein Digestion and Extraction

Initial experiments aimed to repeat the work of Nomura *et al.* that showed the improvement in mass spectrometric data obtained from in-SDS polyacrylamide gel digestion with ALS-I (56). As a prelude to this work digests of proteins in solution were analysed to investigate whether a similar effect also occurred under these conditions.

3.3.1 Solution Digest Experiments

Initial experiments were performed to determine the effect of ALS-I on solution digests. Solution digests were setup in duplicate, with half containing 0.1% ALS-I. Proteins were reduced by DTT and alkylated with IAA and digested overnight at 37 °C. Samples were spotted on the target plate using the dried drop application method co-crystallising with HCCA and analysed by MALDI-TOF MS. A β-galactosidase tryptic digest calibration spot was spotted for every eight samples analysed and Glu-Fibrinogen B (Glu-Fib) was added to sample spots as an internal lockmass calibrant.

Table 6: *MALDI-TOF MS Data for Solution Digest Trials*

Solution Digests	Score	% Coverage	Peptides Matched
1.25pm BSA	no ID		
1.25pm BSA w/ALS	65	18	10
2.5pm BSA	86	22	15
2.5pm BSA w/ALS	105	24	17
1.25pm ADH	no ID		
1.25pm ADH w/ALS	133	54	16

2.5pm ADH	no ID		
2.5pm ADH w/ALS	131	42	12

At 1.25 pmol BSA, the sample without ALS-I did not give enough peptides of sufficient intensity to allow identification. However the sample with ALS-I, all be it with a score that was not significant, was identified correctly and 10 peptides masses were matched. The sample without ALS-I had no identification. At 5 pmol BSA both samples had MOWSE scores that are considered significant. However the sample containing ALS-I matched more peptides and had slightly greater protein coverage. The results for ADH were even more distinct. As can be seen from the table above, two concentrations of ADH, 1.25 pmol and 2.5 pmol, did not yield sufficient peptide signals to result in an identification without ALS-I but had the correct identification on addition of ALS-I. BSA has a peptide of similar mass to the Glu-Fib lock mass, so this BSA data is not lock mass adjusted and so was searched with a wider peptide mass tolerance. Hence the poorer scores seen for BSA solution digests. These tests were repeated several times and the scores varied between the trials, but the solution digests performed in the presence of ALS-I displayed more peptides and so gave better Mascot data.

3.3.2 In-Gel Digestion Experiments

Proteins obtained from Sigma were first separated by 1D SDS-PAGE, stained with Coomassie and destained. The gel bands were then excised and washed with water and acetonitrile to remove excess SDS and then gel pieces were incubated in dithiothreitol (DTT). Incubation with DTT reduces the protein disulfide bonds making it more accessible to tryptic cleavage. The DTT was followed by incubation with

iodoacetamide (IAA). IAA alkylated the free thiols to ensure disulphide bonds do not re-form and that all cysteines are modified identically. Trypsin was added to digest the proteins C-terminal of lysine and arginine residues. Reducing and then alkylating the proteins before digestion augments protein coverage by trypsin.

Previous experiments by Nomura *et al.* showed a marked improvement in peptide recovery from silver stained SDS-polacrylamide gel digestions, particularly digestions of hydrophobic proteins. They showed that pre-treatment of samples with ALS-I at an equal concentration to that which SDS is normally used, peptide recovery for 100 fmol gel loadings were greatly increased. BSA samples pre-treated with ALS had a peptide recovery of 9±1. Analysis of BSA left untreated only recovered 4±1. Similar results were found for a series of proteins including transferrin, SNAP-beta and CRMP-2.

A number of proteins of varying molecular weight and a range of concentrations were subjected to in-gel digestion in the presence of 0.1% ALS-I. Control samples were prepared simultaneously without any detergent in the trypsin solution. These samples were analysed by MALDI-TOF MS and nano LC-ESI MS/MS and the data submitted to the Mascot search engine to determine protein identity.

In order to analyse the samples by MALDI-TOF MS in conjunction with sample spots, calibration samples were spotted containing β-galactosidase tryptic digest. Sample spots contained Glu-Fib to enable lock mass correction, correcting the shift in mass of a peak with a known mass. Once sample data was acquired, spectra were calibrated over the calibration curve of β-galactosidase tryptic digest.

Results are given for samples analysed before ALS-I degradation. Data from Waters Corp. indicates that it is beneficial to degrade ALS-I prior to analysis by MALDI-TOF and LC-ESI MS. However these experiments did not find evidence to substantiate that directive. Samples analysed by MALDI-TOF MS where the ALS-I was decomposed did not identify any proteins tested. It may be that decomposition of the ALS-I stopped peptides from being soluble and allowed for interference of any contaminants located in the in-gel extraction solution thus suppressing peptide signals and reducing the number of correct protein identifications.

It is necessary to note that digestion of proteins and extraction of peptides is not a 100% efficient process. There is no guarantee that a protein will be completely cleaved by trypsin and all the peptides will be successfully extracted from the gel. A study performed by Speicher *et al.* found that the loss of peptides in tryptic digest and extraction procedures amounted to between 14% and 20% (78). Sample was lost during destaining and peptides were lost to the digest tube and the pipette tip and some peptides remained in the gel. The percentage of losses remained fairly constant regardless of the amount of protein loaded on the gel. Samples of weak protein concentration (1.25 pmoles and 2.5 pmoles) that were not treated with ALS-I were often unidentified by MALDI-TOF MS. However samples at lower protein concentrations that were treated with ALS-I had a higher rate of identification, often with a significant MOWSE score.

Samples that are unidentified by initial MALDI-TOF MS analysis can be subjected to sample clean up using ZipTips which contain a small reverse phase column in a micropippette tip. This procedure allows MALDI-TOF MS signals to be obtained

from more dilute peptide extracts as peptide extracts in larger volumes can be concentrated onto the ZipTip resin, contaminants washed away and the peptides eluted in a smaller volume thus concentrating them. ZipTip concentration was not explored with these samples.

The results showed that digests in the presence of ALS-I gave a significantly improved number of peptides, particularly at lower concentrations and with smaller proteins. This therefore improved the Mascot score. Results were good for a broad range of proteins covering a broad range of masses from 7 kDa to 115 kDa. A selection of the MALDI-TOF MS results from manual in-gel digestions are given below. It should be noted that the amount of protein given in the table is the amount of protein that was loaded on the gel. Subsequent digestion and peptide extraction would not have been completely efficient reducing the amount of sample available for analysis. Added to this is the fact that only 0.5 µL of the 20 µL peptide extract is eventually spotted on the target plate. If digestion and extraction were 100% efficient, this would still only be one fortieth of the original amount of protein, and in reality the amount of sample will be less than this. This gives less than 60 fmoles loaded from 2.5 pmol bands and less than 125 fmoles loaded from 5 pmol bands.

Table 7: *MALDI-TOF MS Data for In-gel Digestion Trials*

Manual Digests	Score	% Coverage	Peptides Matched
2.5pm BSA	183	31	16
2.5pm BSA w/ALS	374	52	29
5 pm BSA	339	49	27

5 pm BSA w/ALS	270	45	27
2.5pm ADH	no ID		
2.5pm ADH w/ALS	74	23	6
5pm ADH	93	33	8
5pm ADH w/ALS	89	23	6
2.5 pm CA	no ID		
2.5 pm CA w/ALS	125	32	7
5 pm CA	90	23	6
5 pm CA w/ALS	136	35	8
2.5 pm Aldolase	90	23	7
2.5 pm Aldolase w/ALS	176	43	10
5 pm Aldolase	164	48	15
5 pm Aldolase w/ALS	170	34	12

pm = pmoles

Whilst the data obtained for 5 pmoles of protein loaded on a gel was similar with and without ALS-I in the digests, overall, the samples above benefited from the addition of ALS-I. It can be seen from the data above that lower concentrations of proteins showed improved results upon the addition of ALS-I. The 2.5 pmol samples of carbonic anhydrase and ADH had no identification without ALS-I but significant identifications with ALS-I. Some of the higher molecular weight and higher concentration samples were less affected by the ALS-I and thus the scores varied little between the sample containing ALS-I and the sample without ALS-I.

ALS-I was shown to improve protein identification via in-gel digestion. It is believed this is due to solubilization by ALS-I, which means more peptides are transferred to the target. ALS-I may also compete off any SDS still bound. As previously stated SDS suppresses protein and peptide ionization. Thus solubilization by ALS-I could cause less suppression and therefore improved peptide signals, which would be more likely to lead to sample identification.

Smaller proteins can be harder to identify by MALDI-TOF MS due to the fewer number of peptides available for identification. A rule of thumb suggests it is necessary to have five peptide matches for a valid identification. Some small proteins just do not give five tryptic peptides of suitable mass.

The smallest protein analysed was ubiquitin. This protein is 8 kDa. It has only twelve predicted tryptic peptides if digestion is completely efficient. However of those twelve peptides only seven are within the mass region where peptides are likely to be seen. As it is unusual for all predicted peptides of a suitable mass to ionize, this makes identification difficult, as there will probably be fewer than seven peptide masses obtained by MALDI-TOF MS analysis for mass matching

.

Table 8: *MALDI-TOF MS Data for In-gel Digestion of Ubiquitin*

Manual Digests	Score	% Coverage	Peptides Matched
2.5 pm Ubiquitin	51	50	3
2.5 pm Ubiquitin w/ALS	70	58	4
5pm Ubiquitin	no ID		

5pm Ubiquitin w/ALS	71	47	5

As can be seen from the data above at 2.5 pmol of Ubiquitin the score was more significant for the sample with ALS-I than the sample without ALS-I. This was the difference of matching just one more peptide. At 5 pmol Ubiquitin there was no identification without ALS-I and the identification with ALS-I matched five of the twelve possible peptides overall.

The variation in the sample spot can make protein identification difficult for small proteins as well as large proteins. Some basic proteins give predominantly small peptides upon tryptic cleavage. Peptides below 800 Da fall in the region of the spectrum where matrix signals appear. The MALDI-TOF mass spectrometer is set up so that the matrix suppression lens stops those masses, the region from 0-500 Da, from being transmitted. The use of the suppression lens stops matrix signals from completely saturating the detector at the expense of peptide signals. However it also suppresses the transmission of any peptides that fall in this mass range. Acquiring good data relies on finding a 'sweet spot' where the sample particles fly well, there isn't too much matrix but there is good co-crystalllisation between the sample and the matrix.

Tests were also performed on the proteins transferrin and β-galactosidase (β-gal), results shown below. These proteins are both larger than BSA at 78 kDa and 117 kDa.

Table 9: *MALDI-TOF MS Data for In-gel Digestion of Transferrin and B-gal*

Manual Digests	Score	% Coverage	Peptides Matched
1.25 pm Transferrin	68	9	6
1.25 pm Transferrin w/ALS	100	13	8
2.5pm Transferrin	187	28	19
2.5pm Transferrin w/ALS	255	35	21
5pm Transferrin	145	19	14
5pm Transferrin w/ALS	270	40	26
2.5pm β-gal	no ID		
2.5pm β-gal w/ALS	184	17	16
5 pm β-gal	221	25	21
5 pm β-gal w/ALS	210	22	23

The results for transferrin give all significant MOWSE scores however, the samples with ALS-I have much higher scores and at 5 pmol transferrin loaded on a gel nearly double the number of peptides were matched from the in-gel digest in the presence of ALS, compared to the number of peptides matched without. At 2.5 pmol β-gal the sample without ALS-I gave no identification, whereas with ALS-I a clear identification was obtained. As noted previously, at higher protein concentrations the effect is less pronounced. Thus at 5 pmol β-gal the scores for both samples were comparable and significant.

Some gel slices were tryptically digested using the Genomic Solutions ProGest in-gel digestion robot. The ProGest robot is a high throughput robot capable of digesting up to 96 samples per run. Its dual needle system was utilized in order to perform the in-gel digestions with and without ALS-I. The standard protocol was amended in order to allow the addition of the ALS to the trypsin just prior to the addition of the trypsin to the samples. This was necessary as the trypsin is made up in a solution of 0.01% formic acid at pH less than 2. At this pH ALS-I is readily degraded. Similar results to those obtained from manual in-gel digests were acquired upon MALDI-TOF analysis and are shown in the chart below.

Table 10: *MALDI-TOF MS Data for Robot In-gel Digestions*

Robot Digests	Score	% Coverage	Peptides Matched
2.5pm BSA	189	36	21
2.5pm BSA w/ALS	167	36	20
5pm BSA	170	37	20
5pm BSA w/ALS	236	37	21
2.5pm ADH	83	32	8
2.5pm ADH w/ALS	142	45	14
5pm ADH	114	40	12
5pm ADH w/ALS	96	36	10
5pm Ubiquitin	82	69	7
5pm Ubiquitin w/ALS	88	65	4

This data shows that the presence of ALS-I in the digestion solution has made a significant increase in the MOWSE score for 5 pmol BSA and 2.5 pmol ADH. The results at 2.5 pmol BSA and 5 pmol ADH gave significant MOWSE scores for samples with and without ALS-I but samples without ALS-I gave slightly higher scores. ALS-I made a slight difference in the results given for the two different solutions of ubiquitin, however both were considered significant MOWSE scores. There are many possible explanations for this phenomenon. In MALDI-TOF MS, there is a certain amount of variability in the quality of a sample spot and whether or not a 'sweet spot' is found can depend on the number of areas on the spot that are analysed as well as the extent of co-crystallisation of the matrix and the analyte. Thus the quality of data acquired is dependent on the quality of the sample areas analysed. Another variable in MALDI-TOF analysis of peptides is the peak selection prior to submitting a peak list for MASCOT database searching. The selection of peaks is initially dictated by the Protein Probe software, however manual intervention I used to remove the Glu-Fib reference peaks and add other peaks judged as real to the peak list. The software is also not capable of determining what the first peak in a series is if the first peak happens to fall below the threshold quotient. In order to minimize the variability in sample spots, the same number of spectra is acquired over the same number of areas on the spot. However, this cannot guarantee that equivalent sweet spots are located. Overcoming the variability in peak selection is difficult as it is controlled by the software user and not the software itself.

Use of the ProGest Robot showed ALS-I could be used in high throughput experiments. While manual digestion showed better MALDI-TOF MS results overall,

the samples digested by the robot were only slightly poorer than those achieved by the manual digests, and this could have been due to spot variability.

Samples were also analysed by nano LC-ESI MS/MS. For these experiments the ALS-I was first decomposed, as the samples were loaded in 5% formic acid. Here the peptides were bound first to a reverse phase trap at low acetonitrile concentration, after washing they were eluted off the trap and separated on a 75 μm ID PepMap C18 column with an acetonitrile gradient. Proteins above 40 kDa were analysed using an 85-minute gradient, while those of lower molecular weight were analysed using a 55-minute gradient. The data was acquired in information dependent acquisition mode with a 1 second MS cycle to assess peptides eluting followed by two 3 second MS/MS cycles selecting the two most intense ions for fragmentation, unless they had been analysed in the preceding 60 second. The IDA data file was submitted to Mascot MS/MS Ion Search analysis. The amount of protein given in the table is the amount of protein that was loaded on the gel. The amount of protein from the digestion and extraction that was analysed was half the total. Again the digestion and extraction efficiencies were unknown.

Table 11: *LC-ESI MS/MS Data for In-gel Digestion Trials Upon Decomposition of ALS-I*

Manual Digest	Score	% Coverage	Peptides Matched
0.16pm BSA	127	4	3
0.16pm BSA w/ALS	267	8	5
2.5pm BSA	821	25	24

2.5pm BSA w/ALS	1634	56	61
5pm BSA	1744	53	88
5pm BSA w/ALS	2040	60	106
0.16pm ADH	142	7	3
0.16pm ADH w/ALS	122	4	2
0.625pm ADH	277	14	6
0.625pm ADH w/ALS	315	19	9
2.5 pm Aldolase	373	24	9
2.5 pm Aldolase w/ALS	590	31	20
2.5 pm ubiquitin	keratin		
2.5 pm ubiquitin w/ALS	196	61	5
5pm Ubiquitin	58	30	3
5pm Ubiquitin w/ALS	186	36	3

Note: Some peptides are analysed more than once either because the doubly and triply charged ion is subjected to fragmentation or because its elution profile trails due to its signal intensity past the 60 second time exclusion and so the ion is re-selected for fragmentation.

These results show large differences in the significant scores between the ALS-I treated samples and the untreated samples. Both the BSA samples with ALS-I and without ALS-I had very significant MOWSE scores. Although ALS-I has improved the scores for the 2.5 and 5 pmol BSA data, the data from ubiquitin is actually more

informative. Once a score is highly significant such increases add little to the identification unless one is attempting to identify a post-translational modification, in which case the best coverage possible is required. At lower protein concentrations, e.g. 0.16 pmoles BSA on gel data and weaker samples, the difference between samples digested in the presence of ALS-I and without is the difference between a barely significant score and a definitive score. Likewise, ubiquitin is a small protein with a small number of peptides to identify. At 2.5 pm ubiquitin loaded on the gel, ALS-I was the difference between a significant identification of the protein with ALS-I and no identification for the sample without ALS-I. At a protein amount of 5 pm ALS-I made the difference between a significant identification of ubiquitin and an insignificant score for the sample left untreated. The differences in these scores is much more valuable than the difference in the scores of the BSA samples, which would have been identified regardless of the addition of ALS-I.

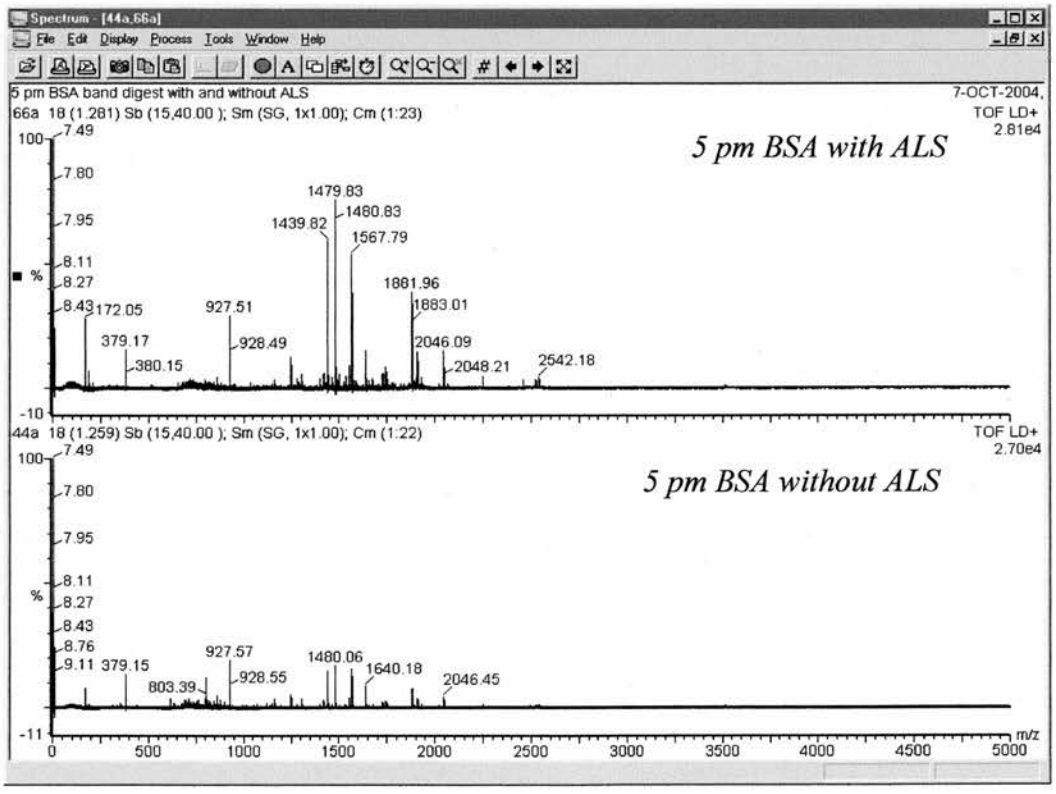
Nano LC-ESI MS/MS analysis gives larger MOWSE scores. This is because the fragmentation data obtained by collision-induced dissociation (CID) of peptides provides sequence information, which is much more specific for a given protein. Indeed one peptide can be enough to identify a protein from MS/MS data. Again the more significant scores obtained from digestions performed in the presence of ALS-I can be attributed to increased solubility of peptides, leading to better extraction from the gel and less losses on plastic ware.

As with the MALDI experiments the effectiveness of ALS-I in improving the coverage from in-gel digestions was tested for a range of protein sizes and concentrations. The larger the protein the more likely it is that there are peptides that

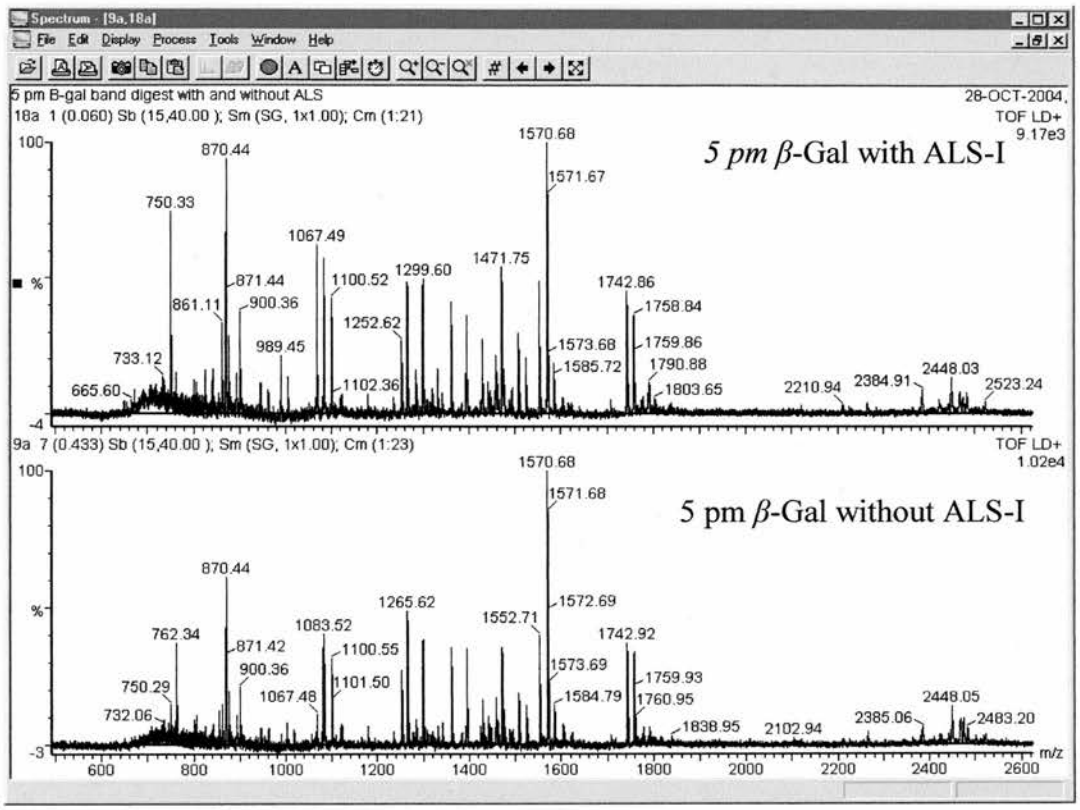
lie in the ideal region between 800 Da and 2500 Da. The implication is that a larger protein was more likely to be identified without the need for ALS-I, except at the lowest protein concentrations. As stated previously, in MALDI-MS analysis peptides that fall below 500 Da will not be seen due to matrix interference. Peptides that fall above 2500 Da are generally unidentified due to poor resolution on our instrument. The first isotope is indistinct and so is not represented in the peak list.

The spectra below demonstrate the difference in peptide signals between samples with ALS and without ALS. The MOWSE score for the sample without ALS-I was 135 and the MOWSE score for the sample with ALS-I was 203.

Spectra 1 and 2: *Spectra of 5 pmoles BSA with and without ALS-I*



***Spectra 3 and 4**: Spectra of 5 pmoles β-Galactosidase with and without ALS-I*



The spectra above depicts how ALS-I is an effective detergent for solubilizing peptides and increasing peptide extraction from SDS-PA gels. The MOWSE scores were 281 for the sample without ALS-I and 355 for the sample with ALS-I.

ESI analysis also has a working mass range of between approximately 800 Da and 2500 Da. Peptides are normally analysed as doubly and triply charged ions. Singly charged ions are normally not peptides and highly charged ions make MS/MS difficult to interpret. As peptides tend to hold one charge per 1000 Daltons of molecular weight, peptides in the mass range 800-2500 Da tend to be doubly or triply charged.

LC-ESI MS/MS data was collected on small concentrations of protein. This is viable as LC-ESI MS/MS data provides sequence information as well as peptide molecular weights. Peptide sequence information is powerful as only a few peptides are required to ionize well enough to obtain some sequence data, and hence an identity. Weaker samples and smaller proteins as well as mixtures of peptides can be successfully analysed by LC-ESI MS/MS. The addition of ALS-I seems to have the most profound effect on weaker samples and smaller proteins where it can make the difference between identification or not.

Analyses by LC-ESI MS/MS showed good, consistent results for proteins as small as 8 kDa (Ubiquitin). These results were more consistent than those given by MALDI-TOF analysis. This is because LC-MS is a much more reproducible technique than MALDI, where signal intensity depends on the quality of the sweet spot located. LC-ESI MS/MS data was gathered on proteins treated with ALS-I ranging in molecular weight between 8 and 78 kDa and analysing half the sample derived from loading between 0.16 pmoles and 5 pmoles loaded on a polyacrylamide gel. Proteins treated with ALS-I above 36 kDa were consistently identified at all concentrations. Protein digests performed without ALS-I did not give consistent correct protein identifications for ubiquitin at these low picomolar amounts.

Further MALDI-TOF MS data for in-gel digestions with and without ALS-I can be found in Appendix I.

Further LC-ESI MS/MS data for in-gel digestions with and without ALS-I can be found in Appendix II.

3.4 Whole Protein Extraction From SDS-Polyacrylamide Gels

Proteins were separated by 1D SDS-PAGE. Bands containing between 50 pmoles and 300 pmoles of protein were excised and subjected to a series of washes, including washes with 10% acetic acid, acetonitrile, methanol, and FWI. FWI was used to destain excised bands stained with Coomassie stain. Following the destaining of the protein bands, a volume of extraction solution was added that just covered the protein slices that contained varying amounts of ALS-I, DTT and tris/glycine. Nine solutions were made comprising of these three components at different concentrations, in order to establish the optimal extraction conditions. The concentrations of each component in the extraction solutions are shown below. The initial extraction time was 1 to 2 hours.

Table 12: *Content of Extraction Solutions*

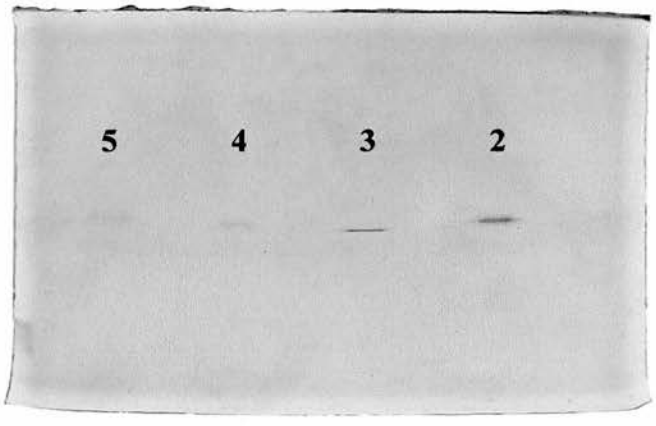
Solution	ALS-I	Tris/glycine*
1	1 mg/mL	None
2	0.5 mg/mL	0.25 M
3	0.25 mg/mL	0.25 M
4	0.1 mg/mL	0.25 M
5	0.05 mg/mL	0.25 M
6	1 mg/mL	0.25 M
7	1 mg/mL	0.1 M
8	1 mg/mL	0.2 M
9	0.25 mg/mL	0.1 M

*Initial extractions were all run with 1M tris/glycine in the extraction solution. It was only upon re-optimisation that the tris/glycine concentration was varied.

Extraction efficiency was tested by running half the total extract volume on a SDS-polyacrylamide gel. The standard proteins tested were myoglobin and superoxide dismutase (SOD).

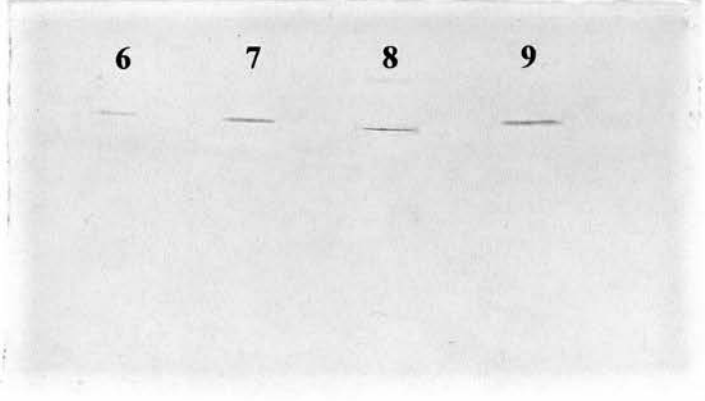
Gel 1: *Myoglobin Whole Protein Extraction Efficiencies for Extraction Solutions 5, 4, 3 and 2.*

10 μL of the total extract of 20 μL was loaded onto the gel.



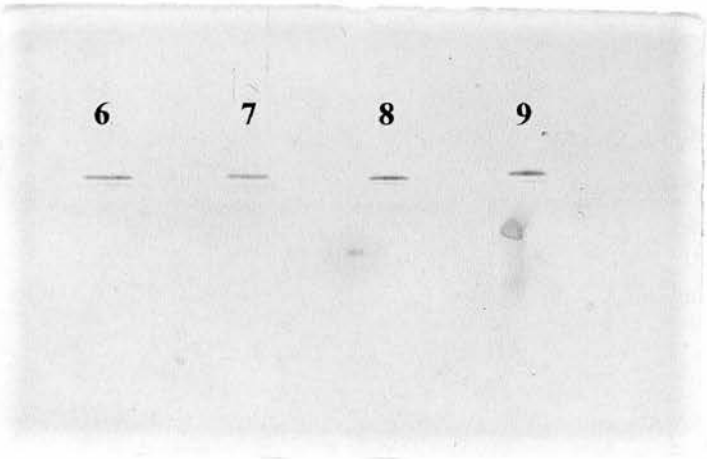
Gel 2: *Myoglobin Whole Protein Extraction Efficiencies for Solutions 6, 7, 8 and 9.*

10 μL of the total extract of 20 μL was loaded onto the gel.



As can be seen from the gels above the most effective extraction solutions for myoglobin were solutions, 2, 8 and 9. Similar results were obtained for the protein SOD.

Gel 3: *SOD Whole Protein Extraction Efficiencies for Solutions 6, 7, 8 and 9.*



It was estimated that between 25 and 30 pmoles of extracted protein was present in the most efficient 50 pmol extractions. This implied that extractions performed with ALS-I appeared to have optimal extraction efficiency between 50% and 60%. Neither of the two major studies of passive elution of proteins for MS presented quantitative data as to the efficiency of their protein extracts (71, 73). Cohen and Chait extracted their protein directly into a solution containing a MALDI matrix and analysed the protein-matrix co-crystals. This was determined to be the most effective method by Cohen and Chait. Thus it was not possible to determine protein extraction levels.

As the gels show, the most effective extraction solutions for both proteins were 8 and 9. However for SOD, extraction solutions 6 and 7 were just slightly less effective than 8 and 9, whereas for myoglobin solution 6 was considerably less effective than solutions 8 and 9. It was decided solution 9 would be used in further studies.

After determining the most effective solutions for extraction, the extraction method was further optimised by adjusting the order of the incubations and lengthening the destaining time in order to determine if this had any effect on the level of protein extraction from the gel. It was determined that the most effective extraction method called for the methanol incubation before the acetonitrile incubation, contrary to the method of Mirza *et al.* However, this only gave a slight improvement in the level of overall extraction. Lengthening the destaining time had no effect past the complete destaining of the gel slice. In order for extraction to be efficient, gel slices had to be completely destained. Destaining times varied over the range and concentration of proteins analysed.

3.4.1 Neutral Matrices and Whole Protein Extraction

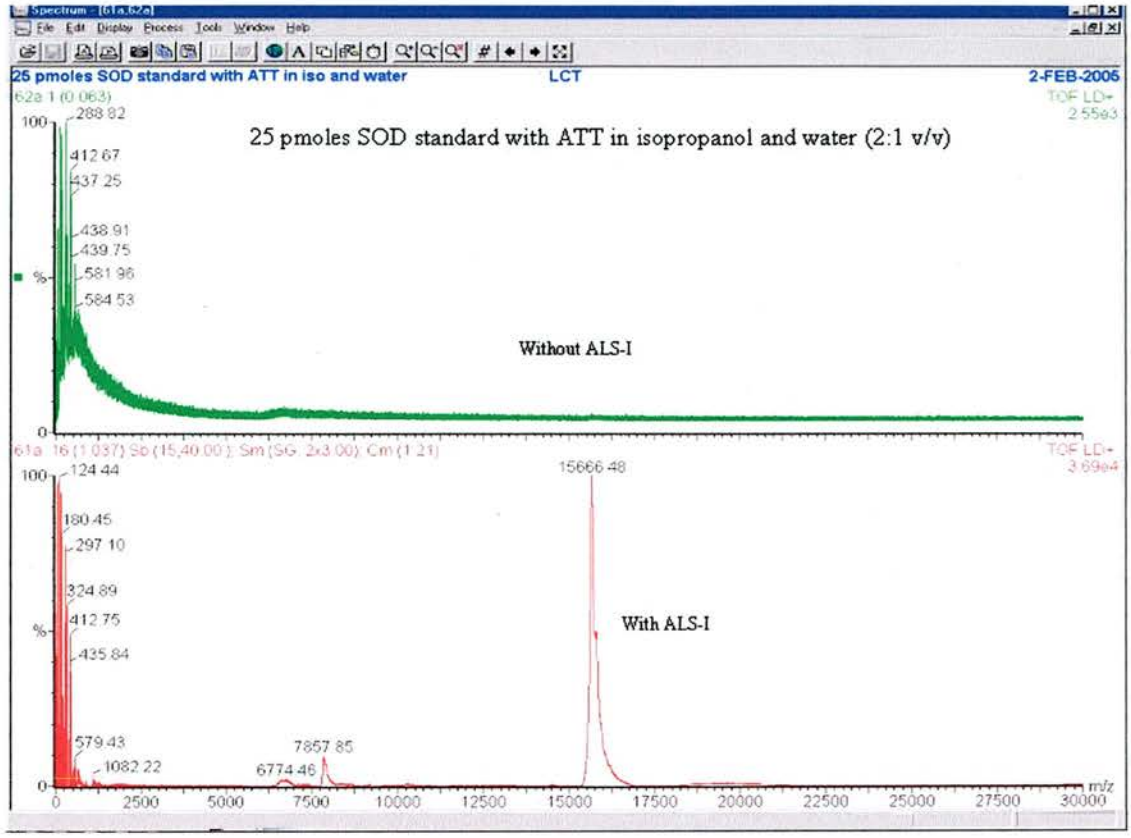
Three neutral matrices were tested for their compatibility with protein extraction aided by ALS-I and subsequent MALDI-TOF MS. They had previously been identified as compatible with MALDI-TOF MS analysis and useful for obtaining protein spectra under neutral conditions (41). However, this work used the dried drop method of crystallisation. The matrices had not been used in a slow crystallisation procedure before.

The three matrices investigated here were 2-amino-4-methyl-5-nitropyridine (AMNP), 4-nitroaniline (NA) and 6-aza-2-thiothymine (ATT). Each matrix was dissolved in several different solutions: formic acid/acetonitrile/isopropanol/water (50:25:15:10 v/v/v/v), formic acid/acetonitrile/isopropanol/water (20:25:15:40 v/v/v/v), isopropanol/water (2:1 and 1:1 v/v), citric acid/water (1:1 v/v) and distilled water. The properties of each matrix were investigated to determine ideal

concentration (saturated or 10 mg/mL), solvent system and compatibility with the dried drop crystallisation method and slow crystallization. The proteins myoglobin and SOD were used as standards.

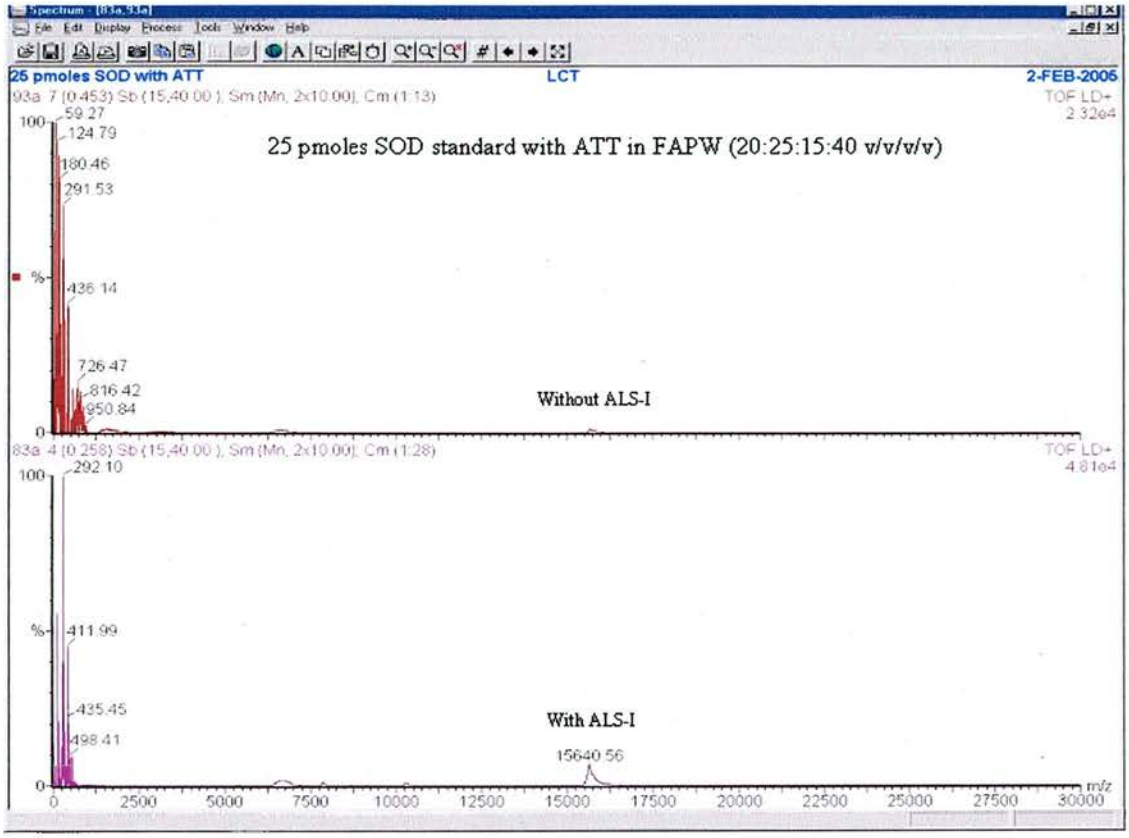
The ideal solvent system for slow crystallisation of 2-amino-4-methyl-5-nitropyridine was determined to be a saturated solution in formic acid/acetonitrile/isopropanol/water (20:25:15:40 v/v/v/v). The ideal solvent system for slow-crystallisation of 6-aza-2-thiothymine was determined to be a saturated solution in isopropanol/water (2:1 v/v). The ideal solvent system for slow crystallisation of 4-nitroaniline was also determined to be a saturated solution in isopropanol/water (2:1 v/v). Each matrix solution was dissolved in the solvent system at concentrations of 10 mg/mL and saturated. It was found that these matrices were readily soluble in the above solvent systems and thus saturated matrix solutions were necessary to promote crystallisation. Initial experiments were performed with an ALS-I concentration of 1 mg/mL.

***Spectra 5 and 6**: Spectra Showing 25 pmoles Slow Crystallisation of SOD with ATT in Isopropanol and Water.*



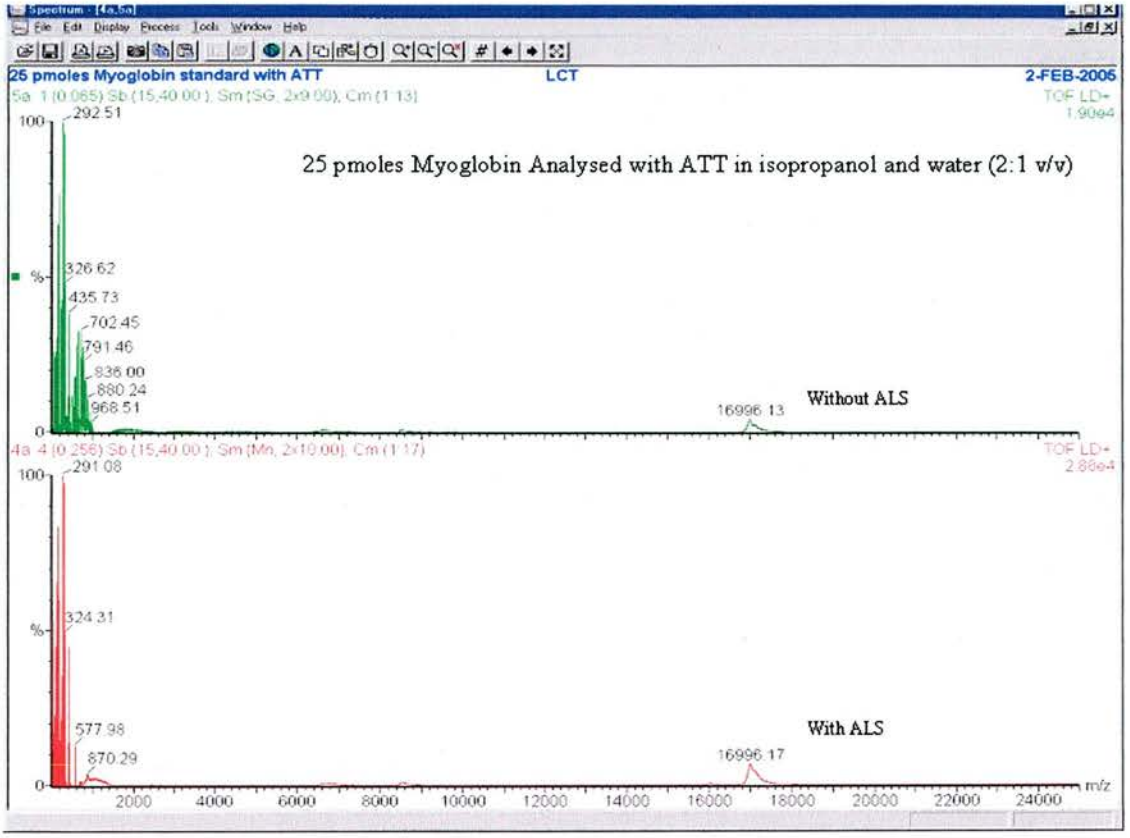
As can be seen from the spectra above and the spectra below, signal intensity can be directly related to the solvent conditions of the matrix. It should also be noted that neither of the standard SOD samples without ALS-I gave signals for either matrix. This was the case in most of the experiments on standard proteins and always the case in experiments on protein extracts.

***Spectra 7 and 8**: Spectra Showing 25 pmoles Slow Crystallisation of SOD with ATT in FAPW*



The spectra of 25 pmoles myoglobin demonstrate the difficulty of MALDI-TOF MS. ATT in isopropanol and water gave very good signals for 25 pmoles SOD but weak signals for 25 pmoles myoglobin. Despite the weak intensity of the signal, the peak is quite sharp allowing for greater accuracy when determining the mass of the protein.

***Spectra 9 and 10**: Spectra Showing 25 pmoles Slow Crystallisation of Myoglobin with ATT in Isopropanol and Water*

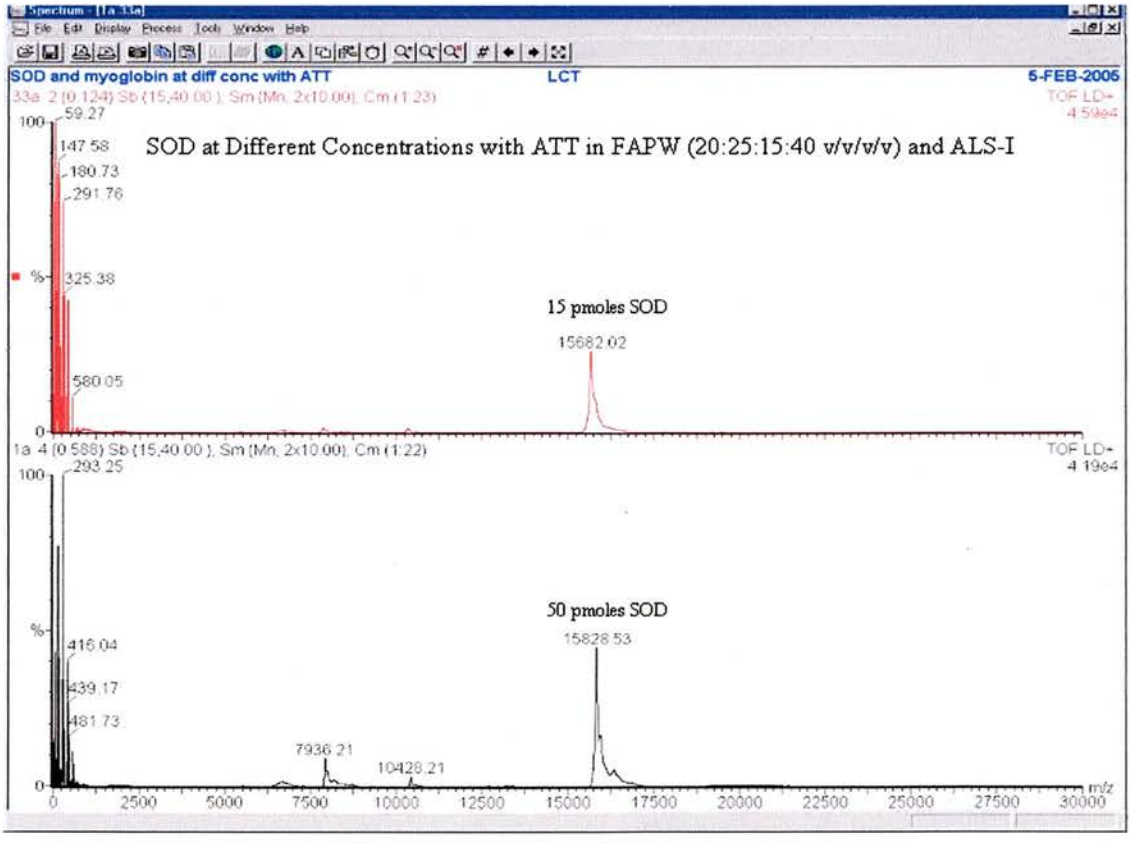


Subsequent experiments were designed to determine the compatibility of these matrices in conjunction with protein standards combined with ALS-I and then to optimize the slow crystallisation process. These matrices were determined to be unsuitable for the dried drop method in conjunction with whole proteins. Fitzgerald *et al.* found the mass limit for these neutral matrices to be 12,000 Da (41). In order to investigate slow crystallisation, fifty picomoles, 5 µL, of each protein were slow crystallised with each matrix, 15 µL in each of the solvent systems. Slow crystallisation involves the addition of a saturated matrix solution to a whole protein sample. The mixture was left open to the air for 2-3 hours by which time co-crystals of matix/analyte were formed. These crystals were washed and then applied directly to the MALDI target plate. The solvent systems giving the sharpest and most intense

peaks were subsequently tested with smaller amounts of protein, ALS-I at different concentrations and finally whole protein extracts from SDS-polyacrylamide gels.

Slow crystallisation was tested with 50, 25, 15 and 10 pmoles of SOD and myoglobin. The lowest amount of standard protein as well as extracted protein that gave good MALDI-TOF MS signals was 10 pmoles. Signals for this series of protein concentrations were achieved for standard myoglobin and SOD by slow crystallisation with AMNP and ATT in their ideal solvent systems. Signals were also achieved for extracted myoglobin and SOD by slow crystallisation with AMNP and ATT in several solvent systems.

***Spectra 11 and 12**: Spectra Showing Different Concentrations of SOD Slow Crystallised with ATT in FAPW*



With smaller amounts of protein, 10 pmoles or less, co-crystals did not always readily form with the three neutral matrices being examined in isopropanol and water. Once the samples had been left on the open bench for up to twelve hours, adding 10 μL of 1% formic acid and leaving the microcentrifuge tubes open on the bench until crystals formed remedied this. This does destroy the ALS, but matrix/analyte co-crystals were formed. Despite this problem, the matrices ATT and NA produced their most ideal whole protein signals under these conditions. The average mass for standard SOD slow crystallised with ATT was calculated at 15,730 Da averaged over ten spectra, which is a mass error of 139 Da on average. The average mass for extracted SOD slow crystallised with ATT was calculated at 15,752 Da over ten spectra, which is an average mass error of 161 Da. The average masses for standard and extracted myoglobin slow crystallised with ATT were calculated at 17,085 Da and 17147 Da respectively. Each was calculated over ten spectra and average mass errors for myoglobin comparable to those of SOD, 134 Da and 196 Da respectively. Both SOD and myoglobin show an increase in average mass error from standard protein to extracted protein. The average mass was not altered extensively by changing the matrix to AMNP or NA.

The average mass errors are not dissimilar to the error calculated for standard solutions of SOD and myoglobin co-crystallised with a supersaturated solution of HCCA in 35% aceotnitrile/65% 0.1% TFA. In order to supersaturate the solution the matrix is heated up to 40 °C prior to addition to the sample solution. Once the two are spotted on the target plate, a line is scratched across the spot in order to encourage crystallisation around a centre. The mass error found for SOD co-crystallised with HCCA in this way was ± 100 Da and the mass error found for myoglobin was ± 200

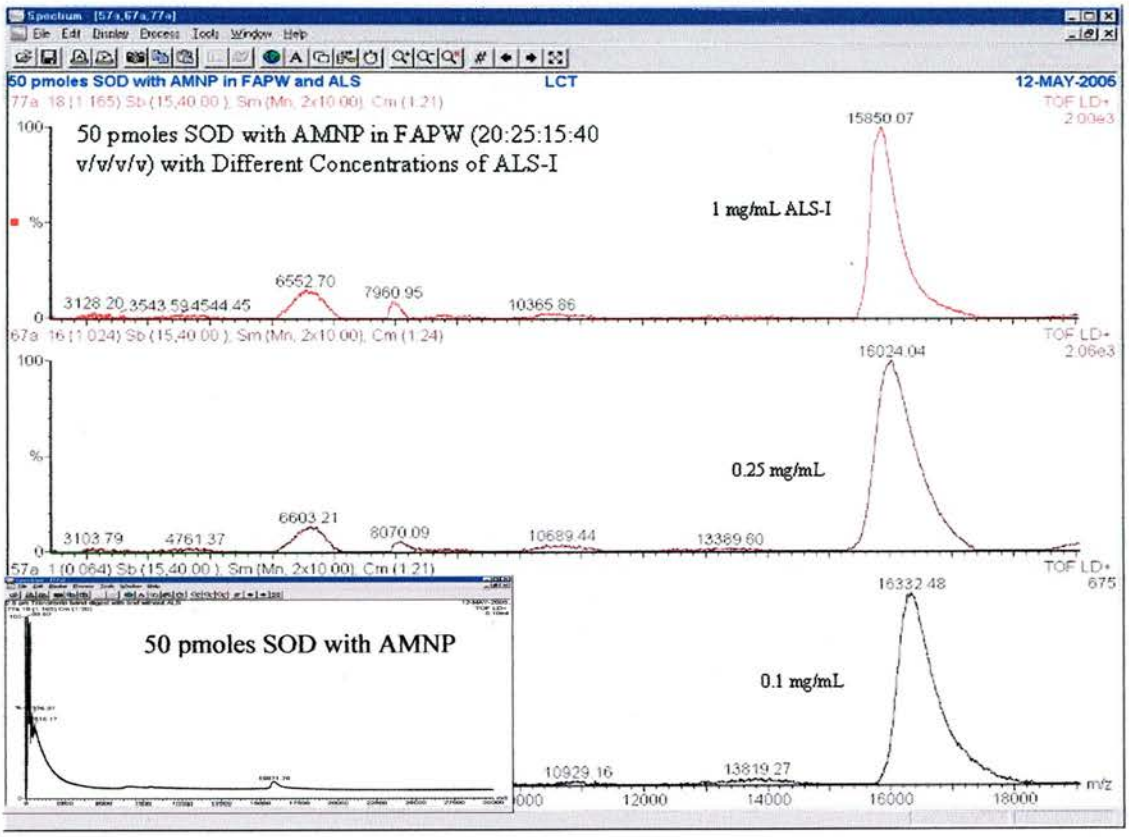
Da. This confirms the problem with whole protein analysis by MALDI-TOF MS; the peak width of the signals is so broad that relatively large discrepancies in mass can come about from minor variations in peak shape. While MALDI-TOF MS is capable of determining the mass to within a few hundred Daltons, this is insufficient accuracy for example to determine post-translational modifications or to determine the site of any proteolytic clipping of the protein.

The next factor investigated in the slow crystallisation of proteins with neutral matrices was the effect of the amount of ALS-I in the protein. This factor was examined in order to determine to what extent ALS-I had an enhancing or detrimental effect on signal intensity. ALS-I may have improved the solubility of the proteins to the extent that co-crystallisation of the matrix and protein was affected or otherwise interfered with the slow crystallisation process. However, one of the strengths of slow crystallisation with HCCA is that it is very tolerant of detergents. It was determined that co-crystallisation was occurring at 0.1 mg/mL, 0.25 mg/mL and 1 mg/mL ALS-I however at 0.1 mg/mL the data obtained gave much better spectra with sharper signals than at the higher concentrations of ALS-I. This can be seen from the spectra of 50 pmoles SOD co-crystallised with AMNP in FAPW at all three concentrations of ALS-I. Comparison of spectra obtained in the presence and absence of ALS-I showed that the addition of ALS-I had a signal enhancing effect on neutral matrix slow crystallisation.

The protein signals in the spectra below are all comparable in size and intensity. Mass resolution for all three peaks was roughly the same. As can be seen in the insert, the peaks that appear at ~6500 Da in the smoothed spectra are insignificant in relation to

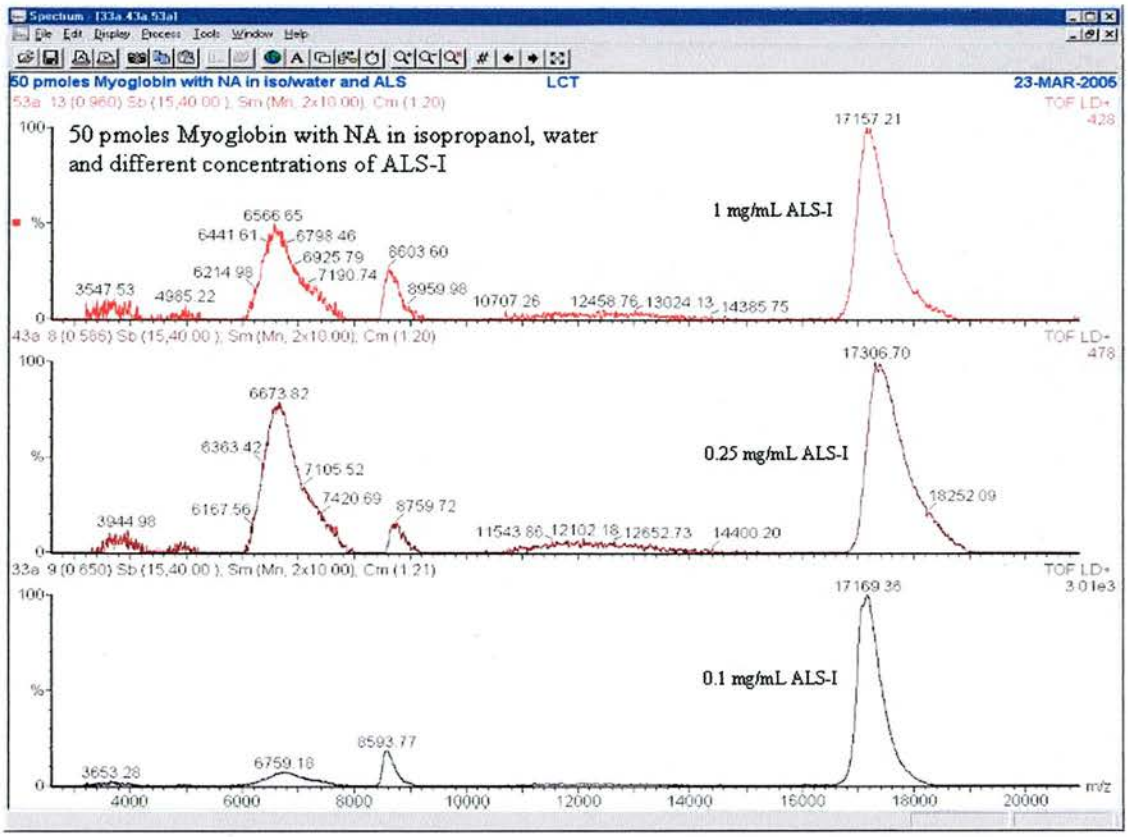
the overall spectrum and most likely appear at that point due to the falling off of suppression, which occurs over a range of 0-5000 Da.

***Spectra 13, 14 and 15**: Spectra Showing Slow Crystallisations of 50 pmoles SOD with AMNP in FAPW Containing Different Concentrations of ALS-I*



However, there is a far greater difference in the intensity of the signals acquired for 50 pmoles myoglobin co-crystallised with NA in isopropanol and water. The peak in the spectrum of the sample containing just 0.1 mg/mL ALS-I is much sharper and intense than the peaks in the spectra of the samples containing higher concentrations of ALS-I.

***Spectra 16, 17 and 18**: Spectra Showing 50 pmoles Myoglobin Slow Crystallised with NA in Isopropanol and Water Containing Different Concentrations of ALS-I*

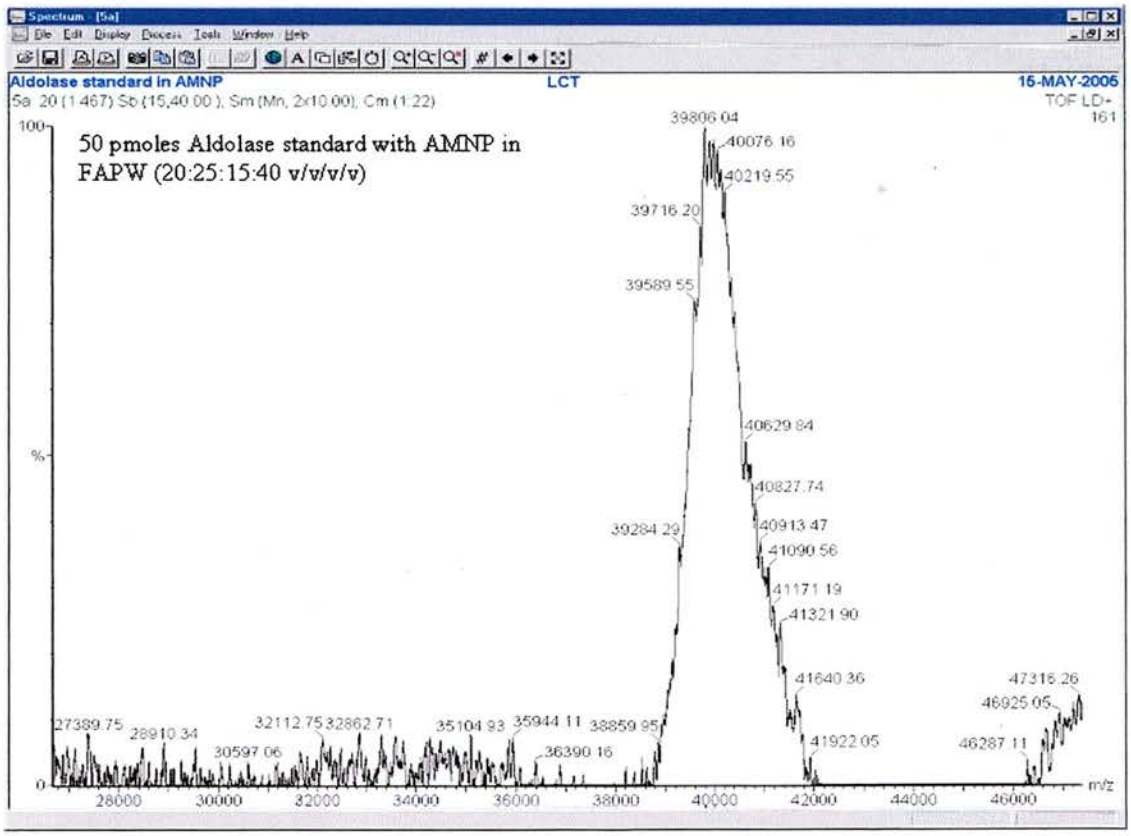


The above spectra show signals for three concentrations of ALS-I in the extraction solution. The best whole protein signal in the spectra above is the sample containing 0.1 mg/mL. The signals obtained for samples containing 0.25 mg/mL and 1 mg/mL ALS-I are still good but the peak resolution at 0.1 mg/mL is much better than either of the others.

The next investigation focused on determining the upper mass limit for which protein signals could be obtained using these three neutral matrices. Previous studies have shown the upper limits for these matrices in conjunction with protein digests to be between 12 and 20 kDa with spectra acquired in reflectron mode (41). A study performed by Farmer and Caprioli demonstrated the use of ATT to investigate leucine

zipper polypeptides with an upper mass limit of m/z 7373 (42). As stated previously, whole protein spectra from MALDI-TOF MS are obtained in linear mode, which has lower mass resolution than reflectron mode. Spectra were obtained of standard solutions of 50, 25, 15 and 10 pmoles aldolase and BSA in a total volume of 20 µL co-crystallised with all three matrices. Signals were achieved for concentrations of standard aldolase as low as 10 pmoles when slow crystallised with AMNP. The spectrum below demonstrates the quality of signal obtained for larger proteins such as aldolase.

***Spectrum 19**: Spectrum Showing 50 pmoles Aldolase Slow Crystallised with AMNP in FAPW*

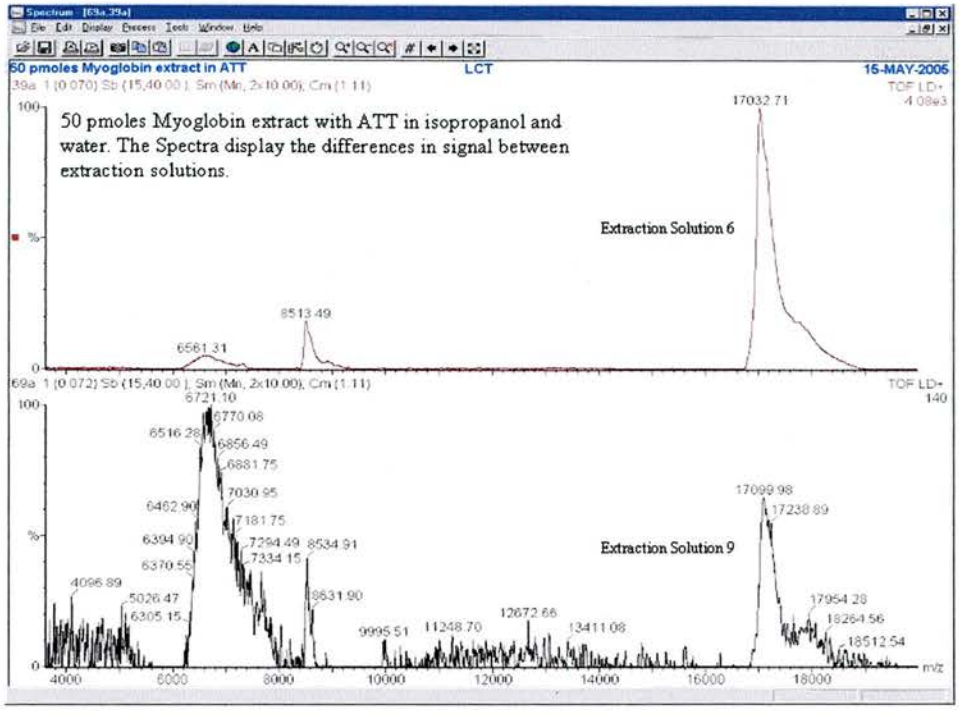


As can be seen from the spectra above, the signal for aldolase is very broad and not very smooth. Determining an accurate mass from peaks such as this is very difficult.

Neither standard nor extracted ADH produced signals when slow crystallised with any of the three neutral matrices. The spectra for BSA were obtained from the slow crystallisation of 50 pmol BSA protein extracts. The spectra show the signals for these larger proteins were much broader and weaker than those obtained for myoglobin and SOD. Thus masses obtained for these proteins were much less accurate. Average masses for extracts of BSA slow crystallised with ATT and NA were 67,662 Da and 68,399 Da respectively, which calculates to mass errors of 1232 Da and 1969 Da respectively. The average mass for standard aldolase slow crystallised with AMNP was determined to be 40,063 for an average of 4 spectra. This is a mass error of 851 Da. No signals were present for aldolase slow crystallised with NA or ATT. The average mass for extracted aldolase slow crystallised with ATT was determined to be 39,733 Da for an average of 3 spectra. This is a mass error of 521 Da. No signals were obtained for extracted aldolase slow crystallised with AMNP or NA.

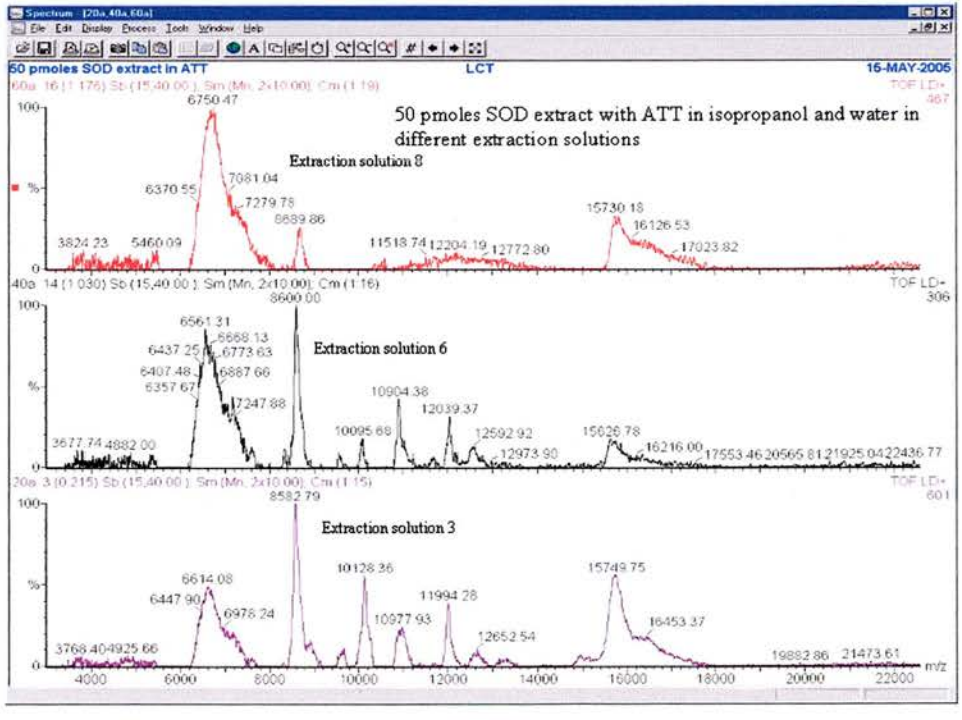
The final experiments performed with the neutral matrices were on whole protein extracts from SDS-polyacrylamide gels of SOD and myoglobin. Bands containing 50 pmoles of SOD and myoglobin were extracted with each of the extraction solutions described in Table 9. These contained differing concentrations of ALS-I and tris/gly. As can be seen from the spectra below, the most compatible extraction solutions for myoglobin were 6 and 9.

***Spectra 20 and 21**: Spectra Showing 50 pmoles Extracted Myoglobin Slow Crystallised with ATT in Isopropanol and Water.*



The most effective extraction solutions for SOD were 8, 6 and 9.

***Spectra 22, 23 and 24**: Spectra Showing 50 pmoles Extracted SOD Slow Crystallised with ATT in Isopropanol and Water*



The most effective extraction solution for further analysis by LC-ESI MS is also solution 9, thus in most cases one whole protein extraction can lead to data for both MALDI-TOF MS and LC-ESI MS.

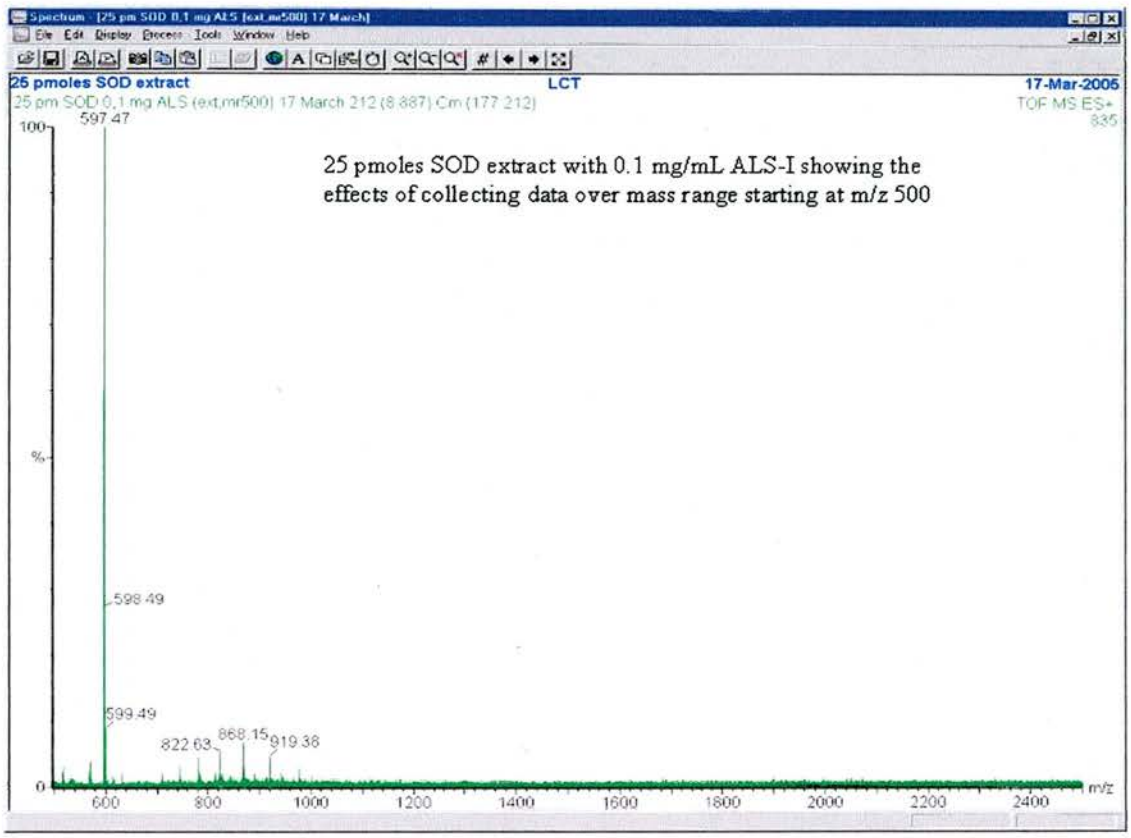
3.4.2 Electrospray Ionisation of Whole Protein Extracts

Once extraction levels had been optimized, it was necessary to determine whether the extracted protein was suitable for analysis by ESI MS. However, further contaminants may be present in the extracts, which might co-elute with protein from a reverse phase column and thus suppress signal. It had been confirmed that standard protein in these buffers could be analysed by ESI MS after on-line desalting through a MassPREP (Waters) On-Line Desalting Cartridge (2.1 x 10 mm). Still it was unclear whether protein extracted from a gel would remain in solution after desalting and hence give signals. Extracts in all nine buffers were tested as the extraction solutions identified as the most effective protein extractors were not necessarily the most ESI MS compatible. Of the extraction solutions shown in the table above, solutions 2, 4, 7, 8 and 9 did indeed give protein signals for extracts of 50 pmoles of SOD and myoglobin. Of those five extraction solutions, 2, 8 and 9 gave the best signals, with solution 9 consistently giving the best protein signals. Fortunately solutions 8 and 9 were also determined to be effective extractors. Protein signals were obtained for as little as 25 pmoles of smaller molecular weight proteins, less than 20 kDa loaded on the gel.

Despite initial promising results, some problems were encountered with the ALS-I. Initial tests were performed using ALS-I synthesised in a first batch. On changing

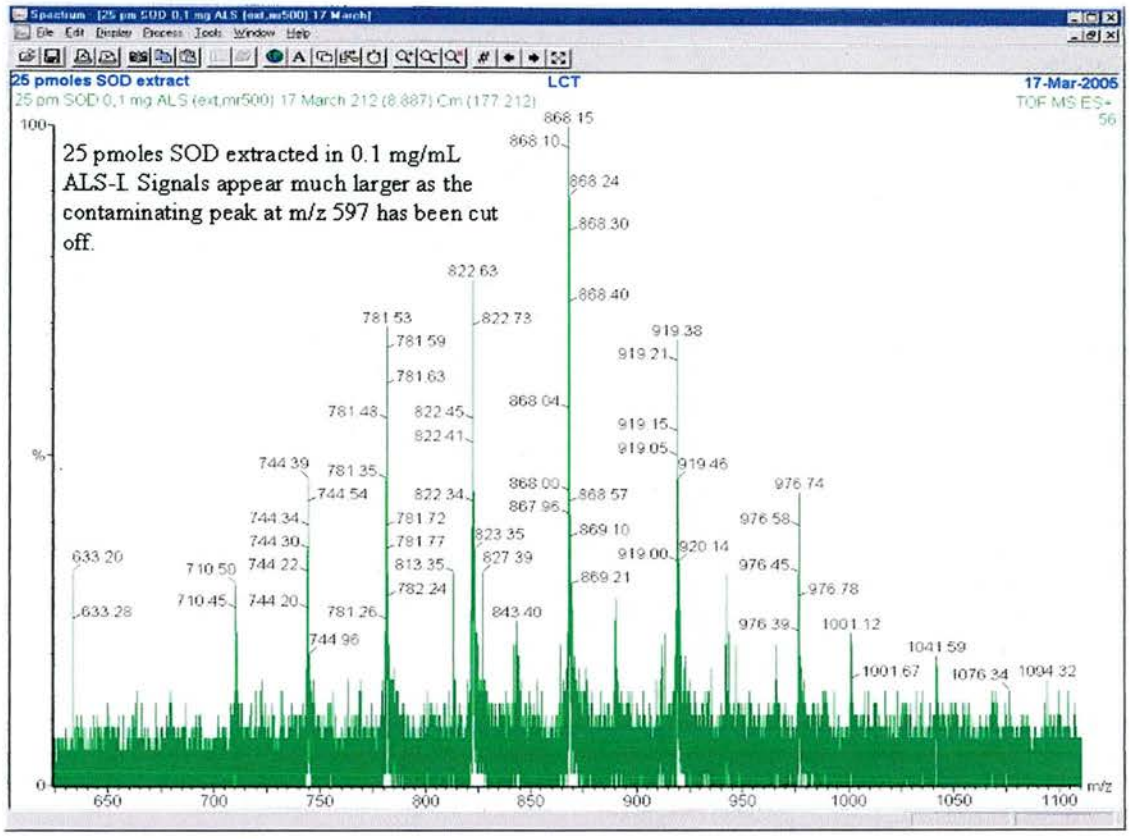
batches of ALS-I it was discovered that the same amount of detergent caused the protein peaks to be obscured by detergent and other contaminant peaks.

***Spectrum 25**: Spectrum Depicting the Effect Contaminant Peaks on Protein Signals*



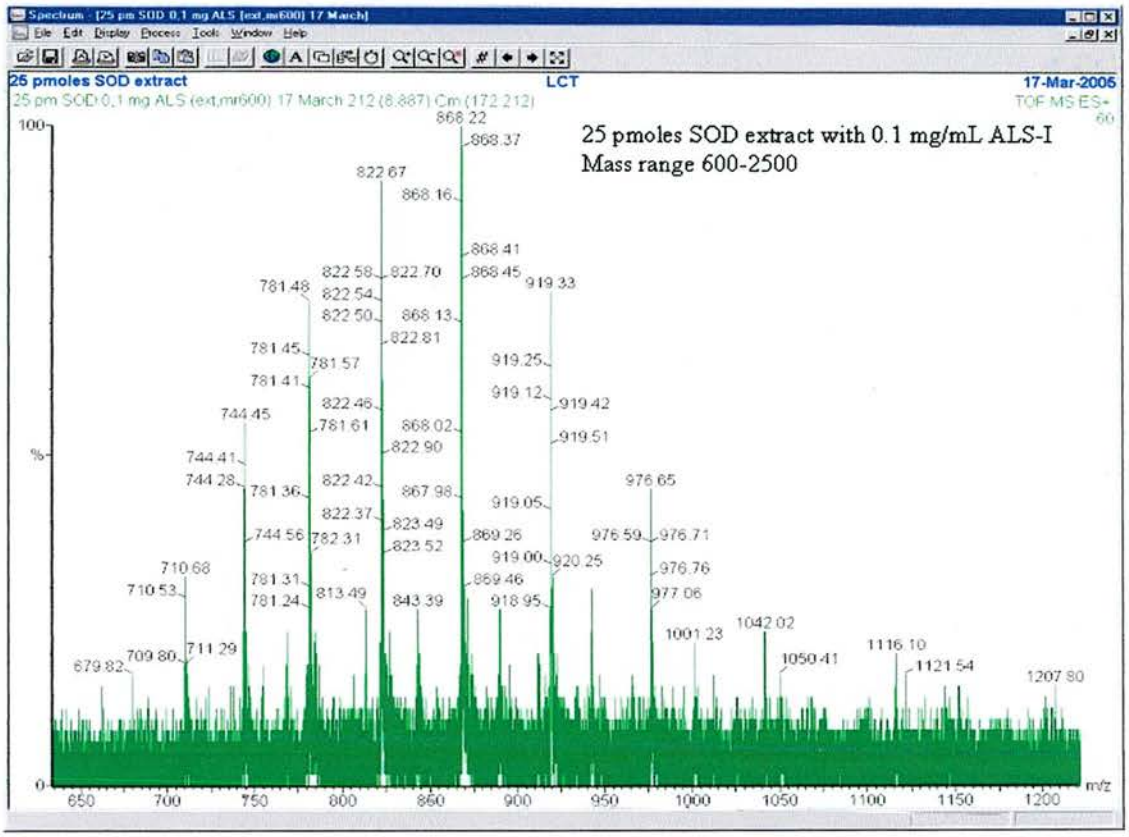
The amount of ALS-I was reduced and the mass range over which data was collected was altered and protein peaks became visible again. The spectrum above however demonstrates the effect of the detergent on the protein signals even at a concentration as low as 0.1 mg/mL. When the spectrum is focused over a smaller mass range (m/z 650-1100), which excludes the major contaminant peak, the signals can be seen clearly.

***Spectrum 26**: Spectrum Depicting the Above Protein Signals Magnified Over a Smaller Mass Range*



The spectrum above represents what occurs when the main contaminating peak is removed from the data collection range. Reducing the mass range over which data is collected does not affect the quality of the spectrum. The mass of SOD extrapolated from the spectrum above is 15,608 Da. The mass of SOD extrapolated from the spectrum acquired over a mass range of 600-2500 Da, as below, is 15,609.5 Da.

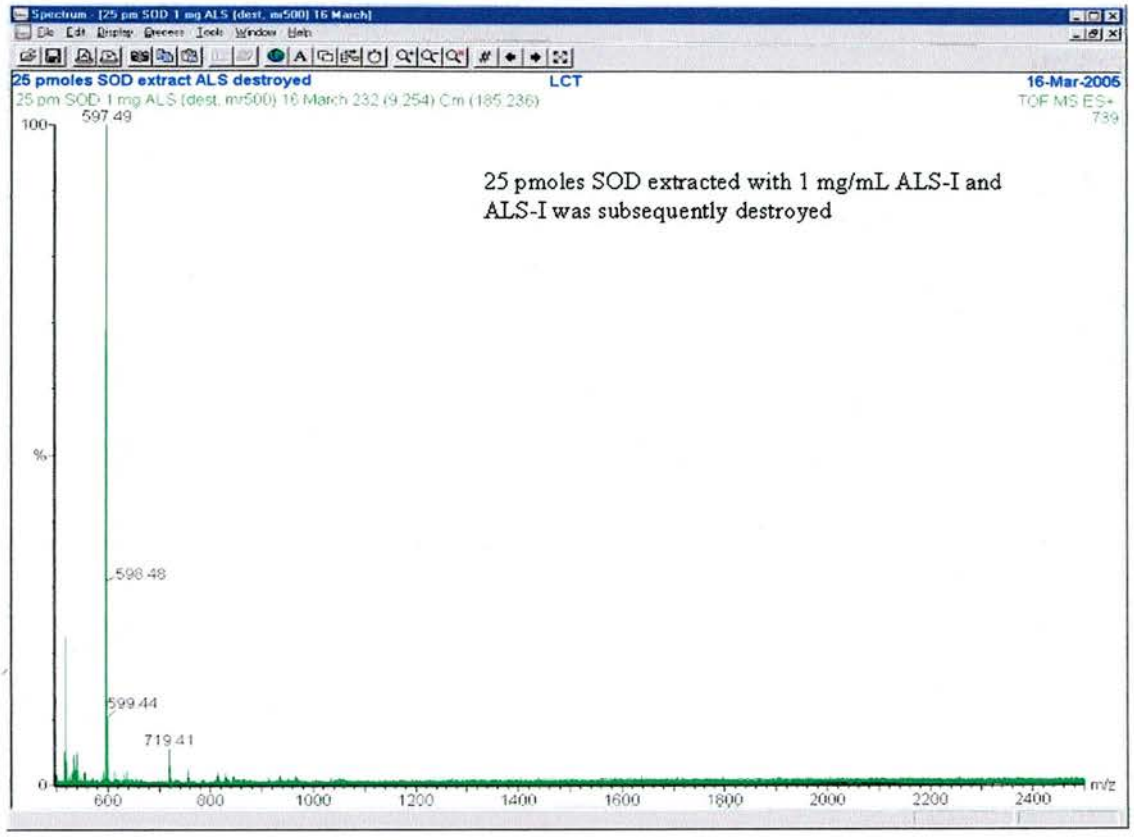
***Spectrum 27**: Spectrum Depicting Protein Signals for SOD Collected Over the Mass Range 600-2500 Da.*



The previous sample of ALS-I was determined to have a higher salt content and was thus at a lower ALS-I concentration than first thought. This was confirmed by obtaining the critical micelle concentration of both ALS-I samples. The CMC for the first sample was calculated to be 7-9 mM whereas the CMC of the second sample was 2.5-4 mM. As the CMC should be the same, this would suggest the initial (small) batch of ALS-I contained a high concentration of salt.

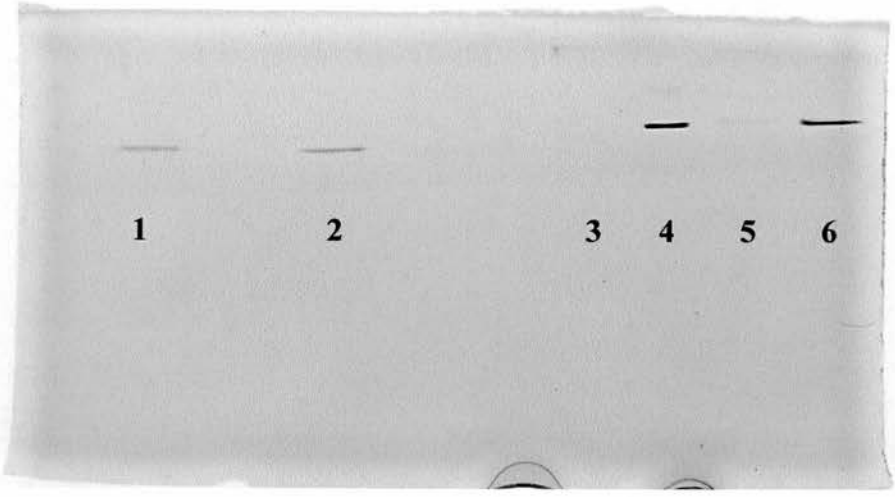
Subsequent experiments were performed to determine whether or not protein signals were achieved after acid hydrolysis of the ALS-I. As can be seen from the spectrum below, protein signals were unapparent in the spectrum. Destroying the ALS-I appeared to have an adverse affect on the signals acquired by LC-MS.

***Spectrum 28**: Spectrum Depicting Lack of Protein Signals After Decomposition of ALS-I*



After several LC-ESI MS spectra were acquired it was found that protein signals were obscured in other instances. Further optimisation of the extraction conditions was therefore undertaken. Several steps were taken in order to combat this. One experiment analysed extraction efficiencies of extraction solutions containing reduced concentrations of tris/glycine including 0.25 M, 0.2 M and 0.1 M. After optimisation the most consistently effective combination of ALS-I and tris/glycine for LC-ESI MS was 0.25 mg/mL ALS-I and 0.1 M tris/glycine.

Gel 4: *Comparison of Extracts containing no Tris/Glycine and 0.25 M Tris/Glycine*



Samples 1 and 2 are myoglobin and SOD extracts of 50 pmol gel bands with 0.25 M tris/glycine. Samples 3 and 5 are barely visible and are also extracts of 50 pmol gel bands; however they contain no tris/glycine. Samples 4 and 6 are 50 pmol SOD and myoglobin standard samples run on the gel for extraction comparisons. As can be seen from the gel, tris/glycine is an important addition to the extraction solution. Results from initial tests proved ALS-I was necessary for optimal whole protein extraction and further tests showed that tris/glycine was necessary but not at a 1M concentration.

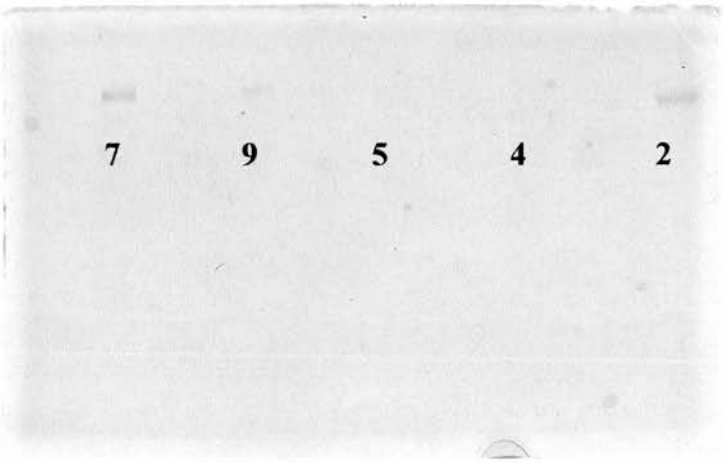
A higher concentration of ALS-I, 1 mg/mL, was more effective at extracting protein from a SDS-polyacrylamide gel, however the same concentration of ALS-I caused obscuring of protein signals. Thus a balance was needed between the amount of protein in the sample and the amount of detergent in the extraction buffer.

As stated above smaller proteins, myoglobin (17 kDa) and SOD (15.6 kDa), were used in the initial extraction experiments. Signals were obtained for proteins extracted

from 50 pmol polyacrylamide gel slices. In order to test the mass limitations of this procedure a range of proteins of different masses were analysed including BSA, ADH, aldolase, Carbonic Anhydrase, Small Ubiquitin like Modifier-1 (SUMO-1), Ubiquitin, Quinolate Phosphoribosyltransferase (QPRTase), NF kappa B P50 subunit, NF kappa B P65 subunit and Fluorinase.

Extraction tests produced disappointing results for extraction of 50 pmoles BSA from a polyacrylamide gel as can be seen in the figure below. Extraction solutions did not appear to extract any protein.

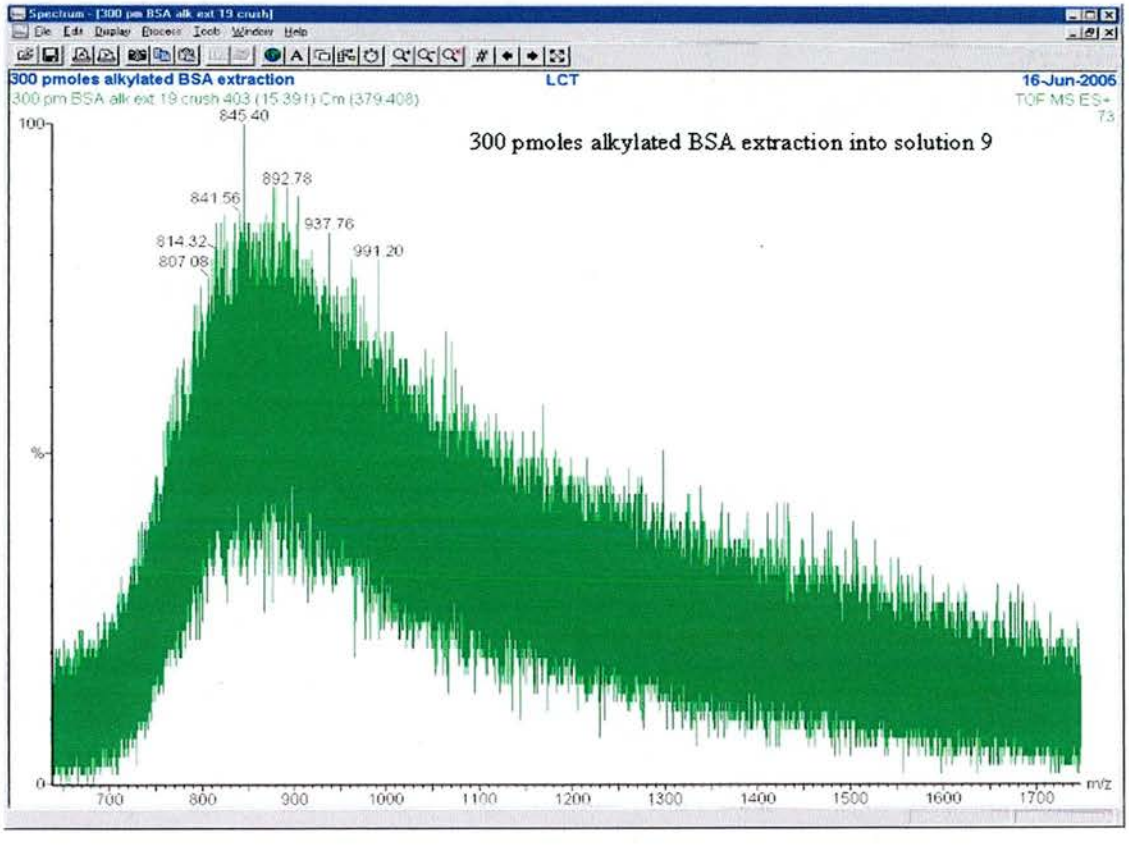
Gel 5: *BSA 2 hr Whole Protein Extraction Efficiencies for Solutions 7, 9, 5, 4 and 2*



Other proteins showed similar results. It was thought that larger proteins were not extracting as efficiently as smaller proteins as a larger protein cannot pass as easily through the small pores of a polyacrylamide gel. In essence, for the same reason larger proteins electrophorese more slowly than smaller proteins, larger proteins are slower to extract from the gel. By loading more protein, up to 300 pmoles total protein, on to the gel more protein was extracted and thus the following spectrum was

acquired for BSA. It should be noted that an accurate mass could not be calculated for any BSA extract.

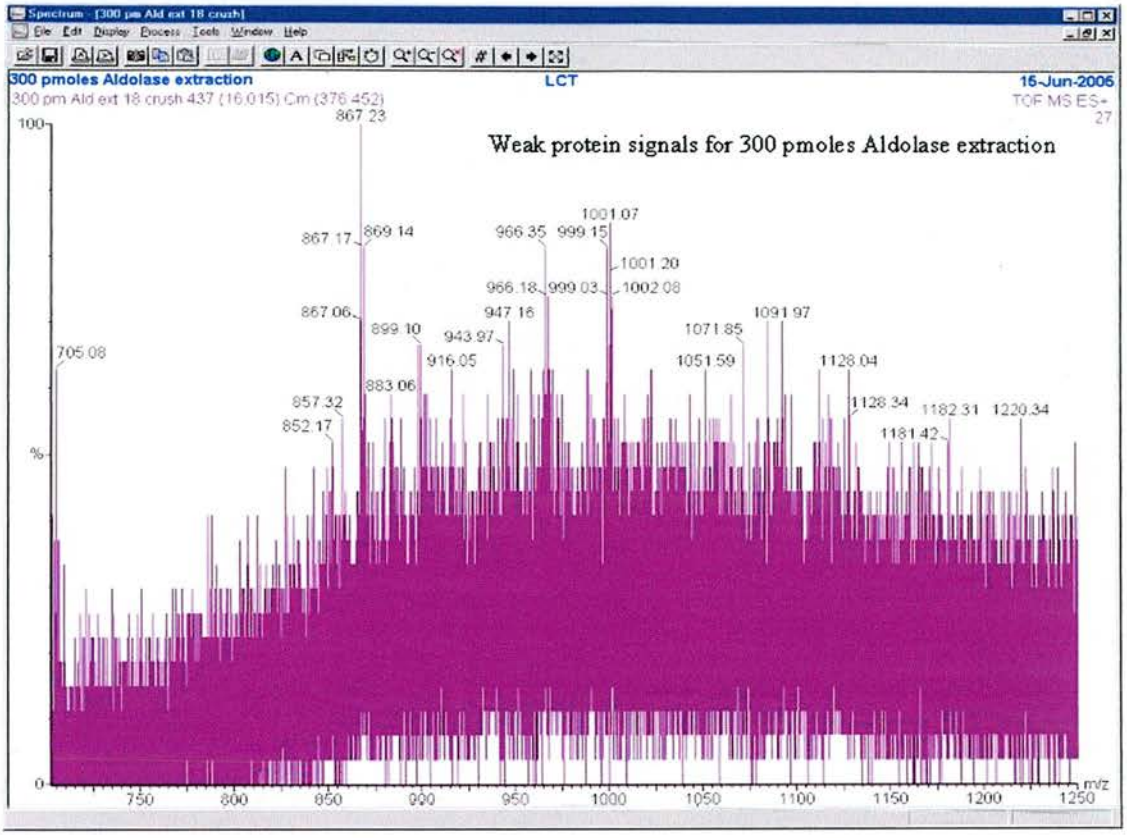
***Spectrum 29**: Spectrum Depicting Protein Signals Acquired for a 300 pmoles BSA Extraction*



Despite not giving an accurate mass, the spectrum above clearly shows protein signals. These signals are only slightly poorer than BSA signals acquired under ideal conditions

The spectra for aldolase also showed weak protein signal that did produce a mass, however confidence in this mass was quite low as the signals in the spectrum are quite weak.

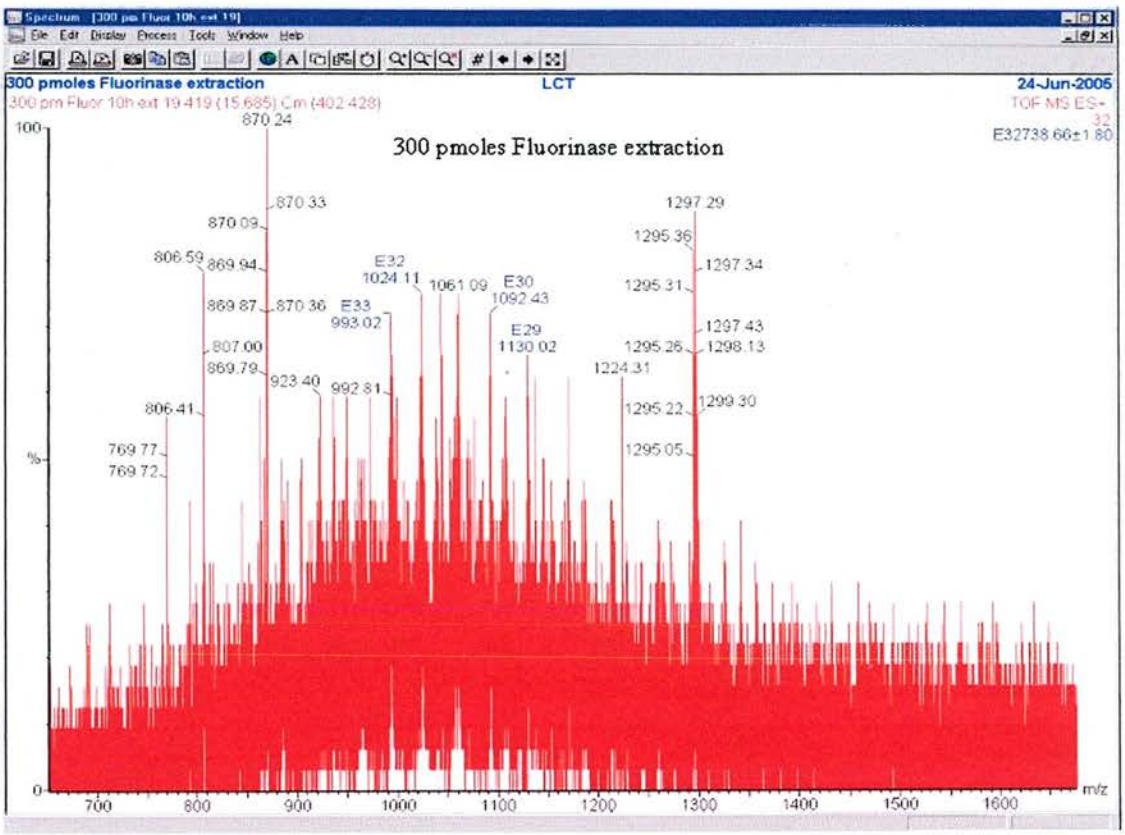
***Spectrum 30**: Spectrum Depicting Protein Signals for a 300 pmoles Aldolase Extraction*



Extraction of fluorinase showed weak signals but gave the correct mass upon transformation of the spectrum. The mass is given in the spectrum below. The mass error was calculated at 109 Da. This mass error is generally unacceptable in LC-ESI MS analyses and ideally this will be optimised in the future. However, as protein signals haven't routinely been acquired for extracts of proteins from SDS-polyacrylamide gels, this mass error is reasonable in comparison to some of the proteins analysed under these conditions, as seen in the table below. The peaks at 870.24 and 1297.29 are small molecule contaminants and thus have ionized very well resulting in their strong intensity. This spectrum also contains peaks at 1061.09 and 1130.02 which could be part of the myoglobin spectrum due to column

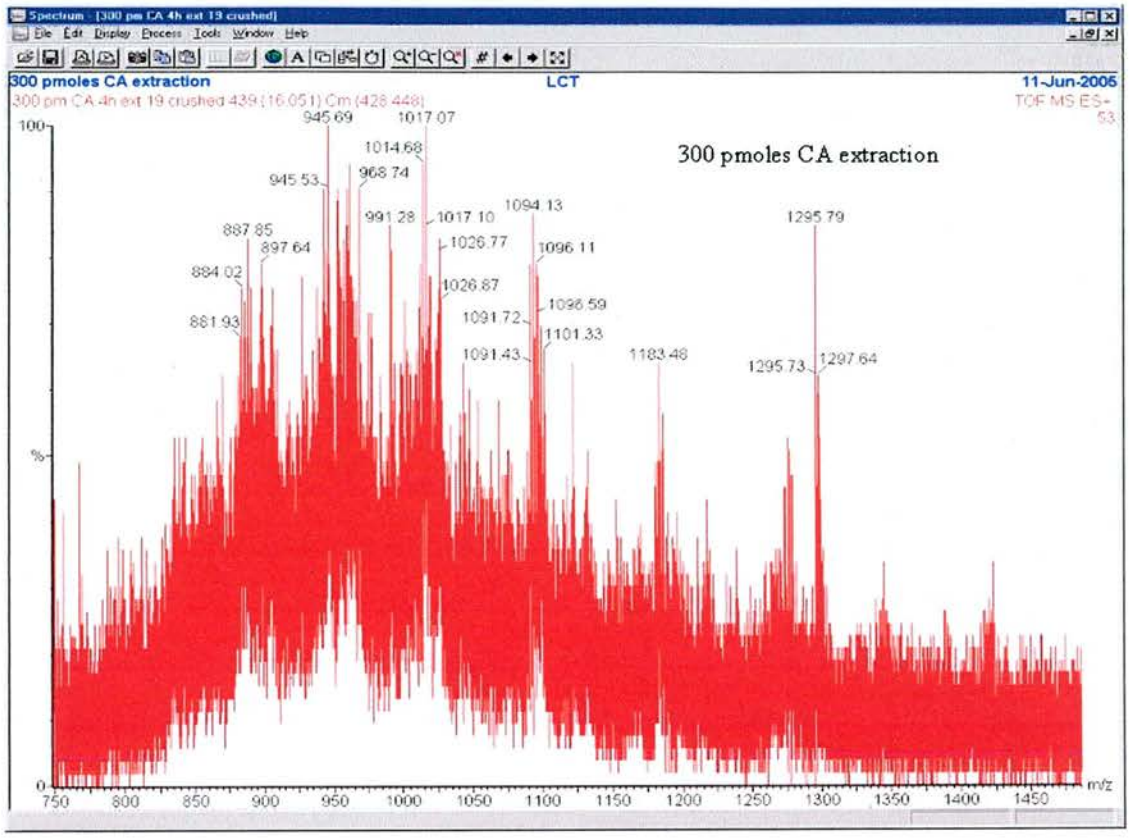
contamination, however the peak at 1130.02 appears to be part of the fluorinase spectrum and its peak shape is similar to others in the fluorinase spectrum.

***Spectrum 31**: Spectrum Showing Protein Signals for a 300 Pmoles Fluorinase Extraction*



Carbonic anhydrase was shown to be efficiently extracted from the gel and gave good protein signals; however the signals are such that a correct mass calculation was difficult. Three hundred picomoles of carbonic anhydrase were necessary to effect clear protein signals. The spectrum below clearly shows the presence of carbonic anhydrase in the sample analysed.

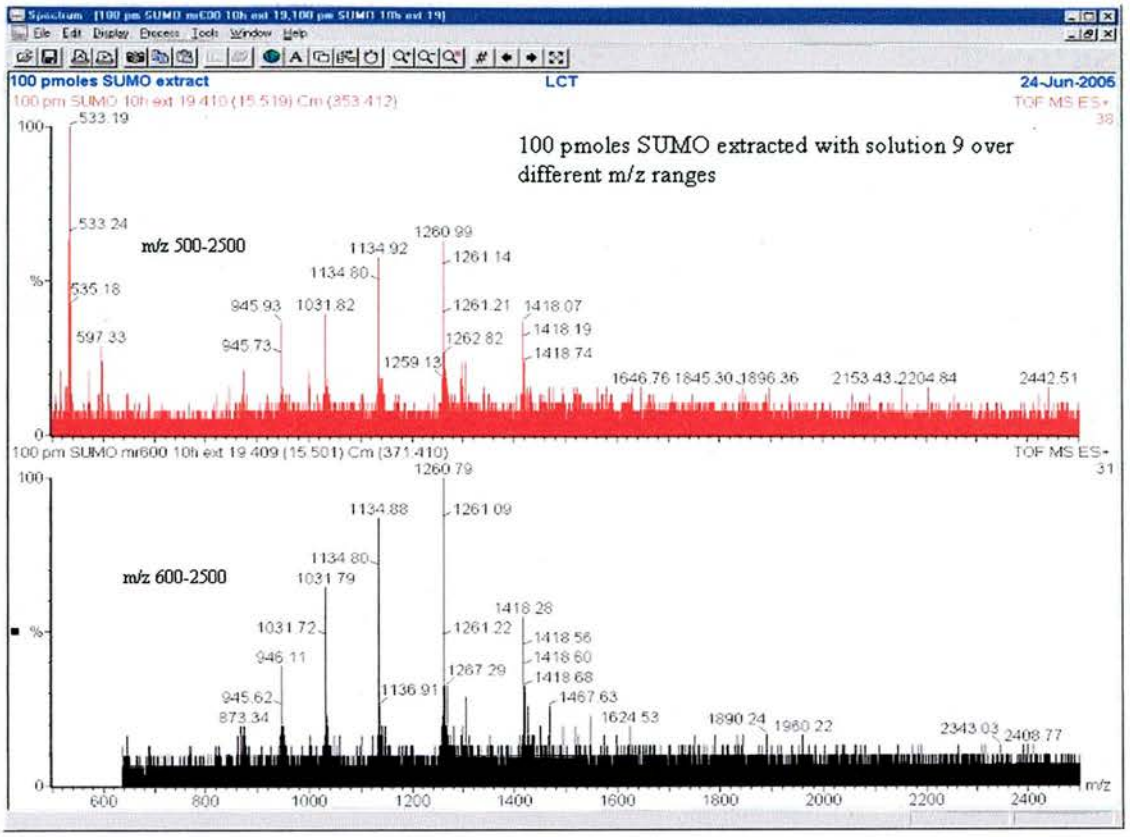
***Spectrum 32**: Spectrum Showing Protein Signals for a 300 pmoles Carbonic Anhydrase Extraction*



It was found that for proteins bigger than 20 kDa, it was necessary to load as much as 300 pmoles of protein on to the polyacrylamide gel in order to extract enough protein to give LC-MS signals. Smaller proteins that showed good extraction levels and good protein signals were ubiquitin, SUMO-1 and carbonic anhydrase. The proteins ubiquitin and SUMO-I were extracted efficiently from polyacrylamide gels from 50 pmol bands of protein.

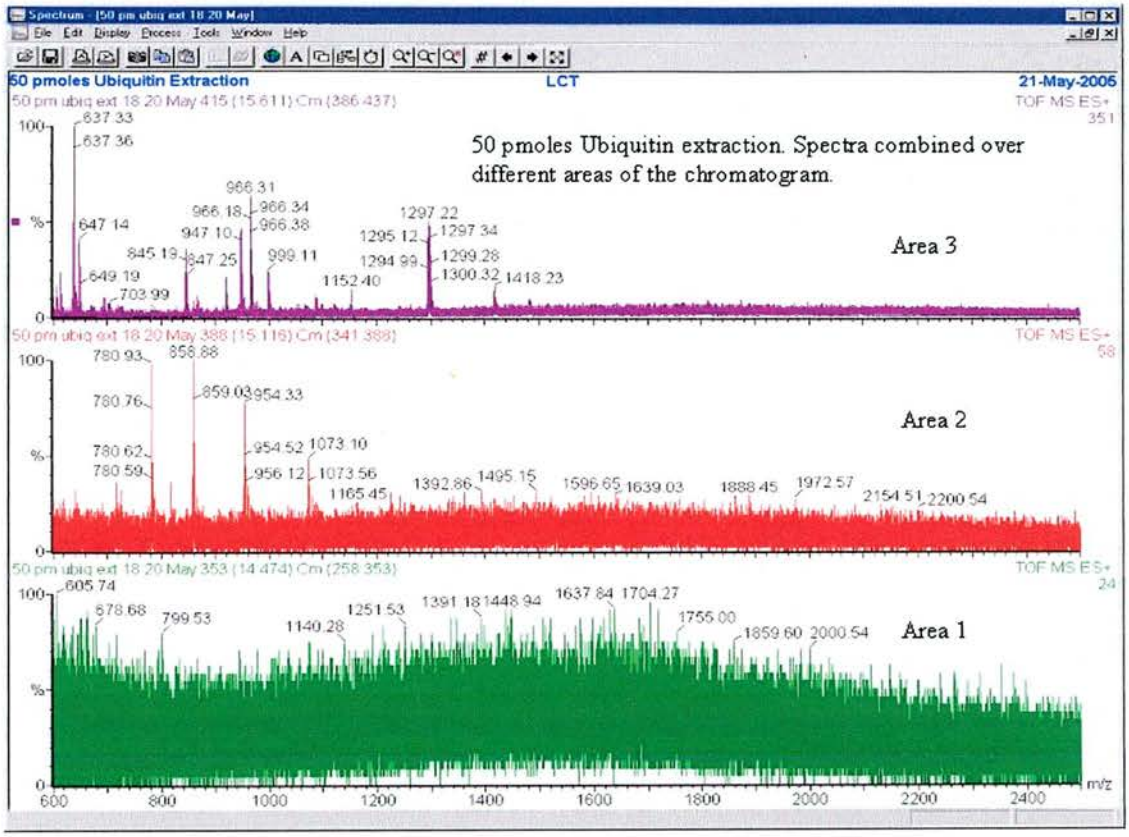
SUMO-1 at 12 kDa showed very good signals and good mass accuracy. As can be seen from the spectra below, changing the mass range over which data was collected made the protein signals clearer.

***Spectra 33 and 34**: Spectra Depicting Protein Signals for a 100 pmoles SUMO Extraction Collected over Different Mass Ranges*



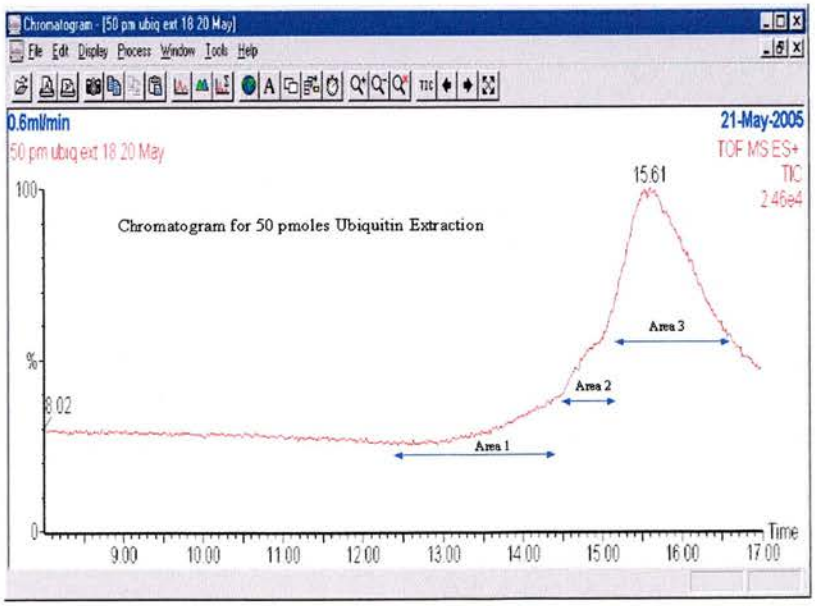
The spectra for ubiquitin below demonstrate an important finding. When combining spectra over a certain area of the chromatogram it is possible to obscure the protein signals by combining too many spectra or by combining the wrong area of the chromatogram as can be seen below.

***Spectra 35, 36 and 37**: Spectra Depicting Ubiquitin Signals and Contaminant Signals Acquired at Different Times in the Chormatogram*



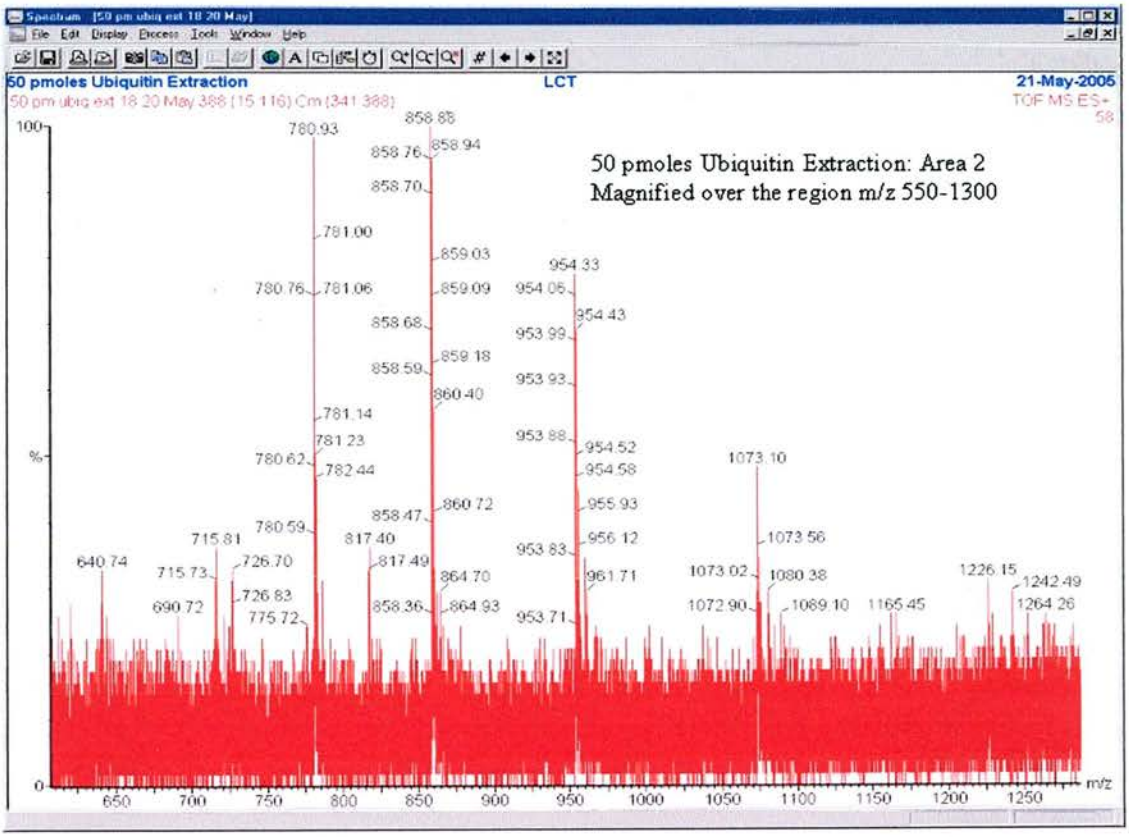
The chromatogram is given below.

***Spectrum 38**: Chromatogram of a 50 pmoles Ubiquitin Extraction Analysed by LC-ESI MS*



As can be seen in the spectra, area two gave the best protein signals. These have been magnified in the spectrum below.

***Spectrum 39**: Spectrum Depicting the Magnified Ubiquitin Signals Found in Area 2 of the Chromatogram Above*



The next set of experiments tested longer extraction times, from three to ten hours, and fine crushing of the gel slices on the same series of proteins listed above in order to improve extraction efficiency. This produced improved signal intensity. It was initially determined that crushing the gel slices up finely improved the extraction efficiency, possibly due to greater surface area exposed to the extraction solution. It is important to note that any gel crushing system that utilizes a membrane causes degradation of the protein signal. Crushing should be carried out using a plastic rod. Secondly, optimum extraction times were linked directly to the size of the protein. For

example, good spectra were obtained for 12 kDa SUMO1 after a 3 hr extraction however 32 kDa Fluorinase required 10 hr extractions before good spectra were acquired. It appeared that QPRTase (34 kDa), P50 (39 kDa), and aldolase (39 kDa) showed efficient extraction after 6 hrs when an aliquot was analysed by SDS-PAGE however, no mass spectrum obtained for these proteins was clear enough to produce an accurate mass. It should be noted that extending the extraction time to 10 hrs for proteins over 25 kDa did not appear to have a detrimental effect on spectra. However for proteins under 25 kDa, extraction times over 6 hrs lead to a weaker protein signal.

Upon acquisition of spectra for each protein the mass accuracy was determined. The mass accuracy is detailed in the table below.

Table 13: *Mass Accuracy of Whole Protein Extraction*

Protein	Actual Mass (Da)	Mass Obtained (Da)	Mass Error
Superoxide Dismutase	15,591	15,704	113
Myoglobin	16,951	17,389	438
Carbonic Anhydrase	29,025	29,832	807
SUMO1	11,694	11,340	354
QPRTase	34,012	/	
P50	38,796	40,805	2009
P65	34,012	32780	1232
Aldolase	39,212	39295	83
Alcohol Dehydrogenase	36,800	/	

Ubiquitin	8,560	8,579	19
BSA	66,430	/	
Fluorinase	32630	32,739	109

A number of in-gel or post-gel modifications may be occurring to give these mass discrepancies. Unlike data acquired by MALDI-TOF MS, the accuracy of the ESI data allows these mass changes to be measured to within one or two Daltons. Therefore we can postulate which modifications may be occurring by the size of the mass shift. The formation of acrylamide adducts from unpolymerized acrylamide in the gel has been reported. Galvani *et al.* were able to locate several sites of cysteine alkylation by unpolymerized acrylamide by delayed extraction MALDI-TOF MS (68). Another set of experiments done by Jeannot *et al.* also identified sites of acrylamide-protein adducts (70). This modification adds 71 Da to the protein. Acrylamide reacts with cysteine residues. Proteins with no cysteines were compared with proteins containing only one cysteine to determine whether cysteine modification was responsible for the mass changes seen. Another possible modification is oxidation of methionine residues, which adds 16 Da to the protein mass.

Although reagents such as iodoacetamide and *N*-ethyl maleimide may be used to protect cysteines from acrylamide attack, an unknown protein (with an unknown number of cysteines) would be mass shifted by X x 71 Da where X is unknown, so the true mass of the protein would not be known.

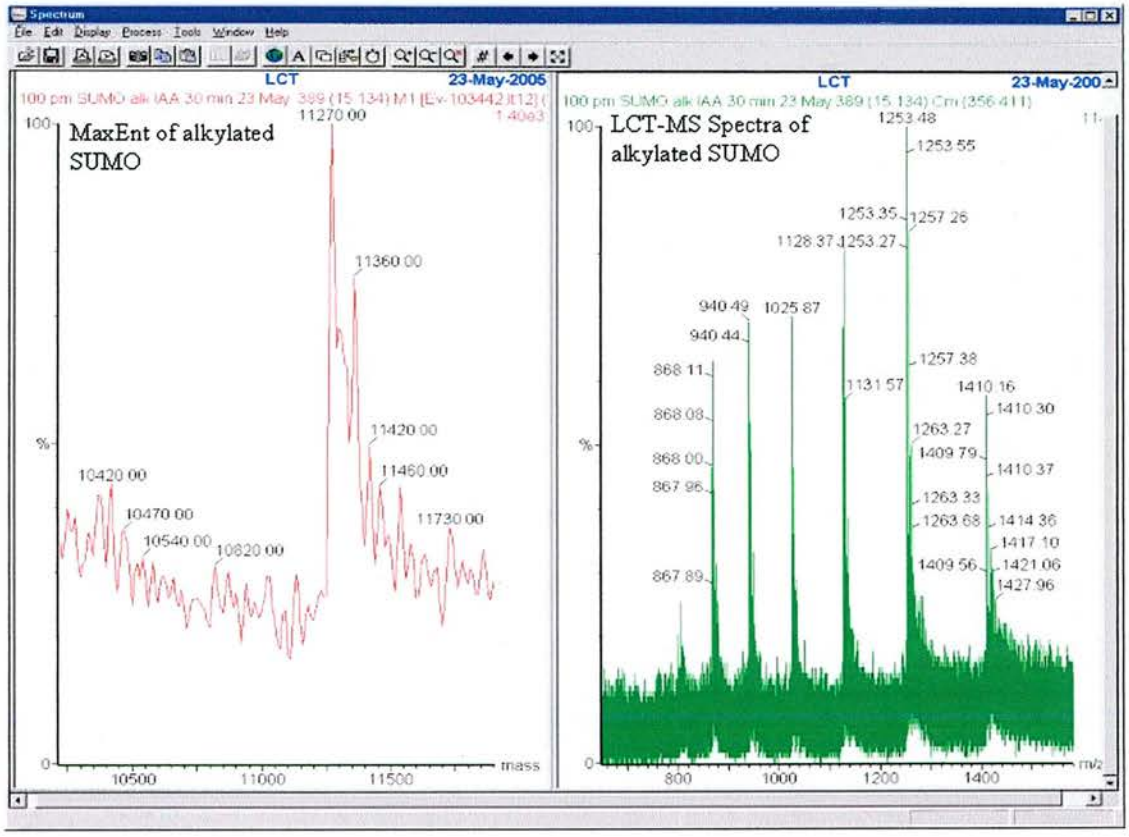
However some related reagents, *N*-ethyl iodoacetamide and *N*-[$^2H_5$]-ethyl iodoacetamide had been previously prepared in the laboratory (79, 80). The light

reagent adds 86 Da to each cysteine modified and the heavy reagent, containing 5 deuterium atoms, adds 91 Da. If half the sample is labelled with the light reagent and half with the heavy reagent, then the mass difference between the two peaks would be proportional to the number of cysteines in the protein and hence the unmodified mass can be calculated.

Mass accuracy is often affected by acrylamide adducts on the protein. It can be seen from the data above that the presence of acrylamide adducts is likely. Hence, steps were taken in order to block sites open to acrylamide attack, such as cysteines.

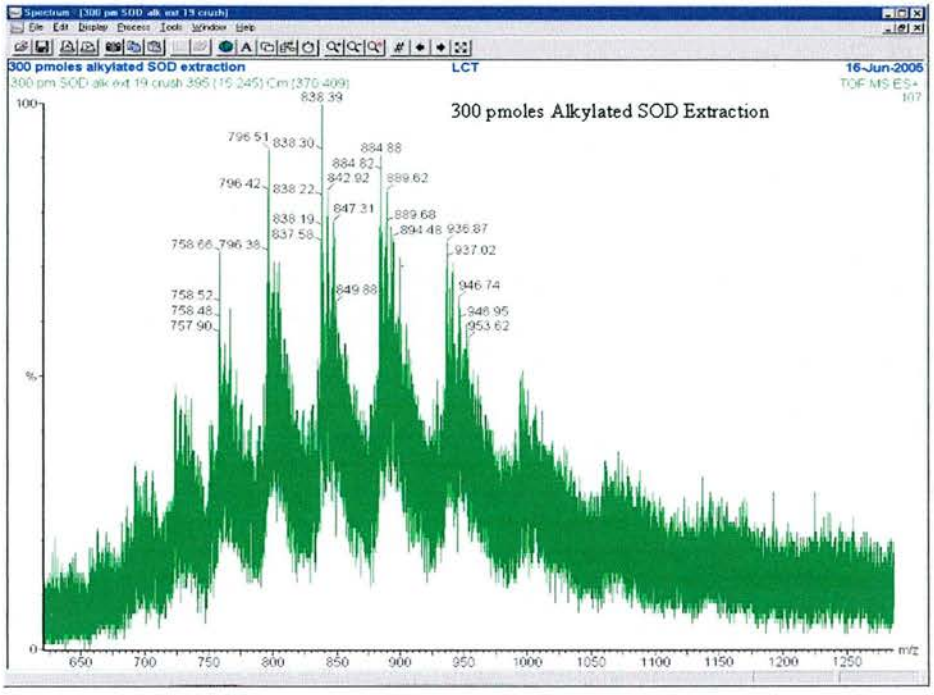
First, proteins in solution were reduced by incubation with 20 mM DTT for 30 minutes at 55 °C and then alkylated with *N*-ethyl-iodoacetamide or *N*-[$^2H_5$]-ethyl-iodoacetamide. Subsequent analysis was performed by LC-ESI MS. Deconvoluted spectra showed that some proteins had been alkylated fully while others, particularly proteins with a high number of cysteines, e.g. BSA, were only partially alkylated. Many samples had mixtures of fully alkylated protein along with partially alkylated protein. Experiments showed that a mixture of alkylated protein and labelled alkylated protein had resolved peaks for two or greater alkylations on the LCT instrument, such as Cytochrome C and SOD. Tests were then performed to determine whether the alkylating tag would block acrylamide adducts from forming and thus allow for the accurate mass of a protein to be calculated. The spectra below represent a 1:1 ratio of unlabelled protein to alkylated protein. As can be seen below the alkylation of the protein solution of SUMO-1 showed a high level of binding of the unlabelled cysteine tag. The mass increase (90 Da) is measured relative to the mass acquired for the non-alkylated whole protein (11, 270 Da).

***Spectra 40 and 41**: Spectra Showing SUMO Protein Signals and the Extrapolated Maximum Entropy*



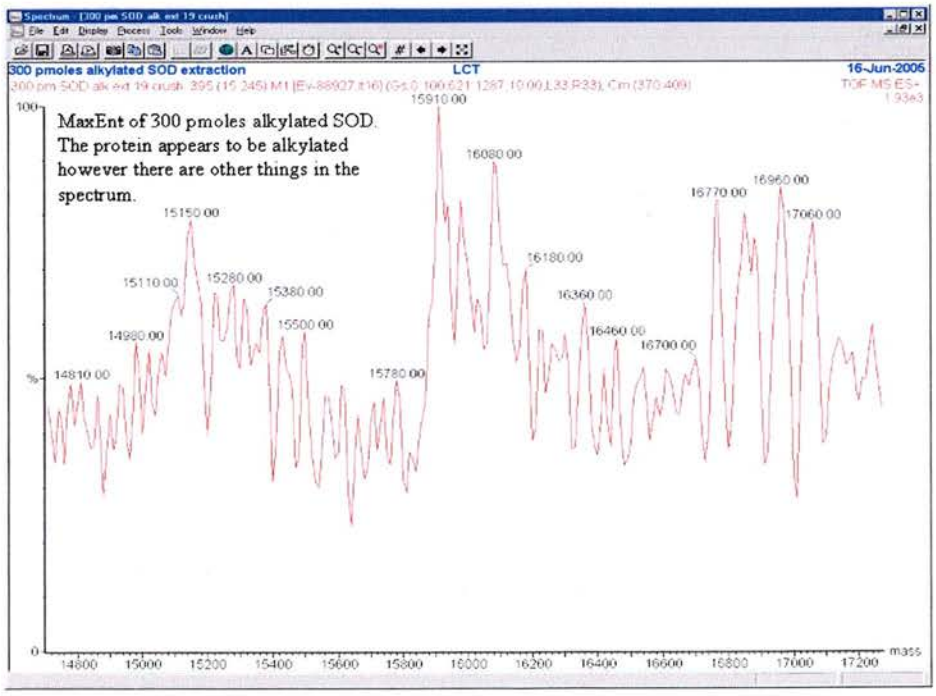
Prior to separation by 1D SDS-PAGE, proteins were alkylated for 30 minutes at 55 °C with *N*-ethyl-iodoacetamide. The protein was then loaded on to a SDS-polyacrylamide gel and electrophoresed for 60-90 minutes. Protein bands were excised and protein extracted as described in the method and materials section. Samples were extracted using extraction solution 9 and were analysed by LC-ESI MS. Data acquired for this experiment was however inconclusive. SOD, which contains three cysteines, appeared to be fully alkylated before electrophoresis but on mass spectrometric analysis of the extract, the mass of the protein was calculated to contain three acrylamide adducts. SOD had previously been found to be unsusceptible to acrylamide attack. This can be seen in the spectra below.

***Spectrum 42**: Spectrum Showing Signals for 300 pmoles Extraction of Alkylated SOD*



While these protein signals appear to be clear the MaxEnt of this spectrum shows otherwise.

***Spectrum 43**: Spectrum of the Maximum Entropy of 300 pmoles Alkylated SOD*



As can be seen in the MaxEnt, the first mass is 15,910 kDa, which is a mass error of 319 Da. However there are other masses of near equal intensity.

Despite these inconclusive results for acrylamide blocking by alkylation with *N*-ethyl-iodoacetamide, there are several possibilities for further experiments. These will be discussed in the following section. Further experimentation is necessary in order to say with certainty the benefits or disadvantages of acrylamide blocking.

4. Conclusions and Further Work

4.1 Synthesis and use of Ruthenium II Bathophenanthroline Disulfonate Chelate (RuBPS)

RuBPS was confirmed as an effective fluorescent stain with equivalent sensitivity to Sypro Ruby. Its optimal staining concentration was 150 nM and its protein detection limit was 10 ng of protein. This stain was determined to be more work intensive than Coomassie Brilliant Blue. As good results were obtained without resorting to this stain it was not used further.

4.2 Synthesis of Acid Labile Surfactant-I (ALS-I)

The synthesis of ALS-I gave a yield of 50% for the first step and a 43% yield for the second step. In comparison to literature values the yields were low; however the scale of the reaction in these experiments was much smaller than those performed in the literature. Characterization of the products showed the desired results had been achieved.

4.3 Critical Micelle Concentration

The CMC value for ALS-I was calculated to fall in the range of 2.5 mM to 4 mM. The CMC determined by Yamamura *et al.* was 0.6 mM obtained by surface tension measurements. Error in the CMC obtained by dye solubilization was attributed to salt content of the ALS-I leading to inaccuracies in solution concentrations. Accurate determination of the CMC concentration range is important in order to allow the detergent to be utilized effectively.

4.4 Acid Labile Surfactants in Protein Digestion and Extraction

The surfactant ALS-I was found to improve the quality of peptide spectra obtained from protein digestion and thus augmented the Mascot score obtained for peptide peak list database searches. Identifications with significant scores were achieved for a wide range of proteins as small as ubiquitin (8 kDa) and as large as β-galactosidase (17 kDa). Both MALDI-TOF MS and LC-ESI MS/MS gave good results for in-gel tryptic digests with ALS-I. At lower protein amounts the addition of ALS-I could make the difference between obtaining an identification or not.

4.5 Whole Protein Extraction

Whole protein extraction from SDS-polyacrylamide gels is a valuable technique. A whole protein extraction method has been developed that has been shown to be 50-60% efficient. Proteins were separated by SDS-PAGE and extracted using a tris/glycine/DTT extraction buffer containing ALS-I. ALS-I was shown to improve protein extraction efficiency.

4.6 Mass Spectrometry of Whole Proteins Extracted From SDS-Polyacrylamide Gels

Proteins extracted from gels were analysed by mass spectrometry. Both LC-ESI MS and MALDI techniques were successfully used.

4.6.1 MALDI-TOF MS Analysis Using Neutral Matrices

The MALDI samples were acquired using a novel neutral matrix system developed in order to analyse acid sensitive samples by slow crystallisation. Neutral matrices were determined to be compatible with whole protein analysis by mass spectrometry where the sample had been slow crystallised. The matrices 2-amino-4-methyl-5-

nitropyridine, 4-nitroaniline and 6-aza-2-thiothymine were shown to give sharp peaks in specific solvent systems for whole protein extracts. The mass accuracy, mass ± 350 Da, was consistent with that reported in previous research. However the mass accuracy achieved was not sufficient enough, due to the poor resolution of MALDI-TOF MS signals in linear mode, to answer the sort of question likely to be posed in further experiments. These include whether a post-translational modification is present or where N or C-terminal clipping is occurring.

4.6.2 LC-ESI MS Analysis of Whole Proteins Extracted From SDS-Polyacrylamide Gels

Proteins above 35 kDa were determined to have good extraction but did not produce any LC-ESI spectra clear enough to identify an accurate mass. Proteins below 35 kDa were more likely to produce clean spectra and thus reveal the proteins accurate mass and proteins below 25 kDa were even more likely to produce good spectra. Proteins such as myoglobin, SOD and SUMO1 consistently gave good spectra and masses with an accuracy of ±438, ±113 and ±354 respectively.

Acrylamide adducts were found to be a mitigating factor in the acquisition of an accurate mass. In an attempt to prevent this alkylation, in a manner that would still allow the accurate mass determination of the unmodified protein, proteins were alkylated with a mixture of *N*-ethyl and *N*-[$^{2}H_{5}$]-ethyl iodoacetamide. Alkylation of whole proteins was successful in protein solution experiments. Proteins for whole protein extraction were alkylated with *N*-ethyl iodoacetamide and then separated by SDS-PAGE. However, data was inconclusive as to the effectiveness of alkylation prior to extraction from an SDS-polyacrylamide gel.

4.7 Further Work

These initial investigations prove that ALS is a mass spectrometry compatible detergent and that it enhances the extraction of whole proteins from polyacrylamide gels and proteins can be analysed in its presence. However extending the HPLC gradient in order to separate away the solution contaminants more effectively might further optimize the analysis of whole protein extracts.

Subsequent experiments utilizing this method will find it crucial to determine a method to effectively block acrylamide adducts such that the mass of the protein can be determined with confidence.

Further exploration of the extracted proteins by digestion and analysis by MALDI-TOF MS, using either slow crystallisation or the dried drop method, and LC-ESI MS/MS following an appropriate clean up method might lead to further information regarding mass shifts. It might prove beneficial to digest extracted proteins and analyse the digest by LC-ESI MS/MS in conjunction with whole protein analysis in order to map sites that are susceptible to acrylamide adducts. This also might involve alkylating the protein with tags, one labelled and one unlabelled, in order to identify how many times the protein has been modified and thus calculate the accurate mass of the protein. *N*-ethyl iodoacetamide and its labelled counterpart, *N*-[$^2H_5$]-ethyl iodoacetamide, gave good results for LC-ESI MS but peaks could only be resolved to a difference of two cysteines, a separation of ten. A greater separation might improve accurate mass acquisition.

Experiments should also be implemented to extend the working mass range of ALS-I in whole protein extractions. Currently, proteins above 40 kDa have been extracted successfully, but accurate mass calculations have been unsuccessful. This might be possible by testing other types of gels including the DATO gel system, which has larger pores than SDS-polyacrylamide gels. This might allow better extraction of larger proteins.

Finally, another area of research that should be pursued is the staining of the polyacrylamide gels with a reverse stain such as Zn-imidazole stain or a fluorescent stain such as RuBPS. These stains might not fix the protein in the gel as much as Coomassie is known to and thus might lead to better extractions and cleaner mass spectra.

References

1) J. Rappsilber, M. Mann, What Does it Mean to Identify a Protein in Proteomics? *Opinion: TRENDS in Biochemical Sciences* (2002) **27**, 74-78.

2) D. R. Goodlett, A. Keller, J. D. Watts, R. Newitt, E. C. Yi, S. Purvine, J. K. Eng, P. von Haller, R. Aebersold, E. Kolker, Differential stable isotope labelling of peptides for quantitation and *de novo* sequence derivation. *Rapid Communications in Mass Spectrometry* (2001) **15**, 1214-1221.

3) O. Stemmann, H. Zhou, S. A. Gerber, S. P. Gygi, M. W. Kirschner, Dual Inhibition of Sister Chromatid Separation at Metaphase. *Cell* (2001) **107**, 715-726.

4) P. D. von Haller, E. Yi, S. Donohoe, K. Vaughn, A. Keller, A. I. Nesvizhskii, J. Eng, X. Li, D. R. Goodlett, R. Aebersold, J. D. Watts, The Application of New Software Tools to Quantitative Protein Profiling Via Isotope-coded Affinity Tag (ICAT) and Tandem Mass Spectrometry. *Molecular Cellular Proteomics* (2003) **2**, 428-442.

5) M. Springer, ICAT Reagent Technology Powers Systems Biology. Retrieved 20 April 2004 from Proteomics SURF [BioBeat] address: http://ww2.proteomicssurf.com:5000/forums/forumdisplay.php?s=&forumid=17

6) Y. Oda, T. Owa, T. Sato, B. Boucher, S. Daniels, H. Yamanaka, Y. Shinohara, A. Yokoi, J. Kuromitsu, T. Nagasu, Quantitative Chemical Proteomics for Identifying Candidate Drug Targets. *Analytical Chemistry* (2003) **75**, 2159-2165.

7) L. Stryer, *Biochemistry*. 3rd Edition, W. H. Freeman and Company, 1988.

8) Retrieved from: http://en.wikipedia.org/wiki/SDS-PAGE. Wikipedia: The Free Encyclopedia, 2005.

9) SDS-PAGE Retrieved from: http://www.bio.davidson.edu/courses/genomics/method/SDSPAGE/SDSPAGE.html. Department of Biology, Davidson College, 2001.

10) V. Neuhoff, N. Arold, D. Taube, W. Ehrhadt, Improved Staining of Proteins in Polyacrylamide Gels Including Isoelectric Focusing Gels with Clear Background at Nanogram Sensitivity Using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* (1988) **9**, 255-262.

11) Laemmli, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* (1970) **227**, 680-685.

12) D. Hamus, One Dimensional Polyacrylamide Gel Electrophoresis of Proteins. A Practical Approach. Edited by B.D. Harris and D. Rickwood. 2nd Edition, IRL Press, 1-48, 1990.

13) A. Schevchenko, M. Wilm, O. Vorm, M. Mann, Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. *Analytical Chemistry* (1996) **68**, 850-858.

14) M. L. Ortiz, M. Calero, C. F. Patron, L. Castellanos, E. Mendez, Imidazole-SDS-Zn Reverse Staining of Proteins in Gels Containing or Not SDS and Microsequence of Individual Unmodified Electroblotted Proteins. *Federation of European Biochemical Societies* (1992) **296**, 300-304.

15) C. Lee, A. Levin, D. Branton, Copper Staining: A Five-Minute Protein Stain for Sodium Dodecyl Sulfate-Polyacrylamide Gels. *Analytical Biochemistry* (1987) **166**, 308-312.

16) Retrieved from Bio-Rad Laboratories EG Bulletin 1820, Detection of Proteins in SDS-PAGE: A Comparison of Gel Staining Methods.

17) I. Valdes, A. Pitarch, C. Gil, A. Bermudez, M. Llorente, C. Nombela, E. Mendez, Novel procedure for the identification of proteins by mass fingerprinting combining two-dimensional electrophoresis with fluorescent SYPRO Red staining. *Journal of Mass Spectrometry* (2000) **35**, 672-682.

18) M. Lebendiker, Characterization Protocols. Retrieved from: http://www.ls.huji.ac.il/~purification/Characterization.html#Gel_Stain The Protein Purification Facility

19) T. Rabilloud, J. Strub, S. Luche, A. van Dorsselaer, J. Lunardi. A Comparison Between Sypro Ruby and Ruthenium II tris (bathophenanthroline disulfonate) as Fluorescent Stains for Protein Detection in Gels. *Proteomics* (2001) **1**, 699-704.

20) Comparison of Deep Purple Total Protein Stain and Sypro Ruby in 1-D and 2-D gel electrophoresis Amersham Biosciences Application Note www.amershambiosciences.com

21) G. B. Smejkal, M. H. Robinson, A. Lazarev, Comparison of Fluorescent Stains: Relative Photostability and Differential Staining of Proteins in Two-Dimensional Gels. *Electrophoresis* (2004) **25**, 2511-2519.

22) A. Gorg, C. Obermaier, G. Boguth, A. Harder, B. Scheibe, R. Wildgruber, W. Weiss. The Current State of Two-dimensional Electrophoresis with Immobilized pH gradients. *Electrophoresis* (2000) **21**, 1037-1053.

23) A. Pundey, M. Mann, Proteomics to Study Genes and Genomes. *Nature* (2000) **405**, 837-846.

24) J. Finchmann, R. Westermeier, 2-D Protein Gel Electrophoresis: An Overview. *Methods in Molecular Biology* (1999) **112**, 1-7.

25) T. Rabilloud, Solubilization of Proteins in 2-D Electrophoresis. *Methods in Molecular Biology* (1999) **112**, 9-19.

26) Amersham pharmacia biotech, *Reversed Phase Chromatography: Principles and Methods*. Amersham pharmacia biotech, 1997.

27) M. P. Washburn, D. Wolters, J. R. Yates, III, Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature* (2001) **19**, 242-247.

28) Isoelectric Focusing Retrieved from: http://www.chemsoc.org/exemplarchem/entries/2003/leeds_chromatography/chromatography/ief.htm. T. Dixon, University of Leeds.

29) E. de Hoffman, V. Stroobant. *Mass Spectrometry: Principles and Applications*. 2nd edition, John Wiley and Sons Ltd., Chichester, 2002

30) G. E. Reid, S. A. McLuckey, 'Top Down' Protein Characterization Via Tandem Mass Spectrometry. *Journal of Mass Spectrometry* (2002) **37**, 663-675.

31) F. Meng, B. J. Cargile, S. M. Patrie, J. R. Johnson, S. M. McLoughlin, N. L. Kelleher, Processing Complex Mixtures of Intact Proteins for Direct Anaysis by Mass Spectrometry. *Analytical Chemistry* (2002) **74**, 2923-2929.

32) H. Steen, M. Mann, The ABC's (and XYZ's) of Peptide Sequencing. *Nature Reviews: Molecular Cell Biology* (2004) **5**, 699-711.

33) R. W. Kriwacki, G. Siuzdak, Combined Use of Proteases and Mass Spectrometry in Structural Biology. *Methods and Reviews* Retrieved from: http://www.abrf.org/JBT/1998/September98/sep98m_r.html. St. Jude Children's Research Hospital and The Scripps Research Insititute, 1998.

34) G. Siuzdak. *Mass Spectrometry for Biotechnology*. Academic Press Limited, London, England, 1996.

35) M. Karas, F. Hillenkamp, Laser Desorption Ionisation of Proteins with Molecular Masses Exceeding 10,000 Da. *Analytical Chemistry* (1988) **60**, 2299-2301.

36) S. L. Cohen, B. T. Chait, Influence of Matrix Solution Conditions on the MALDI-MS Analysis of Peptides and Proteins. *Analytical Chemistry* (1996) **68**, 31-37.

37) 'Matrix Selection' Retrieved from: http://www.srsmaldi.com/Maldi/Step/Mat_Sel.html. 2001.

38) C. H. Botting, The use of the slow crystallisation method to improve matrix-assisted laser desorption/ionisation time-of-flight signals for larger proteins. *Rapid Communications in Mass Spectrometry* (2000) **14**, 2030-2033.

39) F. Xiang, R. C. Beavis, Growing Protein-doped Sinapic Acid Crystals for Laser Desoprtion: an Alternative Preparation Method for Difficult Samples. *Organic Mass Spectrometry* (1993) **28**, 1424-1429.

40) L. F. Marvin, M. A. Roberts, L. B. Fay, Matrix-Assisted Laser Desorption/Ionisation Time of Flight Mass Spectrometry in Clinical Chemistry. *Clinica Chimica Acta* (2003) **337**, 11-21.

41) M. C. Fitzgerald, G. R. Parr, L. M. Smith, Basic Matrices for the Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry of Proteins and Oligonucleotides. *Analytical Chemistry* (1993) **65**, 3204-3211.

42) T. B. Farmer, R. M. Caprioli, Determination of Protein-Protein Interactions by Matrix-assisted Laser Desorption/Ionization Mass Spectrometry. *Journal of Mass Spectrometry* (1998) **33**, 697-704.

43) J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong and C. M. Whitehouse, Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* (1989) **246**, 64-71.

44) Schematic diagram of ESI MS from Mass Spectrometry Resource-Electrospray Ionisation, Retrieved from http://www.chm.bris.ac.uk/ms/theory/esi-ionisation.html.

45) K. F. Medzihradszky, J. M. Campbell, M. A. Baldwin, A. M. Falich, P. Juhaaz, M. L. Vestal, A. L. Burlingame, The Characteristics of Peptide Collision Induced Dissociation Using a High-Performance MALDI-TOF/TOF Tandem Mass Spectrometer. *Analytical Chemistry* (2000) **72**, 552-558.

46) National Centre for Biotechnology Information, U. S. National Library of Medecine. Retrieved from http://www.ncbi.nlm.nih.gov/

47) D. J. C. Pappin, P. Hojrup, A. J. Beasley, Rapid Identificaiton of Proteins by Peptide-Mass Fingerprinting. *Current Biology* (1993) **6**, 327-332.

48) Commonly Used Detergents; Frost, Membrane Lecture. Retrieved from www.med.ufl.edu/biochem/bch6206/DETERGENTS.pdf.

49) R. M. Garavito, S. Ferguson-Miller, Detergents as Tools in Membrane Biochemistry. *Journal of Biological Chemistry* (2001) **276**, 32403-32406.

50) IUPAC Compendium of Chemical Termiology Retrieved from: http://www.iupac.org/goldbook/co1395.pdf. 2nd Edition, 1997.

51) S. M. Bhairi, *Detergents: A Guide to the Properties and Uses of Detergents in Biological Systems*. Calbiochem-Novabiochem Corporation, 2001.

52) S. Reis, C. G. Montinho, C. Matos, B. de Castro, P. Gameiro, J. L. Lima, Noninvasive Methods to Determine the Critical Micelle Concentration of Some Bile Acid Salts. *Analytical Biochemistry* (2004) **334**, 117-126.

53) S. Luche, V. Santoni, T. Rabilloud, Evaluation of Nonionic and Zwitterionic Detergents as Membrane Protein Solubilizers in Two-dimensional electrophoresis. *Proteomics* (2003) **3**, 249-253.

54) R. R. Ogorzalek Loo, N. Dales, P.C. Andrews, Surfactant effects on protein structure examined by electrospray ionization mass spectrometry. *Protein Science* (1994) **3,** 1975-1983.

55) B. Chait, Retrieved from the PROWL website at: http://prowl.rockefeller.edu/contam/deterg.htm and http://prowl.rockefeller.edu/contam/esideter.htm

56) E. Nomura, K. Katsuta, T. Ueda, M. Toriyama, T. Mori, N. Inagaki, Acid-labile Surfactant Improves In-sodium Dodecyl Sulfate Polyacrylamide Gel Protein Digestion for Matrix-assisted Laser Desorption/Ionisation Mass Spectrometrix Peptide Mapping. *Journal of Mass Spectrometry* (2004) **39**, 202-207.

57) S. Konig, O. Schmidt, K. Rose, S. Thanos, M. Besselmann, M. Zeller, Sodium Dodecyl Sulfate Versus Acid-Labile Surfactant Gel Electrophoresis: Comparative proteomic Studies on Rat Retina and Mouse Brain. *Electrophoresis* (2003) **24**, 751-756.

58) M. Zeller, E. K. Brown, E. S. P. Bouvier, S. Konig, Use of an Acid-Labile Surfactant as an SDS substitute for Gel Electrophoresis and Proteomic Analysis. *Journal of Biomolecular Techniques* (2002) **13**, 1-4.

59) Care and Use Instructions for RapiGest™ SF. Waters Corporation.

60) K. O. Bornsen, M. A. S. Gass, G. J. M. Bruin, J. H. M. von Adrichem, M. C. Biro, G. M. Kresbach, M. Ehrat, Influence of Solvents and Detergents on Matrix-assisted Laser Desorption/Ionisation Mass Spectrometry Measurements of Proteins and Oligonucleotides. *Rapid Communications in Mass Spectrometry* (1997) **11**, 603-609.

61) Review article by K. M. Swiderek, A. J. Alpert, A. Heckendorf, K. Nugent, S. D. Patterson Structural Analysis of Proteins and Peptides in the Presence of Detergents: Tricks of the Trade. Retrieved from: http://www.abrf.org/ABRFNews/1997/December1997/dec97detergent.html. Written for The Association of Biomolecular Resource Facilities, 1997.

62) F. Gharahdaghi, M. Kirchner, J. Fernandez, S. M. Mische, Peptide-mass Profiles of Polyvinylidene Difluoride-bound Proteins by Matrix-assisted Laser Desorption/Ionisation Time-of-Flight Mass Spectrometry in the Presence of Nonionic Detergents. *Analytical Biochemistry* (1996) **233**, 94-99.

63) S. Yamamura, M. Nakamura, K. Kasai, H. Sato, T. Takeda, Synthesis and Properties of Destructible Anionic Surfactants with a 1,3-Dioxolane Ring and their use as Emulsifier for Emulsion Polymerization. *Journal of the Japanese Oil Chemistry Society* (1991) **40**, 16-20.

64) A. R. S. Ross, P. J. Lee, D. L. Smith, I. J. Langridge, A. D. Whetton, S. J. Gaskell, Identification of proteins from two-dimensional polyacrylamide gels using a novel acid-labile surfactant. *Proteomics*, (2002) **2**, 928-936.

65) J. L. Norris, N. A. Porter, R. M. Caprioli, Mass Spectrometry of Intracellular and Membrane Proteins Using Cleavable Detergents. *Analytical Chemistry* (2003) **75**, 6642-6647.

66) R. R. Ogorzalek Loo, T. I. Stevenson, C. Mitchell, J. A. Loo, P. C. Andrews, Mass Spectrometry of Proteins Directly from Polyacrylamide Gels. *Analytical Chemistry* (1996) **68**, 1910-1917.

67) L. R. Castellanos-Serra, C. Fernandez-Patron, E. Hardy, H. Santana, V. Huerta, High Yield Elution of Proteins from Sodium Dodecyl Sulfate-Polyacrylamide Gels at the Low-Picomole Level. Application to N-Terminal Sequencing of a Scarce Protein and to In-Solution Biological Activity Analysis of On-Gel Renatured Proteins. *Journal of Protein Chemistry* (1997) **16**, 415-419.

68) M. Galvani, E. Bordini, C. Piubelli, M. Hamdan, Effect of experimental conditions on the analysis of sodium dodecyl sulphate polyacrylamide gel electrophoresis separated proteins by matrix-assisted laser desorption/ionisation

mass spectrometry. *Rapid Commmunications in Mass Spectrometry* (2000) **14**, 18-25.

69) H. Ehring, S. Stromberg, A. Tjernberg, B. Noren, Matrix-assisted Laser Desorption/Ionisation Mass Spectrometry of Proteins Extracted Directly from Sodium Dodecyl Sulphate-Polyacrylamide Gels. *Rapid Communications in Mass Spectrometry* (1997) **11**, 1867-1873.

70) M. A. Jeannot, J. Zheng, L. Li, Observation of Sodium Gel-Induced Protein Modifications in Dodecylsulfate Polyacrylamide Gel Electrophoresis and Its Implications for Accurate Molecular Weight Determination of Gel-Separated Proteins by Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry. *Journal of the American Society of Mass Spectrometry* (1999) **10**, 512-520.

71) S. L. Cohen, B. T. Chait, Mass Spectrometry of Whole Proteins Eluted from Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis Gels. *Analytical Biochemistry* (1997) **247**, 257-267.

72) C. H. Botting, P. R. Antrobus, Development of a robust method for extracting whole proteins from SDS-PAGE gels in order to obtain mass spectrometric signals.

73) U. A. Mirza, Y. Liu, J. T. Tang, F. Porter, L. Bondoc, G. Chen, B. N. Pramanik, T. L. Nagabhushan, Extraction and Characterization of Adenovirus Proteins From Sodium Dodecylsulfate Polyacrylamie Gel Electrophoresis by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *American Society of Mass Spectrometry* (2000) **11**, 356-361.

74) C. H. Botting, Improved Detection of Higher Molecular Weight Proteins by Matrix-Assited Laser Desorption/Ionization Time-of-Flight Mass Spectrometry on

Polytetrafluoroethylene Surfaces. *Rapid Communications in Mass Spectrometry* (2003) **17**, 598-602.

75) A. G. Ferrige, M. J. Seddon, B. N. Green, S. A. Jarvis, J. Skilling, Disentangling Electrospray Spectra with Maximum Entropy. *Rapid Communications in Mass Spectrometry* (1992) **6**, 707-711.

76) K. G. Furton, A. Norelus, Determining the Critical Micelle Concentration of Aqueous Surfactant Solutions Using a Novel Colorimetric Method. *Journal of Chemical Education* (1993) **70**, 254-257.

77) M. M. Bradford, A Rapid and Sensitive Method for Quantitation of Microgram Quantities of Proteins Utilizing the Priniciple of Protein-Dye Binding. *Analytical Biochemistry* (1976) **72**, 248.

78) K. D. Speicher, O. Kolba, S. Harper, D. W. Speicher, Systematic Analysis of Peptide Recoveries from In-gel digestions for Protein Identifications in Proteome Studies. *Journal of Biomolecular Techniques* (2000) **11**, 74-86.

79) Retrieved from the Senior Honours Project Report of C. R. Buckner, Supervisor – Dr. Catherine Botting, The Synthesis of a Cysteine-Specific Mass Tag – Solving The Problem of Quantitation in Proteomics. University of St. Andrews, 2004.

80) W. E. Weaver, W. M. Whaley, The Preparation of Some α-bromoacetamides. *Journal of the American Chemical Society* (1947) **69**, 515-516.

81) L. R. Castellanos-Serra, C. Fernandez-Patron, E. Hardy, V. Huerta, A Procedure for Protein Elution from reverse-stained polyacrylamide gels applicable at the low picomole level: An alternative route to the preparation of low abundance proteins for microanalysis. *Electrophoresis* (1996) **17**, 1564-1572.

82) 'Textbook on High Performance Liquid Chromatography' Retrieved from: http://www.hplc.chem.shu.edu/NEW/HPLC_Book/index.html

83) P. Lecchi, H. M. Le, L. K. Pannell, 6-Aza-2-thiothymine: a matrix for MALDI spectra of oligonucleotides. *Nucleic Acids Research* (1995) **23**, 1276-1277.

Appendices

Appendix 1- MALDI-TOF MS Data

BSA-Solution Digest Data

BSA-Solution Digest	Score	% Coverage	Peptides Matched
1.25pm BSA	60	16	11
1.25pm BSA w/ALS	NO TEST		
1.25pm BSA	no ID		
1.25pm BSA w/ALS	65	18	10
2.5pm BSA	106	24	15
2.5pm BSA w/ALS	105	24	16
2.5pm BSA	74	22	15
2.5pm BSA w/ALS	95	19	14
2.5pm BSA	86	22	15
2.5pm BSA w/ALS	105	24	17
2.5pm BSA	64	18	12
2.5pm BSA w/ALS	no ID		
2.5pm BSA	50	16	13
2.5pm BSA w/ALS	NO TEST		

BSA – Manual In-gel Digest Data

BSA - In gel Manual Digest	Score	% Coverage	Peptides Matched
2.5pm BSA	244	37	21
2.5pm BSA w/ALS	103	21	10
2.5pm BSA	183	31	16

2.5pm BSA ALS	374	52	29
5 pm BSA	339	49	27
5 pm BSA w/ALS	270	45	27
5pm BSA	135	29	17
5pm BSA w/ALS	203	37	21

BSA – Robot In-gel Digest Data

BSA – In gel Robot digest	Score	% Coverage	Peptides Matched
2.5pm BSA	189	36	21
2.5pm BSA w/ALS	167	36	20
5pm BSA	170	37	20
5pm BSA w/ALS	236	37	21

ADH – Solution Digest Data

ADH-Solution Digest	Score	% Coverage	Peptides Matched
1.25pm ADH	no ID		
1.25pm ADH w/ALS	133	54	16*
1.25pm ADH	92	38	10
1.25pm ADH w/ALS	77	35	9
1.25pm ADH	153	45	15
1.25pm ADH w/ALS	124	42	12
2.5pm ADH	no ID		
2.5pm ADH w/ALS	64	49	13*
2.5pm ADH	61	29	7
2.5pm ADH w/ALS	no ID		

2.5pm ADH	142	45	14
2.5pm ADH w/ALS	76	38	10
2.5pm ADH	no ID		
2.5pm ADH w/ALS	131	42	12

ADH – Manual In-gel Digest Data

ADH - In gel Manual Digest	Score	% Coverage	Peptides Matched
2.5pm ADH	no ID		
2.5pm ADH w/ALS	74	23	
5pm ADH	93	33	8
5pm ADH w/ALS	89	23	6

ADH - Robot In-gel Digest Data

ADH – In gel Robot Digest	Score	% Coverage	Peptides Matched
2.5pm ADH	83	32	8
2.5pm ADH w/ALS	142	45	14
5pm ADH	114	40	12
5pm ADH w/ALS	96	36	10

CA - Manual In-gel Digest Data

CA - In gel Manual Digest	Score	% Coverage	Peptides Matched
2.5 pm CA	no ID		
2.5 pm CA w/ALS	125	32	7
5 pm CA	90	23	6

5 pm CA w/ALS	136	35	8
5pm CA no ALS	54		
5pm CA ALS	65	25	5

Ubiquitin - Manual In-gel Digest Data

Ubiquitin In gel Manual digest	Score	% Coverage	Peptides Matched
2.5pm Ubiquitin	no ID		
2.5pm Ubiquitin w/ALS	no ID		
2.5 pm Ubiquitin	51	50	3
2.5 pm Ubiquitin w/ALS	70	58	4
5pm Ubiquitin	no ID		
5pm Ubiquitin w/ALS	71	47	5
5 pm Ubiquitin	no ID		
5 pm Ubiquitin w/ALS	41		2

Ubiquitin - Robot In-gel Digest Data

Ubiquitin In gel Robot Digest	Score	% Coverage	Peptides Matched
5pm Ubiquitin (mine)	62	61	5
5pm Ubiquitin (mine) w/ALS	74	52	4
5pm Ubiquitin (Ro)	82	69	7
5pm Ubiquitin (Ro) w/ALS	88	65	4

Transferrin - Manual In-gel Digest Data

Transferrin In gel Manual digest	Score	% Coverage	Peptides Matched
1.25 pm trans	68	9	6
1.25 pm trans w/ALS	100	13	8
2.5pm Transferrin	187	28	19
2.5pm Transferrin w/ALS	255	35	21
2.5pm Transferrin no ALS	281	35	23
2.5pm Transferrin ALS	355	36	27
2.5pm Transferrin* no ALS	230	34	21
2.5pm Transferrin* ALS	229	30	19
2.5 pm transferrin	179	22	14
2.5 pm transferrin w/ALS	241	28	18
5pm Transferrin	145	19	14
5pm Transferrin w/ALS	270	40	26
5pm Transferrin no ALS	342	43	28
5pm Transferrin ALS	360	45	30
5pm Transferrin* no ALS	321	40	25
5pm Transferrin* ALS	341	40	26

Aldolase - Manual In-gel Digest Data

Aldolase In gel MANUAL digest	Score	% Coverage	Peptides Matched
2.5 pm Aldolase	90	23	7
2.5 pm Aldolase w/ALS	176	43	10
5 pm Aldolase	164	48	15
5 pm Aldolase w/ALS	170	34	12

β-galactosidase - Manual In-gel Digest Data

β-gal In gel Manual digest	Score	% Coverage	Peptides Matched
2.5pm B-gal no ALS	/		
2.5pm B-gal ALS	184	17	16
5 pm B-gal	221	25	21
5 pm B-gal w/ALS	210	22	23

Appendix 2 - LC-ESI MS Data

BSA - Manual In-gel Digest Data

BSA In gel Manual digest	Score	% Coverage	Peptides Matched
0.16pm BSA	127	4	3
0.16pm BSA w/ALS	267	8	5
0.31pm BSA	402	10	7
0.31pm BSA w/ALS	321	9	7
0.625pm BSA	673	19	19
0.625pm BSA w/ALS	345	10	10
2.5pm BSA	821	25	24
2.5pm BSA w/ALS	1634	56	61
5pm BSA	1744	53	88
5pm BSA w/ALS	2040	60	106

ADH - Manual In-gel Digest Data

ADH In gel Manual digest	Score	% Coverage	Peptides Matched
0.16pm ADH	142	7	3
0.16pm ADH w/ALS	122	4	2
0.31pm ADH	196	11	5
0.31pm ADH w/ALS	218	14	5
0.625pm ADH	277	14	6
0.625pm ADH w/ALS	315	19	9

Ubiquitin - Manual In-gel Digest Data

Ubiquitin In gel Manual digest	Score	% Coverage	Peptides Matched
2.5 pm ubiquitin	keratin		
2.5 pm ubiquitin w/ALS	196	61	5
2.5pm Ubiquitin 1	85	17	3
2.5pm Ubiquitin w/ALS 1	183	36	4
2.5pm Ubiquitin 2	78	32	2
2.5pm Ubiquitin w/ALS 2	169	44	3
5 pm ubiquitin	84	28	4
5 pm ubiquitin w/ALS	100	50	3
5pm Ubiquitin	58	30	3
5pm Ubiquitin w/ALS 1	186	36	5
5pm Ubiquitin w/ALS 2	165	40	3
5pm Ubiquitin w/ALS 3	165	40	3

Aldolase – Manual In-gel Digest Data

Aldolase In gel Manual digest	Score	% Coverage	Peptides Matched
2.5 pm Aldolase	373	24	9
2.5 pm Aldolase w/ALS	590	31	20
5 pm Aldolase	471	23	21
5 pm Aldolase w/ALS	617	30	26

β-galactosidase - Manual In-gel Digest Data

β-gal In gel Manual digest	Score	% Coverage	Peptides Matched
1.25 pm B-gal	91	2	2
1.25 pm B-gal w/ALS	keratin		
2.5 pm B-gal	362	7	7
2.5 pm B-gal	274	4	4