Examining the Dynamics of the Proteome Using Stable Isotopes as Metabolic Tracers

Andrea Pennington

Master of Philosophy

University of Liverpool

Faculty of Health and Life Sciences

Examining the Dynamics of the Proteome Using Stable Isotopes as Metabolic Tracers

Andrea Pennington

Abstract

The primary objective of this study was to examine the turnover of different proteins using stable isotopes as metabolic tracers. Two different labelling methods were used during the course of this investigation; $[^2H]_20$ and $[^2H_8]$ valine. Heavy water labelling resulted in the incorporation of deuterium into non-essential amino acids, and was primarily used to assess the turnover of mouse urinary proteins (MUPs). MUP's showed rapid deuterium incorporation, with equilibrium reached at day two, indicating high turnover. This result was anticipated as it had already been established that MUPs are made by the liver and rapidly excreted in the urine. Heavy valine was used to perform a more comprehensive study of turnover. Mice were provided with a diet in which 50% of the valine was in heavy. Heavy valine was incorporated into proteins via protein synthesis and labelling was monitored using mass spectrometry. As with the heavy water labelling, MUP's showed rapid labelling, with other proteins showing a range of different labelling rates. The most interesting observation in the heavy valine study was made when monitoring label incorporation in sperm and seminal vesicle secretion (SVS) proteins. SVS proteins showed rapid label incorporation from day two onwards, indicating high protein turnover. However the sperm proteins showed a distinct delay in labelling relative to the SVS proteins. It is hoped that this delay in the labelling of sperm proteins relative to SVS proteins could be exploited to selectively label specific proteins.

Contents Page

List of Figures – Page Numbers in Brackets

Acknowledgement

I would like to thank BBSRC for funding this study. I would also like to thank my supervisors Robert Beynon and Jane Hurst for their constant support and guidance throughout my MPhil. Finally I would like to thank all members of the Protein Function Group, in particular Amy Claydon and Lynn Mclean, for their help and friendship, without which this work would not have been possible.

1. Introduction

1.1. Aims of the project

This project is an exploration of the use of stable isotopes as metabolic tracers to assess the dynamics of the proteome. Specifically it addresses the protein turnover and differential labelling of proteins that may be involved in sperm competition.

1.2. Background

1.2.1. Proteomics

Proteomics is defined as the study of the protein complement of a cell or organism [1]. Many genomes have now been sequenced including the human genome, however there are still many genes for which the function is as yet unknown. There are several levels at which gene function can be studied; the genome, transcriptome, proteome, metabolome or the interactome level [1]. All of these levels of study can be very informative, however their true potential is not realised in isolation, but rather in conjunction with all the other areas of study. There are several key arguments that justify a protein level study. Firstly, there is a stark disparity between the number of genes in an organism and the number of proteins [1], and the basic assumption that one gene encodes one protein does not hold. Secondly, the number of genes within an organism is not an indicator of complexity. For example, the current estimate of the number of genes in the human genome stands at 20,500 [2], whereas the Arabidopsis genome is estimated to contain 26,500 genes [3]. Clearly humans

are much more complex organisms than plants. For example plants have three different tissue types (dermal, ground and vascular) [4] whereas humans have four tissue types (epithelial, connective, muscle, and nerve) however this is not reflected in genome size. Therefore the difference between organisms may lie in how they express their genes, rather than in merely gene content itself. Indeed, the evolution of more complex organisms appears to be the result of several other mechanisms such as alternative splicing, the interaction of signalling pathways and the diversification of gene regulatory networks [5]. Furthermore, in higher organisms proteins frequently act as a component of a much larger protein macromolecular complex. For example, the mature eukaryotic ribosome is composed of two subunits, the 60S subunit and the smaller 40S subunit, containing approximately 80 proteins in total [6]. For this reason, a direct, global study of the gene products i.e. proteins, and how their expression alters during development and environmental changes is needed to fully understand gene function.

1.2.2. Isotopes

Isotopes are different forms of the same element that differ in mass number (number of neutrons) but not atomic number (the number of protons), and therefore display virtually the same chemical properties. This is because it is the number and arrangement of electrons which determines how an atom behaves [1]. This remains unchanged between isotopes of the same element as the number of electrons equals the number of protons. Although the chemical properties of isotopes of the same element are the same, the physical properties (mass or radioactivity) often differ. This offers a possible way to distinguish between isotopes. Isotopes can exist in both radioactive and stable forms. Stable isotopes have a constant atomic mass whereas the mass of radioactive isotopes

changes during radioactive disintegration, with the emission of various sub-atomic particles [7].

1.2.3. Isotopes as Metabolic Tracers

A metabolic tracer is defined as an identifiable component containing atoms that can be distinguished from their normal counterpart by physical means, and thus can be used to follow the metabolism of the normal substances. In proteomics, the introduction of a tracer into a peptide (also known as a peptide label) is generally used to measure changes in protein expression between various different biological states [8]. Both radioactive and stable isotopes are ideal for use as metabolic tracers as they differ only in physical properties. Traditionally, proteins were labelled using radioactive isotopes (such as [^{14}C], [^{3}H] or [^{35}S]) due to the high specific radioactivity of radioisotopes and ease of detection (scintillation counting), even at low levels of incorporation. The use of radioactive isotopes as tracers was first described by Georg de Hevesy, a Hungarian chemist, in 1913 [9]. This concept originated during an attempt to separate Radium-D (RaD) from lead, when it became apparent that RaD was in fact a radioactive isotope of lead. De Hevesy reported that;

> *"The fourth decay product of radium emanation, RaD, shows all the chemical reactions as lead; if one mixes the RaD with lead or lead salts, the former cannot be separated from the latter by any chemical or physical methods. . . . Since RaD, as a result of its activity, can be detected in incomparably smaller amounts than lead, it can thus serve as a qualitative and quantitative proof [of the presence] of lead to which it was added: RaD becomes an indicator of lead."*

Thus the concept of using radioactive isotopes as biological tracers was born [10]. However the use of radioactive isotopes is not without its problems. Radioisotopes are hazardous in living organisms as they are a source of ionising radiation. This can have stimulatory effects on specific genes that can lead to cancer. Furthermore, it can be difficult to obtain the required degree of incorporation needed to label low abundance proteins or proteins that display a slow rate of turnover. However it is important to point out that the degree of incorporation does not need to be high in order to be detected, providing the specific radioactivity of the precursor pool is high enough [11]. In contrast, the incorporation of stable isotopes does not appear to have any stimulatory effects on specific genes (unlike radioisotopes), and is suitable for use with low abundance or low turnover proteins. For this reason stable isotope labelling techniques can allow for safe and reliable studies of the dynamics of the proteome, and is now the preferred strategy.

1.2.4. Stable Isotopes

The large scale study of the protein complement of a cell or tissue, **proteomics**, has its roots in the methodology of 2D gel electrophoresis [12]. In 2D gel electrophoresis, quantification was determined by recording the differences in staining patterns of proteins derived from two different cell types or tissues [13]. However this had two main problems; inaccuracy and the inability to quantify low abundance proteins that could not be visualised by the staining techniques available. The use of stable isotopes in conjunction with MS has long been used in other fields, for example in pharmaceuticals whereby an isotopic label is added to a compound of interest so that a known amount of sample can be analysed [14]. This is done because MS is not inherently quantitative due to varying detector response and different ionisation yields for different substances. However when MS is used in

conjunction with stable isotopes highly accurate quantification can be achieved using the ratios of isotope analogues, as the different species are chemically identical and are analysed in the same experiment [13].

The use of stable isotopes in conjunction with MS was first developed over 35 years ago, when it was reported that stable isotopes could allow for differentiation between internal standards and analytes during MS analysis [15]. This observation eventually lead to the use of stable isotopes in proteomics, a field that is rapidly evolving. A key experiment in the development of this field was described by Mann and colleagues in 1997. Here ^{18}O atoms were incorporated into the C terminus of a peptide in order to improve sequence assignment in peptide sequencing [16]. In the years following this experiment, stable isotopes have been used in a wide variety of proteomic investigations. The use of stable isotopes in proteomic studies offers a great source of labelling variety. The label can exist in the form on a non-amino acid precursor e.g. $[^{2}H]_2O$ or $[^{13}C]$glucose. Since isotopes display the same chemical behaviour, these heavy counterparts are incorporated into the organism in the same way as their native forms. However, the use of non-amino acid precursors introduces an inherent lack of control, with regards to isotope incorporation. For example, the use of $[^{13}C]$glucose can result in the incorporation of varying numbers of ^{13}C into different amino acids, thereby making subsequent analysis much more complex [17 and 18]. The labelling pattern becomes even more complex with the use of heavy nitrogen containing precursors as the label of choice e.g. $[^{15}N]H_4Cl$, whereby each peptide will incorporate varying levels of label, according to its length and the number of amino acids with nitrogenous side chains [17]. This makes it difficult to determine the absolute rates of protein labelling as the dynamics of labelling of different amino acids pools could vary [19].

A much preferred option in proteomic research is to use a heavy version of an amino acid to incorporate a label into a biological system i.e. an amino acid that has been engineered to contain an isotope analogue in specific regions of its structure. This allows a specific mass shift to be introduced for each labelled amino acid present in a particular peptide, thus simplifying analysis. Using a heavy amino acid offers a great deal of variety as regards to label choice, as a single amino acid can be labelled in variety of ways. This in turn enables researchers to perform multiplexed comparisons that would not be possible if other precursors were used [20]. When using amino acids as the labelling precursor, the choice of an appropriate amino acid is crucial. In animals, the 20 amino acids can be sub-divided into those which must be obtained through diet (essential) and those that the animal is able to synthesise (non-essential). If the labelled precursor is a non-essential amino acid, the relative isotope abundance of the precursor pool will be considerably diluted by synthesis *de novo*. Conversely, if an essential amino acid is chosen as the labelled precursor, the precursor pool will not be subject to the same level of dilution. The relative isotope abundance will however be diluted somewhat due to degradation of existing proteins within the organism. The abundance of the amino acid must also be considered. If an amino acid is highly abundant, it will appear in more peptides than its less abundant counterpart would. Since only peptides that contain a labelled amino acid can be informative, it is advantageous to choose an amino acid that is highly abundant.

1.2.5.Stable isotopes as tags (*in vitro*)

There are various ways that stable isotopes can be used to monitor the proteome. One option is to use a stable isotope as a tag that binds to specific regions of a protein. One such method developed by Gygi and colleagues involves using stable isotopes as affinity tags

(ICAT). ICAT are a group of reagents that consist of three functional elements; a specific chemical reactivity group, an isotopically coded linker and a biotin affinity tag [21]. Two versions of ICAT are used to compare the protein expression of two different cell states. Proteins from one cell state are derivatised with the isotopically light version of the reagent, and proteins from the second cell state are derivitised with the isotopically heavy version of the reagent [22]. After the proteins have bound to the thiol-specific reactive group (selectively binds cysteine residues) the two cell states are combined and the proteins are digested generating peptides, some of which are tagged [23]. The ICAT labelled peptides can then be isolated using the biotin affinity tag, analysed by liquid chromatography MS (LC-MS) and the relative amounts of protein determined [21]. Another option is to use isobaric tags for relative and absolute quantification (iTRAQ), a method developed by Warscheid and colleagues. iTRAQ is an improved approach, analogous to the aforementioned ICAT. The iTRAQ technique was developed in response to one of the major limitations of the ICAT strategy i.e. that this method is not applicable to cysteine free proteins [24]. When one considers the cysteine content of proteins, which tends to be fairly low, the true extent of this limitation is realised [24]. iTRAQ is now a well known method for the relative and absolute quantification of proteins [25-27]. In this technique, proteins are digested and the resulting peptides are labelled with amine specific isobaric tags [25]. The isobaric tag consists of a reporter group (mass 114-117 Da) and a balance group (mass 28-31 Da) [26]. Proteins from different cell states are tagged with isobaric tags that differ in reporter group mass, but not overall isobaric tag mass. The proteins from the different cell states are then combined and subject to MS. Subsequent protein quantification is then achieved by comparing the relative peak areas of the reporter ions, which have m/z values of 114.1, 115.1, 116.1 and 117.1 m/z [25].

1.2.6.Stable isotopes as metabolic tracers (*in vivo*)

An alternative to peptide tagging methods is to introduce a stable isotope as a metabolic tracer. Metabolic incorporation is distinct from the tagging strategies, as the label is incorporated into the protein before cell lysis i.e. *in vivo* [17]. Since the proteome is a very dynamic entity, methods which permit studies *in vivo* are obviously advantageous. As the use of stable isotopes as metabolic tracers became more widespread, several new important methodologies emerged most notably stable isotope labelling with amino acids in cell culture (SILAC). The SILAC method was developed in Mattius Mann's lab in 2002 and is similar to the ICAT method however instead of using affinity tags two cell states are grown in media containing either a heavy or a light version of an essential amino acid. As the cells divide the heavy (or light) amino acid becomes incorporated into all newly synthesised proteins [13]. Like the ICAT method the cells are then combined and digested, and the resultant peptides are analysed and quantified by MS. The value of this technique cannot be dismissed, however when dealing with intact animals, the incorporation of a metabolic label becomes more challenging [17].

Options for metabolic tracer incorporation *in vivo* include via diet or through bolus infusion. As mentioned previously there are a range of different stable isotope labelled compounds available two of which have been used during this study; deuterated water and heavy valine. Although stable isotopes can be incorporated into cells in culture with relative ease, when dealing with intact animals (*in vivo*) this becomes much more difficult. Proteins in complex animal systems may take several days or weeks to turn over. The label must be administered over an extended period of time since a significant amount of label must be

incorporated in order to permit turnover studies. This means that delivering the label by osmotic minipumps or isotope infusion is not an option due to the sheer amount of label that would be required and the resulting cost incurred [11]. A preferred strategy, and one which we have used, is to administer the label orally through diet. Diets that are entirely synthetic tend to be unpalatable to the animal; therefore diets tend to consist of 50% labelled and 50% unlabelled amino acid, or in the case of water 2-8% heavy water. In this study two percentages of heavy water were used; 2% and 8%.

1.2.7.Mass Spectrometry (MS)

Mass spectrometry (MS) is an analytical tool used to measure the mass of a sample. In proteomics we use MS to determine the mass of proteins or peptides. These mass values can then be searched against a database to determine the sequence of the protein and thereby identify it. There are a wide variety of mass spectrometers available but all have three main components; an ionisation source, a mass analyser and an ion detector [28] (Fig. 1A). First, the sample must be introduced into the instrument, this can be done directly or it can undergo chromatography before it reaches the ionisation source. This latter method of introduction usually involves the mass spectrometer being coupled to a chromatography system such as gas chromatography (GC) or high performance liquid chromatography (HPLC). The method of introduction is usually determined by the ionisation method being used and/or the complexity of the sample [28].

1.2.8.Ionisation Source

The ionisation source is responsible for generating a gas phase ion from a liquid or solid, neutral species. Different ionisation methods may be used two of which are electrospray

ionisation (ESI) and matrix assisted laser deopsorption/ionisation (MALDI). ESI creates gas phase ions by the application of an electric potential to a flowing liquid, containing the analyte and polar, volatile, solvent molecules [28]. To increase the detection limit achievable with ESI, the flow rate is usually very slow (nano litre/ minute levels) [29]. The solvent, containing the analyte, is passed through a narrow capillary with a voltage is applied to the capillary tip. As the liquid emerges from the tip, the applied potential causes it to disperse into an aerosol of highly charged droplets [28]. As they move into the mass spectrometer these droplets then diminish in size by solvent evaporation, usually aided by heat or another source of energy such as collisions with an inert gas [28]. Alternatively, MALDI can be used as the ionisation source. In MALDI, excitation with a laser generates gas phase ions [28]. The sample of interest is first spotted onto a target plate and allowed to dry. The sample is then overlaid with a matrix (in excess), which is also allowed to dry. There are several different compounds that can be used as a matrix, for this project α-cyano-4-hydroxycinnamic acid was used. When the matrix is dry, the matrix and the analyte are considered to be co-crystallised. The target is then placed in the mass spectrometer and a laser is fired at the spot of interest, generating gas phase ions. The ions are then guided into the mass spectrometer by electrostatic lenses [28].

It is important to understand that ions of multiple charge states can be generated during ionisation, and that the ionisation source used can dictate the type of ions generated. For example MALDI ionisation sources generate 1+ ions only (protonation results in $[M+H]^+$ molecular ions), whereas ESI is able to generate ions of multiple charge states (protonation can result in a variety of molecular ions such as [M+H]+, $[M+2H]^{2+}$ and $[M+3H]^{3+}$). Generating ions of multiple charge states introduces an additional level of complexity to the

interpretation of the mass spectrum produced. Figure 1B shows the same peptide, with a mass of 1500 at 3 different charge states (1+, 2+ and 3+). It can be seen that although mass of the peptide is consistent throughout (aside from the additional 1Da resulting from each increase in charge state) the mass over charge ratio (m/z) value of the peptide changes dramatically as a direct result of the charge exerted on the ion. The individual peaks in the isotopic envelope of a multiply charged ion display a closer proximity as the charge state increases. This is because the mass of the peptide remains the same however the charge (z) state increases, thus decreasing the observed m/z value [30].

1.2.9.Mass Analyser

After ionisation, the ions pass into the mass analyser. The function of the mass analyser is to separate, or resolve, ions according to their m/z. Various mass analysers exist, with the main types being; time of flight (TOF) analysers, quadrupole analysers, quadrupole ion-trap analysers and Fourier transform-ion cyclotron resonance (FT-ICR) analysers [5]. The decision of which mass analyser to use is usually dependant on the ionisation source, as some are more suitable for use with particular ionisation methods. For example, both of the mass analysers mentioned above can be coupled to ESI, however a quadrupole is not usually coupled to a MALDI ionisation source [5].

When a TOF analyser is used the peptide ions are accelerated into a flight tube by the application of an electric field. During transit the peptides are separated according to their m/z value. Ions with a low m/z value fly much faster, whereas ions with high m/z fly slower. After flight the ions are detected and their time of flight is recorded. The time taken for an ion to traverse a fixed distance can be used to deduce the m/z of the peptide [63].

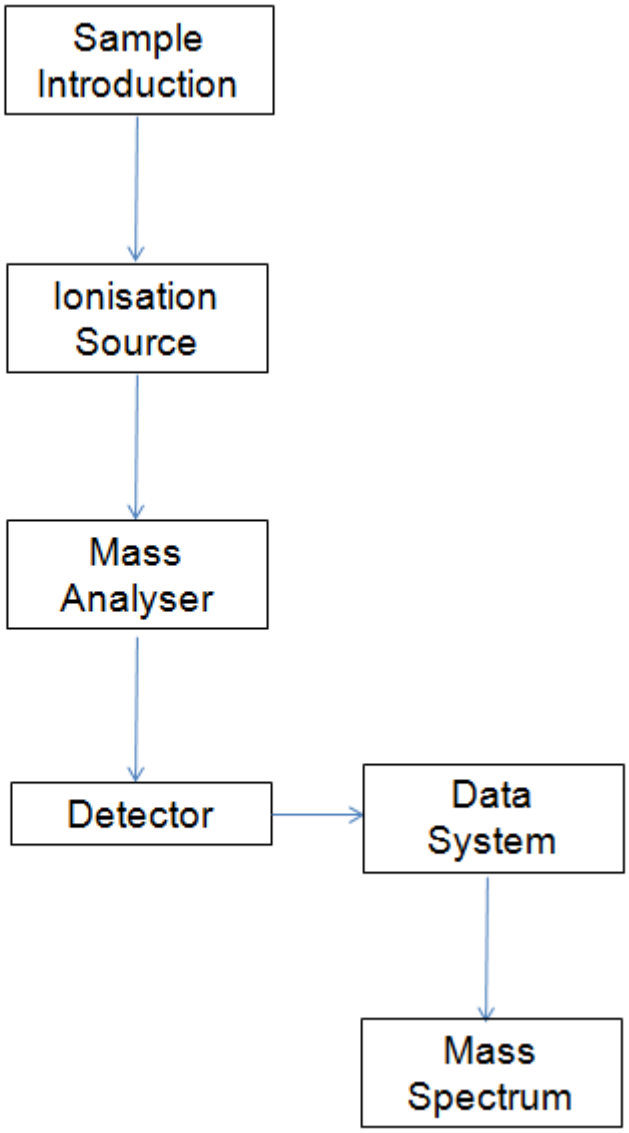


Figure 1A. Idealised schematic view of a mass spectrometer. Outline of the components common to all mass spectrometers. Adapted from Proteins and Proteomics (2003, Cold Springs Harbour Press, first edition).

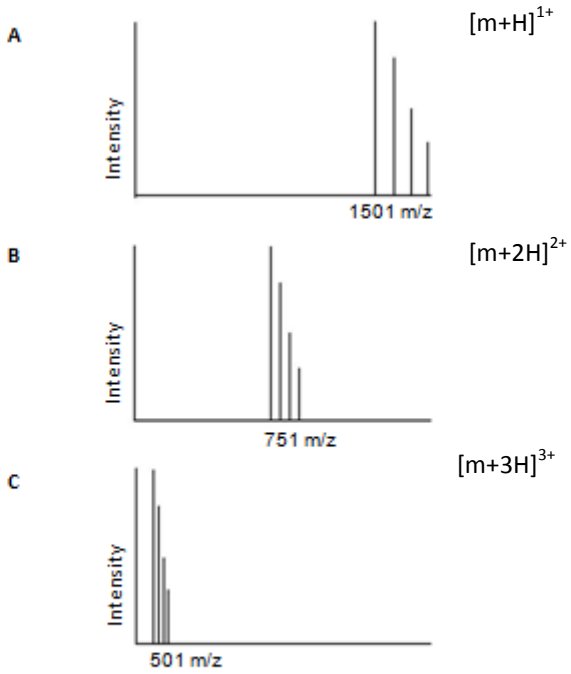


Figure 1. Isotopic distribution of multiply charged ions. A shows a peptide with a mass value of 1500 at a charge state of 1+. **B** and **C** show the same peptide at a charge state of 2+ and +3 respectively. It can be seen that as the charge on the peptide increases the m/z of the peptide and the distance between the peaks in the isotopic envelope decrease.

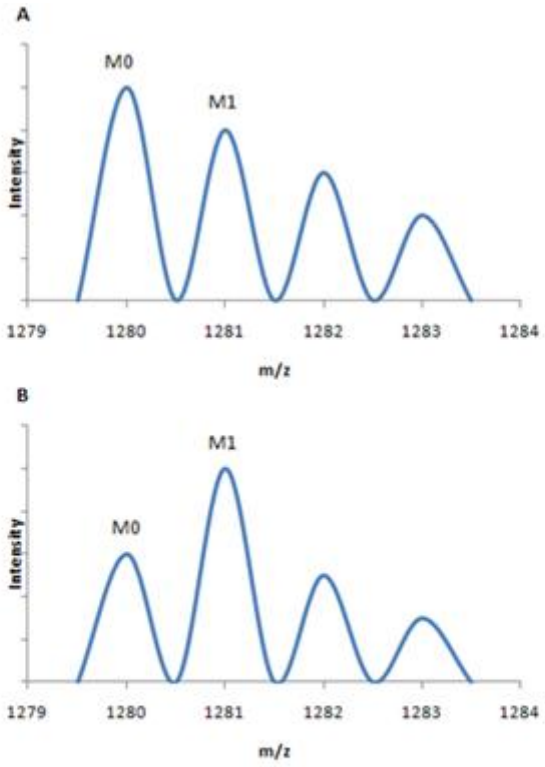


Figure 2. Isotopic distribution of a typical unlabelled peptide compared to a labelled version of the same peptide. A shows the distribution of a unlabelled peptide, with the M0 peak more intense relative to the M1 peak. B shows the distribution of a labelled peptide, with the M1 peak more intense relative to the M0 peak.

This is done by calibrating the instrument before use with known peptides with a known m/z value, and using this as a reference point to infer the m/z value of an ion from its time of flight [63].

Tandem mass spectrometry (MS/MS) has also been used in this study. MS/MS can be defined as a sequence of several events; mass selection of a precursor ion at the first stage of analysis, an intermediate reaction event, finally followed by the analysis of the resultant product ions in a second analysis stage [31]. The intermediate stage involves the fragmentation of selected precursor ions into product ions with different m/z values. Fragmentation can be achieved in a number of ways such as by the collision of the precursor with an inert gas (collision induced dissociation, or CID), collision with a surface (surface induced dissociation, or SID), or electron capture dissociation (ECD), to name a few [31]. MS/MS can be a very useful tool in proteomics as it can reveal more information regarding the structure of the peptide, such as the amino acid sequence [31].

1.2.10.Stable Isotopes in the Context of Mass Spectrometry

As mentioned previously, the two stable isotope precursors that have been used in this study are $[^2H_2]O$ heavy water and $[^2H_8]$valine. Heavy water labelling offers a simple and cost effective way of incorporating a heavy isotope into a biological system. After oral administration, $[^2H]$ is incorporated into non-essential amino acids (AA) via biosynthetic pathways (at which point hydrogen atoms can be exchanged for deuterium atoms) and in turn these AA are incorporated into proteins during translation [17]. The incorporation can be determined at the level of the protein (difficult), peptide (acceptable but analytically challenging) or amino acid (analytically straightforward but technically challenging). Any

exchangeable hydrogen in an AA can potentially be replaced by deuterium, therefore in a protein there are a huge number of possible sites for insertion. If only one ^{2}H is incorporated into a peptide this causes a mass shift of 1Da to give a peak denoted M1 – the monoisotopic peak is referred to as M0. The ^{13}C peak in the isotopic envelope is also 1Da heavier than M0, and therefore the M1 peak and the ^{13}C become indistinguishable (Figure 2). As the peptide becomes more extensively labelled we see a shift, whereby M0 becomes less intense relative to M1. Other peaks such as M2 and M3 also increase in intensity as the labelling continues. Eventually equilibrium is reached, and the peptide cannot be labelled any further unless a higher percentage of heavy water is introduced. Quantification of the degree of labelling present can then be determined by comparing the results from the labelling experiments with theoretical unlabelled spectra generated by mass isotopomer distribution analysis (MIDA) [32].

Heavy valine is also a useful labelling precursor. When using labelled amino acids, selection of the most appropriate amino acid to act as the label is crucial. Heavy valine is an appropriate choice as it is the second most abundant essential amino acid in most proteomes therefore most proteins are likely to contain a valine residue, allowing for monitoring of turnover [17]. Leucine is the most abundant amino acid however there are complications with the use of this amino acid because of the isobaric isoleucine (leucine and isoleucine have the same mass) [19]. Secondly valine is an essential amino acid, therefore there is no chance of the precursor pool being diluted by synthesis *de novo*. The precursor pool will be diluted somewhat, however, due to the release of amino acids from the degradation of existing proteins [19]. The resulting pattern seen in the spectrum when labelling with heavy valine is quite different to that of heavy water. This time when

peptides are labelled we see a 7Da shift, as opposed to the 1Da shift observed in heavy water labelled peptides [64]. Although proteins were labelled with [$^2H_8$]valine, the deuterium on the α-carbon is lost by the reversible process of transamination [17], hence why the mass shift seen when the peptides are analysed is consistent with the incorporation of seven deuterium atoms. Transamination can be tissue specific, thus some valine residues may be incorporated as [$^2H_8$]valine [17].

1.2.11.Applications of Metabolic Tracers

1.2.11.1.Metabolism of Proteins (Protein Turnover)

One of the areas we have applied stable isotope labelling to in this study is protein metabolism. Metabolism can be defined as a group of chemical processes required to maintain life and is usually subdivided into two categories; catabolism (the degradation of biological molecules) and anabolism (the synthesis of biological molecules). This project is specifically looking at the metabolism of proteins, known as protein turnover. Initially proteins were viewed as very stable molecules, only subject to minor "wear and tear" over time [33]; however we now accept proteins to be very dynamic entities that are constantly being synthesised and degraded. It is important to understand that the level of protein within a cell is not just dependant on the rate at which it is synthesised but also the rate at which it is degraded. For this reason, inferring protein concentrations from mRNA abundance levels alone is inappropriate as it assumes the main influence on protein content within a cell to be synthesis, and neglects to take degradation (both regulated and unregulated) into consideration [33]. Furthermore an mRNA focused study cannot detect changes in protein concentration due to post-translational modifications.

Now that the importance of measuring turnover at the protein level (rather than just the mRNA level) is more widely appreciated a new field has started to emerge; dynamic proteomics. As the name suggests the study of proteome dynamics seeks to assess the rate at which the proteins present in a cell or tissue are turned over, rather than just quantifying and identify them. This study can be on a proteome wide scale, or focussed on a particular protein or group of proteins. The study of protein turnover is important for several reasons. Firstly the turnover rate of a particular protein can give us clues about its function. Protein turnover is a very energetically demanding process, therefore one assumes that a cell would not rapidly turnover a protein unless it was a requirement of its function. For this reason, rapidly turned over proteins tend to have regulatory roles, whereby the rapid synthesis and degradation of these proteins is essential to the cell's ability to respond to changing conditions. Despite the recognized importance of this field, only a small minority of the many proteomic studies published to date provide a comprehensive description of the dynamics of proteins present in a particular system [34]. Undoubtedly there has been much advancement in proteomics in recent years such as improved protein separation techniques and the evolution of highly sophisticated mass spectrometers, however identifying and quantifying all the proteins present in a biological system remains technically challenging. One of the reasons for this is that in order to quantify proteins the quality of the data (in terms of information content) needs to be much higher than that required for protein identification [34]. Although MS itself is not inherently quantitative, when used in conjunction with stable isotope labelling, the turnover rate of a protein of interest can be calculated directly from the resulting heavy and light spectrums. Let us take heavy valine labelling as an example. As mentioned previously, when labelling with heavy valine a 7Da mass shift is introduced. This 7Da mass shift is advantageous as it large enough to separate

the C13 peaks of the unlabelled and labelled peptides in m/z space [19]. Although heavy valine is referred to as [$^{2}H_{8}$] valine, one [^{2}H] is frequently lost by transamination during protein synthesis, thus explaining the 7Da shift. The exact appearance of the spectrum is dependent upon how many valine residues are present in a particular peptide and the relative isotopic abundance (RIA) of the precursor pool. Spectra generated by heavy valine labelling can appear quite complex, especially in peptides with multiple valine residues. However with the application of some relatively straightforward equations, as we will see later on, the spectra can be separated into old and new protein and turnover can be calculated.

1.2.11.2.Sperm Competition

A second research area that stable isotope assisted turnover studies can be applied to is sperm competition. Sperm competition is the result of heterospermic insemination (HI) which is defined as the condition whereby a female receives sperm from more than one male within a short period of time, so that a mixture of sperm is available for fertilisation [35]. Sexual selection on males therefore exists past the stage of copulation and extends into the gametic level, with the ejaculates of rival males competing for fertilisation of the ova [36]. The large amount of genetic variation observed between species in the sperm and seminal vesicle secretions (SVS), and in the reproductive anatomy supports this idea of the degree of sperm competition impacting on reproductive investment. Seminal vesicles are accessory glands that secrete a diverse range of proteins that appear to influence the outcome of sperm competition [36]. For example in rodents, species that display a more promiscuous mating system have been shown to have a larger relative seminal vesicle gland and testis size, than that of species where promiscuity is less prevalent [36].

Although rodent sperm is short lived and females do not generally store the sperm for long, ejaculates from multiple males frequently overlap in the female reproductive tract and the importance of adaptations in regards to sperm competition is now widely accepted [37]. It is hoped that this study will lead to the possibility of using stable isotope labelled valine residues as ownership tags to assess the degree of ejaculate investment made by individuals during varying levels of sperm competition.

2. Materials and Methods

2.1 Chemicals

Acetonitrile, methanol and trifluoroacetic acid (all HiPerSolve grade) were obtained from BDH (Poole). Tris and formic acid (both AnalaR grade) were obtained from BDH. SDS and di-sodium-hydrogen phosphate were obtained from BDH. Ammonium persulphate, Tetramethylethylenediamine and all 1D gel equipment was obtained from Biorad (Hemel Hempstead). Coomassie blue, α-cyano-4-hydroxycinnamic acid, ammonium bicarbonate and iodoacetamide were obtained from Sigma (Dorset). Acrylamide (30%) was obtained from Severn Biotech (Kidderminster). Dithiothreitol was obtained from Melford Laboratories (Ipswich). Coomassie protein plus assay and BSA standards were obtained from Pierce and Warriner (Chester). Rapigest was obtained from Waters (Elstree). Trypsin (sequencing grade) was obtained from Roche (Lewes).

2.2 Homogenisation of Tissue Samples

Tissue samples are homogenised to break down tissues and cells and release protein. A homogenisation buffer was made up to the following specifications ; 50mM di-sodium hydrogen phosphate (1420 mg in 200mL) and 1 homogenisation buffer tablet was added per 50mL of buffer. Tissue samples were weighed and homogenisation buffer was added to the tissue at a ratio of 1mL of buffer per 100mg of tissue. The sample was then homogenised on ice for 1 min at increasing speeds (20sec at speed 4, 20sec at 5 and 20sec at 6). After this the samples were spun down (15 min at 13,000g) and the supernatant was extracted and a protein assay was performed. The pellet was also retained in case needed at a later date (-80°C).

2.3 1D SDS Poly Acrylamide Gel Electrophoresis (PAGE)

The purpose of 1D SDS PAGE is to separate proteins by mass. 15% SDS gels were cast in the laboratory using a standard procedure. 10µL (1µg/µL) of sample was added to 10µL of sample buffer and incubated at 95˚C for 20 minutes. When set, samples were loaded onto the gels and the gels were run at 200mV for 45 minutes – 1 hour. Gels was stained for 3 – 5 hours using Coomassie blue and destained over night. After destaining the gels were stored at 4˚C until needed.

2.4 2D Gel Electrophoresis

2D protein separation techniques are used to separate proteins based on molecular weight and charge (iso-electric focusing). This technique is useful when trying to identify low abundance proteins that have the same molecular weight as high abundance proteins. Pre-cast IPG strips (pH 3-10, 13cm) were used to separate proteins in the first dimension, and gels cast in the laboratory 8 x 10cm) were used to separate proteins in the second dimension. A standard 2D gel protocol was followed. Before loading samples, protein assays were performed and 0.5mg total protein was loaded onto each gel.

2.5 In-Gel Digestion

The purpose of an in-gel digest is to obtain peptides from proteins by digestion with a protease. All incubations were performed at 37˚C.

Gel plugs were extracted from bands of interest using a glass pipette and transferred to 500µL eppendorfs. The gel plugs were destained in 20µL of 50mM ammonium bicarbonate (ambic)/ 50% (v/v) acetonitrile and incubated for 15 minutes. This was repeated until the plugs were fully destained. Destain was discarded and 50µL of dithiothreitol (DTT) (20mM) was added. Samples were incubated for 30 minutes, and then the DTT was discarded. 20µL of iodoacetamide (IAN) (55mM) was added and plugs were incubated for 30 minutes. IAN was discarded and 10µL of acetonitrile (ACN) was added, plugs were incubated for 15 minutes, or until they turned white. ACN was discarded and the plugs were allowed to dry for 1 hour. A master mix of 50mM ambic and 100ng/µL trypsin was made up in a 9:1 ratio. 10µL of master mix was added and samples were incubated overnight. After overnight incubation samples were stored at -20˚C until needed. A slight amendment was made to this protocol in later experiments. After treatment with IAN samples were washed three times in 25mM ambic to destroy any excess IAN that might have been interfering with digestion.

2.6 In-Solution Digestion

Like in-gel digests, the purpose of an in-solution digest is to obtain peptides from proteins, however less material is lost as the gel stage is eliminated. In-solution digests were performed using the standard Waters protocol. This protocol has two main features; the use of a digestion enhancer (Rapigest) which is destroyed after digestion, and compulsory reduction and alkylation steps. After overnight incubation samples were stored at -20˚C until needed.

2.7 Zip Tip

Zip Tip treatment is used to concentrate and desalt peptide samples. Zip Tipping was performed according to the manufacturer's instructions. Briefly, 4µL 1% (v/v) TFA was added to the sample. 10µL wetting solution was aspirated and dispensed 10 times to waste. Peptides were bound to the zip tip by aspirating and dispensing the sample at least 15 times into its original tube. 10µL wash solution was aspirated and dispensed to waste. To elute, 4µL of matrix solution was aspirated and dispensed through the zip tip at least 15 times without introducing air into the sample. Sample was eluted into a clean 500µL Eppendorf, ready for MALDI TOF analysis.

2.8 MS Analysis

Sample preparation – 1µL of each of the Zip Tipped samples were spotted onto a MALDI target and allowed to dry. Samples that had not been zip-tipped were spotted on a MALDI target and allowed to dry; they were then overlayed with 1µL matrix which was also allowed to dry. The target was then placed in the mass spectrometer (Axima MALDI TOF instrument obtained from Shimazdu Biotech (Milton Keynes) and a laser was fired at the spot of interest. The matrix then absorbed the energy received from the laser and became ionised. The matrix had two main roles. Firstly to protect the analyte molecules from the high energy of the laser. Secondly to pass on part of the energy generated by the laser to the analyte molecules, resulting in the production of singly charged, gas phase ions [62].

In the middle of every four samples a standard was positioned so that the instrument could be calibrated before use. The instrument was operated in reflectron mode at 50-60 % laser energy.

2.9 Heavy Water Labelling

Three groups of six male mice (wild mice bred in captivity for six generations or less) were provided with a normal diet and heavy water over a 5 day period and allowed to drink *ad libitum*. The first group were given 2% (v/v) heavy water, the second 4% (v/v) and the third 8% (v/v). Urine samples were taken at 24 hour intervals over 6 days (day 0 – day 5). At days 1, 2 and 5 two mice from each group were humanely killed and plasma samples were taken. All samples were stored at -20˚C until needed.

2.10 Heavy Valine Labelling

12 male mice (wild mice bred in captivity for six generations or less) were provided with water and fed a carefully constructed diet for a maximum of 35 days. They were allowed to drink and eat *ad libitum*. The diet consisted of all essential amino acids (at the required amounts according to the nutrient requirements of the mouse) in protein form. To this, a specific amount of heavy valine in crystalline form was added, such that overall there was twice as much valine in the diet (than was required), of which 50% was heavy. The mice were maintained at a 12:12 light dark cycle over the course of the experiment.

At specific time points (2 days, 7 days, 14 days, 25 days and 35 days) two mice were humanely killed. Liver and kidneys were removed and muscle, seminal vesicle secretions, epididymis and sperm samples were taken. While alive, urine samples were taken from each mouse at the same time points. All samples were stored at -80˚C until needed.

2.11 MIDA Data Processing

MIDA was used when analysing the spectra resulting from heavy water labelling, so that the isotopic envelope of a labelled peptide could be compared with that of a theoretically generated unlabelled peptide [67]. The m/z value of a particular peptide of interest was entered into the MIDA computer program and a list of the relative intensities of each peak in the envelope, and the corresponding m/z values for each peak was generated. This data was then used to deconvolute the labelled peptide spectrum, so that it could be determined what proportion of each peak was due to deuterium labelling and what proportion was the result of the presence of the naturally occurring C^{13} isotope within the peptide structure [67].

3. Results and Discussion

3.1 Heavy Water

3.1.1 Heavy Water Labelling in Urine Proteins

As mentioned previously, both urine and plasma samples were collected from animals that were fed a heavy water diet. It was decided to look at the urine samples first, as urine is relatively simple in terms of protein content, compared to plasma [38 and 39]. Furthermore, the most abundant proteins found in mouse urine; major mouse urinary proteins (MUPs), are made by the liver and rapidly excreted (MUPs identified by a Swiss Prot database search). The effect of heavy water can be seen almost immediately in these proteins, making them a good indicator of whether or not labelling is occurring. Firstly 1D SDS gels were run to separate out the urine proteins according to their molecular weight (Fig 3). As can be seen, one band of protein or proteins dominates each lane at approximately 20 kDa. Previously published work indicated that the proteins contained in these bands were major MUPs [40], however to confirm this, the bands were excised, digested using trypsin and analysed by MALDI TOF MS. The resulting spectra were then used to conduct a database search, with each digest indicating the presence of several MUPs. Around ten different MUPs exist, all with very similar sequences and molecular weights (explaining why several MUPs were found in the same 1D band [40]). In these particular samples 4 different MUPs were found to be present consistently; MUP 2, MUP 6, MUP 8 and MUP 1.

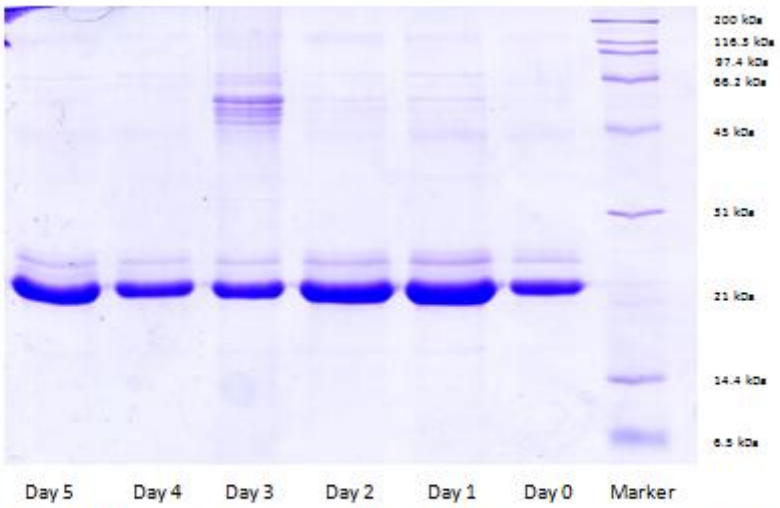


Figure 3 1D SDS gel of urine samples from animal KM6499.5 . Animal was fed 2% heavy water over a five day period, and urine was collected at 24 hour intervals. The urine was then diluted 1 in 5 with distilled water before loading onto the gel.

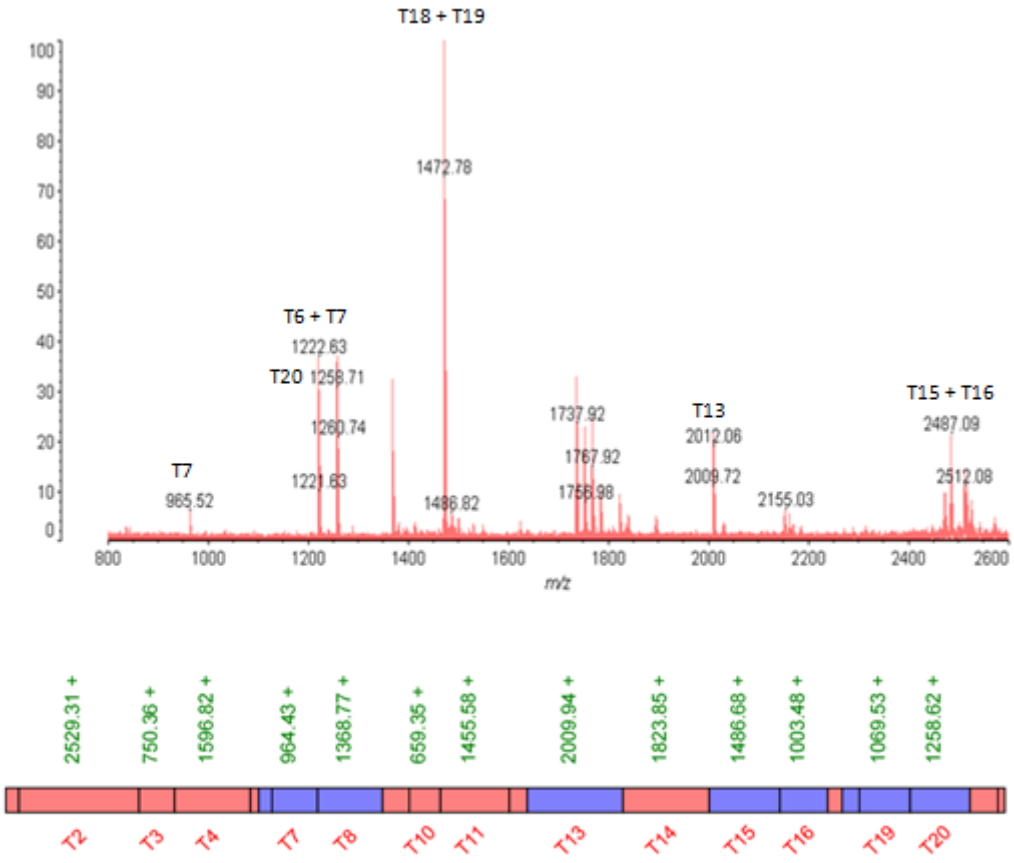


Figure 4. Spectrum generated by MALDI TOF MS using the Axima from an in-gel digest of a mouse urine sample. A database search using Swis-prot identified this sample as MUP 2 protein, with a sequence coverage of 46%. The peptides identified as being from MUP 2 are indicated on the spectrum, three of which are mis-cleaves.

Figure 4 shows a typical spectrum obtained from the digested urine samples. A database search on this particular sample resulted in MUP 2 receiving the highest mouse score. The individual MUP 2 peptides are indicated in the spectrum, and their position within the protein is shown in the peptide map. The sequence coverage obtained for this sample was fairly typical of most of the urine samples, and this was judged to be enough to confidently say that MUPs were present in each band.

After protein identification, work then moved on to looking for evidence of heavy water labelling. To do this, each peptide identified in each sample was looked at in more detail in attempt to track the progression of labelling over the five day period. Although seemingly straightforward, this proved to be quite difficult, as many peptides were present in some samples but not in others. Furthermore, some peptides, although present, displayed spectra that were of low quality, thus making it very difficult to deduce whether or not labelling had occurred, and to what extent. However, after several repeat experiments, and a slight adjustment to the in-gel digest protocol (to try and make the sample cleaner and therefore reduce the noise in the resulting spectrum), higher quality spectra, and increased sequence coverage were obtained (Fig. 5). It was then possible to track several peptides over the course of the labelling experiment. As the peptide became progressively labelled, the distribution of the isotopic envelope changed dramatically. If a peptide is smaller than 1800-2000 kDa, we usually expect the M0 peak to be the most intense, with each successive peak becoming less intense (Fig 6, day 0). However when a peptide is labelled with deuterium, this causes a mass shift of 1Da for every deuterium molecule that replaces a hydrogen molecule. This mass shift occurs because from the body water, [^{2}H] rapidly enters free nonessential amino acids through intermediate metabolic pathways [41].

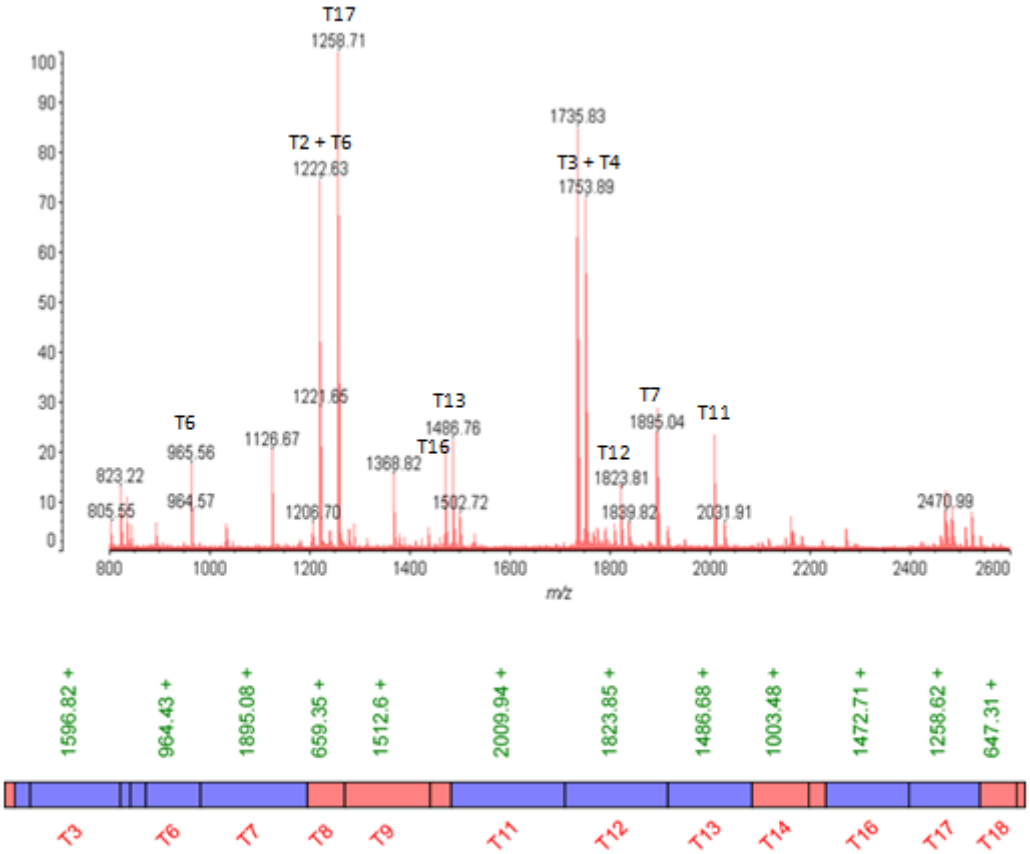


Figure 5. Spectrum generated by MALDI TOF MS using the Axima from an in-gel digest of a mouse urine sample. A database search using Swis-prot identified this sample as MUP 8 protein, with a sequence coverage of 72%. The peptides identified as being from MUP 2 are indicated on the spectrum, three of which are mis-cleaves.

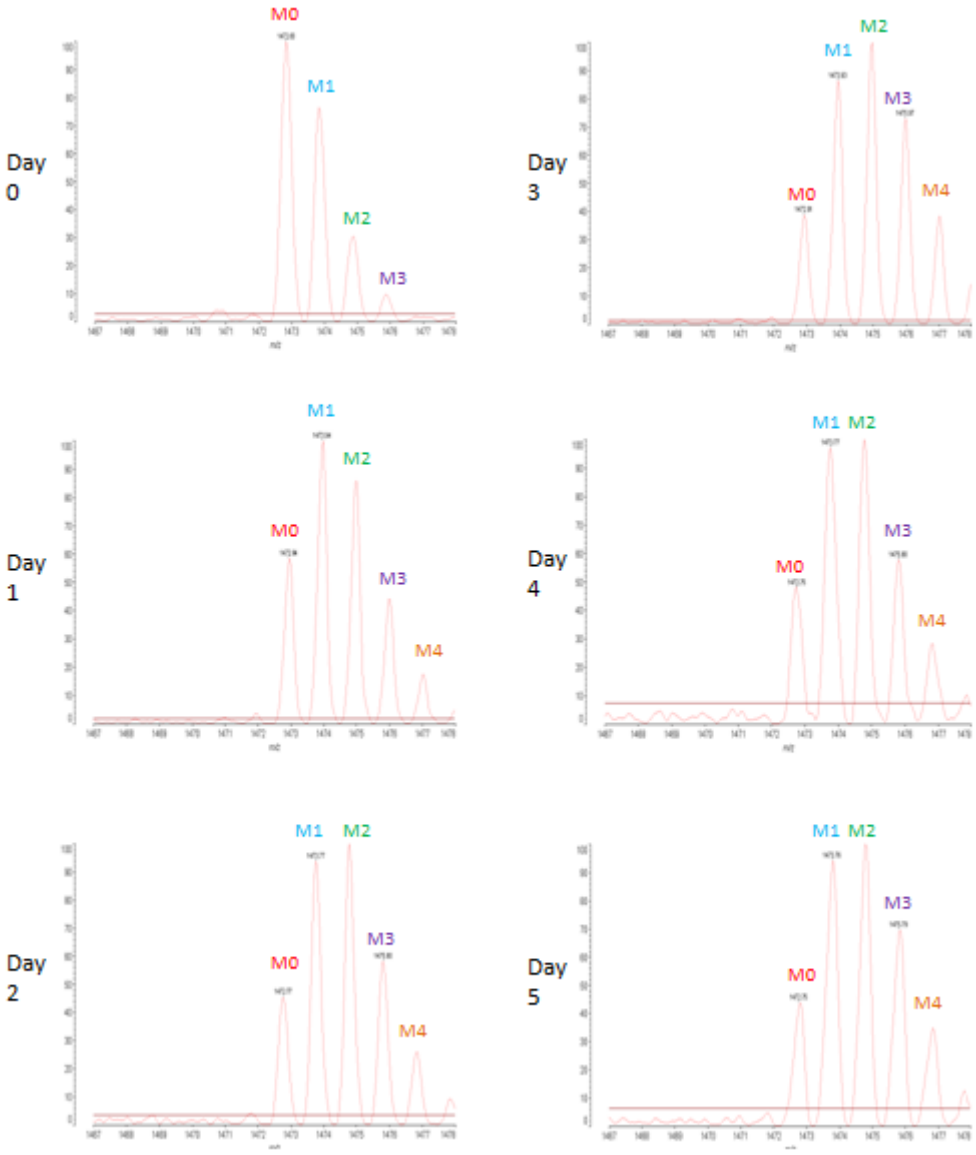


Figure 6. MUP 2 peptide T18 + T19 (mis-cleave) over the course of a 5 day heavy water labelling experiment. Over progressive days distribution moves to the right, so that the M1 peak becomes more intense relative to the M0 peak. Equilibrium appears to have been reached on day 2 – day 3 as the distribution does not change significantly after this point.

In addition all free amino acids are labelled at their alpha carbon by transamination [41]. After 1 day of heavy water exposure, we saw that the isotopic distribution had changed significantly, so that the M1 peak became the most intense (Fig 6, day 1), indicating that the peptide had started to become labelled. By day 2 the peptide showed evidence of further labelling, as the spectrum shifted further to the right, with the M2 peak becoming the most intense. After day 3 the distribution does not seem to alter, suggesting equilibrium has been reached i.e. the peptide will not label any further at 2% heavy water.

Once it had been established that labelling was occurring, focus then moved to calculating the degree of labelling observed in each sample. It is important to realise that the M1 peak does not just represent peptides that contain a substituted deuterium, it also represents peptides that contain a ^{13}C isotope in place of ^{12}C. ^{13}C is a naturally occurring stable isotope with an abundance of 1.1% [42]. The presence of ^{13}C in a peptide causes a mass shift of 1Da for every ^{13}C present, the identical effect that deuterium labelling has, making the M1 peak and the C13 peak indistinguishable. Thus, in order to determine the degree of labelling present, each peak in the spectrum must be separated into light (no deuterium labelling) and heavy (deuterium labelling). To do this, a theoretical isotopic distribution for unlabelled peptide must first be generated. This is done by MIDA. The theoretical distribution can then be used to deconvolute the labelled spectrum, revealing to what degree the peptide is labelled. If we look at figure 7 we can see the spectra of a MUP peptide on day one and day three of the labelling experiment, which both clearly show labelling.

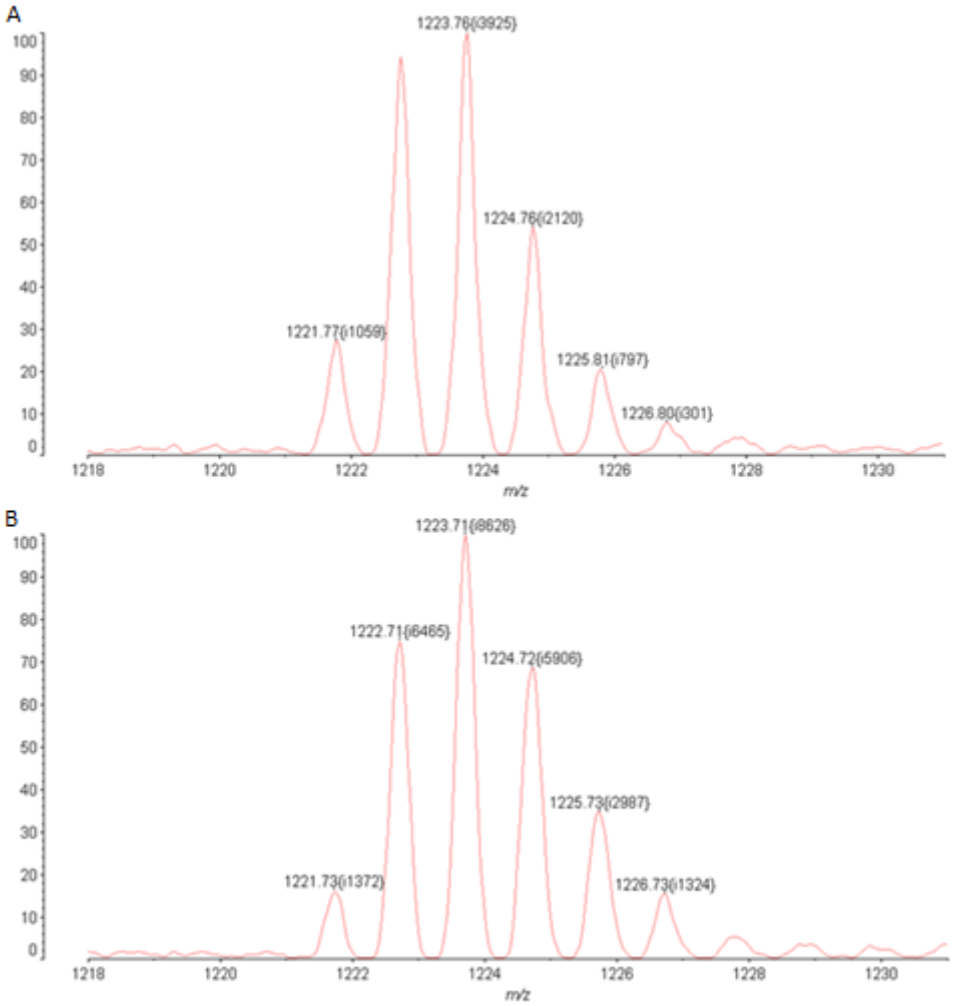


Figure 7. Comparison of the spectra from a MUP peptide at two time points in the labelling experiment. Peptide EKIEDNGNFR was obtained from urine samples from a mouse on a 2% heavy water diet on different days of the labelling experiment. (A) Day 1 (B) Day 3.

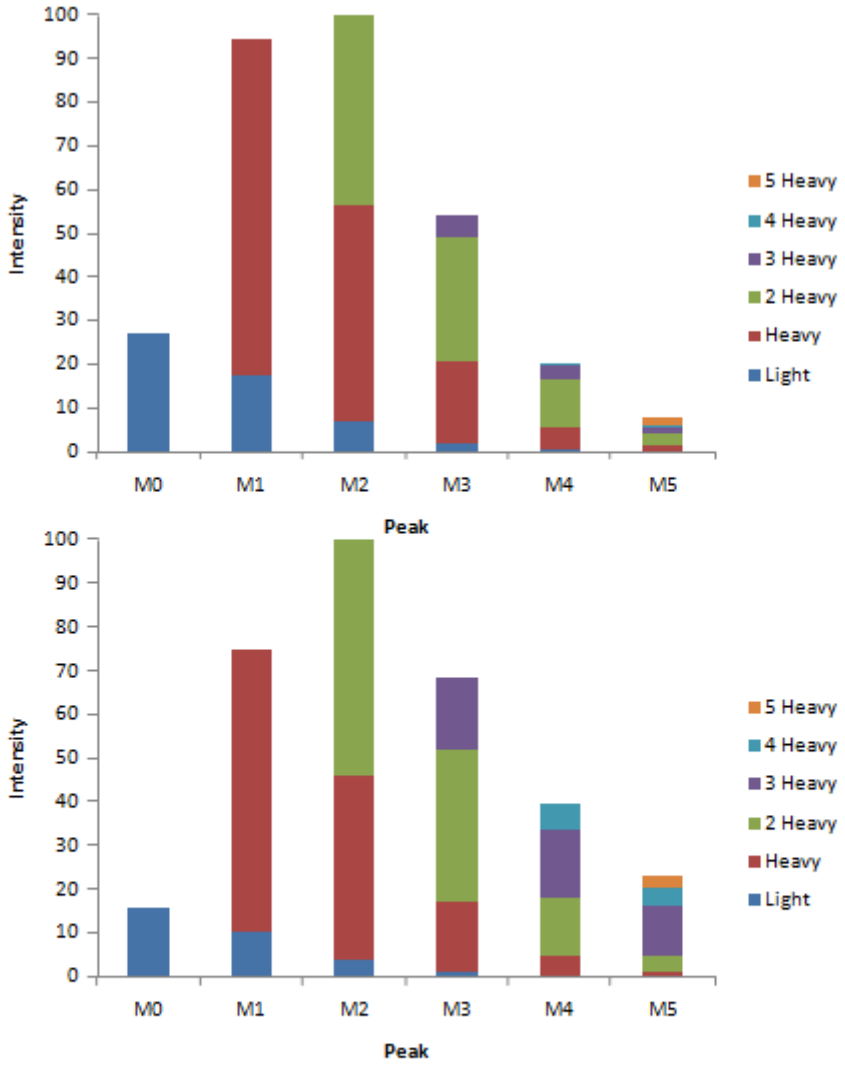


Figure 8. Contribution of labelled peptides to the overall appearance of the labelled spectrum in a MUP 2 peptide. MIDA was used to calculate a theoretical isotope distribution for unlabelled peptide. This model was then used to divide labelled spectra into light and heavy. (A) Day 1, heavy peptide makes up 82% of the spectrum. (B) Day 3, heavy peptide makes up 90% of the spectrum.

In figure 8 the same spectra are shown, but after MIDA has been used to separate each peak into light and heavy. On the graph, blue represents the proportion of the spectrum that is due to light peptide, it can be seen that there is much less light peptide on day 3 compared to day one, as would be expected. MIDA analysis was repeated for all peptides, on all days taken from mice fed both a 2% and a 4% heavy water. We also wanted to determine whether there was a difference in the extent to which peptides labelled when the animals were supplied with 2% or 4% heavy water. To do this, the intensity of the M1 peak was divided by the intensity of the M0 peak, and this was repeated for multiple peptides with different molecular weights (Fig. 9). M1/M0 values were also calculated for unlabelled peptide distributions generated by MIDA. It can be seen that the unlabelled peptide shows the lowest M1/M0 values, and the 4% heavy water peptides show the highest M1/M0 values. Therefore a higher percentage of heavy water produces a greater extent of labelling, as expected. After establishing the degree of labelling in the urine samples, we then moved on to look at the plasma samples.

3.1.2 Heavy Water Labelling in Plasma Proteins

Samples were first run on a 15% SDS gel (Fig. 10A), however this resulted in unsatisfactory separation of the higher molecular weight proteins. To resolve this, the samples were re-run on a lower percentage SDS gel (Fig. 10B), giving better separation. Due to its position on the gel, the most intense band was assumed to be albumin. To confirm this, this band was excised from each lane, digested, subject to MALDI-TOF MS, and a database search was performed.

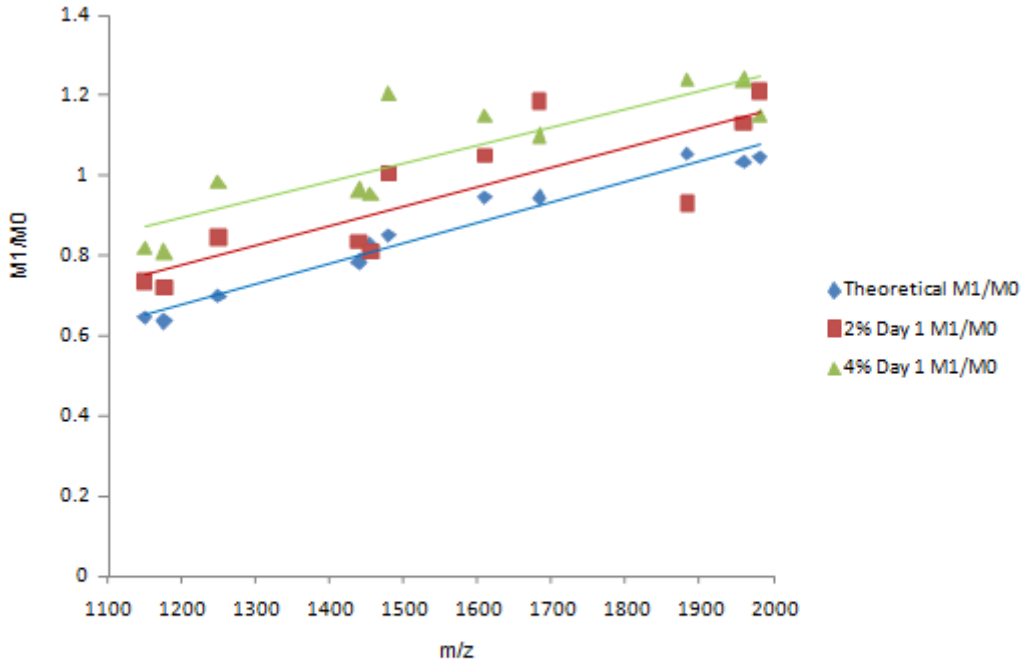


Figure 9. M1/M0 values for multiple MUP peptides at three different experimental conditions. (Blue) Expected M1/M0 values for unlabelled peptides. (Red) Actual M1/M0 values observed in a mouse provided with 2% heavy water (day 3). (Green) Actual M1/M0 values observed in a mouse provided with a 4% heavy water diet (day 3). Both the 2% and the 4% samples were taken from day 2 of the labelling experiment, at which point equilibrium had been reached.

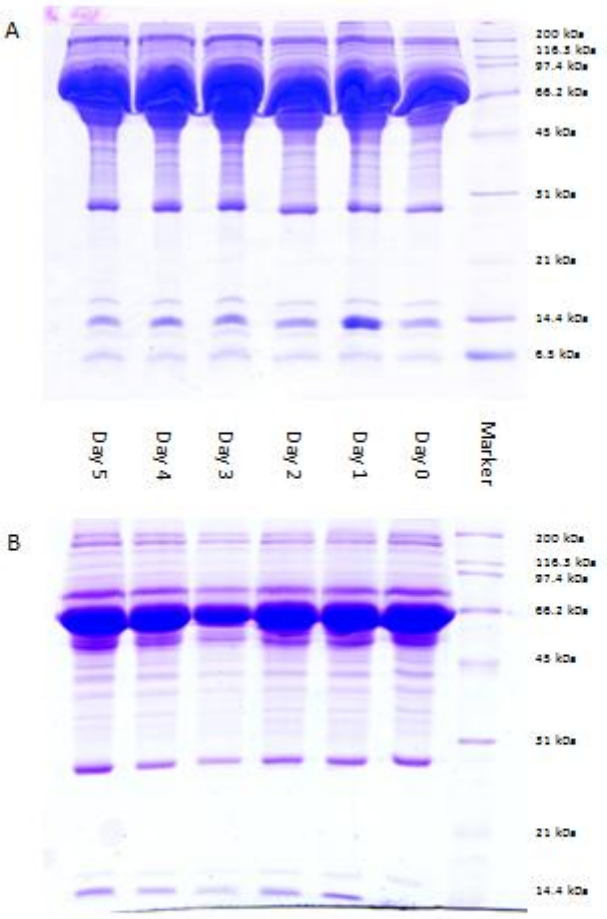


Figure 10. 1D SDS gel of plasma samples collected at 24 hour intervals over 6 days. The mice that these samples were collected from were fed a diet of 2% heavy water between day 1 and day 5. The samples were first run on a 15% SDS gel (A), and then a 12% SDS gel (B) to try and improve protein separation. In each instance 5µL of urine sample was loaded into each well.

The database search indicated, as expected, that the bands at this position contained albumin (Fig. 11). Although it was very easy to identify albumin, several problems were encountered with the plasma samples. Firstly, it was very difficult to identify any other proteins apart from albumin. Plasma samples are very complex, containing many proteins in extremely varied concentrations. The most abundant protein in plasma is albumin, accounting for approximately 50% of total protein. The remaining plasma is dominated by a further 8 considerately abundant proteins; IgG, haptoglobin, transferrin, transthyretin, α_1-antitrypsin, α_1-acid glycoprotein, hemopexin and α_2 – macroglobulin [43]. For this reason the lower abundance proteins are difficult to identify as the spectrum is often dominated by the peptides of the more abundant proteins [65]. After 1D gel separation and in-gel digest, six of the proteins present in plasma (including albumin) were identified. All of the six proteins identified were considered abundant. To try and identify the less abundant proteins, it was decided to run a 2D gel.

2D gels separate proteins first according to their charge (Iso-electric focusing) and then according to their molecular weight [44]. This, in theory should give a better separation of the proteins present in plasma and therefore allow for more identifications, in particular, the identification of less abundant proteins. Two plasma samples were chosen at random to run on two 2D gels. The two gels were then stained with Coomassie (Fig. 12, a and b). Clearly Coomassie is not sensitive enough to detect the less abundant proteins present in plasma, therefore the two gels were silver stained to see if this would result in the visualization of more proteins (Fig 12, c and d). Silver staining is much more sensitive than Coomassie staining, and is able to detect as low as 2-4 ng of protein [45].

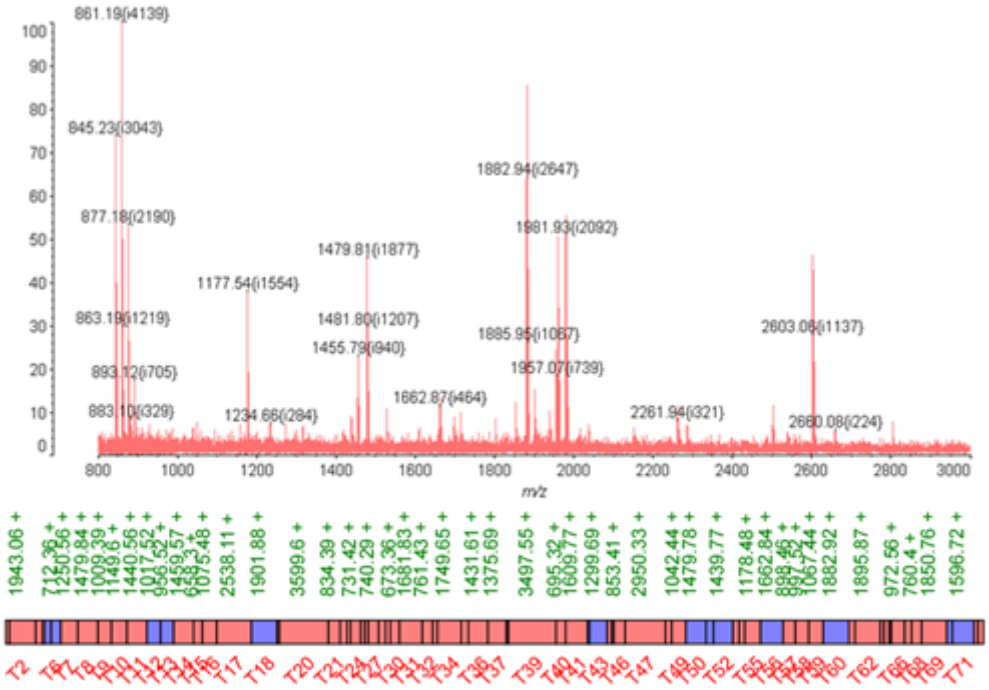


Figure 11. Spectrum generated by MALDI TOF MS using the Axima from an in-gel digest of a mouse plasma sample from an animal fed a 2% heavy water diet. A database search using Swiss-prot identified this sample as serum albumin, with a sequence coverage of 22% and a Mowse score of 62.

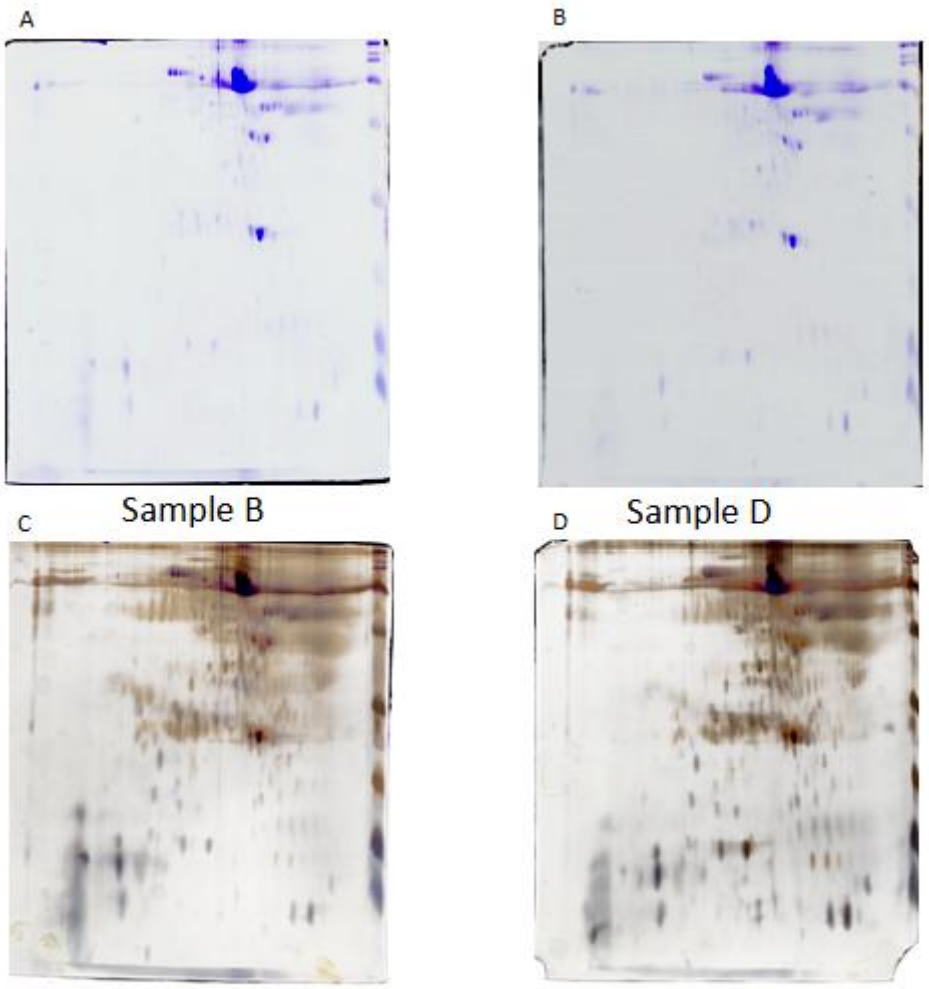


Figure 12. 2D gels of plasma samples collected from two mice, visualised using two different staining methods. The proteins were isoelectrically focused with pH 3-10 IPG strips in the first dimension and then resolved by 15% SDS PAGE in the second. Firstly the gels were visualised using Coomassie (A and B), they were then stained again using silver stain (C and D) to try and visualise more proteins.

Two spots were taken from one of the silver stained gels, digested and subject to MS analysis on the Axima. One of the spots was visible after the first staining (coomassie) and one only became visible after the second staining (silver). After running a database search, the first spot gave a positive ID for Apolipoprotein A, one of the more abundant plasma proteins. However the second spot did not result in any positive identifications. It was suspected this was due to an insufficient amount of protein being present in this spot. There are several things that, if the identification of plasma had been pertinent to this study, we could have tried in an order to identify the lower abundance proteins in plasma. Firstly, 2D gels could have been run again, but this time with a higher concentration of protein loaded than previously. This may have allowed for the identification of the lower abundance proteins when the protein spots were digested. Secondly Equalizer beads could have been used (Fig.13). Equalizer beads are comprised of beads attached to a combinatorial library of ligands. Each bead has a specific protein binding partner, and the binding capacity for each protein present in a sample is equal [46]. Thus, when a complex sample (such as plasma) is incubated with the bead library, each protein will find its binding partner. High abundance proteins, such as albumin, saturate the beads, and the excess protein that is not bound to the beads is washed away [46]. Conversely, low abundance proteins are concentrated on their specific affinity ligand, relative to the high abundance proteins. Thus Equalizer beads act to normalise the protein sample, allowing for the identification of low abundance proteins [46].

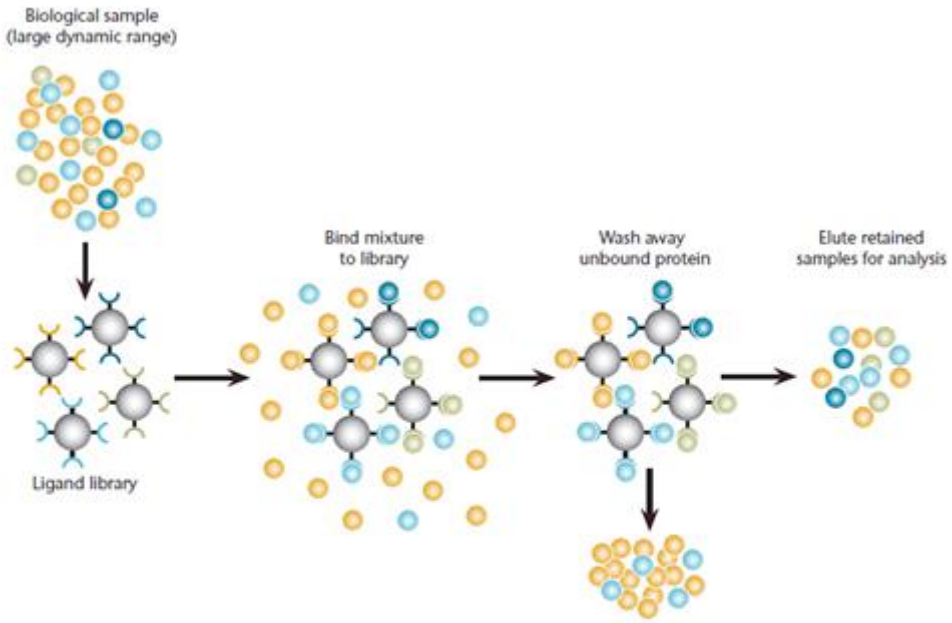


Figure 13. The use of equalizer beads. Each bead has a different ligand with affinity for a specific protein. (From Product Note, Ciphergen's Deep Proteome). Equaliser beads act to normalise a sample with a large dynamic range, thus assisting the identification of low abundance proteins.

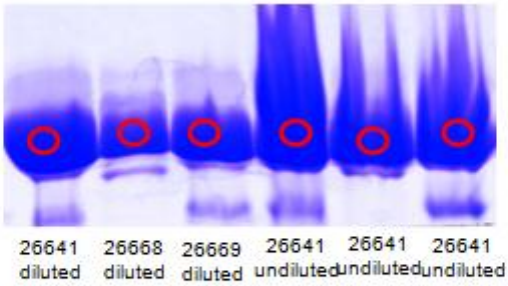


Figure 14. 1D SDS gel of urine samples from mice fed an unknown diet. Samples were collected on days 2 and 3 of feeding. Each sample was loaded onto the gel twice, once in an undiluted form and once in a diluted form (dilution factor 1 in 2.5). The MUP bands were then exised (as indicated by the red circles) and digested. The 5 digit numbers represent different animals.

A second problem encountered with the plasma samples was the quality of the spectra obtained. Although it was possible to identify albumin, when the spectra was looked at in more detail, the isotopic envelope showed poor resolution, and it was not possible to deduce whether or not the peptides were labelled, or the degree of labelling present. To try and overcome this problem, after digestion the samples were cleaned up using a C18 zip tip. Zip tips are used to de-salt and concentrate up samples, hopefully improving the spectrum obtained [47]. After zip-tipping, the samples were analysed by MALDI TOF MS, and each spectra looked at in more detail for evidence of labelling. Despite the clean up step, the spectra were still of poor quality, and it was not possible to determine labelling.

3.2. Heavy Valine

3.2.1. Deducing Whether or Not Labelling Has Occurred

As an initial experiment to ensure the method was working two mice (5798 and 5770) were fed a diet that contained heavy valine or [$^{2}H_{8}$] valine. In order to determine whether or not labelling had occurred, urine samples were taken on day 2 and day 3 of feeding (urine could not be obtained from mouse 5770 on day 3, so only 3 samples were obtained). All samples taken in the course of this study were stored at -80°C until subsequent analysis was conducted. These samples were then analysed for evidence of labelling. If labelling had occurred each valine containing peptide would show a second isotopic envelope 7 Da downstream of the monoisotopic peak of the peptide in question. If the peptide contained two valines there would also be another envelope 14 Da downstream of the monoisotopic peak. It is important to understand that although the peptide is labelled with [$^{2}H_{8}$] valine,

an isotopic envelope appears 7 Da downstream of the original isotopic envelope, not 8 Da downstream as may be initially expected. This is due to the loss of one deuteron by transamination in a large proportion of the heavy valine, resulting in [$^{2}H_{8}$] becoming [$^{2}H_{7}$]. A very small proportion of the heavy valine does not lose one deuteron through transamination [11], therefore the second peak in a heavy isotopic envelope is a combination of peptides with one carbon isotope [^{13}C] and one [$^{2}H_{7}$] valine in their structure, and peptides with one [$^{2}H_{8}$] in their structure. Both of these possibilities result in the 8 Da mass shift observed in the second peak of the heavy isotopic envelope. Similarly, consecutive peaks are also the result of a combination of [$^{2}H_{8}$] valine and [$^{2}H_{7}$] valine labelled peptides with increasing numbers of carbon isotopes in their structure.

Firstly the samples were defrosted and subject to 1D SDS gel electrophoresis (Fig. 14). Each sample was diluted 1 in 2.5 (the dilution used in previous experiments with mouse urine) before 5μL of each sample was loaded onto the gel.

Because the exact concentration of each sample was unknown an undiluted form of each sample was also loaded onto the gel. A protein assay was not performed as this would result in loss of material. The gel was allowed to run for an hour at 200mv, after which the gel was stained in Coomassie for 4 hours (until the bands became visible). An initial visual assessment of the 1D gel suggested that too much protein had been loaded onto the gel, as the bands were very thick and somewhat smeared. It appears that the protein concentration in these urine samples is much higher than the urine samples used previously, as a 1 in 2.5 dilution was insufficient. Therefore, in subsequent experiments with urine samples from this group of mice the urine was diluted more heavily (1 in 10). Although the

amount of protein present on the gel was too high, the band assumed to contain MUPs was still distinguishable. Since the purpose of this experiment was only to determine if labelling has occurred, and therefore no in-depth analysis was needed, it was decided to proceed with digestion after destaining the gel overnight in 40% Methanol 10% Acetic Acid.

After destaining, plugs were excised from the gel, as indicated by the red circles (Fig. 14). These bands were selected for digestion as both their position and intensity suggested they contained MUP's. The bands were digested following the usual protocol; however two times as much trypsin was added to account for the possible saturation of the gel. Concentration of trypsin was doubled but the same volume was added, therefore keeping the overall volume of the digest the same. After leaving to digest overnight the samples were plated onto an Axima target and subject to MALDI TOF MS (Fig. 15 A). First inspection suggested that MUPs were present in this sample, as possible MUP peptides can be seen immediately (964.53 m/z, 1258.69 m/z, 1472.72 m/z and 1486.77 m/z). A database search was then conducted and as expected it gave hits for several MUPs (MUP8, MUP6, MUP1 and MUP2). Two of the peptides in the samples contained valine (MUP8 T7 and MUP8 T11) and it was decided to look at these peptides more closely. MUP8 T7 is a monovaline peptide and it was decided to look at this spectrum first (Fig 15 B). As expected the isotopic envelope for the light version of the peptide was present at 2010.04 m/z, however there was also another isotopic envelope present at 2017.11 m/z, where we would expect to see the heavy version of this peptide. We would not expect to see anything at 2014 m/z as this peptide only contains one valine. Bearing in mind that this sample was only taken on day 2 of the labelling experiment, the spectrum suggests that most of the T11 peptides were still light. It is also possible that the envelope at 2017.11 m/z is not a heavy version of T11 but

rather a different peptide all together. One way to try and rule this out is to look at the same peptide on day 3 of labelling. On day 3 we would expect there to be more heavy peptide than on day 2, and therefore the heavy envelope would be more intense. Figure 15 C shows the same peptide, from a sample taken from the same mouse on day 3 of labelling.

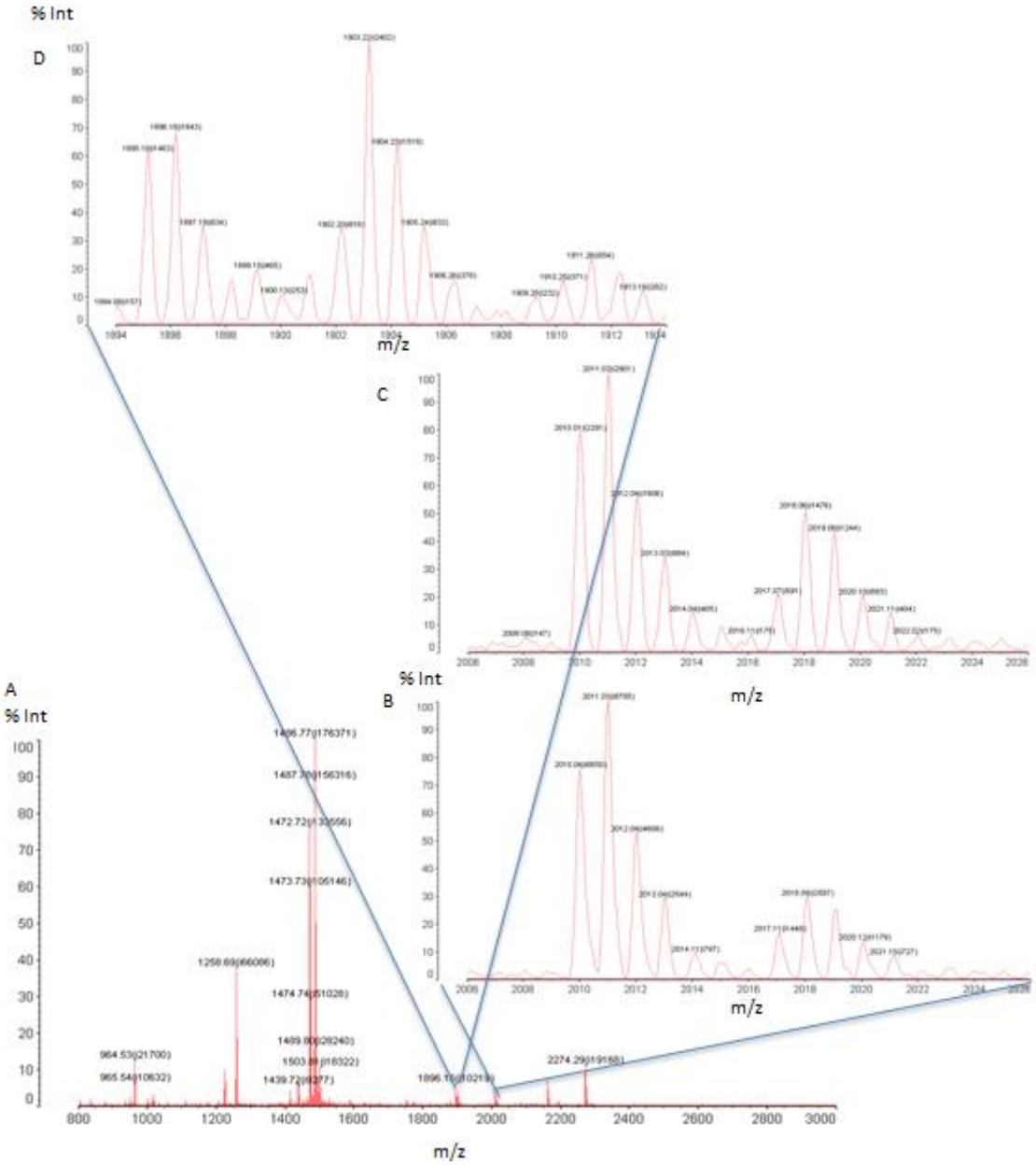


Figure 15. Heavy valine labelled mass spectrum, generated by MALDI-TOF MS using the Axima. A shows the mass spectrum in its entireity, from an in-gel digest taken on day 2 of labelling. B, C and D are individual peptides. B shows a labelled monovaline peptide found in spectrum A. C shows the same peptide from a urine sample taken on day 3 of labelling. D shows a labelled divaline peptide also found in spectrum A.

It can clearly be seen that the profile representing the heavy version of the peptide has increased in intensity relative to the profile of the light version of the peptide. This further suggests that the envelope at 2018.06 is heavy T11 rather than a separate peptide. The second valine containing peptide, MUP 8, T7, also suggested that labelling had occurred. Since this peptide contained eight valines, we would expect to see a profile at 7Da heavier, and a second at 14Da heavier, as is observed in figure 15 D. The evidence of labelling in peptides that had been identified as valine containing, lead to the conclusion that labelling had indeed taken place.

3.2.2. Urine Samples from Mice Fed a Heavy Valine Diet

Urine samples were collected at specific time intervals (day 0, 2, 7, 14, 25 and 35) from mice supplied with a 50% heavy valine diet. A 100% heavy diet was not used as this would be unpalatable, possibly resulting in the mice not eating enough to introduce significant labelling [19]. After 14 days of labelling, all the samples collected thus far were run on a 1D SDS gel [Fig. 16]. The MUP bands were excised, digested and analysed by MALDI TOF MS. A database search identified the proteins present in these bands to be MUPs, the identifications were then used to monitor the labelling of valine containing peptides over the 14 day period. Figure 17 shows the spectra of the same divaline peptide, identified in four different samples on day 0 of the labelling experiment. From these spectra it can be seen that there was no labelling on day 0 of the experiment, as would be expected since the mice had not yet had heavy valine introduced into their diet.

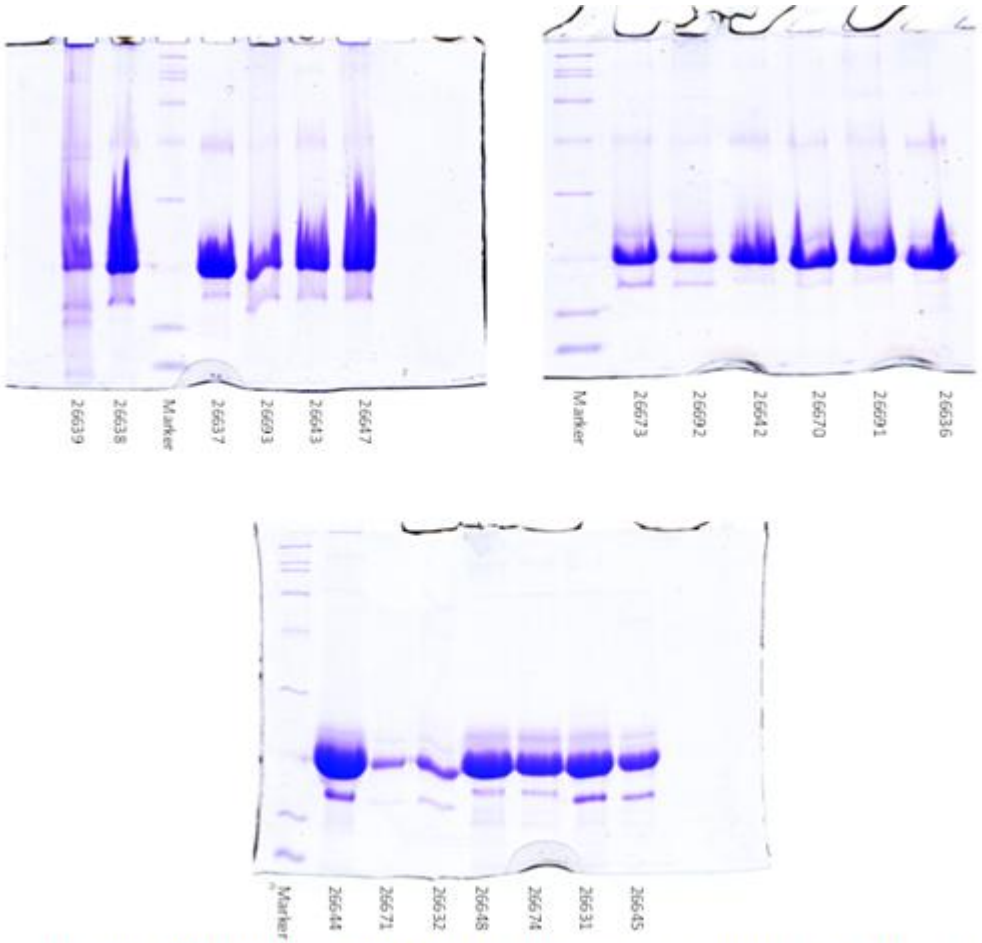
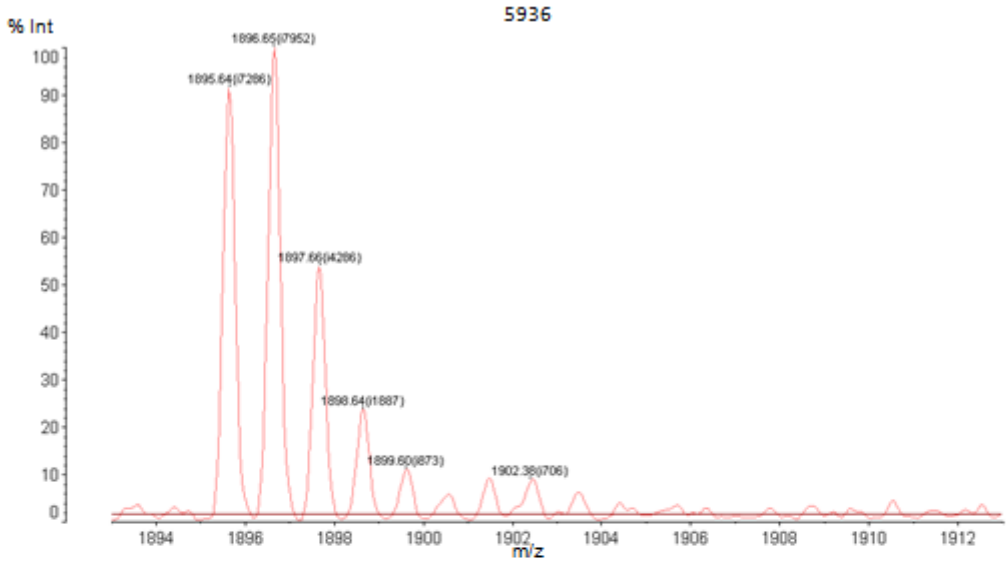
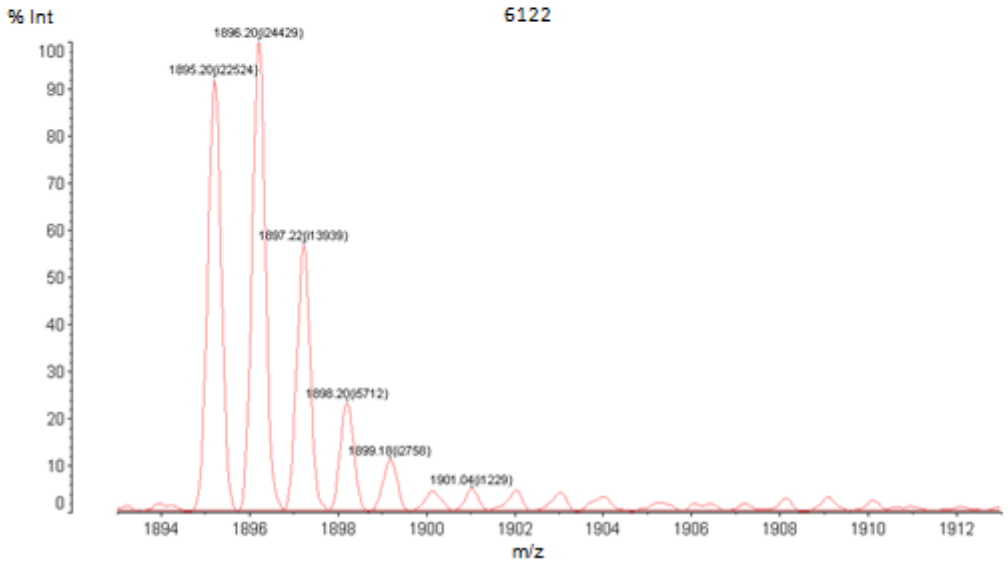


Figure 16. 1D SDS gel (15%) of mouse urine samples. The animals that were sampled were fed a heavy valine diet for 35 days. Urine samples were collected at specific time intervals. The numbers under the three gels correspond to individual mouse identification numbers.

However these spectra were still useful as they indicated that there did not appear to be a peptide 7 Da downstream of the divaline peptide. Therefore, if during the experiment we started to see a peptide envelope at 1902 m/z, we could be fairly confident that this was a labelled version of the original peptide, and not an entirely different peptide.

If we compare these spectra to spectra obtained from day two (Fig. 18), immediately we can see that a peptide appeared at 1902 m/z. Since this peptide contained two valine residues, the peak at 1902 m/z represents peptides containing one heavy valine and one light valine residue i.e. HL or LH. We can also see that a peak was starting to emerge at 1909. This represents peptides containing two heavy valine residues i.e. HH. However since these samples were taken on day two of labelling the HH peak is of low intensity. If we look at the spectra of peptides on day seven (Fig. 19), the HH peak appears to have increased in intensity slightly, although the increase is not as dramatic as was observed between days 0 and two. Again on day 14 we saw a slight increase in labelling, although it appeared to be levelling off, and possibly reaching a plateau (Fig. 20). Since the mice were fed a 50% heavy diet, when the proteins are fully labelled we would expect both the LL and HH peaks to be half the intensity of the HL peak. However, as mentioned previously, since the pre-cursor pool is being diluted by the degradation of existing proteins, this theoretical distribution will not be reached and the LL peak will always be over represented. The relative isotopic abundance (RIA or r) of the precursor pool can be calculated using the intensities of each peak in each isotopic envelope. Conversely, a hypothetical r value can be used to determine the resulting isotopic distribution.


5936
% Int
1896.65(7952)
1895.64(7286)
1897.66(4286)
1898.64(1887)
1899.60(873)
1902.38(706)
m/z


6122
% Int
1896.20(24429)
1895.20(22524)
1897.22(13939)
1898.20(5712)
1899.18(2758)
1901.04(1229)
m/z

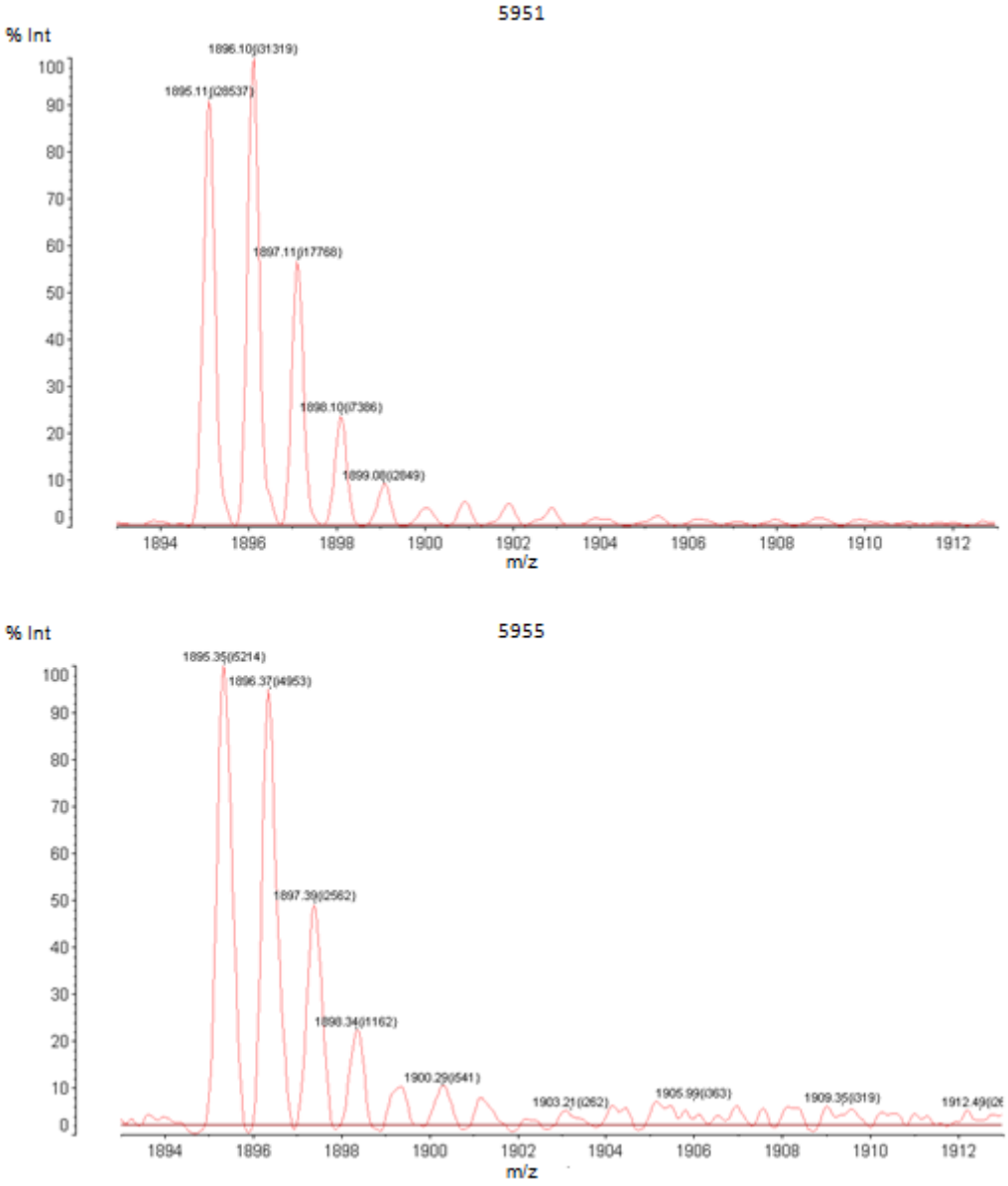
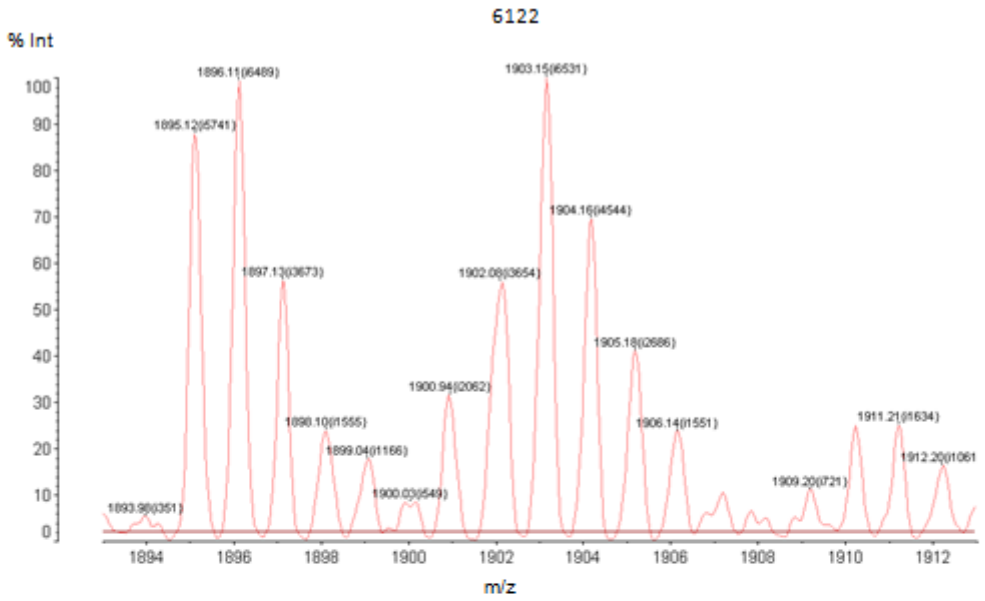
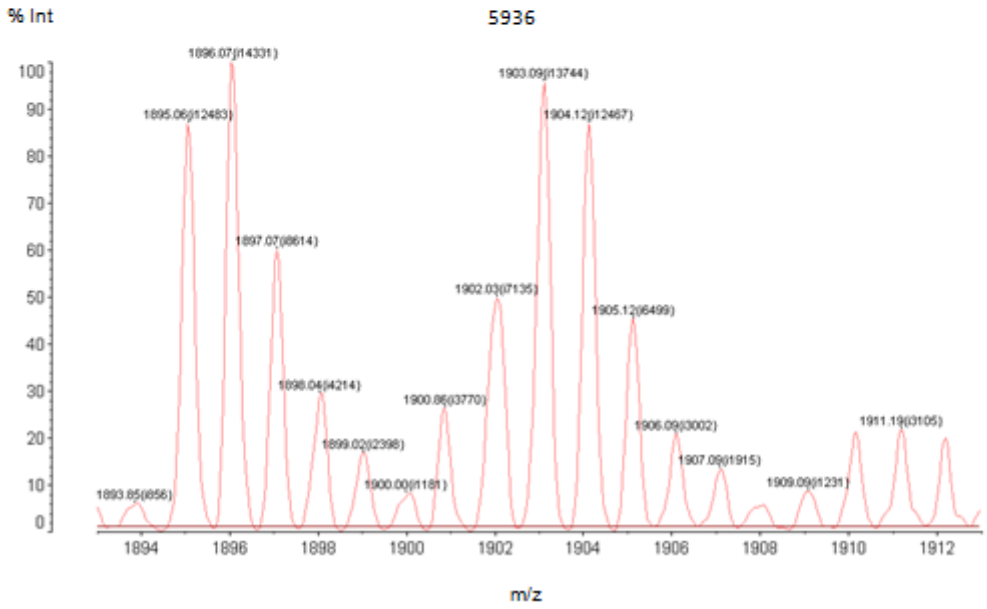


Figure 17. A divaline MUP peptide on day 0 of the labelling experiment. The four spectra are the same peptide obtained from different animals. At this point the animals are still being fed a normal (light) diet.


6122
% Int
1896.11(6489)
1903.15(6531)
1895.12(5741)
1904.16(4544)
1897.13(3673)
1902.08(3654)
1905.18(2686)
1900.94(2062)
1906.14(1551)
1898.10(1555)
1911.21(1634)
1899.04(1166)
1912.20(1061)
1909.20(721)
1900.03(549)
1893.98(351)
100
90
80
70
60
50
40
30
20
10
0
1894
1896
1898
1900
1902
1904
1906
1908
1910
1912
m/z


5936
% Int
1896.07(14331)
1903.09(13744)
1895.06(12483)
1904.12(12467)
1897.07(8614)
1902.03(7135)
1905.12(6499)
1898.04(4214)
1900.86(3770)
1906.09(3002)
1911.19(3105)
1899.02(2398)
1907.09(1915)
1909.09(1231)
1900.00(1181)
1893.85(656)
100
90
80
70
60
50
40
30
20
10
0
1894
1896
1898
1900
1902
1904
1906
1908
1910
1912
m/z

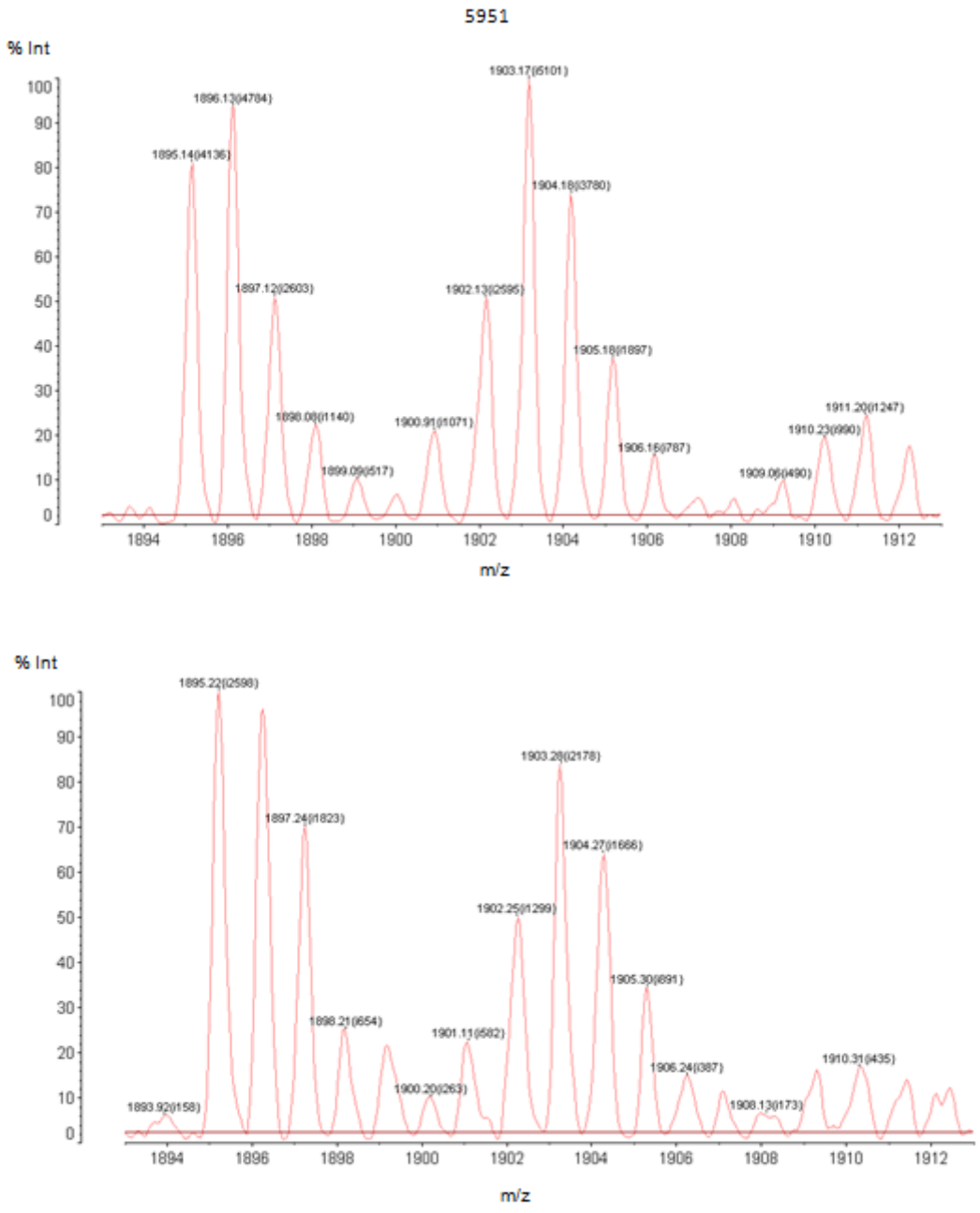
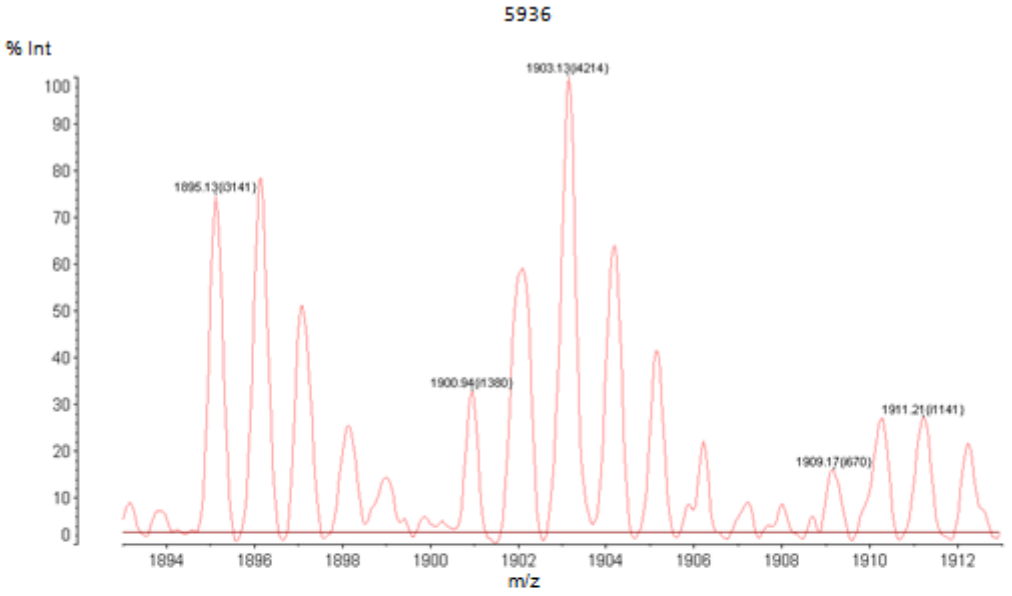
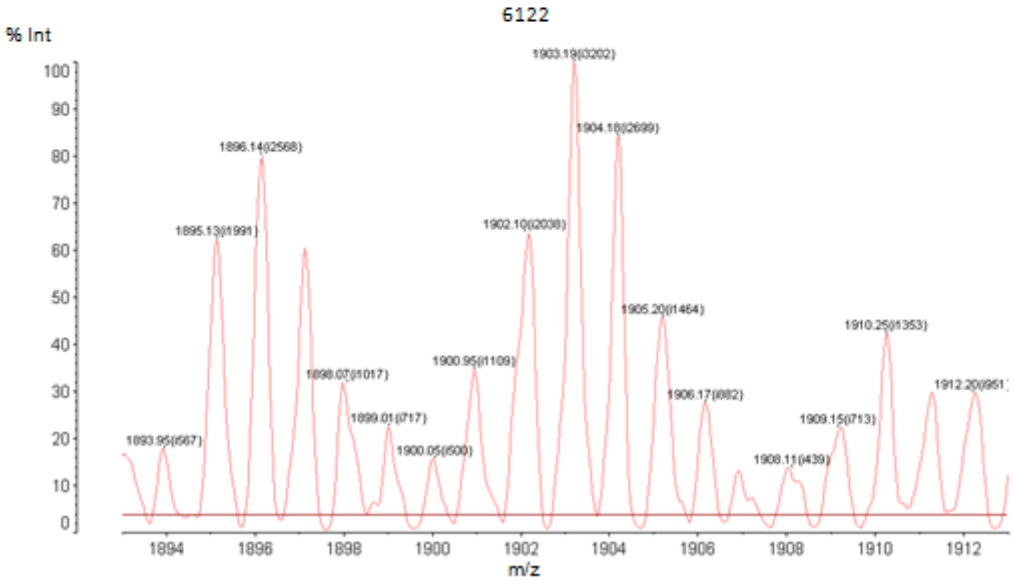


Figure 18. A divaline MUP peptide on day 2 of the labelling experiment. All four spectra now show evidence of labelling.


5936
% Int
1903.13(4214)
1895.13(3141)
1900.94(1380)
1911.21(1141)
1909.17(670)
m/z


6122
% Int
1903.19(3202)
1904.18(2699)
1896.14(2568)
1895.13(1991)
1902.10(2038)
1905.20(1464)
1910.25(1353)
1898.07(1017)
1900.95(1109)
1912.20(951)
1906.17(882)
1899.01(717)
1909.15(713)
1893.95(567)
1900.05(500)
1908.11(439)
m/z

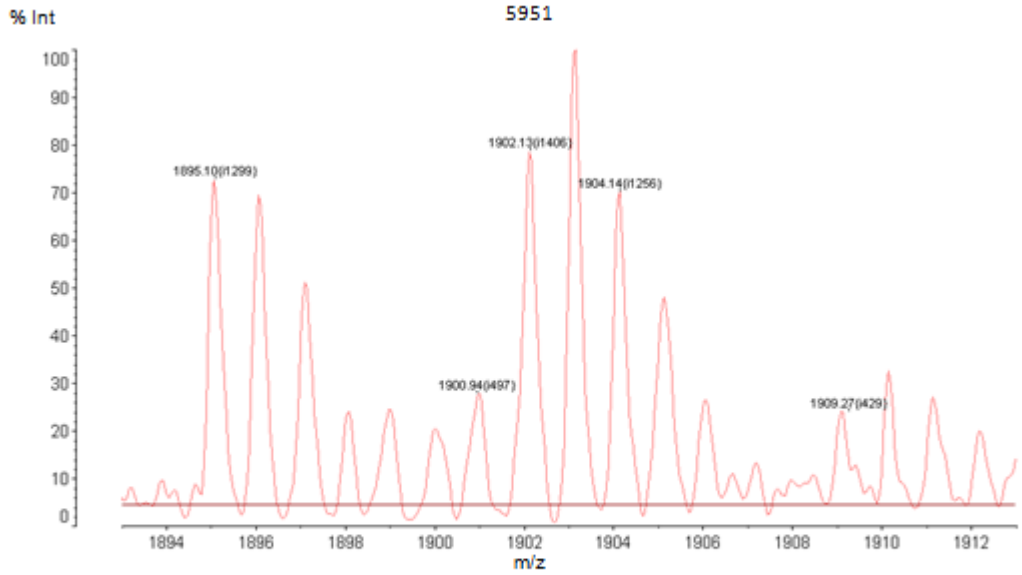
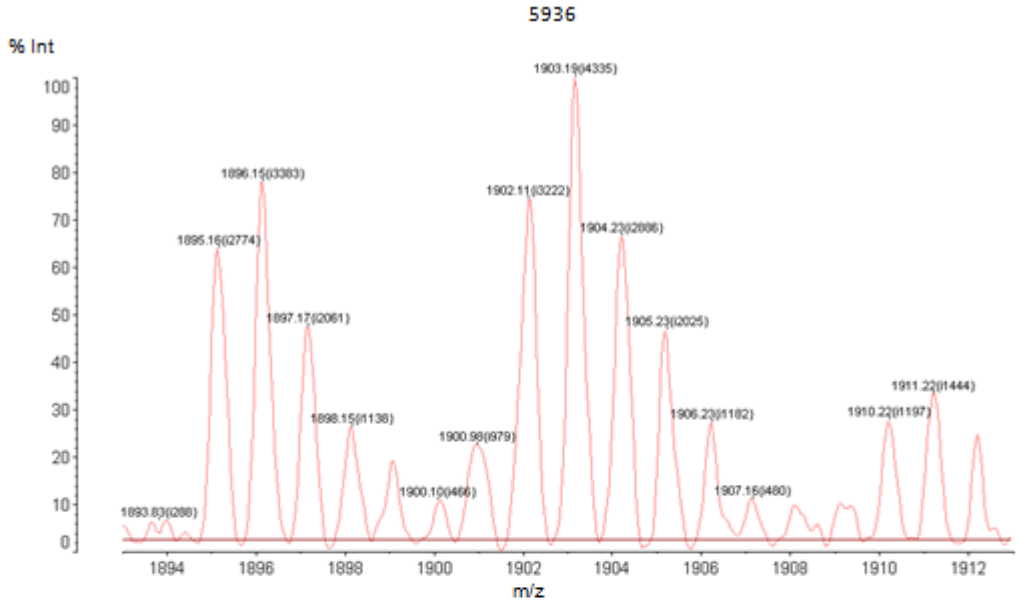
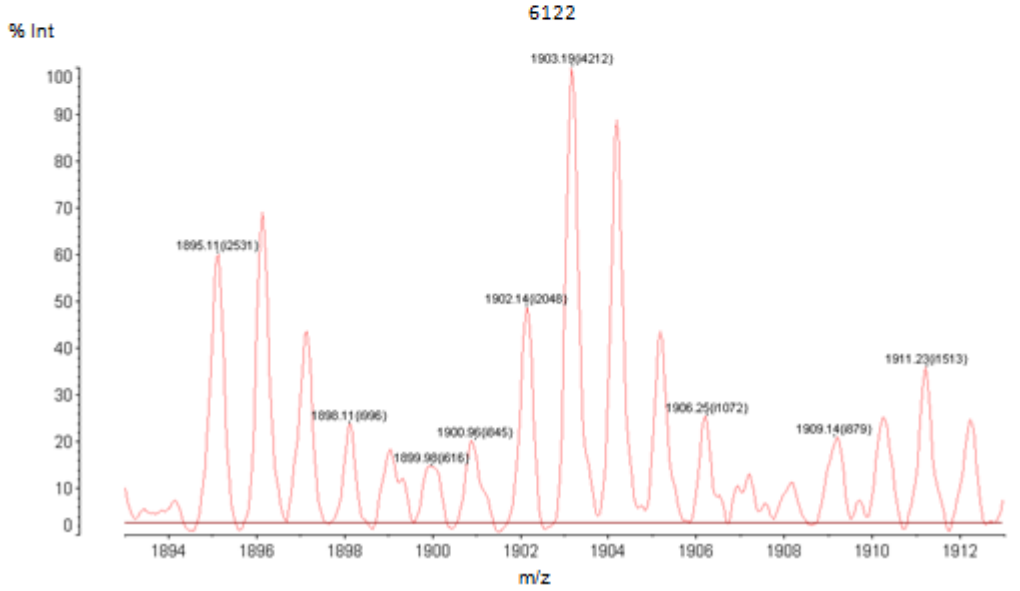


Figure 19. A divaline MUP peptide on day 7 of the labelling experiment. The three spectra show evidence of progressive labelling. Only 3 spectra are shown because a urine sample could not be obtained from the fourth animal.


5936
% Int
1903.19(4335)
1896.15(3383)
1902.11(3222)
1895.16(2774)
1904.23(2686)
1897.17(2061)
1905.23(2025)
1911.22(1444)
1898.15(1138)
1906.23(1182)
1910.22(1197)
1900.98(979)
1900.10(466)
1907.16(480)
1893.83(288)
100
90
80
70
60
50
40
30
20
10
0
1894
1896
1898
1900
1902
1904
1906
1908
1910
1912
m/z


6122
% Int
1903.19(4212)
1895.11(2531)
1902.14(2048)
1911.23(1513)
1898.11(996)
1906.25(1072)
1900.96(645)
1909.14(879)
1899.98(616)
100
90
80
70
60
50
40
30
20
10
0
1894
1896
1898
1900
1902
1904
1906
1908
1910
1912
m/z

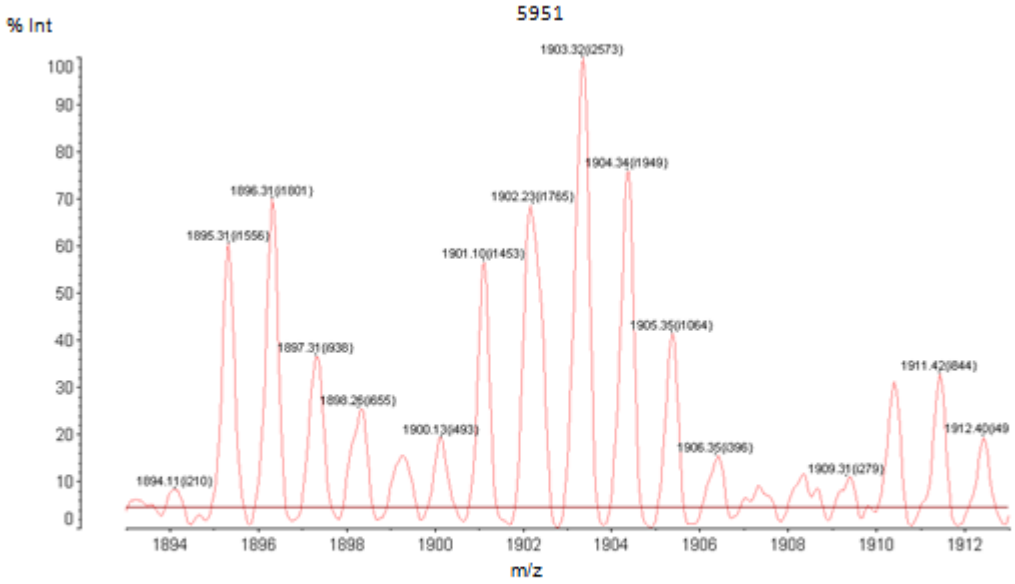


Figure 20. A divaline MUP peptide on day 14 of the labelling experiment. The three spectra show evidence of labelling, a plateau appears to have been reached. Only 3 spectra are shown because a urine sample could not be obtained from the fourth animal.

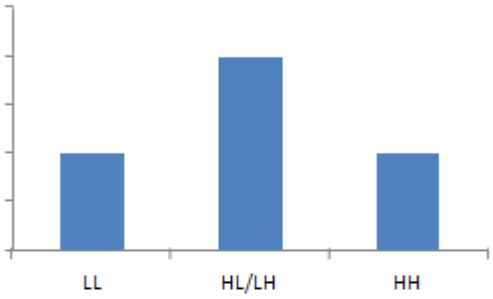
In figure 21 the isotopic distribution of a divaline peptide has been calculated for two different r values (0.5 and 0.25). These calculations were made using observations made and calculations developed by Beynon and colleagues. If we consider a peptide containing two valine residues, when exposed to heavy valine, three variants are possible; light light (LL), light heavy (LH) and heavy heavy (HH). If the relative isotopic abundance (r) of the heavy species is 0.5 i.e. it is equally likely that a heavy or a light valine residue will be incorporated, an abundance ratio of 1:2:1 (LL:LH:HH) will result [19] (Fig. 21A). These observations were then used to calculate the theoretical isotopic distribution if r was equal to 0.25. If we compare these distributions, to that of the observed isotopic distribution, we see that the r is likely to be somewhere between 0.25 and 0.5. In order to calculate the r from the observed distribution seen in figure 21, the intensity for each peak in each species must first be totalled (Fig. 21). Since the intensity of the HH species is equal to r^2, it follows that r is equal to the square root of the intensity (Fig. 21). Using this calculation, RIA values for the two mice that gave samples at all four time points (mouse 6122 and mouse 5936) were calculated (Fig. 22). Both mice display rapid labelling, with a sharp increase in the first two days of labelling. After day two, the RIA values begin to level off, stabilising at around 0.4. The rapid labelling observed in these samples was expected, due to the nature of MUPs. Once it had been established that labelling was occurring, focus then moved to heavy valine incorporation into the various tissue proteins.

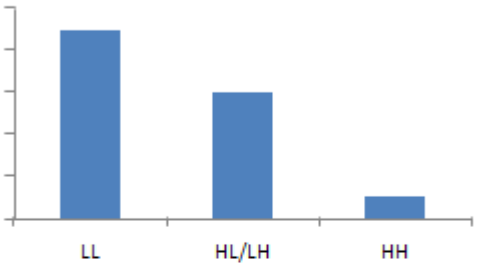
A

Isotopic Species	Formula	Sum	Relative Value
LL	(1-r)(1-r)	0.5 x 0.5	0.25
LH, HL	2 (r(1-r))	0.5 x 0.5 x 2	0.5
HH	r^2	0.5 x 0.5	0.25

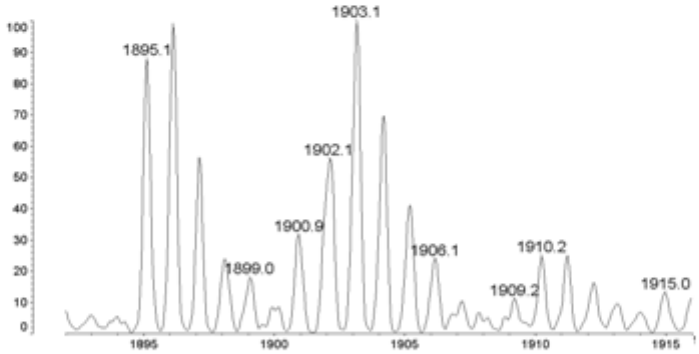
B

Isotopic Species	Formula	Sum	Relative Value
LL	(1-r)(1-r)	0.75 x 0.75	0.5625
LH, HL	2 (r(1-r))	0.25 x 0.75 x 2	0.375
HH	r^2	0.25 x 0.25	0.0625





C



D

Species	M0	M1	M2	M3	M4	Total	Relative Value
LL	3141	3301	2160	1072	600	10274	0.39
HL/LH	2488	4214	2698	1747	927	12074	0.45
HH	670	1134	1141	915	429	4289	0.16

Figure 21. Using peak intensities to calculate relative isotope abundance (RIA or r). A and B show the theoretical distribution we would expect to see for different RIA values; 0.5 and 0.25 respectively, and how these distributions are calculated. C shows an actual distribution for a labelled divaline peptide, generated by MALDI TOF MS using the Axima. D demonstrates how totalled peak intensities can be used to calculate r. Since the relative value of the HH species is equal to r^2, therefore r equals 0.401. Equation developed by Beynon and colleagues [19].

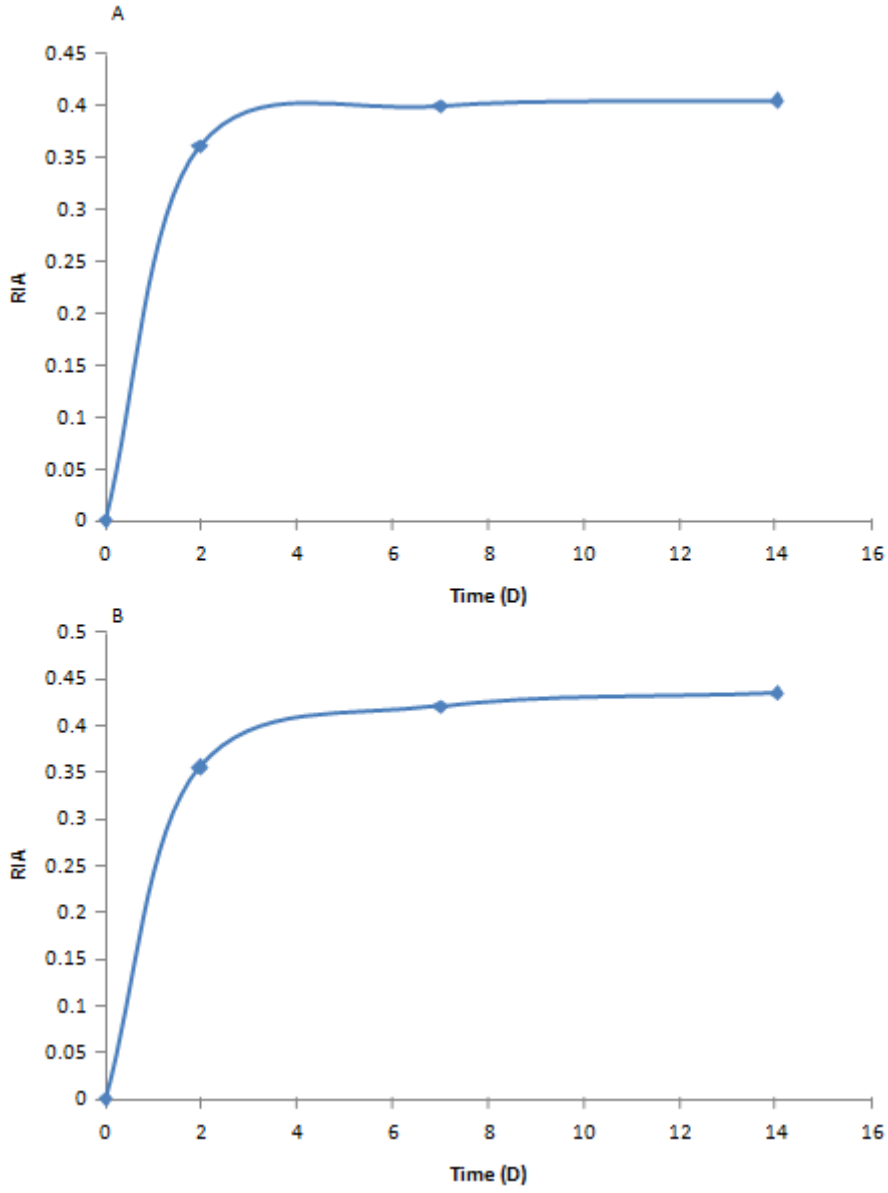


Figure 22. RIA values over a 14 day period. RIA values were observed in two different animals, they were obtained from MUP peptides found in digested urine samples. A shows RIA values obtained from animal 5936, B shows RIA values obtained from animal 6122.

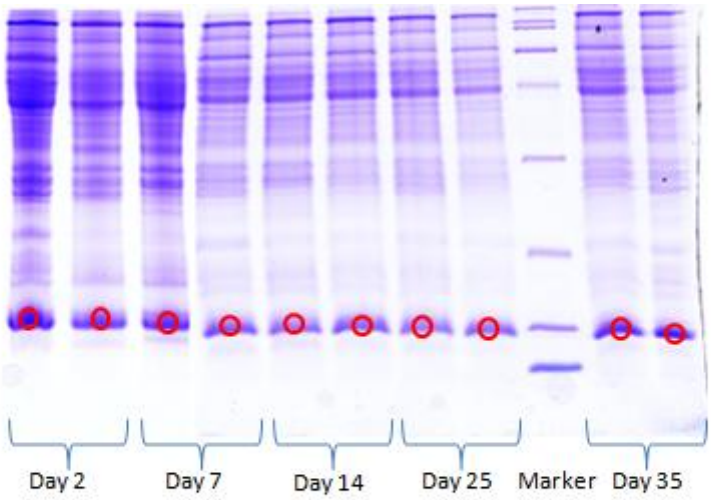


Figure 23. 15% SDS 1D gel of homogenised liver samples. 10µg of protein was loaded onto each lane, and the gel was allowed to run for 45 mins. A portion of the lower bands were extracted from the gel, as indicated, and digested with trypsin using a standard in-gel digest protocol

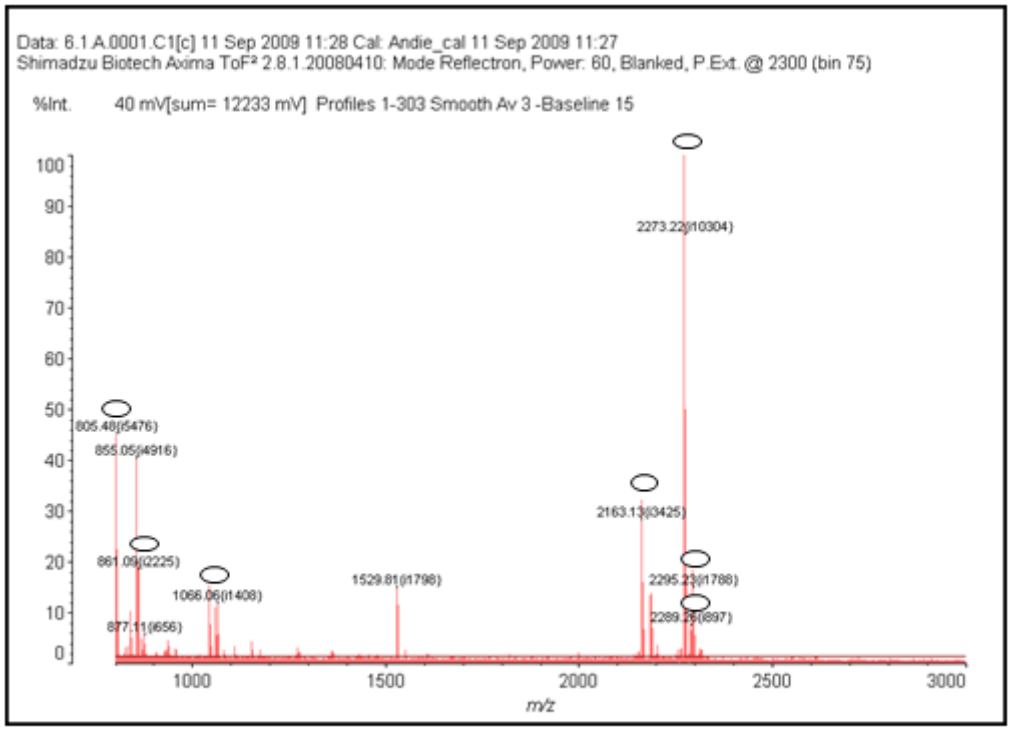


Figure 24. Spectrum from an in-gel digest of liver sample, generated by MALDI-TOF MS using the Axima. Circles represent trypsin autolysis peaks.

3.2.3. Turnover studies in tissue samples from mice fed heavy valine diet

The tissue samples (liver, skeletal muscle and kidney) were homogenised according to a standard protocol and 10µg of each sample was run on a 1D SDS gel (Fig. 23). All the tissue samples were subject to the same protocols, with similar results, however only the liver samples will be discussed at length here. Bands were first excised from the gel containing the liver samples, as indicated (Fig. 23). These bands were digested, analysed by MALDI TOF MS, and a database search was performed. Unfortunately, on this occasion no proteins were identified, and looking at the spectra it became clear why. Figure 24 shows a typical spectrum obtained from this digest. It is dominated by trypsin autolysis peaks, as indicated. Trypsin autolysis peaks occur when trypsin cleaves itself, this occurs because trypsin catalyses the hydrolysis of amide bonds whose carbonyls are part of lysine or arginine [48-50]. Since the structure of trypsin contains 14 lysines and two arginines, it is a substrate for its own action [51]. Since the intensity of trypsin autolysis peaks can increase with low concentrations of analyte, long digestion times and large quantities of trypsin [52], digestion was repeated, this time using less trypsin (enzyme to substrate ratio of 1:10 instead of 1:5), and allowing digestion to proceed for 14 hours instead of 16 hours (Fig. 25B). An extra protein sample was also digested using the old protocol, to act as a control (Fig. 25A). It can be seen that using less trypsin and decreasing the incubation time has greatly decreased the intensity of the trypsin autolysis peaks. Furthermore a database search on this spectrum resulted in the identification of beta haemoglobin (Fig 25C). Once the protein present in this band had been indentified, attention turned to the valine containing peptides present in beta haemoglobin.

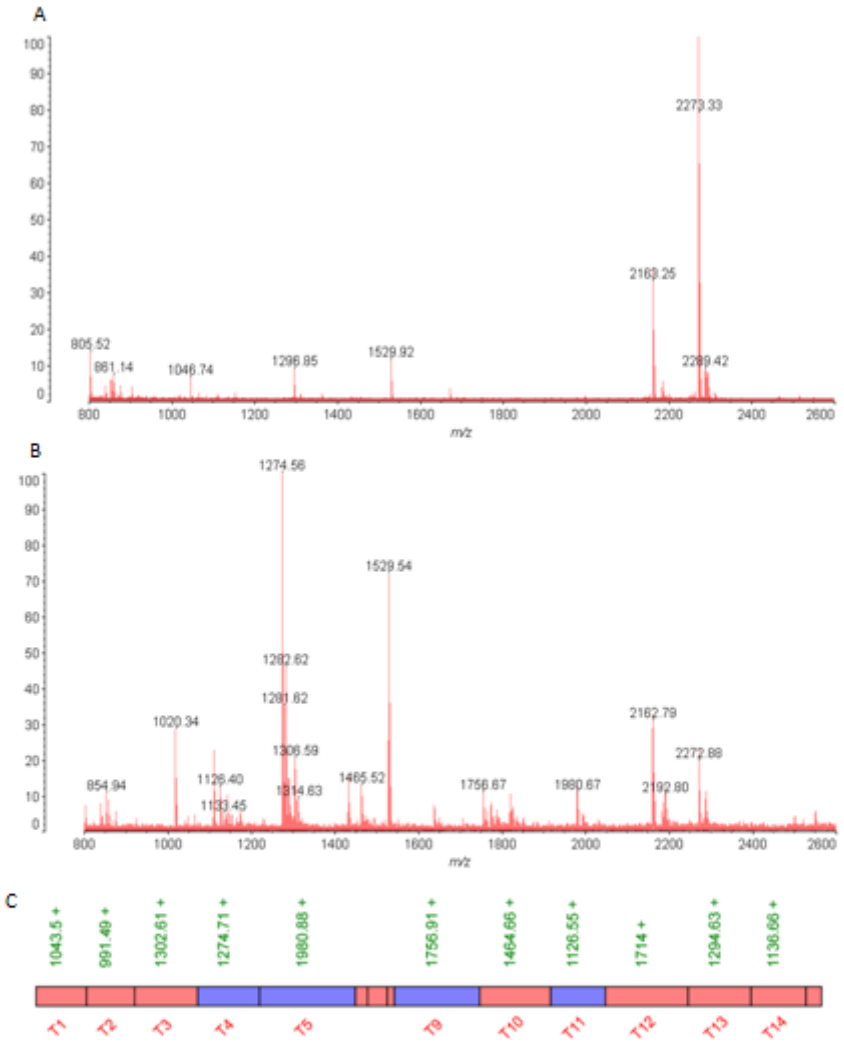


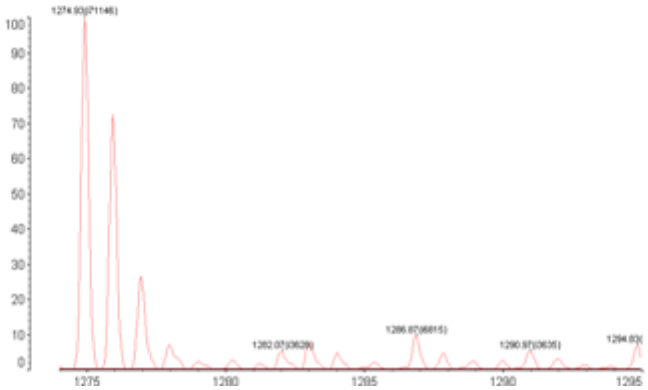
Figure 25. In-gel digests of liver samples, analysed by MALDI TOF MS on the Axima. (A) Digested with trypsin at an enzyme to substrate ratio of 1:5 for 16 hours. (B) Digested with trypsin at an enzyme to substrate ratio of 1:10 for 14 hours. (C) Peptide map for beta haemoglobin, generated using a database search on the peptides present in spectrum B (36% sequence coverage).

Of the four beta haemoglobin peptides identified, three contained valine (two monovaline peptides and one divaline peptide). Figure 26 shows the divaline peptide T4 over the course of the 35 day labelling experiment. It can be seen that labelling emerges much slower in haemoglobin, compared to the MUPs looked at earlier. In fact, it is not until approximately 14 days that evidence of labelling is apparent. Since the peptide in figure 26 is divaline, it can be used to calculate the RIA value of the precursor pool. Although not impossible, using a monovaline peptide to calculate RIA is much more complex as the light species contains an unknown mixture of peptides derived from pre-existing and newly synthesised protein [19]. RIA was calculated at each interval of the labelling experiment and plotted against time (Fig. 27). If we compare this to the RIA values observed in the MUP peptides (Fig. 22), there appears to be a delay in the labelling of haemoglobin. This was not the result that was expected, as RIA usually shows a sharp increase to a plateau phase. However, since haemoglobin is synthesised in the bone marrow, incorporated into erythrocytes and transported to organs after erythrocyte maturation [66], this may explain the delay in the emergence of labelling observed in haemoglobin found in the liver. Because haemoglobin is not a liver tissue protein, but rather it is present in the liver due to blood in the liver, an attempt was made to identify the proteins present in the other bands, in the hope that this would inform on heavy valine labelling of liver proteins. Despite repeat digestions of the bands present on the 1D gel of liver samples, no further identifications could be made. This could possibly be due to the complexity of the sample (mouse liver contains in excess of 2000 different proteins) and the dynamic range [52].

2 Days

1 Week

2 Weeks



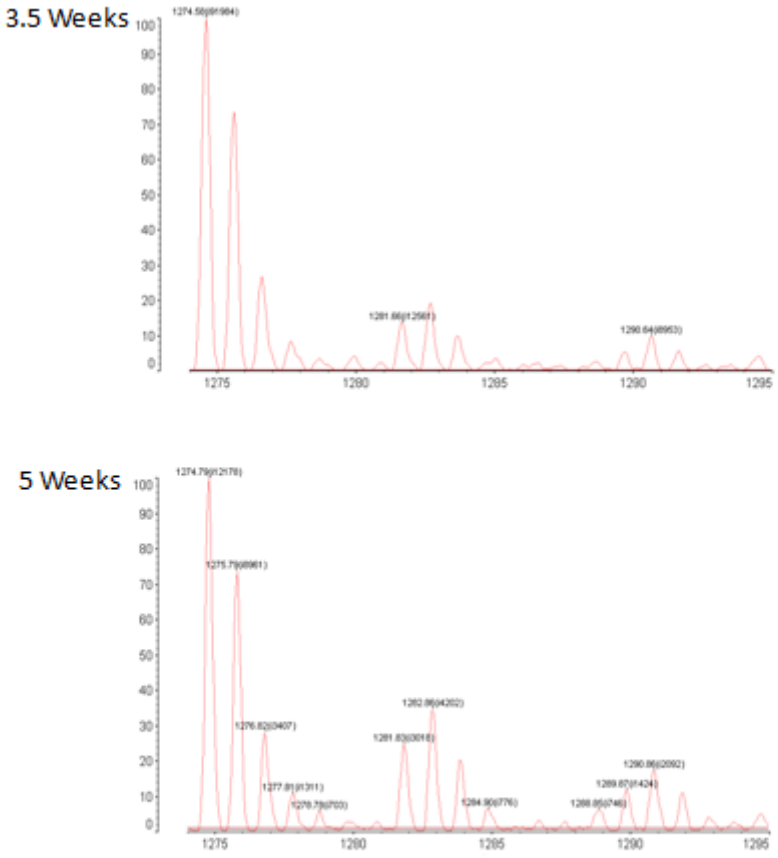


Figure 26. A beta haemoglobin (gene ref: Hbb-b1) peptide (T4 - LLVVYPWTQR) over the course of the labelling experiment. Spectra were generated by MALDI TOF MS of in-gel digests. Each spectrum is from a different animal, because in order to look at liver proteins the animal must be humanly killed. Database download date : 29 Sept 2009.

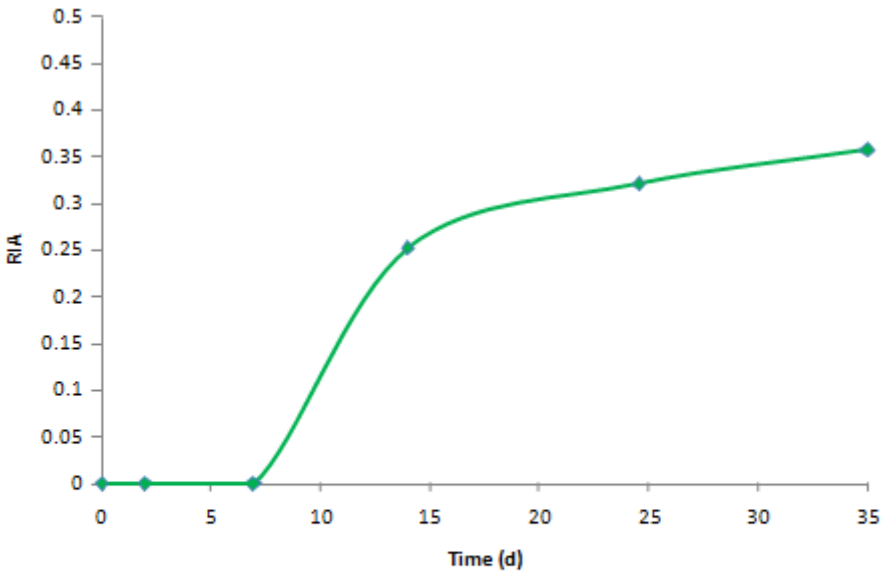


Figure 27. RIA of beta haemoglobin over 35 days. RIA was calculated from the divaline peptide T4 (LLVVYPWTQR),. obtained from liver samples. The intensities of the three different peptide species (LL, LH/HL and HH) were used to determine RIA at each interval of the labelling experiment

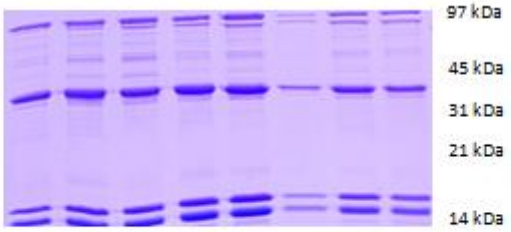


Figure 28. Reproducibility of SVS samples. Samples were collected from 8 different animals over the course of the labelling experiment. They were then loaded onto a 1D SDS gel and run at 200mV for 1 hour.

Thus each band will contain a considerable number of proteins, all at varying concentrations. Clearly separation by 1D gel is insufficient for the liver tissue samples, therefore, in order to reduce sample complexity, and increase protein identifications, more rigorous separation techniques must be employed.

There are several techniques that could be used for future analysis of complex tissue samples, such as liver. The first technique that may be used is GeLC MS. GeLC MS first involves running a 1D gel of the sample/samples of interest. Gel slices are then excised all the way down the track of the gel, and an in-gel digest is performed on all slices. Nano LC MS/MS is then performed on the fractions to obtain peptide sequences and thus protein identifications [53]. Since there are two separation steps in this technique (1D gel and LC), it is hoped that this will lead to a greater number of protein identifications in complex samples [54 and 55]. Another option is the use of LC-MALDI to increase protein identifications. As the name suggests, this technique combines separation by LC with MALDI TOF MS. The sample is initially separated by 1D SDS PAGE, bands of interest are then excised and digested, using a standard digestion protocol. After digestion, the peptides are separated by off-line LC, and the fractions are eluted directly onto a MALDI target, and analysed by MALDI TOF MS [56]. Again, like GeLC MS, since this technique employs two separation steps, sample complexity is greatly reduced. In addition, because the fractions are eluted onto a MALDI target, the time allowed for MS analysis is greatly increased. It is hoped that the use of one, or both, of these techniques, will lead to increased protein identifications in complex samples. This in turn will make the identification of valine containing peptides much easier and more reliable i.e. if the identity of a peptide is known, we can be more

confident regarding its valine content, as opposed to an unknown peptide that we assume to contain valine due to its isotopic profile.

3.2.4. Differential Labelling of Sperm and SVS Proteins

One application of this work is to use metabolic tracers as ownership tags in sperm and seminal vesicle proteins. As mentioned previously, litters of mice frequently have multiple paternities due to heterospermic insemination (HI) [35]. A particular area of interest for behavioural biologists the degree of investment made by each male during a HI, and if their position in mating order has any bearing on this investment. With a view to investigating this, first the labelling patterns of both sperm and seminal vesicle proteins were examined.

3.2.4.1. Protein Identification

The SVS samples were first run on a 12% 1D SDS gel to separate the protein mixture. It can be seen that all 8 samples showed a high level of reproducibility, with each lane displaying 4 distinct protein bands (Fig. 28). A small portion of each of these bands were excised and digested. The samples were then analysed by MALDI TOF MS and a database search was performed in the hope that the proteins present in each band could be identified. Figure 29 shows how the mass spectrum obtained from a digest (Fig. 29A) was used to identify SVS 1 and with the peptide map indicating exactly which peptides contributed to this identification (Fig. 29B). The results obtained from the database search gave a high level of confidence that the identification obtained was correct (Fig. 29C).

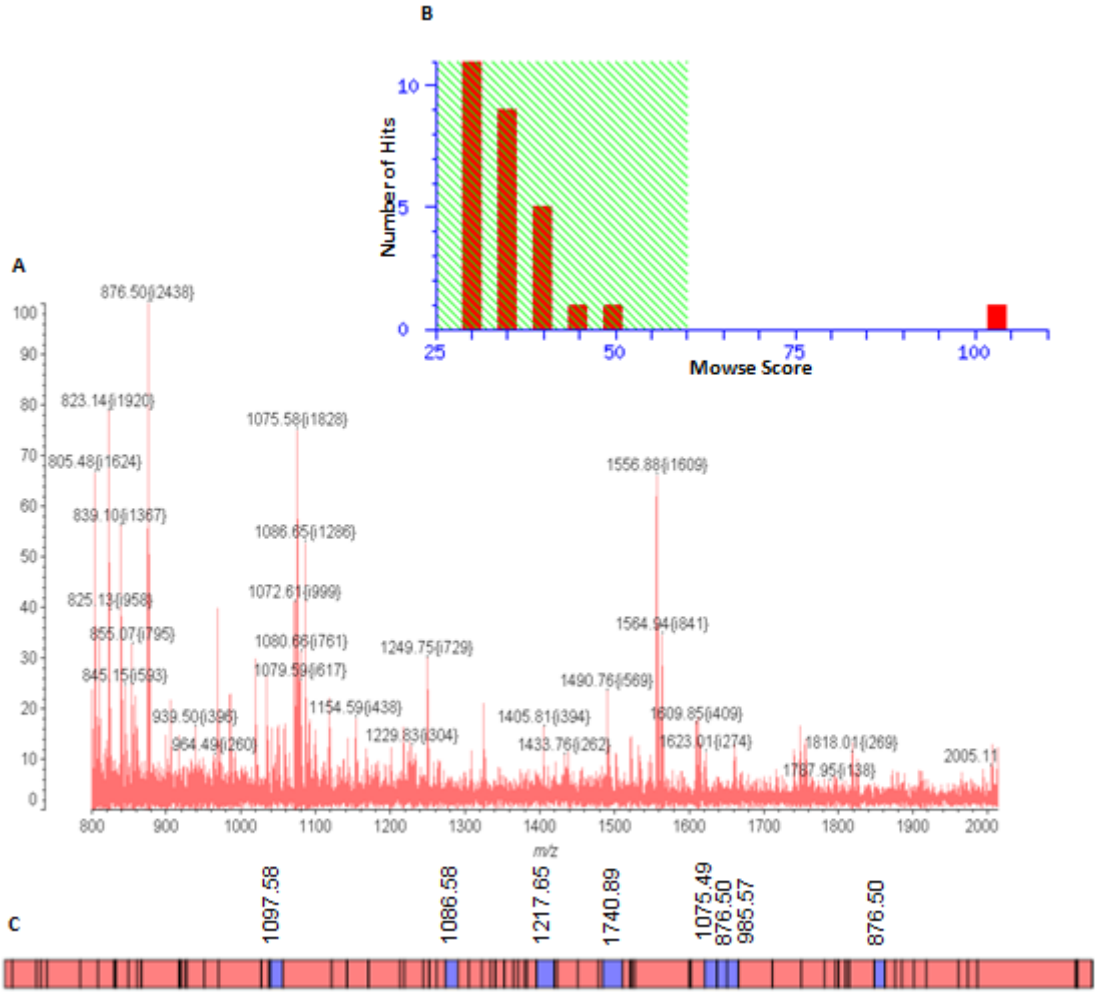


Figure 29. Identification of SVS 1 using a MALDI TOF MS derived spectrum (A). Peptide M+H values were used to determine the identification of an unknown protein. A search was performed using the MASCOT database, and the protein was identified as SVS 1 (B). C is a peptide map showing the peptides identified by the database search (identified peptides are in blue).

Because the Mowse probability score was high (103), as was the sequence coverage (13%), and 9 peptides were matched (one miscleave) it is unlikely that the identification was incorrect. The identification of the proteins present in the SVS samples was fairly straightforward, as the samples were dominated by four, very abundant proteins; SVS 1, SVS 2, SVS 4 and SVS 5 (Fig. 30). Identifying the proteins present in the sperm samples was much more difficult, as there were many proteins present in each sample. One study identified 858 proteins present in rodent sperm [57].

Like the SVS samples, when the sperm samples were subject to 1D PAGE a level of reproducibility was observed (Fig.31A). However there were a great deal more bands present in the sperm samples, than would be expected given the known complexity of these samples. Figure 31B shows the proteins identified in the sperm samples, by 1D SDS PAGE, in gel digestion and MALDI TOF MS using the Axima. Clearly a greater degree of separation is required to identify more sperm proteins, therefore the next step was to use LC-MS to try and identify more proteins. Instead of performing an in-gel digestion before LC-MS, an in-solution digestion was used. This method of digestion was chosen for two main reasons. Firstly, although in-gel digestion is a useful tool, it can result in the loss of a substantial amount of material (protein) during digestion itself, or before digestion, during SDS PAGE. For example, protein loss may occur during destaining, by adsorption on surfaces of pipette tips and digestion tubes, or due to incomplete extraction of peptides from the gel after digestion [58]. Secondly, since the samples would be subject to chromatographic separation prior to MS analysis, a pre-digestion separation method (PAGE) resulting in unnecessary loss of protein could be avoided.

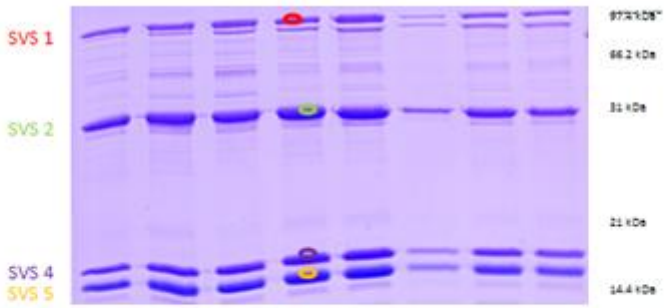


Figure 30. Identification of proteins present in a 1D SDS gel of SVS samples. The coloured circles indicate which bands were excised, although this is only shown in one lane, all lanes were excised. Each lane represents a sample taken from a different animal over the course of the labelling experiment. The column on the left-hand side indicates the proteins identified in each band. SVS 1 was identified by MS and a database search, the remaining SVS proteins were identified by deductive labelling using previously published data [59].

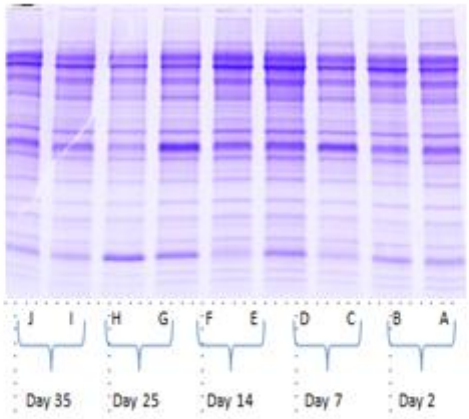


Figure 31A. Reproducibility of sperm samples. Samples were collected from 8 different animals over the course of the labelling experiment. They were then loaded onto a 1D SDS gel and run at 200mV for 1 hour.

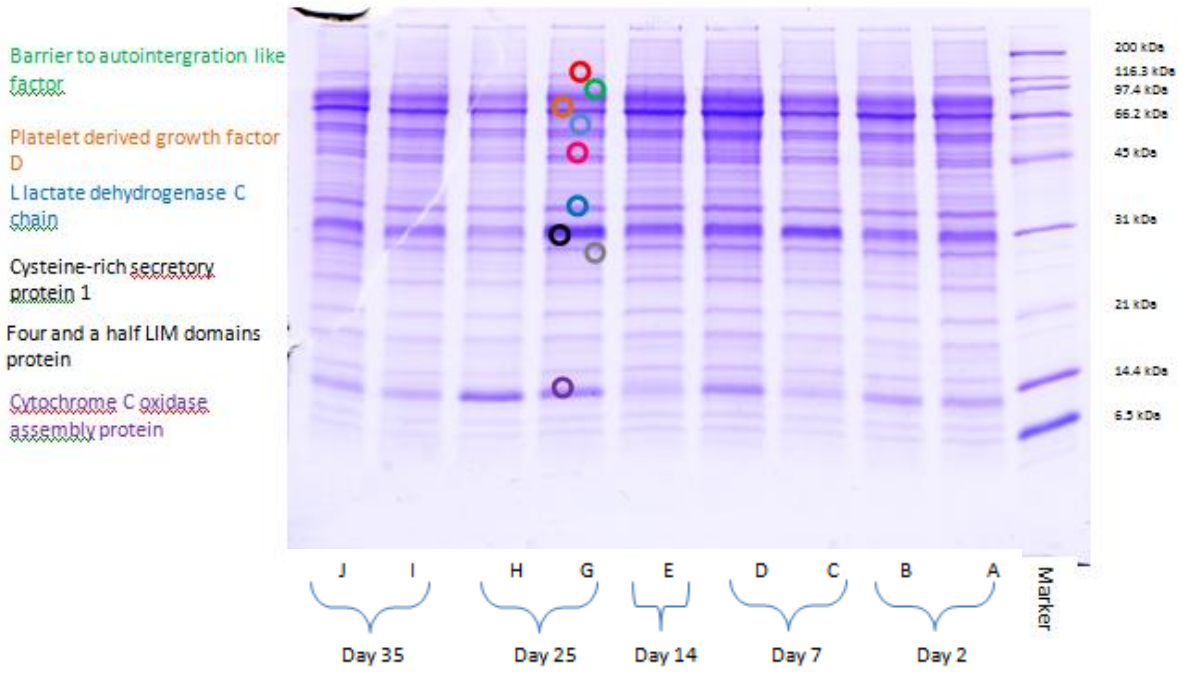


Figure 31B. 1D SDS of sperm samples. The coloured circles indicate which bands were excised, although this is only shown in one lane, all lanes were excised. The column on the left-hand side indicates the proteins identified in each band. Sample F was low in volume and therefore was omitted from this gel to conserve sample for in solution digests.

Furthermore, with complex samples such as the sperm samples, although PAGE will offer some degree of protein separation, there will still be many proteins in each band, making the identification of these proteins a more difficult task.

After in-solution digestion samples were run on a 1D SDS gel to check for complete digestion (Fig. 32). An undigested sample ran alongside the digested samples to act as a control. It was appropriately diluted so that the amount of protein loaded in the control lane was the same as that loaded in the digestion lanes (5µg- protein concentration determined by a Bradford assay). Since bands were observed in the control lane, this suggested that the amount of protein loaded was sufficient enough to be detected by the staining method used (Coomassie). Furthermore the banding pattern observed in the control lane was very similar to that seen in figure 31B, suggesting it was infact sperm proteins that were being observed and not a contaminant. Therefore any undigested material, if present, would have been detected. In the lanes containing the digested material no bands were observed, suggesting that digestion had reached completion. After confirmation of complete digestion was obtained, samples could then be subject to MS analysis on the LTQ. This data was then used to interrogate a database (Swis-Prot). 47 proteins in total were identified, of which 21 were found to be either sperm specific or highly expressed in the testis (Table. 1). 20 of these sperm proteins contained at least one valine residue somewhere in their sequence, and therefore could potentially inform on the labelling pattern of sperm specific proteins. Sperm protamine-P1 did not contain any valine residues and therefore could not be of use in this labelling study. After identifying a substantial number of sperm specific proteins, the next step was to look at the labelling pattern of valine peptides within these proteins.

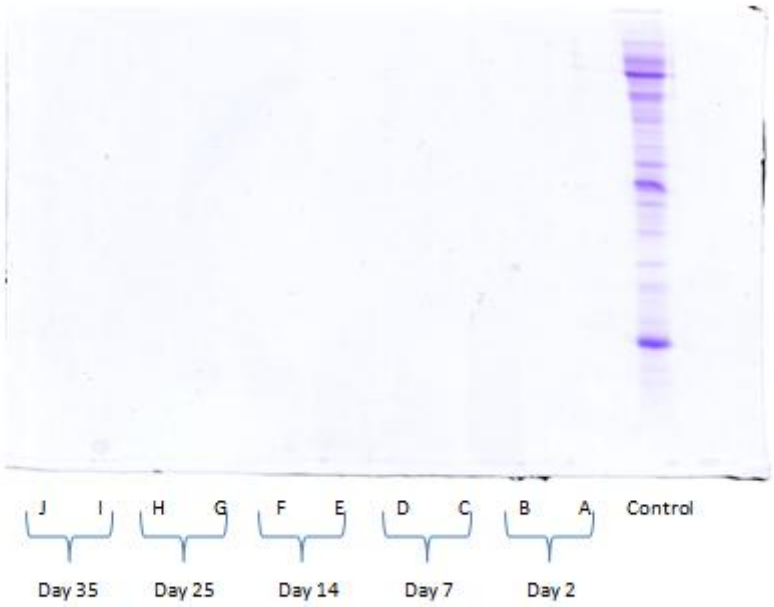


Figure 32. ID SDS gel of in-solution digested sperm samples. Samples were digested with trypsin overnight. 10µL of the digest was loaded onto each lane, the equivalent of approximately 5µg of protein. The control was an original, un-digested sperm sample, appropriately diluted so that 5µg of protein was loaded onto the gel.

Protein	Score	Coverage	Sperm Specific?
Serum albumin precursor	767	54%	No
Intergral membrane associated protein	393	16%	No
A-kinase anchor protein 4 precursor	308	17%	Yes
L-lactate dehydrogenase C chain	201	26%	Yes
Tubulin beta-2C chain	192	20%	No
Cysteine-rich secretory protein 1 precursor	161	28%	Yes
Alpha-tubulin 3/7	151	14%	Yes
Tubulin beta-3 chain	135	14%	No
Outer dense fiber protein	133	27%	Yes
Phosphoglycerate kinase	127	14%	Yes
Alpha-tubulin 8	116	10%	No
Tubulin beta-2A	111	13%	No
Lipocalin-type prostaglandin-D synthase	102	20%	No
PHGPx	100	21%	Yes
Pyruvate dehydrogenase E1 componant	96	7%	No
Haemoglobin alpha chain	90	29%	No
Testis lipid-binding protein	87	10%	Yes
Contrapsin	82	10%	No
Phosphatidylethanolamine-binding protein	70	26%	No
Glutathione S-Transferase	70	13%	Yes
ATP synthase subunit alpha	64	6%	No
Hexokinase-1	63	3%	Yes
ATP synthase subunit beta	60	6%	No
Adenylate kinase isoenzyme 1	58	13%	No
Alpha-enolase	57	2%	No
Clusterin precursor	56	3%	Yes

Table 1. Proteins identified from a database search using Swis-Prot of an in-solution digest of a sperm sample. The sample was collected on day 2 of the labelling experiment. 20µg of protein was digested in a total volume of 40µL using a standard Rapigest protocol. 500ng of the digest was loaded onto the LTQ for MS analysis. A mowse score of 56 and above is considered significant, therefore only proteins that scored above this figure have been shown. Determination of whether or not a protein was sperm specific was done using the information provided by Mascot during identification.

Although useful for identification purposes, the LTQ was not suitable for the collection of heavy valine labelling data. This was because the heavy and light species of each peptide would elute from the chromatography system at different times, making subsequent analysis very difficult. Instead it was decided to use an MS^E approach on the Synapt.

3.2.4.2. Deduction of Labelling Pattern

Several peptides identified in both the SVS and sperm samples contained valine, and these peptides were used to track and compare labelling. The peptides obtained from the SVS samples showed very rapid labelling, with evidence of labelling present at as early as two days in some samples (Fig. 33). In contrast, the peptides from a protein identified in the sperm samples (lactate dehydrogenase C chain) labelled much more slowly, generally not showing any evidence of labelling until day 25 of the labelling experiment (Fig. 34). The RIA for both the SVS and the sperm peptides were calculated, and the results compared (Fig. 35). It is very apparent that there is a delay in the labelling of sperm proteins relative to SVS proteins. By day 35 the SVS 1 peptides showed an average RIA of 4.9, indicating that equilibrium had been reached at this point, since the mice were fed a 50% heavy diet. However if we look at the sperm protein (lactate dehydrogenase) on day 35, an RIA of 0.43 is observed, suggesting that equilibrium has not yet been reached and the sperm protein is labelling at a much slower rate.

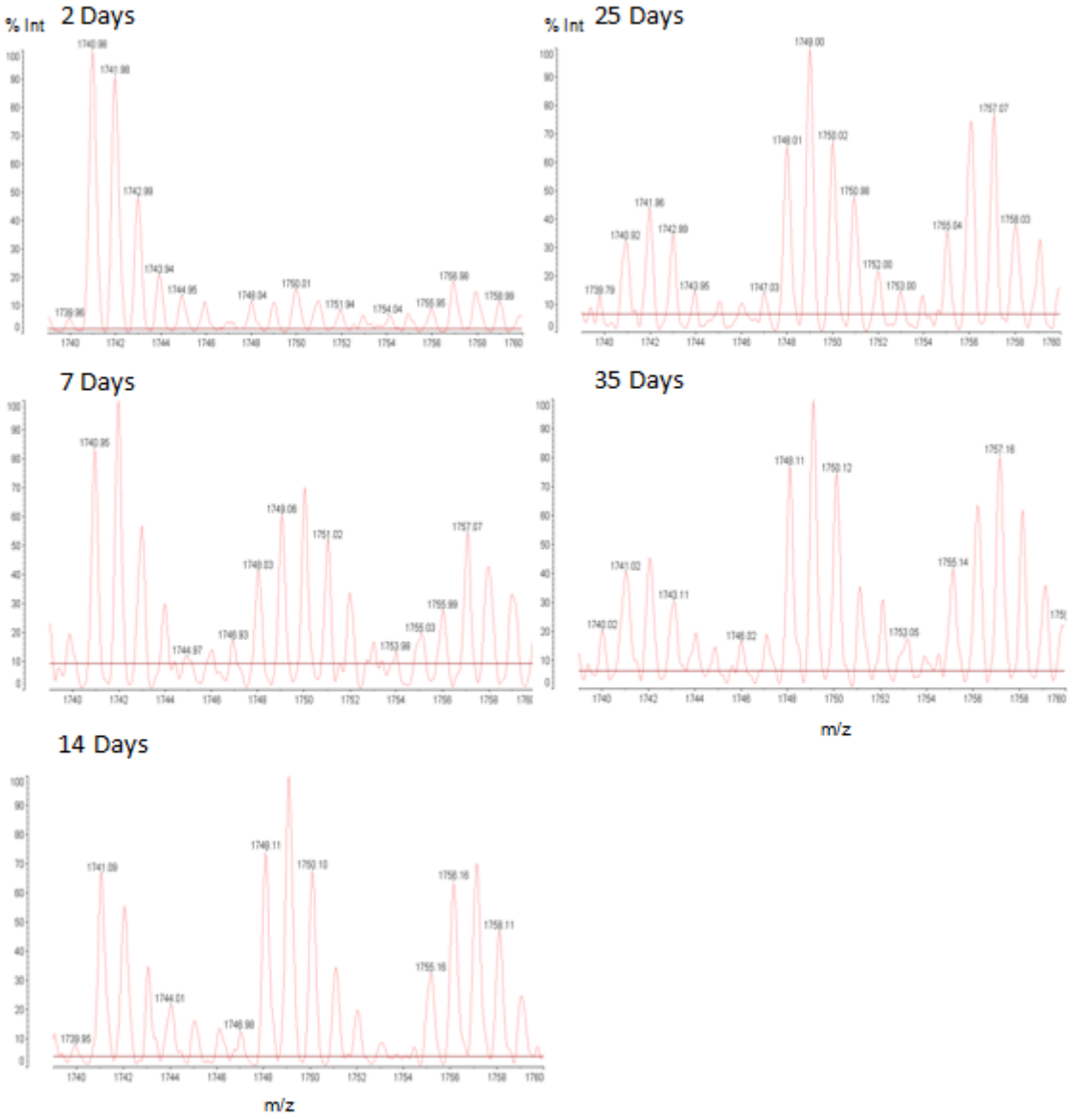


Figure 33. Svs 1 Peptide T47 (IAYEVGVQEVMALYR) over the course of the labelling experiment. Samples were analysed by MALDI-TOF MS using the Axima. Each spectrum is from a different animal.

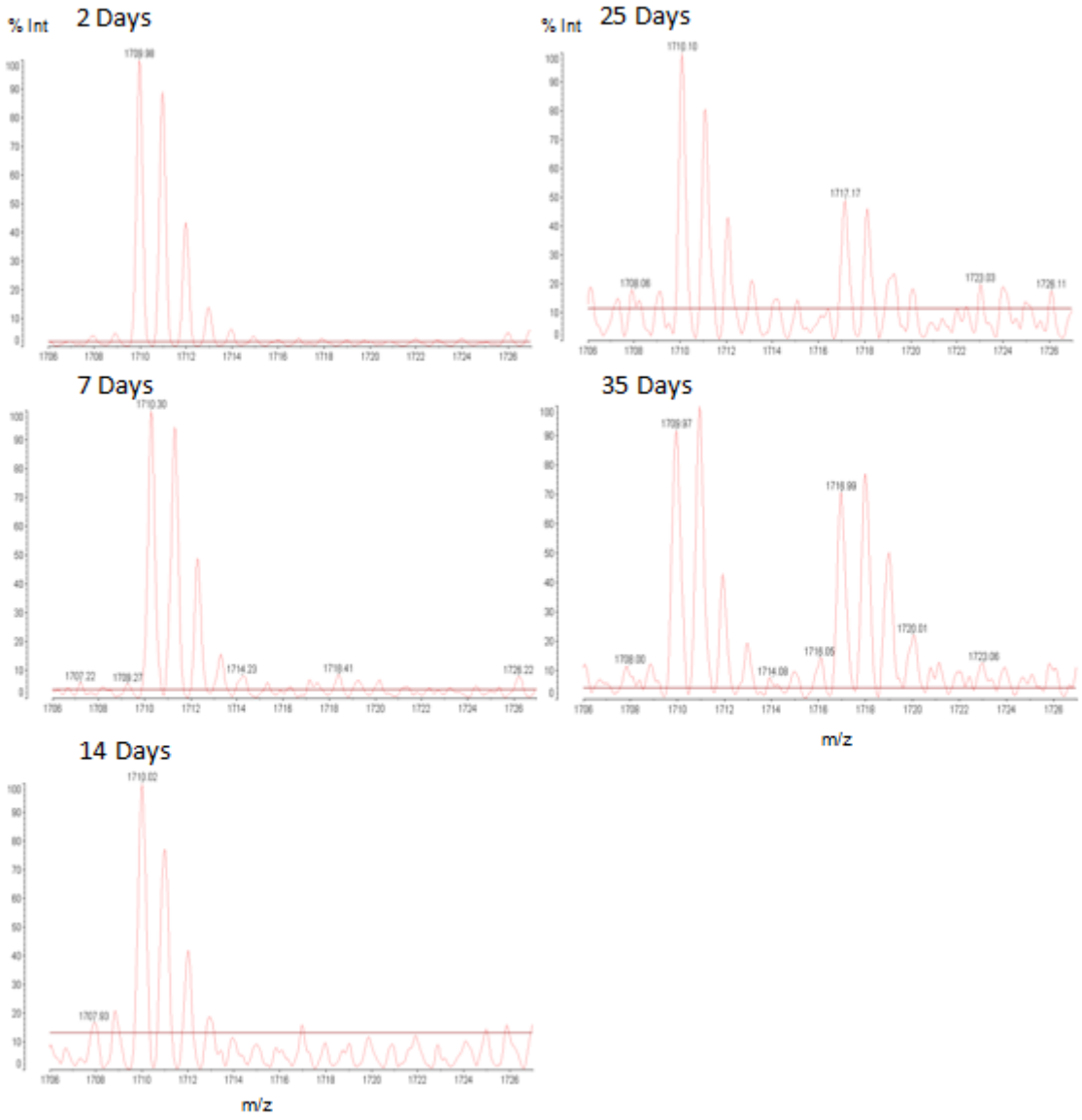


Figure 34. Lactate dehydrogenase C chain (LDH-X) peptide T26 (GYTSWAIGLSVTDLAR) over the course of the labelling experiment. Samples were analysed by MALDI-TOF MS using the Axima. LDH-X is an isoenzyme of lactate dehydrogenase, specific for germinal epithelium activity.

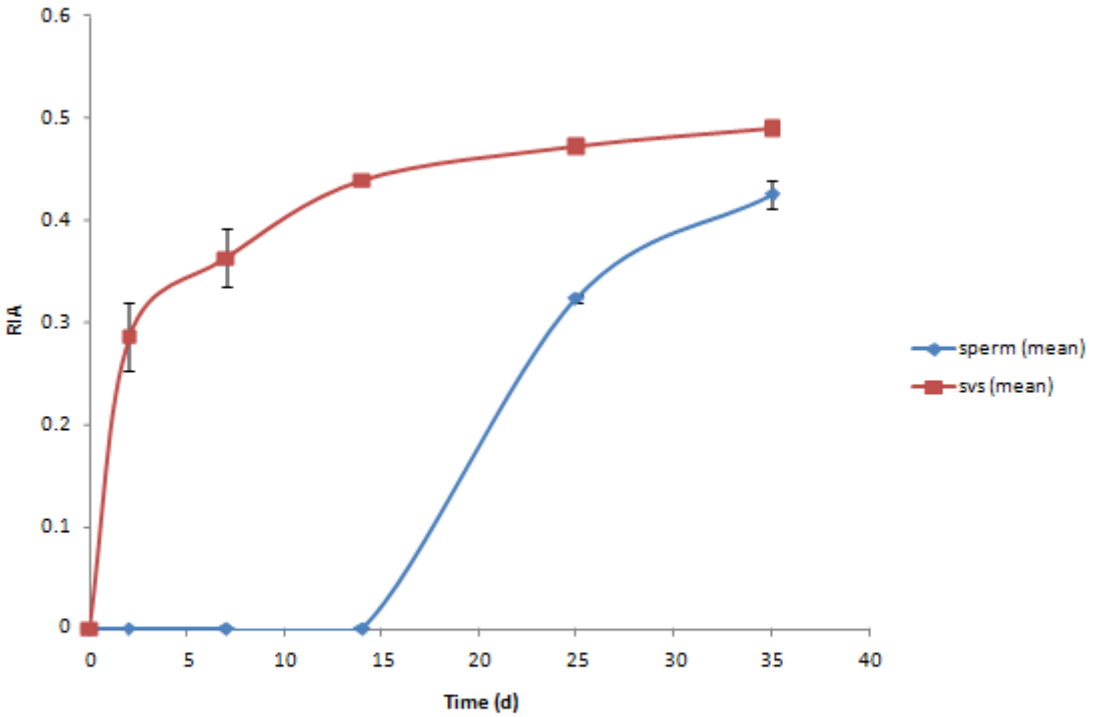


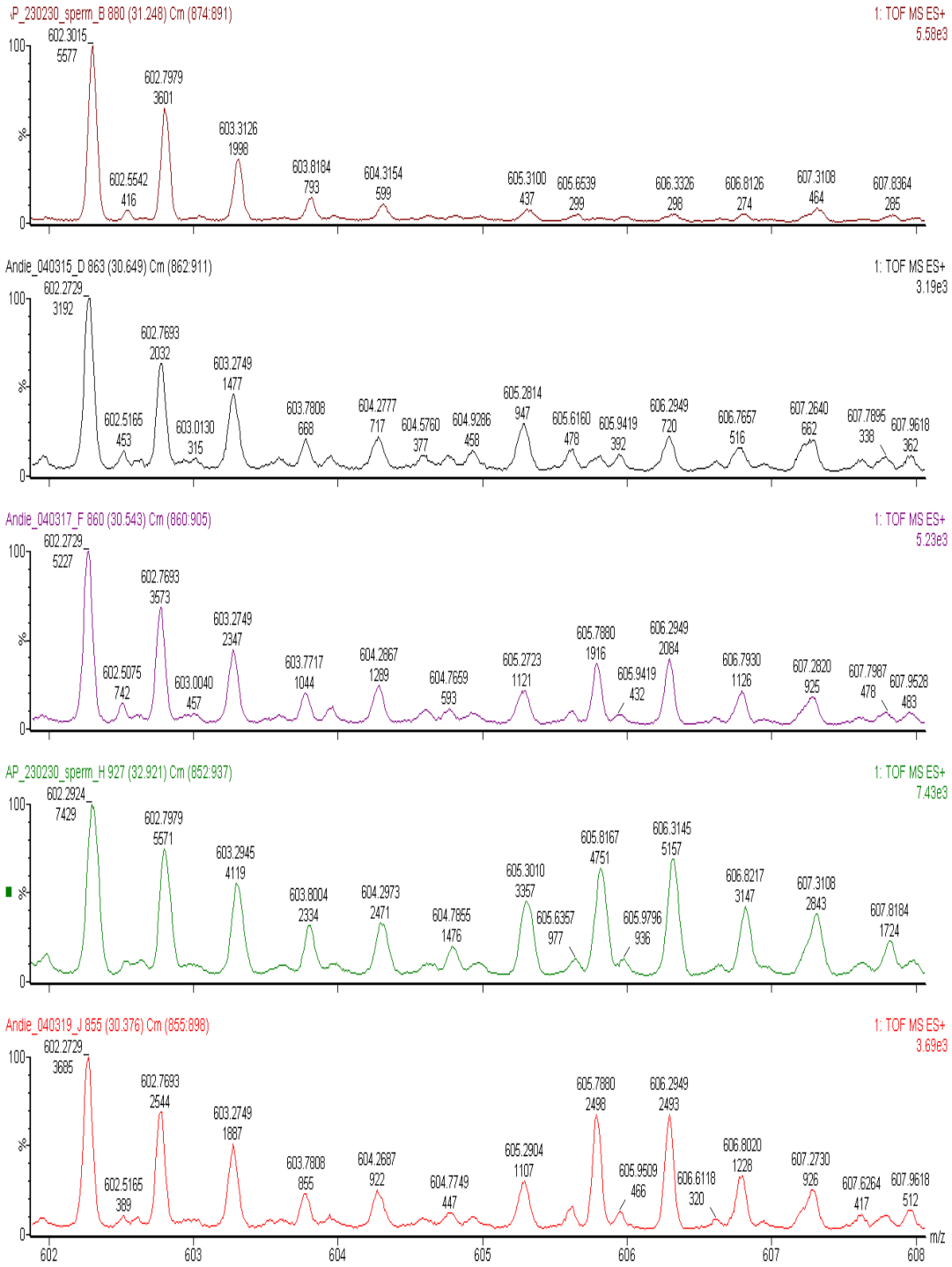
Figure 35. Differential labelling of sperm proteins and seminal vesicle secretion proteins. Observed in mice exposed to a 50% heavy valine diet over a 35 day period. The sperm protein used was lactate dehyrogenase and the SVS protein used was SVS 1 RIA was calculated at each interval of the labelling experiment for each valine peptide that was identified. The results were then averaged for each time point.

This concurs with what we know about spermatogenesis. In mice, the total duration of spermatogenesis is 35 days, after which the sperm remain in the epididymis for seven days [60]. This would explain a delay in the emergence of labelled peptide, as for at least seven days we are looking at sperm that have matured prior to the introduction of heavy valine.

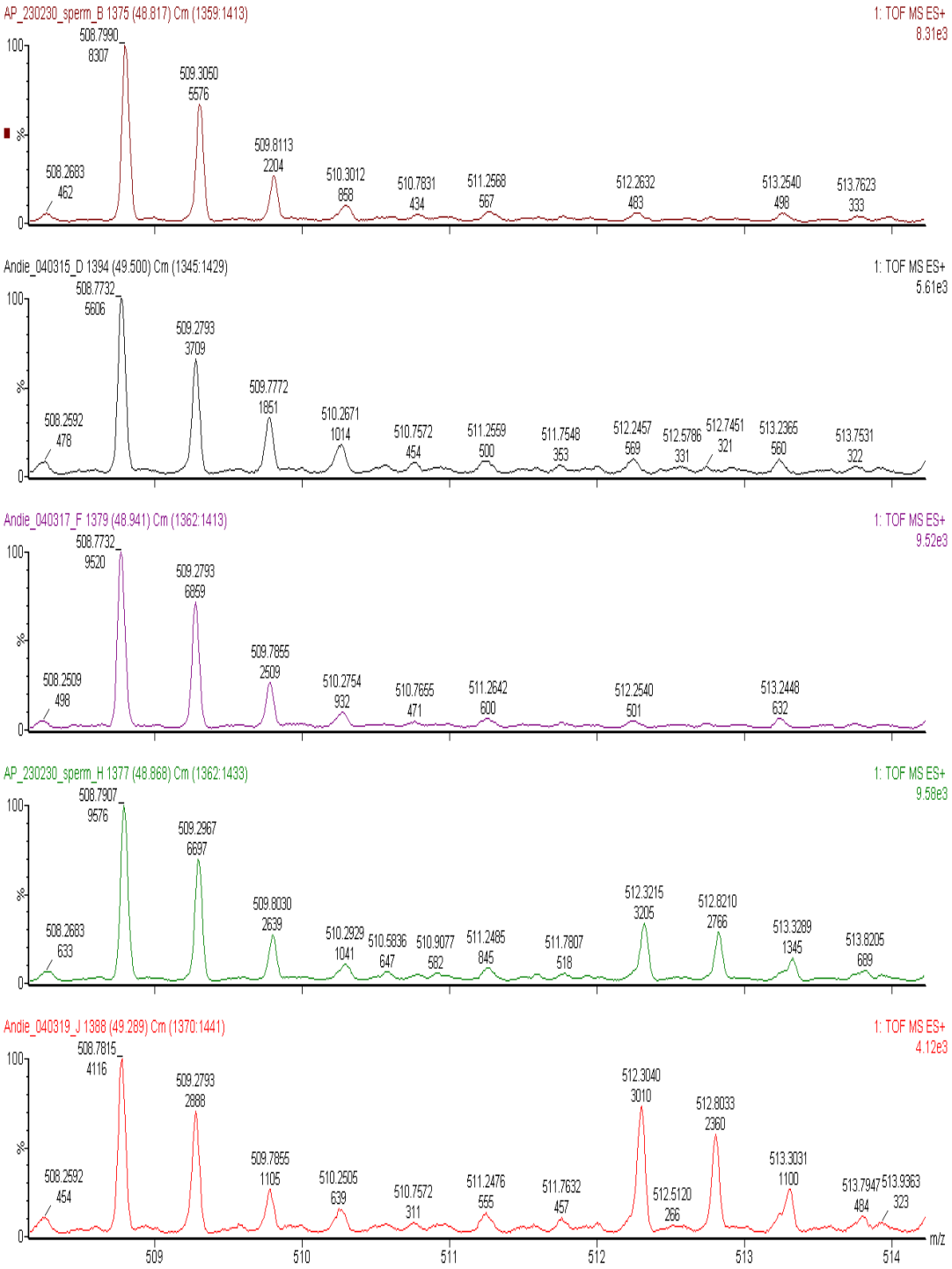
Although spermatogenesis lasts for 35 days, protein synthesis is stage specific. Therefore after the introduction of heavy valine, proteins that are expressed in the later stages of maturation will show evidence of labelling earlier (as they will be released from the testis earlier) whereas proteins that are expressed in the early stages of maturation will show evidence of labelling later (as they will be released from the testis later). One study suggested that lactate dehydrogenase is highly expressed during pachetyne, and to a lesser extent post-mitotically [61]. The combination of these two factors could explain the delay in the emergence of heavy valine in the lactate dehydrogenase peptides.

It was decided to use MS^E to look at the emergence of heavy label in more sperm proteins, to determine whether the delay in labelling was characteristic of all sperm proteins. Figure 36 shows the emergence of the heavy species in four different proteins; A kinase anchor protein, voltage dependent anion selective channel protein 2, lactate dehydrogenase C chain and CUB and zona pellucida like domain containing protein. With the exception of CUB and zona pellucida like domain containing protein all the sperm proteins looked at showed a delay in labelling relative to SVS proteins (Fig.37). However the level of delay observed is not consistent across sperm proteins and shows much variation.

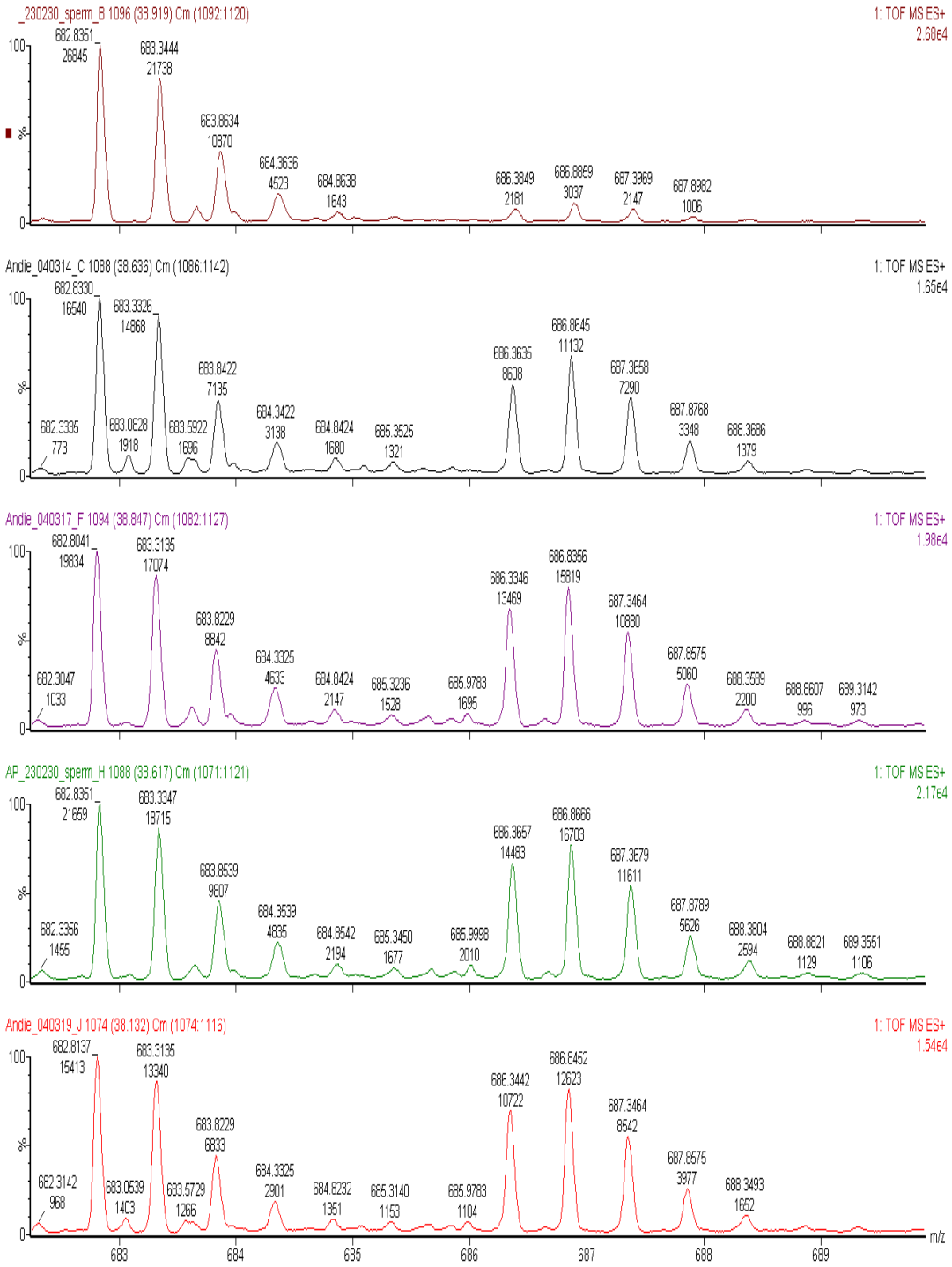
A


AP_230230_sperm_B 880 (31.248) Cm (874:891)
1: TOF MS ES+
5.58e3
602.3015 5577
602.7979 3601
603.3126 1998
602.5542 416
603.8184 793
604.3154 599
605.3100 437
605.6539 299
606.3326 298
606.8126 274
607.3108 464
607.8364 285
Andie_040315_D 863 (30.649) Cm (862:911)
1: TOF MS ES+
3.19e3
602.2729 3192
602.7693 2032
603.2749 1477
602.5165 453
603.0130 315
603.7808 668
604.2777 717
604.5760 377
604.9286 458
605.2814 947
605.6160 478
605.9419 392
606.2949 720
606.7657 516
607.2640 662
607.7895 338
607.9618 362
Andie_040317_F 860 (30.543) Cm (860:905)
1: TOF MS ES+
5.23e3
602.2729 5227
602.7693 3573
603.2749 2347
602.5075 742
603.0040 457
603.7717 1044
604.2867 1289
604.7659 593
605.2723 1121
605.7680 1916
605.9419 432
606.2949 2084
606.7930 1126
607.2820 925
607.7987 478
607.9528 483
AP_230230_sperm_H 927 (32.921) Cm (852:937)
1: TOF MS ES+
7.43e3
602.2924 7429
602.7979 5571
603.2945 4119
603.8004 2334
604.2973 2471
604.7855 1476
605.3010 3357
605.6357 977
605.8167 4751
605.9796 936
606.3145 5157
606.8217 3147
607.3108 2843
607.8184 1724
Andie_040319_J 855 (30.376) Cm (855:898)
1: TOF MS ES+
3.69e3
602.2729 3685
602.7693 2544
603.2749 1887
602.5165 389
603.7808 855
604.2687 922
604.7749 447
605.2904 1107
605.7680 2498
605.9509 466
606.2949 2493
606.6118 320
606.8020 1228
607.2730 926
607.6264 417
607.9618 512
100
%
0
602
603
604
605
606
607
608
m/z

B


AP_230230_sperm_B 1375 (48.817) Cm (1359:1413)
1: TOF MS ES+
8.31e3
508.7990 8307
509.3050 5576
509.8113 2204
508.2683 462
510.3012 858
510.7831 434
511.2568 567
512.2632 483
513.2540 498
513.7623 333
Andie_040315_D 1394 (49.500) Cm (1345:1429)
1: TOF MS ES+
5.61e3
508.7732 5606
509.2793 3709
509.7772 1851
508.2592 478
510.2671 1014
510.7572 454
511.2559 500
511.7548 363
512.2457 569
512.5786 331
512.7451 321
513.2365 560
513.7531 322
Andie_040317_F 1379 (48.941) Cm (1362:1413)
1: TOF MS ES+
9.52e3
508.7732 9520
509.2793 6859
509.7855 2509
508.2509 498
510.2754 932
510.7655 471
511.2642 600
512.2540 501
513.2448 632
AP_230230_sperm_H 1377 (48.868) Cm (1362:1433)
1: TOF MS ES+
9.58e3
508.7907 9576
509.2967 6697
509.8030 2639
508.2683 633
510.2929 1041
510.5836 647
510.9077 582
511.2485 845
511.7807 518
512.3215 3205
512.8210 2766
513.3289 1345
513.8205 689
Andie_040319_J 1388 (49.289) Cm (1370:1441)
1: TOF MS ES+
4.12e3
508.7815 4116
509.2793 2888
509.7855 1105
508.2592 454
510.2505 639
510.7572 311
511.2476 555
511.7632 457
512.3040 3010
512.8033 2360
512.5120 266
513.3031 1100
513.7947 484
513.9363 323
m/z

C


'_230230_sperm_B 1096 (38.919) Cm (1092:1120)
1: TOF MS ES+
2.68e4
682.8351 26845
683.3444 21738
683.8634 10870
684.3636 4523
684.8638 1643
686.3849 2181
686.8859 3037
687.3969 2147
687.8982 1006
Andie_040314_C 1088 (38.636) Cm (1086:1142)
1: TOF MS ES+
1.65e4
682.8390 16540
683.3326 14868
683.8422 7135
684.3422 3138
684.8424 1680
685.3525 1321
686.3635 8608
686.8645 11132
687.3658 7290
687.8768 3348
688.3686 1379
682.3335 773
683.0828 1918
683.5922 1696
Andie_040317_F 1094 (38.847) Cm (1082:1127)
1: TOF MS ES+
1.98e4
682.8041 19834
683.3135 17074
683.8229 8842
684.3325 4633
684.8424 2147
685.3236 1528
685.9783 1695
686.3346 13469
686.8356 15819
687.3464 10880
687.8575 5060
688.3589 2200
688.8607 996
689.3142 973
682.3047 1033
AP_230230_sperm_H 1088 (38.617) Cm (1071:1121)
1: TOF MS ES+
2.17e4
682.8351 21659
683.3347 18715
683.8539 9807
684.3539 4835
684.8542 2194
685.3450 1677
685.9998 2010
686.3657 14483
686.8666 16703
687.3679 11611
687.8789 5626
688.3804 2594
688.8821 1129
689.3551 1106
682.3356 1455
Andie_040319_J 1074 (38.132) Cm (1074:1116)
1: TOF MS ES+
1.54e4
682.8137 15413
683.3135 13340
683.8229 6833
684.3325 2901
684.8232 1351
685.3140 1153
685.9783 1104
686.3442 10722
686.8452 12623
687.3464 8542
687.8575 3977
688.3493 1652
682.3142 968
683.0539 1403
683.5729 1266
683
684
685
686
687
688
689
m/z

D

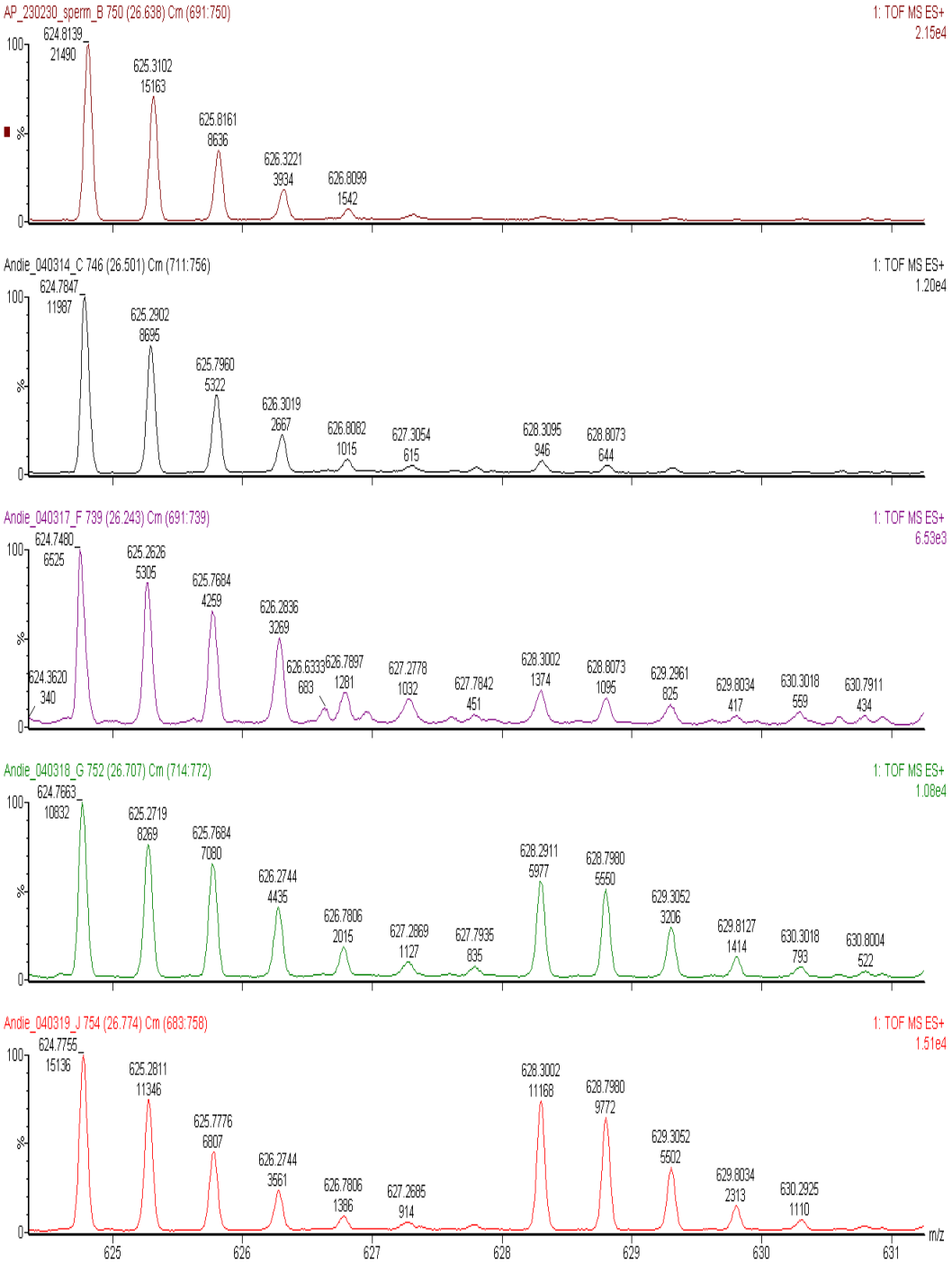


Figure 36. The emergence of heavy labelled peptide in 4 different proteins found in sperm. Observed in mice exposed to a 50% heavy valine diet over a 35 day period. Sperm samples were digested in solution using trypsin, then analysed using the Synapt, an LC-MS instrument. A is A kinase anchor protein, B is voltage dependant anion exchange protein, C is CUB and zona pellucinda like domain containing protein and D is Lactate dehydrogenase C chain.

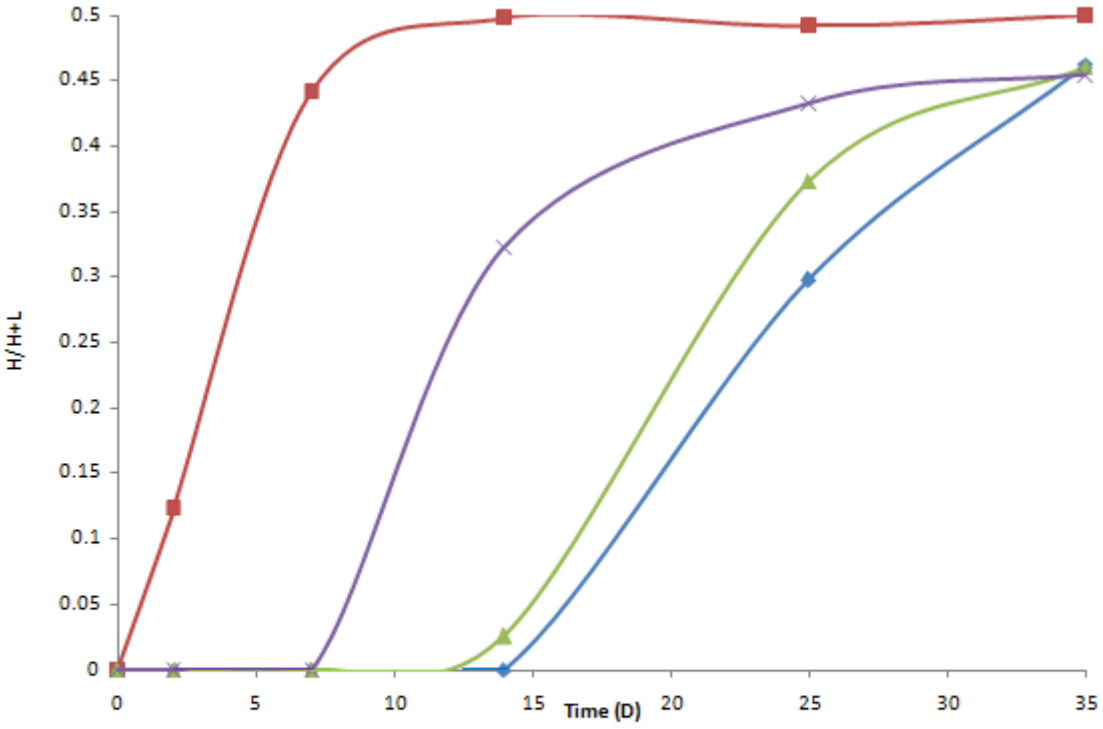


Figure 37. The labelling pattern of four sperm specific proteins. Observed in mice exposed to a 50% heavy valine diet over a 35 day period. RIA was calculated at each interval of the labelling experiment using all visible peaks in the isotopic envelope. The different colours represent different proteins; red represents Cub and zona pellucinda like domain containing protein, purple represents A kinase anchor protein, green represents lactase dehyrogenase C chain and blue represents voltage dependant anion selective channel protein 2.

For example A kinase anchor protein shows evidence of labelling after day seven, however dependent anion selective channel protein 2 does not show evidence of labelling until after day 14. This would suggest that A kinase anchor protein is turned over at a faster rate than dependent anion selective channel protein 2 and therefore labels quicker. The delay in labelling of some sperm proteins relative to SVS proteins could be exploited to selectively label specific proteins. This in turn could be used in sperm competition studies, as different animals could be exposed to different labels for a specific length of time, resulting in a label being incorporated into protein of interest. This label would be unique to the animal and therefore could be used as an ownership tag during heterospermic insemination.

4. Summary and Conclusions

In conclusion, the use of stable isotopes as metabolic tracers is an important tool in proteomic research. However in order to realise the full potential of this tool, labelling with stable isotopes must be coupled with rigorous separation techniques such as GeLC MS, LC-MS and LC MALDI and high resolution MS analysis. All of the SVS proteins indentified share a characteristic labelling pattern i.e. rapid labelling commencing as early as day 2. Sperm proteins show a lot of variation in their labelling pattern, with some proteins showing evidence of rapid labelling, while others display delayed labelling. These findings are particularly interesting as it is never before been shown that SVS proteins label much faster relative to sperm proteins, or that sperm proteins show such a great deal of variation in labelling. The results of this study could possibly be exploited in order to gain a greater understanding of the role protein expression and regulation of protein expression in the male mouse reproductive system. It is possible that this variation in labelling pattern could be exploited in sperm competition studies. Animals could be exposed to a heavy isotope label for a specific length of time, so that a sperm protein could be selectively labelled. This label could then act as an ownership tag in sperm competition studies. This will hopefully enable us to gain a deep understanding of both protein turnover and investment during HI.

References

1. Simpson, R. J. (2003). Proteins and Proteomics. New York: Cold Spring Harbour Laboratory Press. 5-6.
2. Clamp, M., Fry, B., et. al. (2007) Distinguishing protein-coding and noncoding genes in the human genome. Proceedings of the National Academy of Sciences of the United States of America. **104,** 19428-19433.
3. Sterck, L., Rombauts S., Vandepoele, K., Rouze P., and Van de Peer, Y. (2007) How many genes are there in plants (... and why are they there?). Current Opinion in Plant Biology. **10,** 199-203.
4. Purves, W. K., Sadava, D., Orians G. H., and Heller H. C. (2003) Life: The Science of Biology. W. H Freeman. 122-143.
5. Simpson, R. J. (2003). Proteins and Proteomics. New York: Cold Spring Harbour Laboratory Press. 7-8.
6. Doudna, J. A., and Rath, V. L. (2002) Structure and function of the eukaryotic ribosome: the next frontier. Cell. **109,** 153-156.
7. Rennie, M. J. (1999) An introduction to the use of tracers in nutrition and metabolism. Cambridge Journals. **58**, 935–944.
8. Ryu, S., Gallis, Y. A. G., Shaffer, S. A., Radulovic, D., and Goodlett D. R. (2008) Comparison of a Label-Free Quantitative Proteomic Method Based on Peptide Ion Current Area to the Isotope Coded Affinity Tag Method. Cancer Informatics. **6,** 243-255.
9. Roberts, R. M. (1989) Serendipity/Accidental Discoveries in Science. John Wiley, New York. 9-23
10. Gest, H. (2006) The Early History of 32P as a Radioactive Tracer in Biochemical Research. A Personal Memoir. Biochemistry and Molecular Biology Education. **33,** 159-164.
11. Beynon, R. J. (2005) The dynamics of the proteome: Strategies for measuring protein turnover on a proteome-wide scale. Briefings in Functional Genomics and Proteomics. **3,** 382-390.
12. O'Farrell, P. H. (1975) High resolution two-dimensional electrophoresis of proteins. Journal of Biological Chemistry. **250,** 4007–4021
13. Ong, S-E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labelling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Molecular and Cellular Proteomics. **1,** 376-386.
14. Browne, T. R., Van Langenhove, A., Costello, C. E., Biemann, K., and Greenblatt, D. J. (1981) Kinetic equivalence of stable-isotope-labeled and unlabeled phenytoin. Clinical Pharmacology and Therapeutics. **29,** 511–515
15. Cho, A. K., Lindeke, B., Hodshon, B. J., and Jenden, D. J. (1973) Deuterium substituted amphetamine as an internal standard in gas chromographic/mass spectrometric (GC/MS) assay for amphetamine. Analytical Chemistry. **45,** 570-574.

16. Shevchenko, A., Chernushevich, I., Standing, K. G., Thompson, B., Wilm, M., and Mann, M. (1997) Rapid "de novo" peptide sequencing by a combination of nanoelectrospray, isotopic labeling and a quadrupole/time-of-flight mass spectrometer. Rapid Communications in Mass Spectrometry. **11,** 1015–1024
17. Beyon, R. J., and Pratt, J. M. (2005) Metabolic labelling of proteins for proteomics. Molecular and Cellular Proteomics. **4,** 857-872.
18. Cargile, B. J., Bundy, J. L., Grunden, A. M., and Stephenson, J. L. (2004) Synthesis/degradation ratio mass spectrometry for measuring relative dynamic protein turnover. Analytical Chemistry. **76,** 86-97.
19. Doherty, M. K., Whitehead, C., McCormack, H., Gaskell, S. J., and Beynon, R. J. (2005) Proteome dynamics in complex organisms: Using stable isotopes to monitor individual protein turnover rates. Proteomics. **5,** 522-533.
20. Blagoev, B., Ong, S. E., Kratchmarova, I., and Mann, M. (2004) Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. Nature Biotechnology. **22,** 1139–1145
21. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M.H., and Aebersold, A. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nature Biotechnology. **17,** 994-999.
22. Roth, F.P., Hughes, J.D., Estep, P.W. and Church, G.M. (1998). Finding DNA regulatory motifs within unaligned noncoding sequences clustered by whole-genome mRNA quantitation. Nature Biotechnology. **16**, 939–945.
23. Velculescu, V.E. (1997). Characterization of the yeast transcriptome. Cell. **88**, 243–251.
24. Wiese S, Reidegeld KA, Meyer HE, and Warscheid B. (2007) Protein labelling by iTRAQ: a new tool for quantitative mass spectrometry in proteomic research. Proteomics. **7,** 1004
25. Chong, P. K.; Gan, C. S.; Pham, T. K.; Wright, P. C. (2006) Isobaric tags for relative and absolute quantification Journal of Proteome Research. **5,** 1232-1240.
26. Ross, P.L. et al. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Molecular and Cellular Proteomics. **3**, 1154–1169.
27. Gan, C. S., Chong, P. K., Pham, T. K., and Wright, P. C. (2007) Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). Journal of Proteome Research. **6,** 821–27
28. Simpson, R. J. (2003). Proteins and Proteomics. New York: Cold Spring Harbour Laboratory Press. 220-222
29. Shen, Y., Page, J. S., and Smith, R. D.. (2009). Advanced capillary liquid chromatography - mass spectrometry for proteomics. In: Grushka, E., and Grinberg, N. Advances in chromatography. Boca Raton, Florida: CRC Press. 37-38.
30. Mann, M., Meng, C. K., and Fenn, J. B. (1989) Analytical Chemistry. **61**, 1702.
31. Simpson, R. J. (2003). Proteins and Proteomics. New York: Cold Spring Harbour Laboratory Press. 441-447.
32. Hellerstein, M. K., and Neese, R. A. (1992) Mass isotopomer distribution analysis: a technique for measuring biosynthesis and turnover of polymers. American Journal of Physiology – Endocrinology and Metabolism. **263,** E988-E1001.

33. Ciechanover, A. (2007) Intracellular protein degradation : from a vague idea thru the lysosome and the ubiquitin-proteosome system and onto human diseases and drug targeting. New York Academy of Sciences. **1116,** 1-28.
34. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007) Quantitative mass spectrometry in proteomics : a critical review. Analytical and Bioanalytical Chemistry. **389,** 1017-1031.
35. Dziuk, P. J. (1996) Review : Factors that influence the proportion of offspring sired by a male following heterospermic insemination. Animal Reproductive Science. **43,** 65-88.
36. Ramm, S. A., Parker, G. A., and Stockley, P. (2005) Sperm competition and the evolution of male reproductive anatomy in rodents. Proceedings B. **272,** 949-955.
37. Stockley, P. (2004) Sperm competition in mammals. Human Fertility. **7,** 91-97.
38. Mudge, J. M., Armstrong, S. D., Mclaren, K., Beynon, R. J., Hurst, J. L., Nicholson, C., Robertson, D. H., Wilming, L. G., and Harrow, J. L. (2008) Dynamic instability of the major urinary protein gene family revealed by genomic and phenotypic comparisons between C57 and 129 strain mice. Genome Biology. **9,** R91.
39. Bernhard, O. K., E. A. Kapp, and Simpson R. J**.** (2007) Enhanced analysis of the mouse plasma proteome using cysteine-containing tryptic glycopeptides. Journal of Proteome Research. **6,** 987-995.
40. Duncan, R. Matthai, R. Huppi, K. Roderieck, T. Potter, M. (1988) Genes that modify expression of MUPs in mice, Molecular and Cellular Biology. **8** 2705-2712.
41. Busch, R., Kin, Y-K., Neese, R. A., et.al. (2006) Measurement of protein turnover rates by heavy water labelling of nonessential amino acids. Biochimica et biophysica acta. **1760,** 730-744.
42. Xiao, G. G., Garg, M., Lim, S., Wong, D., Go, V. L., and Lee, W. N. P. (2008) Determination of Protein Synthesis in Vivo Using Labeling from Deuterated Water and Analysis of MALDI-TOF Spectrum. Journal of Applied Physiology. **104,** 828–836.
43. Barrett, W., **Zolotarjova**, N., Chen, H., et.al. (2004) Removal of multiple high-abundant proteins from mouse plasma using the agilent multiple affinity removal system for mouse. Application note: https://www.chem.agilent.com/Library/applications/5989-1347EN.pdf
44. Simpson, R. J. (2003). Proteins and Proteomics. New York: Cold Spring Harbour Laboratory Press. 144-147.
45. Simpson, R. J. (2003). Proteins and Proteomics. New York: Cold Spring Harbour Laboratory Press. 192-194.
46. Product note: Ciphergen Deep Proteome™ Research Services Explore the low abundance proteome. Ciphergen Biosystems, Inc.
47. Tiss, A., Smith, C., Timms, J., et.al. (2007) Serum peptide profiling using MALDI mass spectrometry; avoiding the pitfalls of coated magnetic beads using well-established zip tip technology. Proteomics. **7,** 77-89
48. Walsh. K. A. (1970) Trypsinogens and trypsins of various species. Methods in Enzymology. **19,** 41-63.
49. KeiCDlouha, V., Zylber, N., Tong, N-T., and Keil, B. (1971) Cleavage of glucagon by α- and β-trypsin. FEBS Letters. **16,** 287-290.

50. Poncz L., Dearborn, D. G. (1983) The resistance to tryptic hydrolysis of peptide bonds adjacent to N^{ε}, N-dimethyllsyl residues. Journal of Biological Chemistry. **258,** 1844-1850.
51. Vestling, M. M., Murphy, C. M., and Fenselau, C. (1990) Recognition of trypsin autolysis products by high-performance liquid chromatography and mass spectrometry. Analytical Chemistry. **62,** 2391-2394.
52. Shi, R., Kumar, C., Zougman, A., Zhang, Y., Podtelejnikov, A., Cox, J., Wisniewski, J. R., and Mann, M. (2007) Analysis of the mouse liver proteome using advanced mass spectrometry. Journal of Proteome Research. **6,** 2963-2972.
53. Schirle, M., Heurtier M-A., and Kuster B. (2003) Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography-tandem mass spectrometry. Molecular and Cellular Proteomics. **2,** 1297-1305.
54. Lasonder, E., Ishihama, Y., Anderson, J. S., et.al. (2002) Analysis of the Plasmodium falciparum proteome by high-accuracy mass spectrometry. Nature. **419,** 537-542.
55. Anderson, J. S., Lyon, C. E., Fox, A. H., et.al. (2002) Directed proteomic analysis of the human nucleolus. Current Biology. **12,** 1-11.
56. Montgomery, H., Francis, S., Sekiya, S., Gaskell, S. J., and Tanka, K. (2004) Comparative study of protein identification by off-line LC-MALDI-MS and 1D-gel electrophoresis MALDI-MS. Presented at ASMS.
57. Baker, M. A., Hetherington, L., Reeves, G. M., and Aitken, R. J. (2008) The mouse sperm proteome characterized via IPG strip prefractionation and LC-MS/MS identification. Proteomics. **8,** 1720-1730.
58. Granvogl, B., Plöscher, M., and Eichacker, L. A. (2007) Sample preparation by in-gel digestion for mass spectrometry-based proteomics. Analytical and Bioanalytical Chemistry. **389,** 991-1002.
59. Ramm, S.A., McDonald, L., Hurst, J.L., Beynon, R.J., and Stockley, P. (2009). Comparative Proteomics Reveals Evidence for Evolutionary Diversification of Rodent Seminal Fluid and Its Functional Significance in Sperm Competition. Molecular Biology and Evolution. **26**, 189-198.
60. Adler, I-D. (1996) Comparison of the duration of spermatogenesis between male rodents and humans. Mutation Research. **352,** 169-172.
61. Kramer, J. M., and Erikson, R. P. (1982) Analysis of stage-specific protein expression during spermatogenesis of the mouse by two-dimensional gel electrophoresis. Journal of Reproduction and Fertility. **64,** 139-144.
62. Karas, M., Bachmann, D., Bahr, U., and Hillenkamp, F. (1987) Matrix assisted ultraviolet laser desorption of non-volatile compounds. International Journal of Mass Spectrometry and Ion Processes. **78,** 53-68
63. Simpson, R. J. (2003). Proteins and Proteomics. New York: Cold Spring Harbour Laboratory Press. 444.
64. Pratt, J. M., Petty, J., Riba-Garcia, I., Robertson, D. H. *et al.* (2002) Dynamics of protein turnover, a missing dimension in proteomics. Molecular and Cellular Proteomics. **1,** 579–591.

65. Anderson, N. L., Polanski, M., Pieper, R., Gatlin, T., Tirumalai, R. S., Conrads, T. P., Veenstra, T. D., Adkins, J. N., Pounds, J. G., Fagan, R., and Lobley, A. (2004) The human plasma proteome: a non-redundant list developed by combination of four separate sources. Molecular and Cellular Proteomics. **3,** 311– 326.
66. Sherwood, L. (2010) Human Physiology: From Cells to Systems. Belmont: Brooks/Cole Cengage Learning. 391-416.
67. Hellerstein, M. K., and Neese, R. A. (1999) Mass isotopomer distribution analysis at eight years: theoretical, analytic, and experimental consideration. American Journal of Physiology, Endocrinology and Metabolism. **6,** 1146-1170.