Analytical methods for the identification of MHC restricted peptides by Mass Spectrometry

by

Joanne Elizabeth Clayton

A thesis submitted in partial fulfilment of the requirements of The Nottingham Trent University for the degree of Doctor of Philosophy

August 2006

This thesis is, to the best of my knowledge, original except where due reference is made.

Joanne Elizabeth Clayton, August 2006

Acknowledgements

I would like to thank my supervisors Professor Colin Creaser, Dr. Phil Bonner, Dr. Stephanie McArdle and Professor Robert Rees for their advice support and encouragement during the course of this study.

I would also like to mention the good friends I have made in the research group both past and present, Becky, Lucy, Claire, David, James, Michelle, Sadaf, Dan and Amanda. I would like to add a special mention for Dani Barry who provided me with a lot of samples and helped me to understand what they meant, she has been invaluable. She will be missed.

I thank my parents Ron and Christine Clayton and my brother and sister, Simon and Gemma and the boys (Jess, Barney, Callum, Trooper and Toby) for always encouraging and supporting me and most of all my husband Pete who is my soul mate.

Finally I would like to thank the Kay Kendall Leukaemia Fund for financial support.

Presentations

Immobilised metal affinity chromatography combined with nano-electrospray-LC/MS for the analysis of histidine containing peptides.

Joanne E. Clayton, Colin S Creaser, Emma Sheridan, Stephane Chambrey, Phillip L R Bonner, Stephanie McCardle and Robert C Rees. Proceedings of the 2nd Annual Analytical Research Forum, Kingston University, 15-17 July 2002.

Immobilised metal affinity chromatography combined with nano-electrospray-LC/MS for the analysis of histidine containing peptides.

Joanne E. Clayton, Colin S Creaser, Emma Sheridan, Stephane Chambrey, Phillip L R Bonner, Stephanie McCardle and Robert C Rees. Proceedings of the 26th Annual meeting of the BMSS, Loughborough University, 8-11th September 2002.

Immobilised metal affinity chromatography combined with capillary LC/electrospray MS for the recovery of MHC associated peptides.

Joanne E. Clayton, Emma Sheridan, Stephane Chambrey, Phillip L R Bonner, Stephanie McCardle, Robert C Rees and Colin S Creaser. Proceedings of the 3rd Annual Analytical Research Forum, Sunderland University, 21-23 July 2003.

Development of a procedure for the identification of MHC restricted peptides isolated from cell surfaces by capillary LC/nano-electrospray mass spectrometry.

Joanne E. Clayton; Stephanie McArdle; Robert C. Rees; Philip L.R. Bonner; Richard E. Clark; Anthony I. Dodi; Colin S. Creaser. Proceedings of the 52nd annual ASMS conference on mass spectrometry and applied topics, Nashville, 23-27 May 2004.

Development of a method for the identification of cell surface MHC restricted peptides by capillary LC/nano-electrospray mass spectrometry.

Joanne E. Clayton; Stephanie McArdle; Robert C. Rees; Philip L.R. Bonner; Richard E. Clark; Anthony I. Dodi; Colin S. Creaser. Proceedings of the 4th Annual Analytical Research Forum, University of Central Lancashire, Preston, 19-21 July 2004.

Development of a procedure for the clean-up and identification of cell surface MHC restricted peptides by capillary LC/nano-electrospray mass spectrometry.

Joanne E. Clayton; Stephanie McArdle; Robert C. Rees; Philip L.R. Bonner; Richard E. Clark; Anthony I. Dodi; Colin S. Creaser. Proceedings of the 27th annual meeting of the BMSS, University of Derby, 5-8th September 2004.

Abbreviations

APC	Antigen Presenting Cell
β_2m	β_2-microglobulin
CEA	Carcinoembryonic Antigen
CID	Collision Induced Dissociation
CLIP	Class II invariant chain protein
CML	Chronic Myeloid Leukaemia
CTL	Cytotoxic T-cell
ESI	Electrospray ionisation
FAB	Fast Atom Bombardment
HLA	Human Leukocyte Antigen
IDA	Imminodiacetic acid
IMAC	Immobilised metal ion affinity chromatography
LSIMS	Liquid secondary ionisation mass spectrometry
MALDI	Matrix assisted laser desorption/ionisation
MHC	Major Histocompatibility Complex
NTA	Nitrilotriacetic acid
QITMS	Quadrupole ion trap mass spectrometer
ODS	Octadecyl siloxane
PSD	Post Source Decay
RP-HPLC	Reverse phase high performance liquid chromatography
SCX	Strong cation exchange
SPE	Solid phase extraction
TAA	Tumour Associated Antigen
TCA	Trichloroacetic acid
TCR	T-cell receptor
TSA	Tumour Specific Antigen
TFA	Trifluouroacetic acid

Abstract

HPLC and mass spectrometry are versatile analytical techniques which have been widely applied to the analysis of naturally occurring proteins and peptides. Mass spectrometric conditions have been investigated for the analysis of MHC class I and II restricted peptides. A micro-electrospray ionisation source has been constructed for coupling with a capillary HPLC system and investigations have been performed to increase the sample loading volume onto the system.

A robust method has been developed for the clean-up and identification of MHC class I and class II restricted peptides from surface eluted cells by mass spectrometric techniques. The procedure utilises solid phase extraction to desalt and concentrate the cell surface eluates followed by fractionation using strong cation exchange chromatography and analysis by on-line capillary HPLC/μESI/MS/MS using a quadrupole ion trap. The method has been shown to recover model peptides at an immunologically significant level. The procedure has been applied to the analysis of cell surface eluates of patients suffering from chronic myeloid leukaemia and to eluates from k562 transfected cell lines. The purpose of these analyses was to acertain if a selection of HLA.A2 predicted peptides, originating from the bcr/abl fusion protein, are expressed on the surface of leukaemic cells. The clean-up method has also been applied to the identification of the class II invariant chain protein (CLIP) from immunopurified and cell surface eluted class II molecules.

Immobilised metal ion affinity chromatography (IMAC) also known as metal chelate chromatography has been shown to be a powerful tool for the purification of proteins. Using chelated and immobilised copper ions, conditions have been investigated for the selective and non-selective binding of peptides. A histidine selective method has been developed, using a HEPES buffer at pH 7.2, which is ~90 % selective. This method has been applied to the purification of cellular extracts and CML patient eluates prior to analysis by HPLC/MS/MS. The method has been shown to remain highly selective when subjected to these complex samples.

CONTENTS

CHAPTER 2: Development of MS and capillary HPLC/MS for the analysis of peptides

CHAPTER 3: Development of a method for the identification of MHC class I and class II restricted peptides

Overall and Specific Research Aims

Cancer is a collective term for a large number of diseases that exhibit common factors with regard to the loss of the biochemical mechanism that controls cellular growth. Such disease are particularly difficult to treat because the immune response mechanisms of the human body do not recognise cancerous cells as they are not foreign in nature. Leukaemia is a disease, simply described as a cancer of the blood cells, see Chapter 1.

The study of cancer diseases is a vast, ongoing, area of scientific research. Although a great deal of progress has been made in recent decades in utilising surgical, radiochemical and chemotherapeutic treatments, many researchers feel that the body's own immune response represents a potentially powerful ally in this fight if it could be activated effectively. The objective of this field was and continues to be the development of tailor made vaccines for the treatment of individual patients suffering from these diseases. The overall aim of the work reported in this thesis was the development of methods for the identification of MHC restricted peptides associated with chronic myeloid leukaemia (CML) by high performance liquid chromatography mass spectrometry (HPLC-MS).

These MHC restricted peptides, expressed on the surface of cells, dictate the body's immune response. By employing HPLC-MS to study cell surface peptides it was believed that peptides unique to leukaemia patients could be identified. Such unique

peptide molecules could then be incorporated into a leukaemia vaccine to bring about an immune response against cells expressing those identified peptides.

Individual Objectives

The work that was carried out to meet the aim originally established can be separated into three sections, each with their own particular objective. The objectives of these individual sections are described as follows.

Chapter 2

In order to allow the efficient analysis of peptides by HPLC-MS, initial work was carried out with the aim of establishing and optimising on-line electrospray/mass spectrometric conditions for the separation and identification of peptides. To fulfil this objective a number of instrumental parameters were studied, including variation of capillary temperature and tube lens offset voltage. The effect of parent ion charge state on the energy required to bring about fragmentation and the resultant fragment pattern generated was also studied. Since it has been shown that these two parameters can have a powerful influence on the nature of ions transmitted into the mass spectrometer.[1,2] Following this system optimisation, the optimised instrumental conditions would then be used for the analysis of MHC class I and II restricted peptides.

An on-line miniature electrospray (microspray) source was constructed in-house and evaluated with the aim of enhancing the detection capability of the instrument. In addition, the reduction in flow rate allows the on-line coupling of nano-scale (75 μm internal diameter) HPLC columns which themselves offer a sensitivity advantage. This work is described in Chapter 2.

Chapter 3 Objectives

As a continuation of the objectives described above in chapter 2, employing the optimum instrumental conditions established, analysis was to be carried out with the aim of developing a versatile mass spectrometry based approach to allow the identification of both MHC class I and class II peptides. Within the immunological community it is the current opinion, of a majority of researchers, that the identification of both types of peptide is necessary for the successful development of immunotherapeutic vaccines.[3-5]

Such analyses would then to be compared with the established electrospray ionisation (ESI)-MS procedures for the study of such peptides and the procedure previously developed at the Nottingham Trent University for the identification of the A3 class I bcr/abl peptide (KQSSKALQR) on the surface of leucophoresed white blood cells from patients diagnosed with CML, described in Chapter 1.

The surface elution method, as opposed to the cell lysis method, has formed the basis of a procedure previously developed at Nottingham Trent which was used to identify a peptide expressed on the surface of HLA-A3 CML patient cells, KQSSKALQR, which originates from the bcr/abl fusion protein.[6] Although this method was successfully used to identify this peptide it suffered from a number of deficiencies which this research aimed to improve. The problems associated with this procedure are discussed, and their potential remedies addressed, in detail in Chapter 3.

Aims-To develop a procedure for the analysis of cell surface peptides including;

1. Identify procedures for the initial clean-up of extraction buffer, removal of buffer salts and high molecular weight proteins and pre-concentration of peptide fractions required for the large volumes of elution buffer required compared to immunoaffinity extraction methods.
2. Develop and orthogonal multidimensional chromatographic fraction procedure for the analysis of the crude extracts prior to analysis by ESI/MS/MS.
3. Evaluate the procedure for the analysis of standard MHC class I peptides and internal standard peptides spiked into cell eluates and CML patient samples

4. Apply the procedure for the identification of bcr/abl MHC class I peptides expressed on the surface of leukaemic cells and particularly to establish if the bcr/abl peptide SSKALQRPV is expressed on the surface of HLA-A2 cells.
5. Explore the application of the cell surface method to the analysis of MHC class II associated peptides.

Chapter 4 Objectives

A parallel avenue of research was also introduced to those aims already described. Immobilised metal ion affinity chromatography (IMAC) was studied as a potential approach to aid in the analysis of biological samples. IMAC is a technique which has been widely used for the purification of proteins as described in section 1.4.1. The most widespread application to the analysis of peptides involves the use of histidine tags rather than the naturally occurring peptide.

The potential of IMAC exists in its inherent ability to distinguish between a molecule that contains a particular moiety and one that does not. Such distinction can then allow the isolation of desired species. In the case of the IDA-Cu(II) technique which was explored here the target moiety is one or more histidine residues.

In conjunction with the objectives discussed in the first two research sections, it is the aim of this avenue of research to explore the potential application of IMAC to the analysis of biological samples.

References

1. Lill J. R., *PhD Thesis*, **2000**, Nottingham Trent University.
2. Reynolds J. C., *PhD Thesis*, **2002**, Nottingham Trent University.
3. Smith C. M., Wilson N. S., Waithman J., Villadangos J. A., Carbone F. R., Heath W. R., Belz G. T., *Nat. Immunol.*, **2003**, 5, 1143-1148.
4. Steinaa L., Rasmussen P. B., Wegener A. M., Sonderbye L., Leach D. R., Rygaard J., Mouritsen S., Gautam A. M., *J. Immunol.*, **2005**, 175, 329-334.
5. Shedlock D. J., Shen H., *Science*, **2003**, 337-339.
6. Clark R. E., Dodi I. A., Hill S. C., Lill J. R., Aubert G., Macintyre A. R., Rojas J., Bourdan A., Bonner P. L. R., Wang L., Christmas S. E., Travers P. J., Creaser C. S., Rees R. C., Madrigal J. A., *Blood*, **2001**, 2887-2893.

CHAPTER 1

INTRODUCTION

1.1.0. Mass Spectrometry

Mass spectrometry is an analytical technique for the study of gas phase ions. The first mass spectrometer was constructed in 1912 by J.J. Thompson who used it to study O_2, N_2 CO, CO_2 molecular ions and later neon isotopes 20 and 22.[1] Since then it has found use in structural determination and characterisation of unknowns, quantification at trace levels and mixture analysis with sensitivities as low as the attomole/zeptomole region.[2]

A mass spectrometer produces ions from a sample and separates them according their mass to charge ratio. A mass spectrum is produced which is a record of the ion intensity at each mass to charge ratio for all ions that reach the detector. All mass spectrometers consist of three major components, an ionisation source, a mass analyser and a detector (Figure 1.1.0.1). Each of these components will be discussed in detail below.

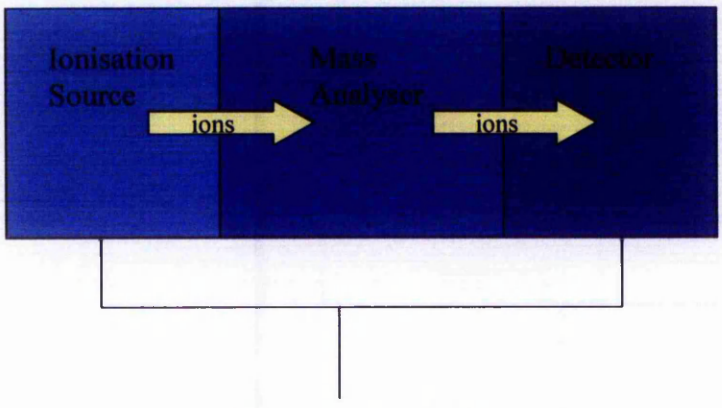


Figure 1.1.0.1. A schematic of a typical mass spectrometer.

1.1.1. Ionisation Sources

There are a variety of different ionisation techniques in use, the purpose of which is to ionise the sample prior to entrance into the mass analyser. The technique used is dependent on the state of the sample and its physicochemical properties. Three techniques, Fast atom bombardment (FAB), matrix assisted laser desorption ionisation (MALDI) and electrospray ionisation (ESI) have been used widely in biological mass spectrometry for the ionisation of peptides and proteins and these techniques are described in this section.

1.1.1.1. Fast Atom Bombardment

Fast atom bombardment (FAB), also known as liquid secondary ionisation mass spectrometry (LSIMS) is a low fragmentation ionization technique which does not require prior evaporation of the sample. It can be applied to polar compounds and those with acidic/basic functionality and so is useful for smaller biomolecules. A schematic of the method is shown in Figure 1.1.1.1.

Developed in 1980 by the Barber group,[3] fast atom bombardment was developed from an earlier technique, secondary ionization mass spectrometry (SIMS), in which a beam of high energy ions, typically Ar^+ were used to bombard the sample surface leading to a short lived secondary current of ions. By dissolving the sample in a viscous polar liquid matrix, such as glycerol, and replacing the initial ion beam with a high energy (6-10 KeV) beam of ions and neutral atoms, Barber increased the duration of the secondary ion current allowing it to be observed for periods of several minutes. As with SIMS the high energy incident

beam damages the sample surface, however, convection and diffusion inside the liquid matrix result in the replenishment of the surface from which new ions can be ejected.[4]

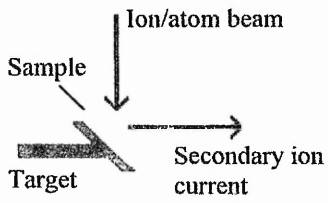


Figure 1.1.1.1. A schematic diagram of FAB/SIMS ionization technique. The ion/atom beam bombards the sample on the target surface leading to a secondary ion current which is directed into the mass analyser.

1.1.1.2. Matrix assisted laser desorption/ionization

Matrix assisted laser desorption/ionization (MALDI) was developed simultaneously in the late 1980s by two groups; Karas and Hillenkamp[5] and Tanaka *et al*,[6] the latter receiving the Nobel Prize for chemistry in 2002 for the work which has revolutionised the analysis of large biomolecules in excess of 200 KDa.

The sample is deposited on a target and mixed with an excess of a matrix. Typically an organic matrix is mixed with the sample so to absorb energy at the same wavelength as that of the laser, most commonly a pulsed nitrogen laser (λ = 337 nm). The laser beam is directed onto the target energising the matrix molecules which are evaporated as a plume, from the target, carrying with them sample molecules. Gas-phase proton transfer leads to the formation of protonated

molecules and singly charged Na^+ and K^+ adducts are also commonly observed as a result of cationisation.[7] MALDI is a 'soft ionisation' technique which allows the desorption of very large biomolecules into the gas phase intact. A schematic of the MALDI process can be seen in Figure 1.1.1.2.

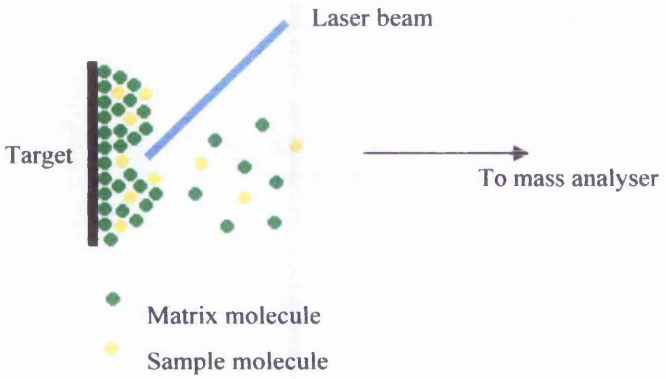


Figure 1.1.1.2. A schematic diagram of the MALDI process. The sample is deposited on a target in an excess of matrix. An incident laser beam excites the matrix molecules causing a plume of matrix and sample molecules to be desorbed from the target and ionised before transfer into the mass analyser.

1.1.1.3. Electrospray Ionization

Electrospray ionisation (ESI) is a 'soft ionisation' technique well suited to the analysis of biochemical molecules as large non-volatile molecules can be analysed directly from the gas phase. ESI was first described by Dole *et al* in 1968,[8] but it was the work of Fenn *et al* and their application of ESI to the analysis of proteins and peptides[9,10] which allowed the full potential of this technique to be realised. Table 1.1.1.1 exhibits a list of examples illustrating the

types of molecules which have been analysed by electrospray ionisation mass spectrometry.

Table 1.1.1.1. Applications of the types of molecules investigated by electrospray ionisation mass spectrometry.

Application Examples	Reference
Proteins	11-15
Oligonucleotides	16-20
Oligosaccharides	21, 22
Drugs/Natural Products	23, 24
Polymers	25
Inorganic metal complexes	26, 27

Initial interest arose from the ability of ESI to form multiply charged analytes of the type $[M+nH]^{n+}$ as well as the potential for multiply charged fragment ions.[28] This enabled the molecular weight of large molecules to be determined using instruments whose mass range was limited to as low as 2000 Th. Today, with its ease of coupling to high performance liquid chromatography (HPLC), it has become the ionisation technique of choice for a range of analytes from small drug molecules to peptides and polymers.

The mechanism of electrospray is discussed in the literature.[29, 30] In conventional ESI the sample solution is passed through a conductive needle at a flow rate of 1-5 μL min^{-1}. An electric field gradient is obtained by applying a potential difference between the needle and a counter electrode. The application of this potential causes a separation of charges in the flowing liquid stream at the needle tip, which leads to the formation of a Taylor cone (Figure 1.1.1.3). The emerging liquid is then dispersed as a fine spray of charged droplets. A flow of drying gas,

typically nitrogen, may also be used to aid in the formation and vaporisation of droplets. The evaporation of sample solvent reduces the droplet size whilst the overall charge remains the same until the Rayleigh limit[31] (equation 1.1.1.1) is approached.

$$q^2 = 8\pi^2 \epsilon_o \gamma D^3 \qquad \textbf{(Equation 1.1.1.1)}$$

q = charge

ϵ_o = permittivity of environment

γ = surface tension

D = diameter of droplet

At the Rayleigh limit the repulsion of charges on the surface of the droplet is equal to the surface tension and "Coulomb explosion" follows resulting in the formation of smaller droplets. There have been two mechanisms proposed for the subsequent formation of gas phase ions from these droplets, these can be seen in Figure 1.1.1.3.

- *Charge residue model*

In the charge residue model proposed by Dole *et al*[8] the repetition of the coulomb explosion cycle proceeds until the formation of very small droplets containing only one ion. Evaporation of the remaining solvent then reveals the naked ion.

- *Ion evaporation model*

Proposed by Iribarne and Thomson[32] this model suggests ion evaporation or emission of ions from the surface of small highly charged droplets. This is the result of surface electric field effects causing the ion to be ejected from the droplet.

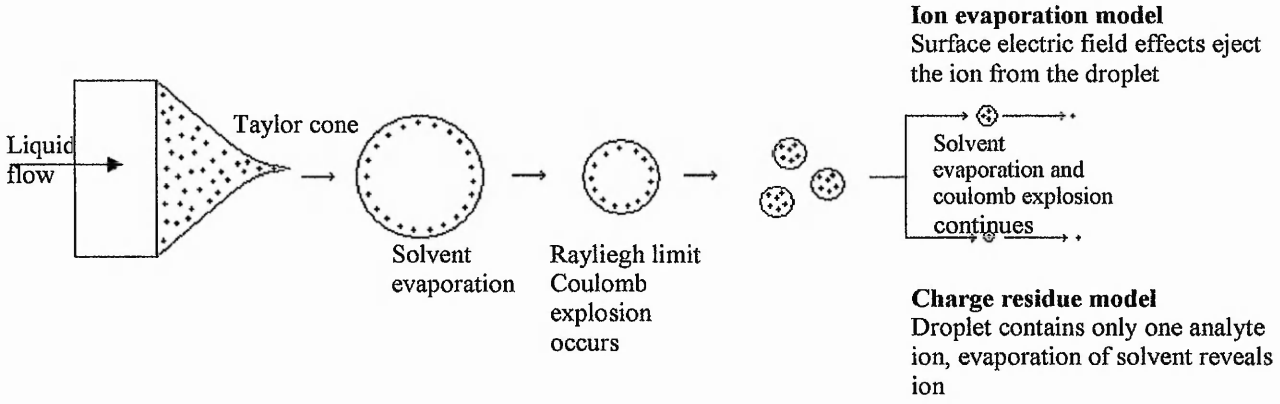


Figure 1.1.1.3. The electrospray process incorporating the two models of ion evaporation and charge residue.

1.1.1.4. Microspray and Nanospray

Microspray and nanospray are ionisation techniques which operate on a similar principle to electrospray, but with smaller spray tip internal diameters and lower, sub μl min^{-1}, flow rates. Microspray typically refers to the source developed by Caprioli,[33] which relied on a pumped liquid flow of 0.3-0.8 μL min^{-1} into the needle, ideal for coupling with liquid chromatography. Wilm and Mann reported a static nanospray source utilising pulled capillary tips with 1-5 μm i.d. and flow rates of 20-40 nl min^{-1} allowing sequencing of proteins at femtomole concentrations.[34, 35]

Reduction of the flow rate and tip internal diameter eliminate the requirement of a nebulising sheath gas. Mass spectrometers are mass flow dependent detectors, so reduction in the flow would reduce the sensitivity. However ESI/MS behaves as a concentration sensitive detector. Analyte desorption efficiency is increased and greater transmission of the analyte from the spray tip into the mass spectrometer is observed as the flow rate decreases leading to an increase in sensitivity.[36, 37]

1.1.2. Mass Analysers

Once the sample has been ionised it is necessary to sort the ions according to their mass-to-charge ratio. This function is performed by the mass analyser of which there are several different types. Those commonly used in peptide analysis are discussed below, particularly with regard to the three main characteristics associated with them (upper mass limit, transmission efficiency and mass resolution).

1.1.2.1. Quadrupole mass analysers

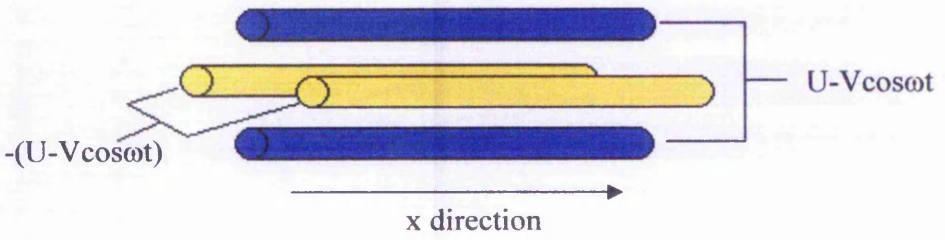
Quadrupole mass analysers exist in two common configurations, the linear quadrupole analyser and the quadrupole ion trap.

1.1.2.2. Linear quadrupole analysers

Linear quadrupole mass analysers are relatively simple, compact and easy to operate devices. The analyser uses the stability of trajectories in an oscillating

electric field to separate ions according to their m/z ratio. A quadrupole analyser consists of four circular and perfectly parallel rods. An electric field is created within the region by applying a positive dc potential (U) and a superimposed rf potential ($V\cos\omega t$), to one pair of parallel rods, the other pair of rods carry a dc potential of –U and rf potential ($V\cos\omega t$) which is out of phase with the first pair of rods.[6, 37-40] A linear quadrupole analyser is shown in Figure 1.1.2.1.

By adjusting the dc and rf voltages applied to the rods it is possible to create stable trajectories for ions of different m/z ratio. At a given set of parameters ions of a very narrow m/z range will have a stable trajectory between the rods and will reach the detector. The remainder of ions will have unstable trajectories, be rejected from the quadrupole and not be detected. A mass spectrum is acquired by ramping the applied fields forcing other ions to follow the stable pathway with increasing m/z value. The upper mass limit of this analyser is limited to a few thousand m/z due to the limits of the rf amplitude which may be applied.



ω = angular frequency (related to frequency, f by $\omega = 2\pi f$)

t = time

Figure 1.1.2.1. A quadrupole mass analyser schematic with applied potentials to create an electric field between the rods.

The mass resolution of a linear quadrupole depends on the number of rf cycles. By applying an rf voltage only, a quadrupole may be used as a focusing device to transmit all m/z ions. Structural information can be obtained with quadrupole mass spectrometers by employing three in series, this is known as a triple quadrupole device and is shown in Figure 1.1.2.2.[41] This configuration allows tandem mass spectrometry to be performed. This is discussed in Section 1.1.3.

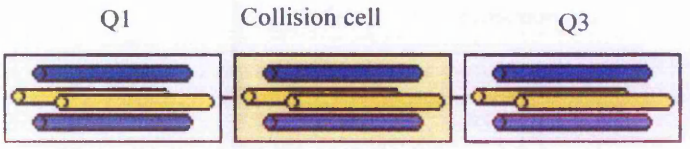


Figure 1.1.2.2. A schematic diagram of a triple quadrupole device.

1.1.2.3. Quadrupole Ion trap

The quadrupole ion trap (QIT) is a device which can function as both an ion store, where gaseous ions may be confined, and as a mass analyser.[4, 42, 43] This permits the study of gas phase ion chemistry and ion structure elucidation by tandem mass spectrometry (discussed in Section 1.1.3). The mode of operation of a quadrupole ion trap is similar to that of a quadrupole mass filter. The trap itself consists of three electrodes, a ring electrode and two end-cap electrodes, which have a hyperbolic cross-section (see Figure 1.1.2.4).

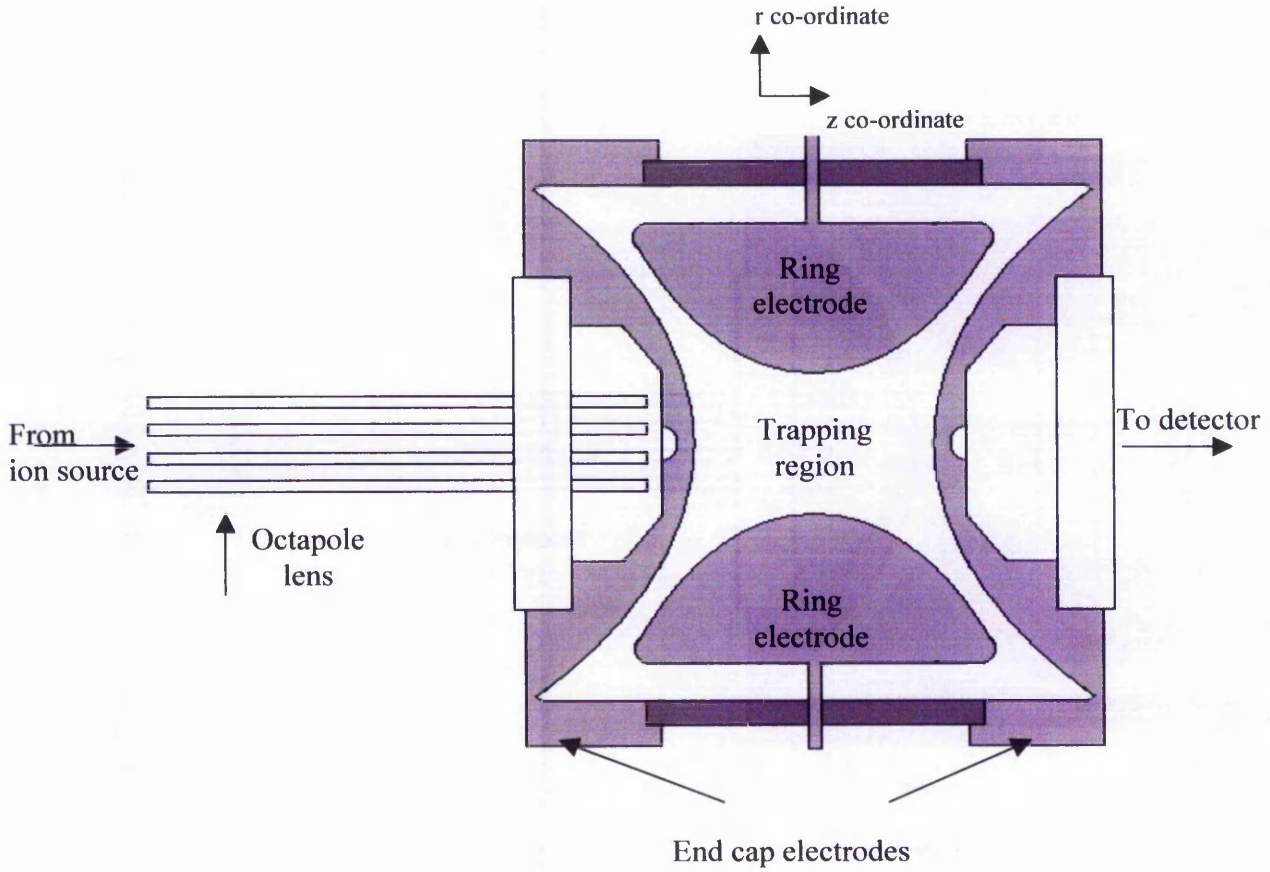


Figure 1.1.2.4. A schematic representation of the quadrupole ion trap.

There are small openings in the end-caps to allow ions in and out of the trap which is typically held at 10^{-3} Torr helium, this acts to reduce the expanding trajectories of the ions caused by ion repulsion and motion within the trap. The QIT is a dynamic mass analyser in which ion trajectories are influenced by time dependent forces.

A quadrupolar field is created by the application of a dc voltage and rf voltage to the electrodes, in practice this is applied to the ring electrode only. The motion of ions within the quadrupolar field can be described mathematically by the Mathieu equation and a stability diagram for ions in a pure quadrupole field can be constructed, as in Figure 1.1.2.5.[44] This assumes that the relationship between the distance from the centre of the trap to the end cap (z) and the radius of the ring electrode (r) is $r_o^2 = 2z_o^2$. The axes of the stability diagram, a and q, are defined by the following relationships.

$$q_z = \frac{4eV}{mr^2{}_o\Omega^2} \qquad a_z = \frac{-8eU}{mr^2{}_o\Omega^2}$$ **(Equations 1.1.2.1 & 1.1.2.2)**

V = amplitude of applied rf voltage

U = applied dc voltage

r_o = radius of ring electrode

Ω = angular frequency of the rf potential applied to the ring electrode ($= 2\pi f$ where f is the frequency in Hz).

Ions within the trap have their own secular frequency (ω_o) given by $\omega_o = \beta\Omega/2$ where β_r and β_z are determined by the working point of the ion or β value at a particular a and q. The working point of an ion is defined by the a_z and q_z values and can be altered by changing the amplitude of the applied dc and rf voltages. The a_z and q_z values are chosen to ensure stability or instability of an ion of interest. Ions are stable when their trajectories remain within the dimensions of the trap z_o and r_o.

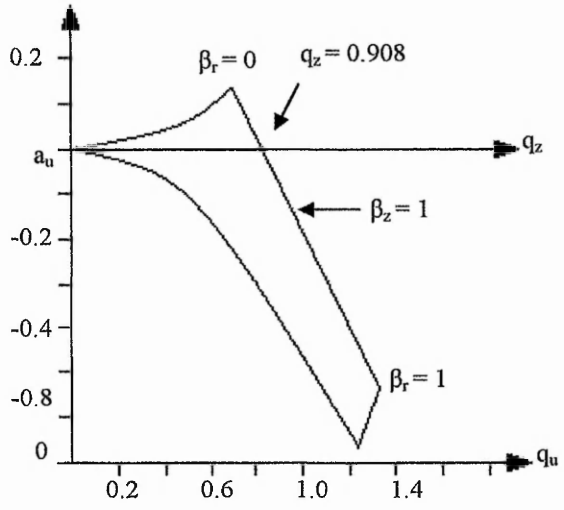


Figure 1.1.2.5. A stability diagram for a quadrupole ion trap. Ions within the boundary are stable in the trapping field.

Generally the trap is operated with the rf voltage applied only to the ring electrode (i.e. $a_z = 0$). Under these conditions the working point of an ion is described by the q_z parameter as illustrated in Figure 1.1.2.6. As q_z for each ion species reaches a value of 0.908 the ions are ejected axially through the end cap electrodes in order of increasing m/z value. This is known as mass selective axial

instability and occurs at the boundary $\beta_z = 1$ of the stability diagram where $\omega_o = \Omega/2$.

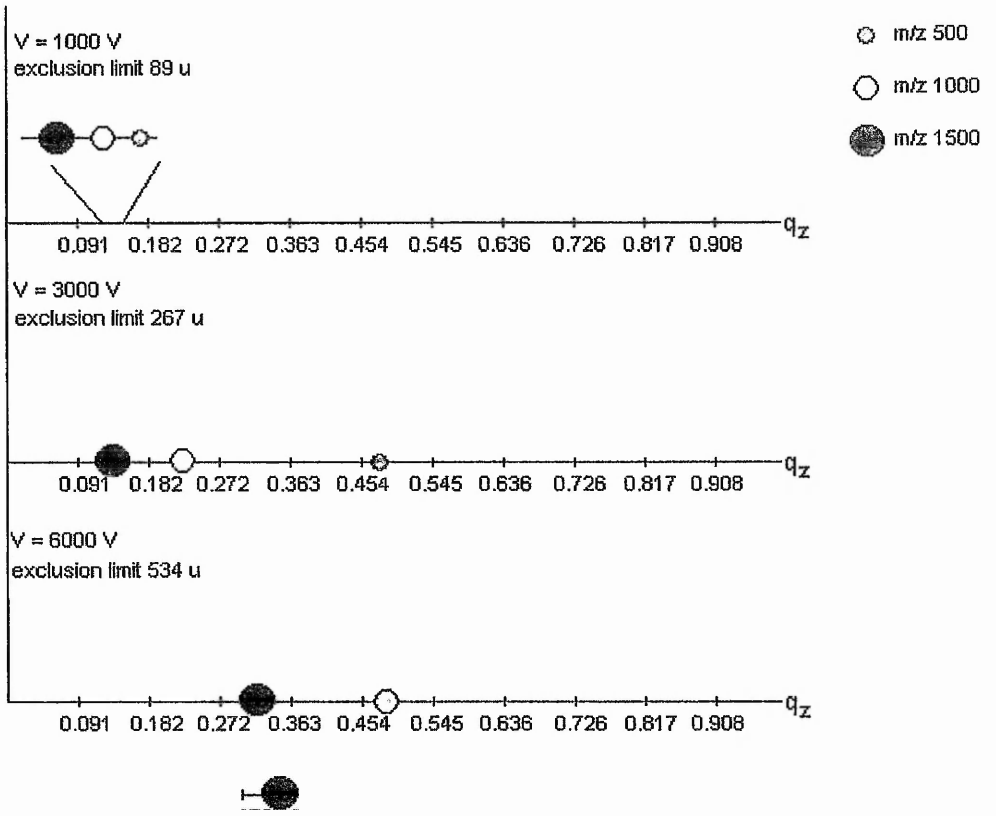


Figure 1.1.2.6. A simplified stability diagram for three ions of different mass to charge ratio showing the effect of increasing the applied rf voltage.

1.1.2.4. Time-of-flight

A time-of-flight (ToF) mass analyser operates on the principle of the separation of ions on the basis of their difference in velocity as they travel down an evacuated flight tube.[7, 40, 45] After acceleration to a constant kinetic energy (see equation 1.1.2.3) by an electric field, ions travel at velocities that have an inverse relationship with their m/z value (see equation 1.1.2.4) with lighter ions

travelling faster than heavier ions. The time-of-flight of an ion down a flight tube of length d, is t = d/V where V is the potential by which the ions are accelerated. Wiley and McLaren published the design for the linear ToF which became the first commercial instrument of its type.[46]

$$zeV = \tfrac{1}{2} mv^2 \qquad \textbf{(Equation 1.1.2.3)}$$

$$v = (2zeV/m)^{\frac{1}{2}} \qquad \textbf{(Equation 1.1.2.4)}$$

ToF analysers have enjoyed a renewed popularity since the development of MALDI and progress in electronic technology to allow the handling of the data flow.[47]

A linear ToF consists of an acceleration region after the ionisation source and a flight tube culminating in a detector. All ions must enter the flight tube simultaneously and so ToF is well suited to pulsed ionisation techniques such as MALDI. The mass resolution of these linear ToF instruments, however, is poor due to factors that create a distribution in flight times among ions with the same m/z. These include length of the ion formation pulse, the size of the volume where the ions are formed and the variation of the initial kinetic energy of the ions prior to the acceleration phase, which result in a distribution of flight times for ions of the same m/z. One method of improving resolution involves a delay of up to a few microseconds between ionisation and acceleration to allow

dissipation of the excess kinetic energy, this is known as delayed pulsed extraction.

Reflectron ToF is another way to improve resolution by correcting for any initial position and energy dispersions. The ions are accelerated through an initial field free region before being reflected by an electrostatic mirror, the ions are slowed and brought to rest, the direction of motion is then reversed and the ions are accelerated through another field free region to the detector. Ions with greater initial kinetic energy penetrate further into the reflectron field and so spend a longer time in the reflectron lens. Mass resolution is consequently improved. A schematic diagram of a reflectron ToF analyser, which can be operated in linear or reflectron mode, is shown in Figure 1.1.2.3.

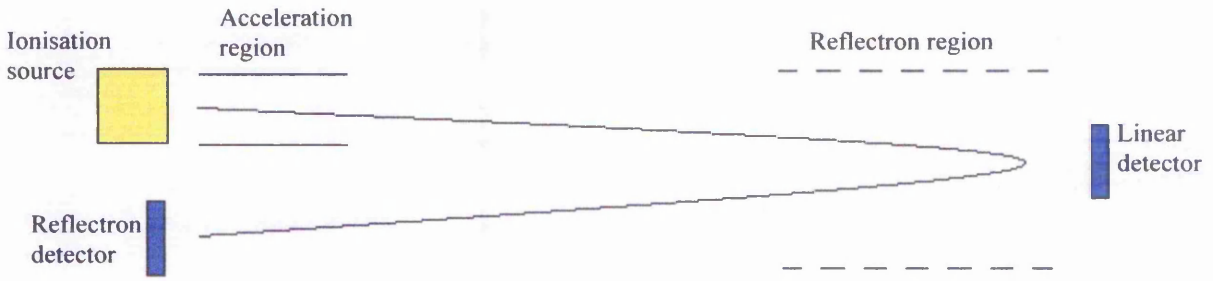


Figure 1.1.2.3. A Schematic of a reflectron ToF. Ions are produced as a pulse in the source and accelerated by an electric field applied in the acceleration region. The ions continue down the flight tube, which is a field free region, where ions are separated according to their m/z. In linear mode the ions are then detected, in reflectron mode they are reflected back down another field free region for detection.

Orthogonal acceleration is a modified ToF configuration for continuous mode ionisation techniques. It consists of an additional field free sampling region in which ions are stored after ionisation before a short pulse of an orthogonal accelerating field is applied to eject those ions into the flight tube.

1.1.2.5. Post source decay

Structural information can be obtained with ToF instruments using post source decay (PSD). PSD is a metastable decay which occurs within the flight tube. A precursor ion will fragment and its daughter ions have the same velocity but different kinetic energy to the parent. In linear mode the time-of–flight of both parent and daughter ions will be the same, however, in reflectron mode this is not the case. So on comparison of spectra, ions that appear under reflectron mode but not linear mode are the result of PSD.

An alternative approach to elucidating structure of an unknown molecule by ToF mass spectrometry is to add a collision cell between tandem ToF analysers as described in Section 1.1.3.

1.1.3. Tandem mass spectrometry

Tandem mass spectrometry or MS/MS is a method which involves multiple stages of mass analysis separated by ion dissociation to yield structural information. This process is shown schematically in Figure 1.1.3.1. The most commonly used technique for ion dissociation is collisionally induced

dissociation (CID). This is a powerful technique for the determination of ion structures of isolated ions. CID is a two stage process in which the precursor ion (M^+) is first excited by collision with a neutral atom or molecule (N) (equation 1.1.3.1), it then dissociates as a result of gaining excessive internal energy following the collision (equation 1.1.3.2).

$$M^+ + N \longrightarrow M^* \quad \textbf{(Equation 1.1.3.1)}$$

$$M^* \longrightarrow F_1 + F_2 \quad \textbf{(Equation 1.1.3.2)}$$

There are two ways that CID can be carried out, either by coupling multiple analysers together, as with the triple quadrupole mass spectrometer (Figure 1.1.2.2), or by trapping ions in an ion storage device, such as a quadrupole ion trap. In the former case, the central quadrupole is filled with gas and acts as a collision cell. The first quad is used to select specific masses which are passed into the collision cell to undergo CID and the final separates the fragments by m/z prior to detection. In the latter case, a selected ion is isolated in the trap by applying an auxiliary rf voltage in resonance with the secular frequency of all ions except the one to be isolated. This causes the ions to become excited and develop an unstable ion trajectory causing them to be ejected from the trap. The selected ion remains in the trap and can be subjected to CID by applying a different auxiliary rf voltage in resonance with its secular frequency such that excitation is achieved but not ejection. On application of this voltage the excited ion moves away from the centre of the trap, colliding with the helium buffer gas and thus causing dissociation.

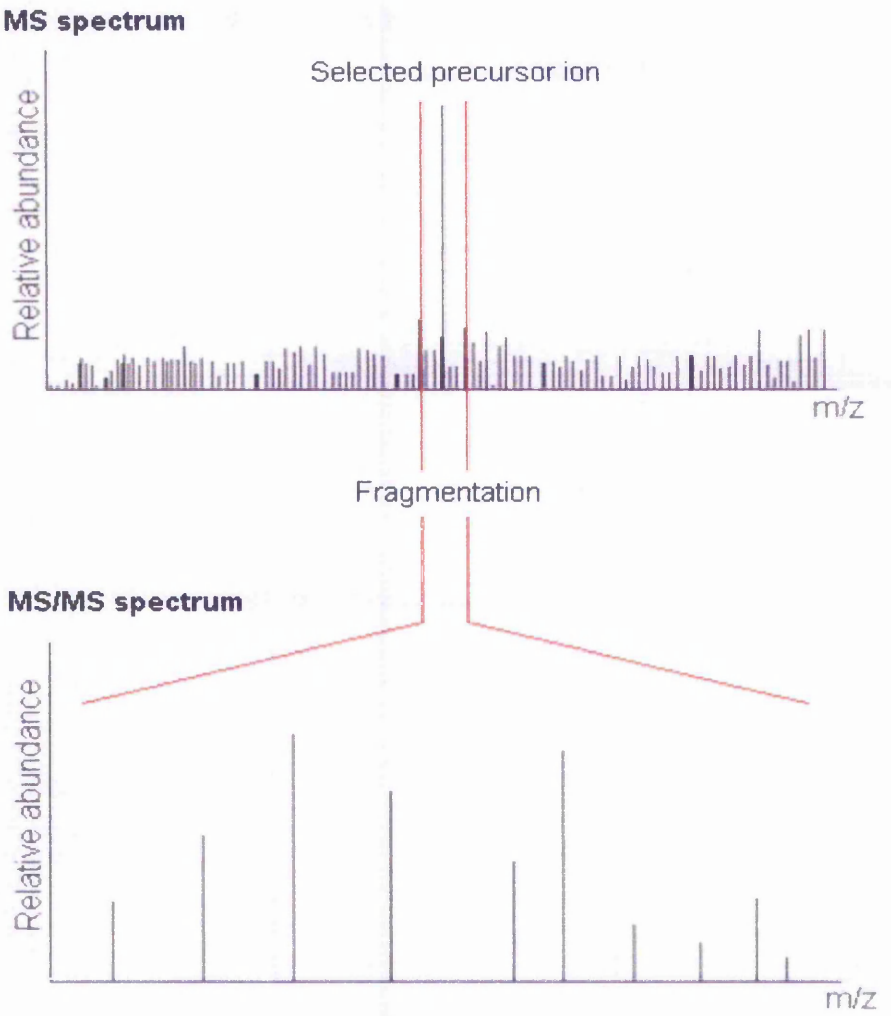


Figure 1.1.3.1. A schematic of the MS/MS process. The precursor ion is selected and isolated, it is then subjected to dissociation to produce fragment ions which can be used to garner structural information.

1.1.3.1. Tandem mass spectrometry of peptides

Peptides are formed by the combination of two or more amino acids. Peptide ions, when subjected to CID, fragment predominantly at the amide bond along the peptide backbone generating a ladder sequence of ions as shown in Figure 1.1.3.2.

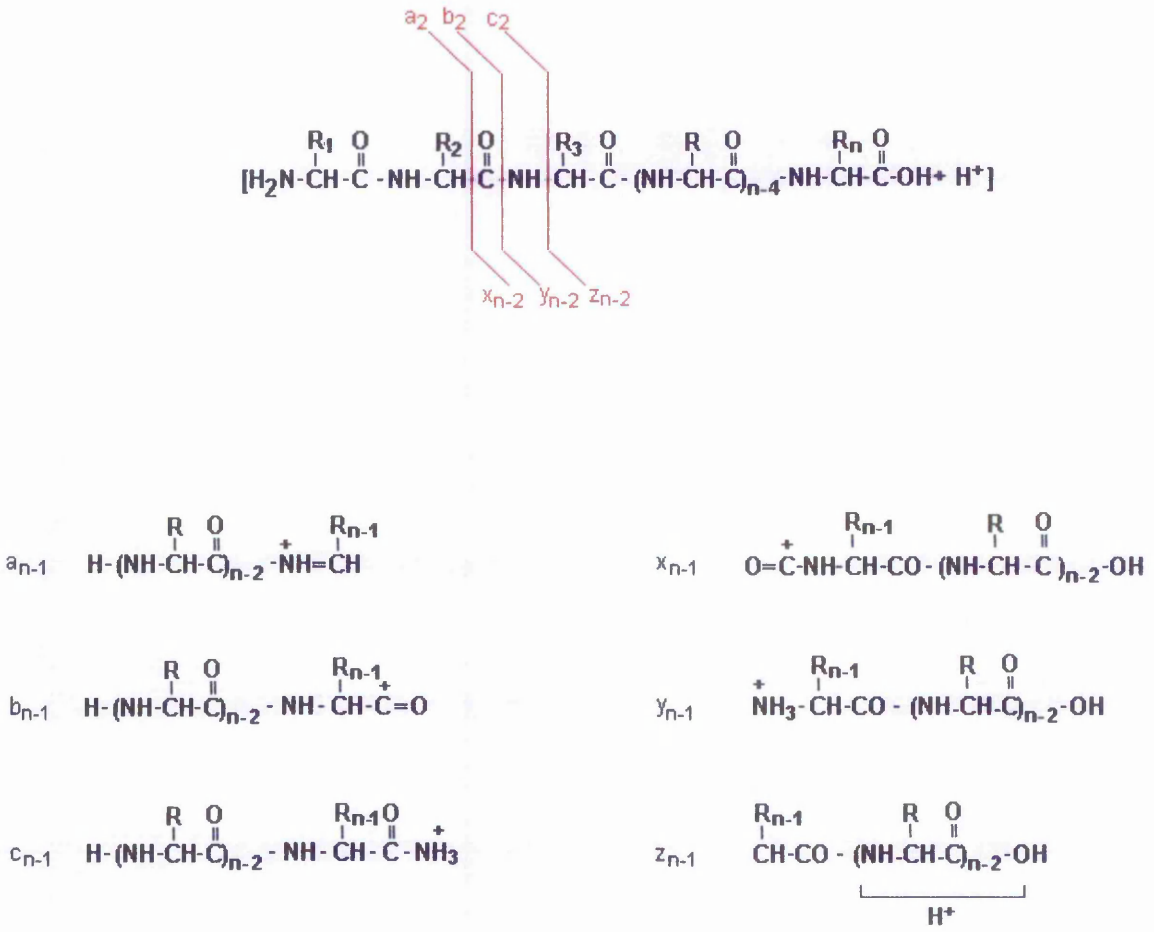


Figure 1.1.3.2. A peptide fragment ion series produced when ions undergo collisionally induced dissociation. Cleavage of the alkyl-carbonyl (CHR-CO), the peptide-amide (CO-NH), and amino-alkyl (NH-CHR) bonds are shown. Both the N- and C-terminus fragments can retain the charge resulting in 6 possible sequence specific fragment ions. The nomenclature used is shown in this Figure.

The b and y ions are most frequently observed, with loss of H_2O and NH_3 from b and y ions also commonly found. A series of either type of fragment ions allows the amino acid sequence to be determined by differences in masses of adjacent sequence ions.[48-50]

During ionisation of peptides protonation will occur at all the strongly basic sites in the molecule, the N-terminal amine and the side group of any of the basic amino acid residues (e.g. lysine, arginine, histidine). A proton associated with the more basic sites is strongly attached and remains fixed during CID whereas a proton on the less basic N-terminus is free to move along the length of the peptide. A population of peptide ions with the mobile proton associated at different points along the sequence results. The protonation site then directs the fragmentation reaction producing a series of product ions which are collectively used to determine the peptide structure. Examples of this 'mobile proton' are shown in Figures 1.1.3.3 and 1.1.3.4.

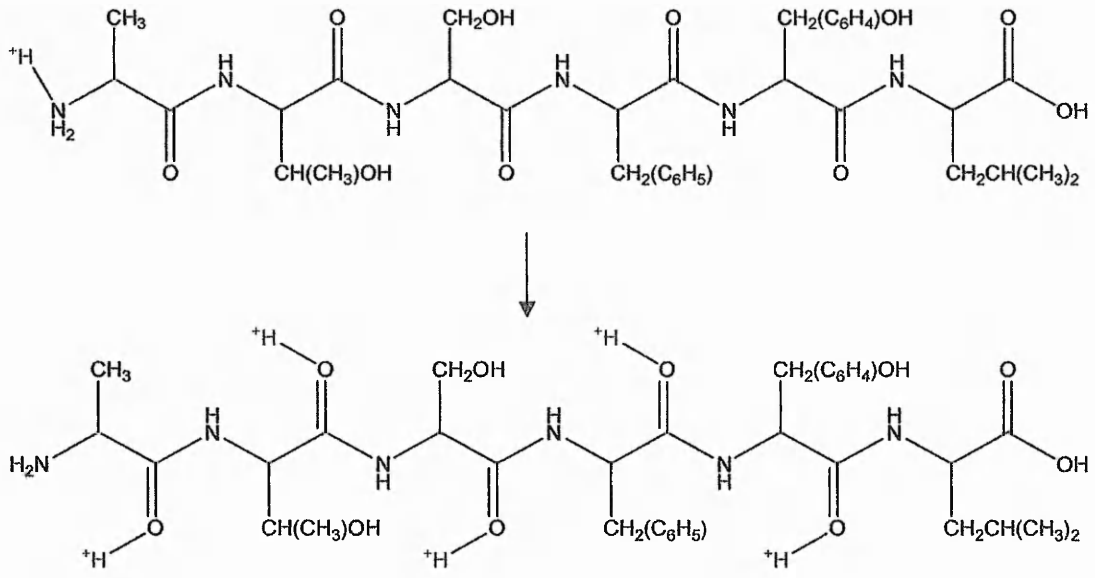


Figure 1.1.3.3. The singly protonated peptide ATSFYL. The single basic site, the N-terminus, is protonated. The proton (shown in red) is then able to migrate along the peptide producing a pool of peptides with differing sites of protonation. (Adapted from reference 45).

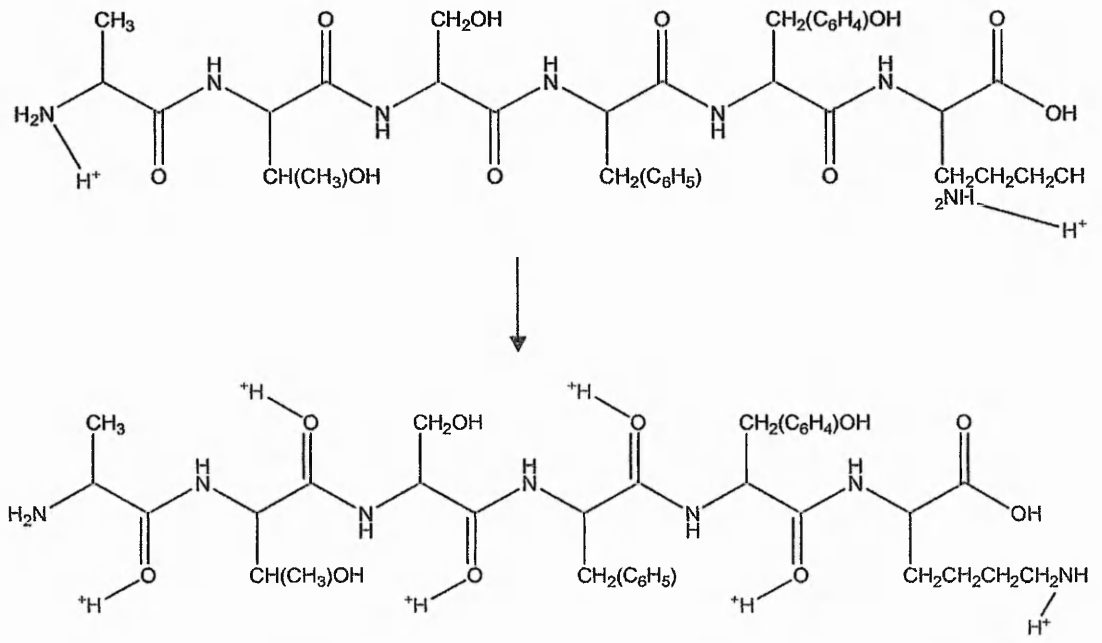
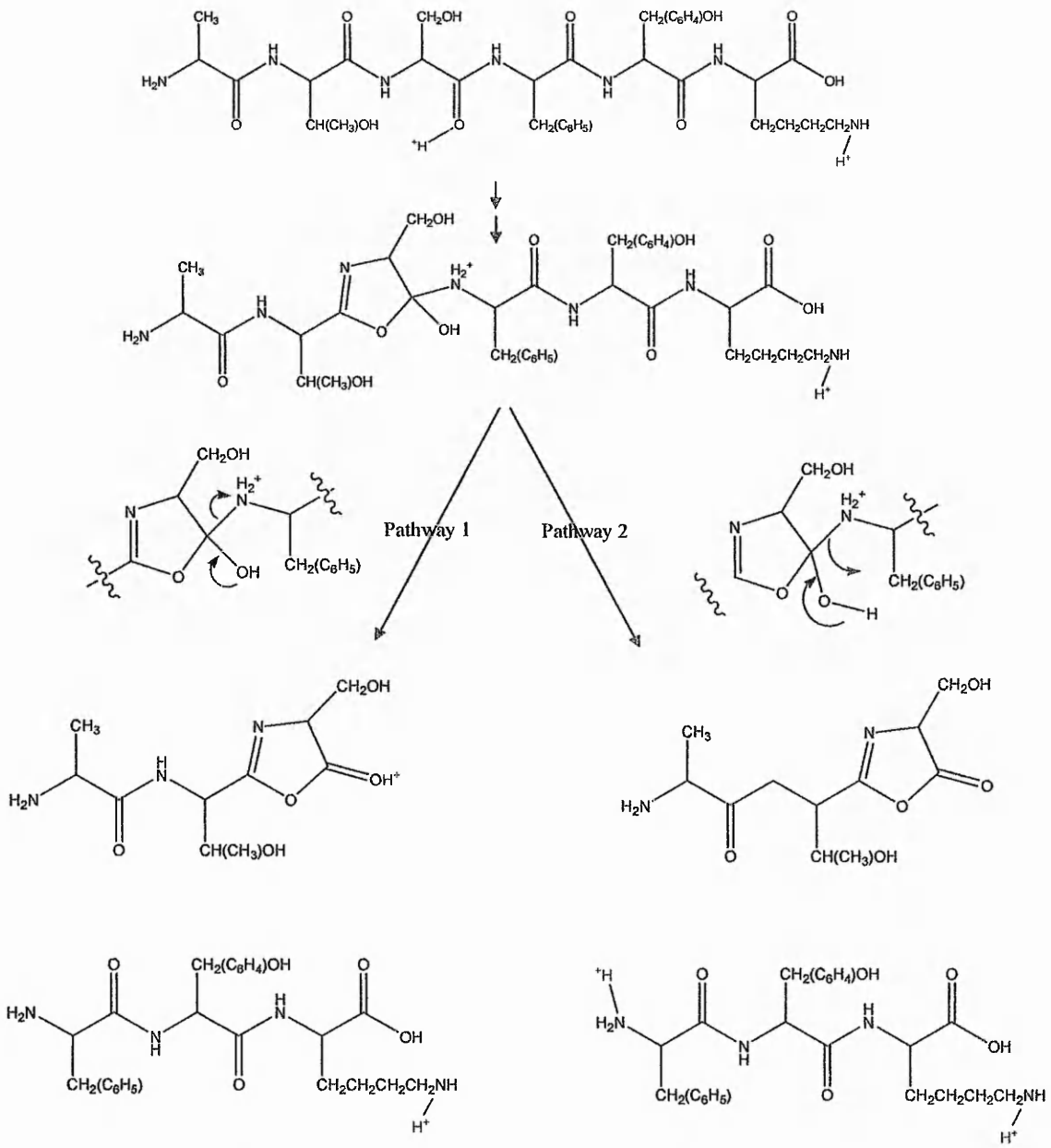


Figure 1.1.3.4. The doubly protonated peptide ATSFYK is protonated at the N-terminus and the basic lysine (K) residue. The proton associated with the lysine residue remains fixed whilst the N-terminal proton is the mobile proton and is able to migrate along the peptide directing fragmentation. (Adapted from reference 45).

The pathway of collisionally activated dissociation for a doubly charged peptide ion is shown in Figure 1.1.3.5. The fragmentation reaction occurs to release energy gained by the peptide ion during its collision. The protonation site initiates a cyclisation reaction leading to two possible pathways. Pathway 1 is favoured and results in the formation of a complementary pair of b and y ions. Pathway 2 results in the formation of a doubly charged y ion.



Singly charged y_3 ion (m/z 457)

Doubly charged y_3 ion (m/z 229)

Figure 1.1.3.5. Collisionally induced fragmentation pathway of a doubly charged peptide ion (ATSFYK). Fragmentation occurs via a cyclisation reaction. Pathway 1 is favoured, pathway 2 results in a doubly charged product ion. (Adapted from reference 45).

1.1.4. Chromatography

Chromatography is a separation method in which analyte components are selectively distributed between two immiscible phases, one stationary and one mobile. The stationary phase may be packed in a column or spread as a layer or film. The mobile phase may be gaseous or liquid.[51] Separation occurs as a result of differences in an individual analytes affinity for the stationary phase and the mobile phase.

1.1.4.1. Reverse phase high performance liquid chromatography

Reverse phase high performance liquid chromatography (RP-HPLC) is a liquid chromatography technique in which a polar mobile phase is mechanically pumped through a column which contains a non-polar stationary phase. One of the first of these systems was constructed by Horvath *et al* in the late 1960's.[52] The most common stationary phase used in RP-HPLC consists of octadecyl siloxane (ODS) groups covalently attached to spherical silica particles with diameters in the range 3-10 μm. The mobile phase is usually a mixture of aqueous and organic solvents (e.g. methanol, acetonitrile) and separation is a result of hydrophobic interactions of the analyte. Hydrophobic compounds are strongly retained by the stationary phase and elute later requiring greater organic mobile phase content than hydrophilic compounds. The mobile phase can be pumped through the column in two modes; isocratic elution, which refers to a flow of unchanging mobile phase composition or gradient elution, which involves continuous change in the composition of the mobile phase. Gradient elution allows separation of sample components of widely varying affinity for

the stationary phase. There are three major properties that determine the resolution of two peaks in a chromatographic separation. These are capacity factor, selectivity and efficiency.

Capacity factor(k_R')

This is a measure of the retention of an analyte as a result of the strength of interaction between the analyte and the stationary phase and is given by the expression in equation 1.1.4.1.

$$k_R' = (t_R - t_o) / t_o \qquad \textbf{(Equation 1.1.4.1)}$$

t_R = retention time of the analyte

t_o = time taken for a non-retained analyte to pass through the system

Selectivity (α)

This is a measure of the difference in retention times between two peaks and is given by the equation 1.1.4.2.

$$\alpha = (t_{RB} - t_o) / (t_{RA} - t_o) \qquad \textbf{(Equation 1.1.4.2)}$$

t_{RA} = retention time of first analyte

t_{RB} = retention time of second analyte

Efficiency

This factor is obtained from the number of theoretical plates and is given by equation 1.1.4.3.

$$N = 16 (t_R/W)^2 \qquad \textbf{(Equation 1.1.4.3)}$$

W = peak width at baseline.

Resolution (R)

The resolution parameter concerns the degree of separation between two peaks and is given by equation 1.1.4.4.

$$R = 2(t_{RB} - t_{RA}) / (W_B + W_A) \qquad \textbf{(Equation 1.1.4.4)}$$

Complete resolution of two peaks is achieved when $R \geq 1.5$ and for accurate quantification of two peaks a resolution of greater than 0.8 is required.

The relationship between resolution and capacity factor, selectivity and efficiency is given by equation 1.4.5.

$$R = N/4 (\alpha - 1 / \alpha) (k_R' / 1 + k_R') \qquad \textbf{(Equation 1.1.4.5)}$$

1.1.4.2. Miniaturisation

Conventional HPLC is performed on columns with internal diameters of 2.0-4.6 mm. Microcolumn HPLC uses columns with internal diameters <1.0 mm and has some advantages over conventional HPLC.[53] One such advantage is an improvement in sensitivity when coupled to a concentration sensitive detector, such as ESI/MS, due to reduced chromatographic dilution.[53-55] McLuckey *et al* first recognised the benefits of miniaturisation when coupling a microbore LC with an ion trap mass spectrometer.[37]

Chromatographic dilution (D) occurs because a compound will be subjected to dilution, within the mobile phase, during the chromatographic process and is given by equation 1.1.4.6. [56]

$$\frac{D = c_o}{c_{max}} = \frac{\epsilon \pi r^2 (1+k) \sqrt{2\pi LH}}{V_{inj}} \qquad \textbf{(Equation 1.1.4.6)}$$

c_o = original compound comcentration

c_{max} = final compound concentration at peak maximum

r = column radius

L = column length

ϵ = column porosity

V_{inj} = injected sample volume

H = column plate height is given by H=L/N, where $N=(t_r/W_b)^{0.5}$

Chromatographic dilution increases proportionally with the square of the column radius and to the square root of the length and plate height of the column. It is possible to calculate that a reduction in column diameter from 4.6mm to 75μm will result, theoretically, in approximately a 5000 fold relative concentration increase at the detector, if the peak widths are the same for both columns, as shown in Table 1.1.4.1.[57]

Table 1.1.4.1. Effect of column diameter on relative capacity and concentration at the detector.

column	i.d. /mm	Typical flow rate/μL min^{-1}	Relative concentration at the detector	Relative loading capacity
Conventional	4.6	~1000	1	5000
Microbore	1.0	~50	20	400
Packed capillary	0.1-0.32	~5.0	200	40
Nanoscale	<0.1	~0.2	5000	1

One important consideration of miniaturisation is that extra-column band broadening processes must be reduced to prevent considerable loss of efficiency which would result in decreased resolution of the column. Connecting tubing in the system is also particularly important. The three most commonly used tubing materials are stainless steel, polyetheretherketone (PEEK) or fused silica. The presence of tubing in the system configuration contributes to extra column band broadening linearly with length and to the power of four of the radius, so for 500μm – 1.0 mm i.d. columns, tubing i.d. should be 0.25 mm to 0.125 mm. For smaller diameter columns the connecting tubing should be avoided, where possible, with connections made directly. When this cannot be done tubing

should be ≤ 25 μm i.d. and only short lengths should be used with connections made carefully to ensure a zero dead volume. One of the major drawbacks of reducing the column diameter, and hence flow rate through the system, is the corresponding reduction in the injection volume which can be applied to the column in a reasonable amount of time. This can be overcome, however, by preconcentration of analytes at the head of the column.

1.1.4.3. Immobilised metal ion affinity chromatography

Immobilised metal ion affinity chromatography (IMAC) was first reported by Porath in 1975 under the name of metal chelate chromatography.[58] IMAC has since become a powerful separation tool employed in the purification of proteins and peptides.[58-62] In IMAC a chelating ligand is covalently bound to a solid chromatographic support, metal ions are then complexed with the ligand and immobilised on the support. The metal ion acts as a lewis acid and can bind to electron donor groups on the surface of a peptide or protein allowing separation from other sample components. The commonly used complexation metals include the transition metals (e.g. Cu (II) and Ni (II)) to which the electron donating atoms may be nitrogen, oxygen or sulphur. Hard lewis metal ions (such as Al (III)) have also been used in IMAC due to their high affinity for binding with oxygen rich groups, such as phosphate groups.[63]

The Cu (II) metal ion is commonly used in the isolation of proteins due to its high affinity for histidine residues (Histidine pKa ~6.5).[64] Histidine residues occur rather infrequently within protein structure, in terms of natural abundance,

appearing on average as 2 % of the molecule. Only half of such histidine residues will be surface exposed, i.e. available for binding to the Cu (II) metal ion. Ni, Zn, Co, Ca and Mg may also be used to isolate histidine residues, however such metal ions differ in their ability to bind. Although Ni may be used for the isolation of peptides, it exhibits more stringent binding requirements such as the need for two residues to be close to one another.[65, 66]

The analyst can take advantage of the high affinity between Cu (II) ions and histidine residues by employing chelating agents, or 'tags', to enhance the association between metal ion and peptide or protein. Such chelating agents broaden the range of target molecules that can be isolated via this approach. Iminodiacetic acid (IDA), a tridentate chelator, is the most commonly used chelating agent but others may also be employed.[67] Nitrilotriacetic acid (NTA) is an alternative tetradentate chelator developed by Hochuli.[68] A tetradentate chelator may provide enhanced chelating strength, however can exhibit weaker protein retention power. For example when Cu^{2+} is chelated with IDA there is only one coordination site free for protein binding, whereas when Ni^{2+} is chelated with IDA three coordination sites are available but binding with the target is not as strong and leakage can occur. The relative stability of complexes formed with an IDA derivative of cellulose and a divalent metal ion are Cu(II)>Ni(II)>Zn(II)≥Co(II)>>Ca(II), Mg(II).[69,70] Alternatively when Ni^{2+} is chelated with NTA two sites are free for binding, such an approach has been shown to be selective for neighbouring histidine residues on peptides and proteins.[67]

Hochuli *et al* demonstrated the use of IMAC in the purification of recombinant proteins with engineered histidine affinity tags in conjunction with a Ni (II)-nitrotriaceticacid (Ni-NTA) stationary phase which selectively bound adjacent histidine residues, these are relatively uncommon in nature.[68, 72] Such tags have become widely popular recently.[73-76]

In a peptide/protein molecule of interest if only one histidine residue is present the most appropriate metal ion for best retention is Cu (II). If more than one residue is present in the molecule then Ni (II) may also be used.[77] Table 1.1.4.4 highlights the appropriate metal ions required for the isolation of target molecules relative to the component number of histidine residues.

Table 1.1.4.4. Illustrates the most appropriate metal ion for use relative to the number of histidine residues present on the protein surface.[77]

Number of His residues	Metal ions providing adsorption
No His	-
One His	Cu^{2+}
>One His	Cu^{2+} (stronger ads), Ni^{2+}
Cluster of His	Cu^{2+}, Ni^{2+}, Zn^{2+},Co^{2+}

Ni-NTA IMAC has been successfully employed in the isolation of a range of proteins using an affinity tail, typically consisting of multiple histidine residues. Phenacrylic acid decarboxylase, GroES and rat finger protein are examples of sample molecules isolated using a hexahistidine tag on the N- or C- terminus of the protein.[78-80] The penicillin binding protein has also been isolated using a heptapeptide containing four histidine residues.[81]

Figure 1.1.4.1 is a reaction scheme highlighting the binding process for a histidine residue employing an IDA chelating agent.

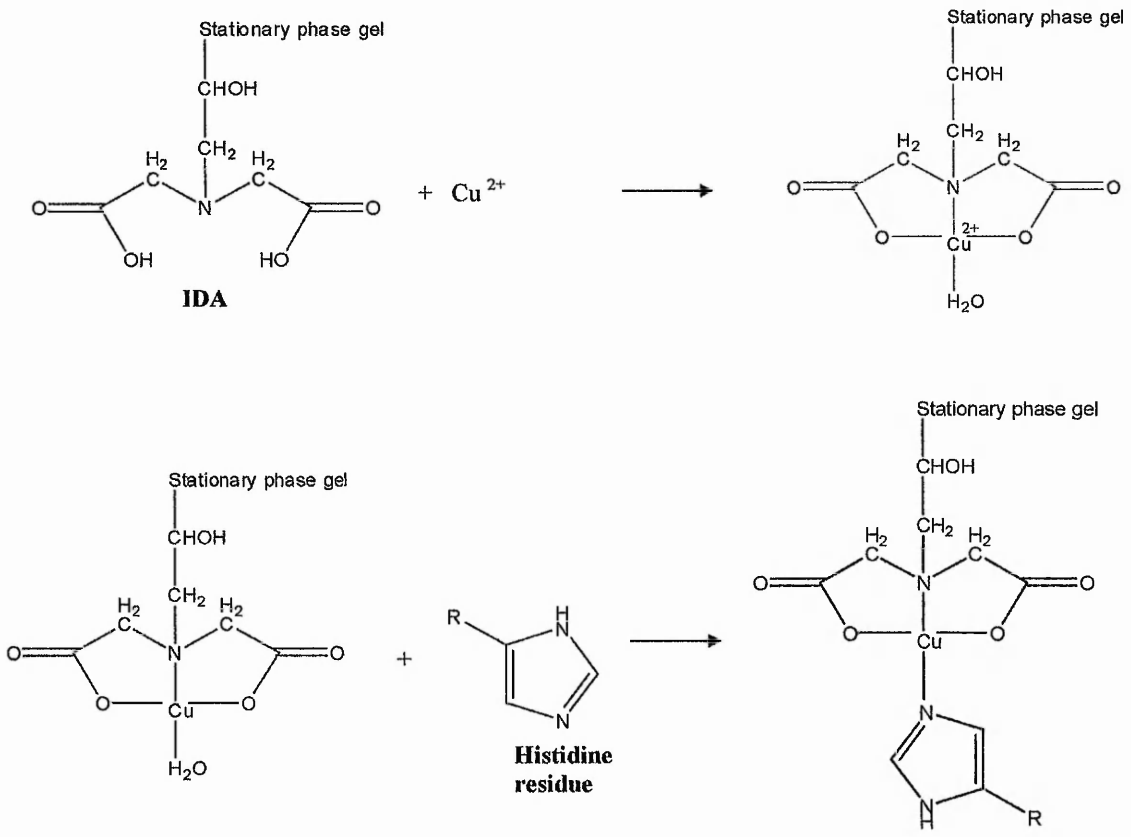


Figure 1.1.4.1. A reaction scheme of the complexation of Cu(II) with IDA and the binding of the histidine residue to the immobilised metal ion.

In Figure 1.1.4.1 the metal/histidine binding occurs at neutral to basic pH (6-8), when the histidine residues are in the non-protonated form under more alkaline conditions coordination with amino functional groups can also occur thus decreasing selectivity.[82] Following binding, the release of the analyte molecule can be achieved by lowering the pH or by ligand exchange using, for example, imidazole which exhibits the same side chain structure as that in histidine so

competes for the binding site. The imidazole has a greater affinity for the copper therefore resulting in displacement of the histidine residues.

IMAC has high ligand stability and uses mild elution conditions, experiencing complete ligand recovery following column regeneration. This technique is also a low cost approach to peptide/protein isolation.[77]

1.1.4.4. Ion exchange chromatography

Ion exchange chromatography is a form of adsorption chromatography in which one ionic species is exchanged for another.[85,86] An ion exchanger consists of a matrix to which a stationary phase can be bound for presentation to a mobile phase carrying the analyte molecules of interest. Basic (anionic) exchangers present positively charged groups whilst in acidic (cationic) exchangers the groups are negatively charged. Stationary phases are classified as weak or strong with reference to the pk_a of their charged groups similar to that of acids and bases, however, this classification is not representative of the binding strength of analyte molecule to stationary phase functional groups. The stationary phase of an ion exchanger exhibits an overall positive or negative charge providing sites for ion exchange with counter ions of opposite charge. The functional groups of an anion exchange stationary phase are primarily amines, quarternary amines are employed when a stronger phase is required. For cation exchangers carboxylic acid groups are normally used for the interaction whilst sulfonate groups are presented in stronger phases. All cation exchangers have a limiting pH below

which they are not functional. Some examples of stationary phase functional groups can be seen in Table 1.1.4.2.

Table 1.1.4.2. Examples of cation and anion exchange functional groups and their structures.[84]

	Designation	pk	Structure
Anion			
Diethylaminoethyl	DEAE	9.-9.5	$-OCH_2N^+H(C_2H_5)_2$
Trimethylaminoethyl	TMAE		$-OCH_2CH_2N^+(CH_3)_3$
Dimethylaminoethyl	DMAE	ca. 10	$-OCH_2\ CH_2N^+H(CH_3)_2$
Trimethylhydroxypropyl	QA		$-OCH_2CH(OH)N^+H(C_2H_5)_2$
Quaternary aminoethyl	QAE		$-OCH_2N^+(C_2H_5)_2CH_2CH(OH)CH_3$
Quaternary Amine	Q		$-OCH_2N^+(CH_3)_3$
Triethylamine	TEAE	9.5	$-OCH_2N^+(C_2H_5)_3$
Cation			
Methacrylate		6.5	$CH_2=CH(CH_3)COOH$
Carboxymethyl	CM	3.5-4.	$-OCH_2COOH$
Orthophosphate	P	3, 6	$-OPO_3H_2$
Sulfoxyethyl	SE	2	$-OCH_2CH_2SO_3H$
Sulfopropyl	SP	2-2.5	$-OCH_2CH_2CH_2SO_3H$
Sulfonate	S	2	$-OCH_2SO_3H$

Ion exchange chromatography is a technique that is well suited to the separation of proteins and peptides. It has a high resolving power and high protein/peptide binding capacity. The technique is versatile in terms of the variety of exchangers available and the ability to vary both pH and buffer composition.

Proteins and peptides are composed of amino acids that carry individually charged groups, such as carboxy or amino groups, resulting in an overall net charge. The strength of adsorption to stationary phase functional groups increases as this net charge increases, therefore the degree of adsorption is a function of pH. At low pH carboxy groups lose their negative charge whilst

amino groups become protonated. At high pH carboxy groups dissociate to become negatively charged whilst amino acids are deprotonated to become neutral. pH is one of the most important parameters in this separation technique as it determines the charges on both the analyte and ion exchanger. Almost all proteins exhibit both positive and negative charges at pH values used in ion exchange chromatography. Of the 20 naturally occurring amino acids 7 contain functional groups that are weakly acidic or basic. The pk_a values and % occurrence of these 7 amino acids can be seen in Table 1.1.4.3.[84]

Table 1.1.4.3. The naturally occurring amino acids with weakly acidic or basic side chains.

Amino acid group	pk_a in protein	Occurrence in proteins (%)
Arginine	12.0	5.7
Lysine	10.4-11.1	5.7
Tyrosine	10.0-10.3	3.2
Cysteine	9.0-9.5	1.73
Histidine	6.0-7.0	2.2
Aspartic acid	3.9-4.0	5.3
Glutamic acid	4.3-4.5	6.2
α-amino	6.8-8.0	
α-carboxy	3.5-4.3	

The interaction between stationary and mobile phase species is a dynamic equilibrium and the elution of adsorbed species can be achieved in two ways. Firstly a change in pH can lead to the reduction in net charge of both the stationary phase and analyte functional groups allowing desorption of the sample molecule. Secondly the addition of a salt solution can lead to competitive binding and blocking of sample molecules and ultimately their elution. Sample molecules

are eluted in order of increasing net charge and elution can be performed in a stepwise fashion or via a gradient.

Ion exchange chromatography utilises an orthogonal mode of separation to RP-HPLC and combination of the two techniques enables two dimensional separation to be performed.

1.1.4.5. Multidimensional Chromatography for protein and peptide identification

Through the application of orthogonal modes of separation, multidimensional chromatography can be a powerful approach to the resolution of complex samples. Traditionally the identification and quantification of proteins, in the field of proteomics, has been carried out by making use of 2-dimensional gel electrophoresis coupled to mass spectrometry. More recently soft ionization techniques such as electrospray and MALDI have been employed during identification and quantification analysis.[87,88] However, for proteins of low abundance these techniques are not sensitive enough, particularly when studying whole cell lysates and proteome digests.[89] The dominant multidimensional technique has, therefore, become ion exchange chromatography (typically cation exchange) followed by RP-HPLC. Yates *et al* developed a technique incorporating these two phases, which they call MudPIT (multidimensional protein identification technology), in which a biphasic column consisting of mixed strong cation exchange and C18 RP materials are packed into the same column. The MudPIT strategy has been used to identify almost 1500 proteins

from the yeast proteome.[90-92] Table 1.1.4.4 details various examples of samples analysed by multidimensional chromatography.

Table 1.1.4.4. Examples of some uses of multidimensional chromatography.

Type of multidimensional analysis	Sample Type	Reference
SCX/SAX-RP-HPLC	Ribosomes (Tryptic digest)	93
SCX-RP-HPLC	Urine	94
SCX-RP-HPLC	Blood ultrafiltrate	95, 96
SCX-RP-HPLC	E-coli	97
SCX-RP-HPLC	Saccharomyces Cerevisiae (Yeast)	90
SCX-RP-HPLC	BSA in-gel digest	98
MudPIT	Saccharoyces Cerevisiae (Yeast)	92
MudPIT	Plasmodium Falciparum (Yeast)	99
RP-mPC-CE	MHC peptides in EL4 cells	100
SE-RP-HPLC	E-coli	101

The majority of examples given utilise strong cation exchange and RP-HPLC however there are also instances of the use of other orthogonal modes of separation, for example size exclusion chromatography followed by RP-HPLC.

1.1.5. Immunology

Immunity can be divided into two basic types; innate or natural immunity and acquired or adaptive immunity.[102-106]

Innate immunity is the basic immune response we are all born with and consists of a range of cells which are available for immediate response to a wide range of pathogens without the need for previous exposure. An innate response is achieved in several ways including anatomic barriers such as skin, physiologic response to pathogens, such as fever which inhibits their growth, and inflammation in response to tissue damage with secretion of antibiotic serum proteins and delivery of phagocytes and endocytes, which engulf and digest invading macromolecules and micro organisms.

An acquired immune response is able to specifically recognise and selectively eliminate foreign micro organisms and molecules. Acquired immunity possesses four distinct characteristics which separate it from innate immunity. These are antigenic specificity and diversity, which allows the body to distinguish subtle differences among antigens; immunologic memory, which means that the immune response will be more effective on re-exposure to a particular antigen; and self and non-self recognition, which allows the immune system to distinguish between the two and only attack the non-self.

Together the innate and acquired immune responses have evolved to become a powerful defence against infection and disease. All cells of the immune system develop from pluripotent stem cells in the bone marrow and can be divided into two groups; lymphocytes and antigen presenting cells.

1.1.5.1. Lymphocytes

Lymphocytes are types of white blood cells that make up approximately 20 % of their complement and are found in the lymphoid organs. The different types of lymphocytes are described as follows.

B lymphocytes

B lymphocytes mature in the bone marrow and each possess a unique antigen binding receptor on its surface known as the B-cell receptor. This is a membrane bound antibody consisting of two identical light peptide chains and two non-identical heavy peptide chains. A light chain is bound to a heavy chain by disulphide linkages and non-covalently by hydrogen bonds and salt linkages to form heterodimers, these in turn are similarly bound together as shown in Figure 1.1.5.1. The antigen specificity of the antibodies is conferred by differences in the 'v' chains of the molecule's antigen binding site.

Upon encounter with the antigen for which the receptor is specific, the naïve B cell divides rapidly into memory B cells and effector cells, also called plasma cells, which produce a secretable form of antibody. The antibody circulates in the blood seeking out and neutralising antigens or marking them for elimination. B-cells recognise antigen on the surface of cells which have been released by an invading pathogen or altered self cells, for example that have become cancerous. The process of B-cell maturation is illustrated in Figure 1.1.5.2.

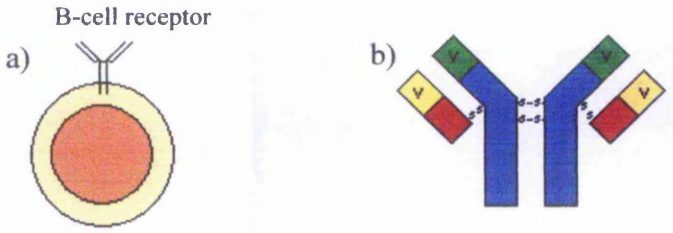


Figure 1.1.5.1. a) Representation of a B lymphocyte with B-cell receptor on its surface. b) An enlarged view of the structure of the surface bound antibody molecule. (Adapted from reference 106).

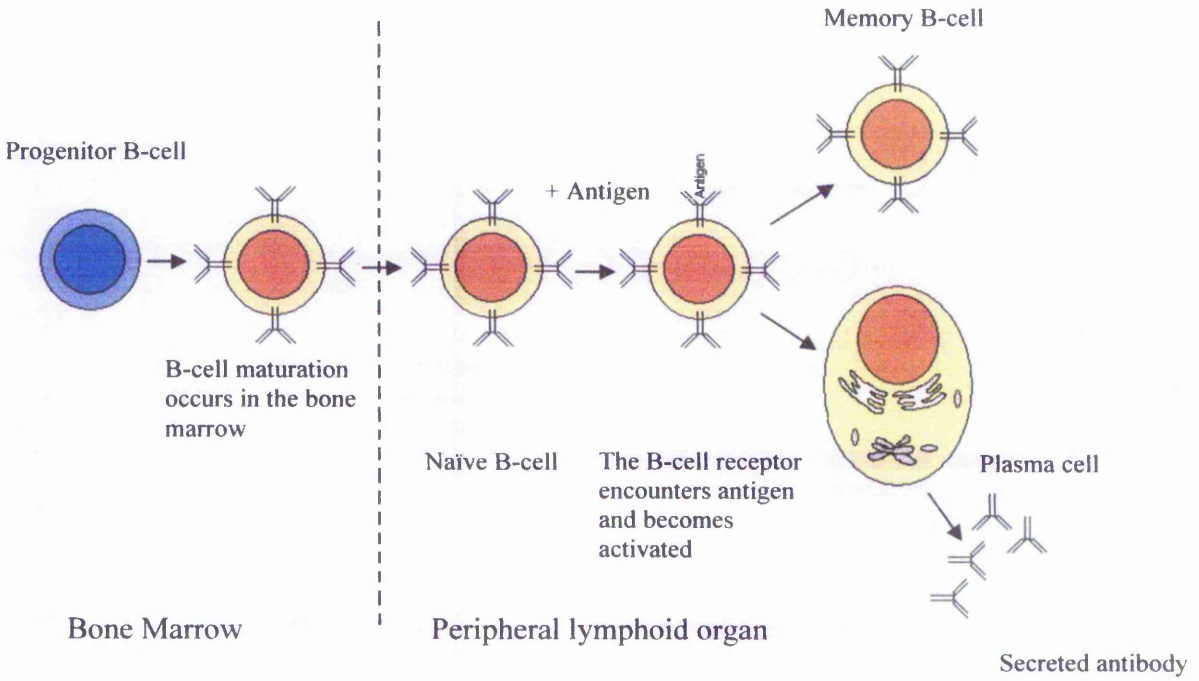


Figure 1.1.5.2. B-cell maturation and proliferation scheme after encounter with an antigen. (Adapted from reference 106).

T lymphocytes

In contrast to the B lymphocytes, T lymphocytes mature in the thymus and express a unique antigen binding receptor. After maturation, the naïve T-cells circulate through the bloodstream and peripheral lymphoid organs in search of foreign antigens. T lymphocytes further differ from their B counterparts in that they are only able to recognise an antigen when it is associated with the cell membrane protein called the major histocompatibility complex (MHC), see section 1.1.6. When the T-cell encounters the MHC associated antigens it also proliferates and differentiates into memory and effector cells.

There are two kinds of T lymphocyte, these are distinguishable by the presence on their surface of a glycoprotein. T-helper cells possess a molecule known as CD4 and bind to antigen peptides associated with MHC class II. Once bound they become effector cells which stimulate the release of cytokines, which are important in the activation of B-cells and further immune responses. Cytotoxic T cells are identified by the presence of a molecule known as CD8 on their surface. They bind antigen peptides associated with MHC class I. Upon interaction with the complex the cytotoxic lymphocytes (CTLs) proliferate and differentiate into effector cells which possess cytotoxic activity and destroy any cells which display that antigen.

Figure 1.1.5.3 a) T-helper cell, characterised by the presence of T-cell receptor and CD4 membrane glycoprotein. b) Cytotoxic T-cell characterised by the presence of T-cell receptor and CD8 membrane glycoprotein.

The process of T-lymphocyte activation for both T-cells and T-helper cells is shown schematically in Figure 1.1.5.4.

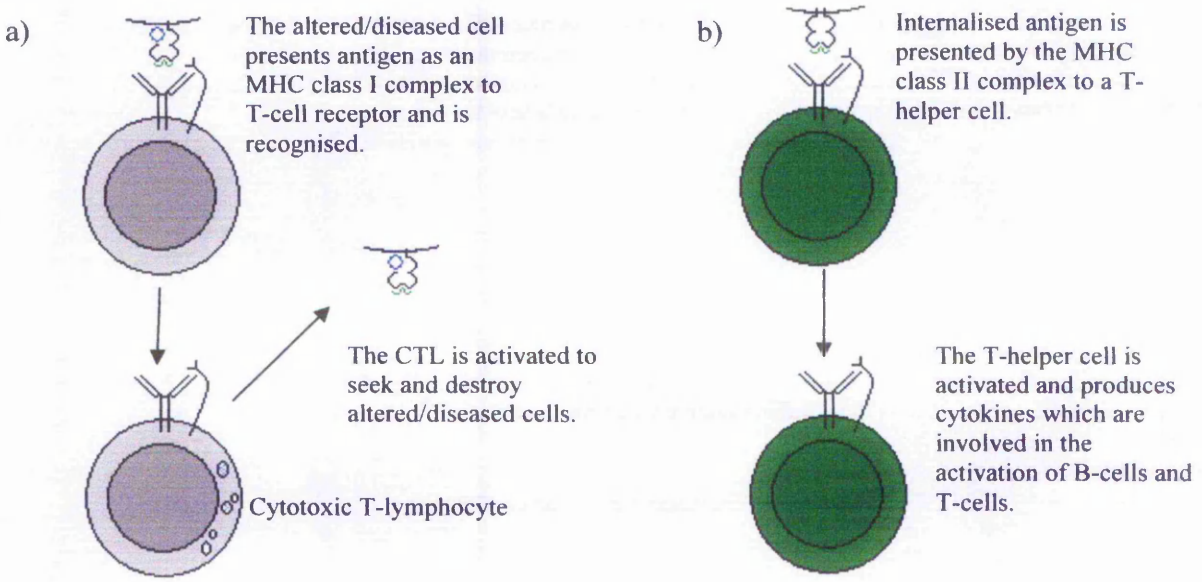


Figure 1.1.5.4 a) Scheme showing the activation of cytotoxic T-lymphocytes after recognition of T-cell associated with CD8 and antigen-MHC class I complex. b) Scheme showing the resulting release of cytokines and activation of CTL and B-cells upon

recognition of the T-helper cell associated with CD4 with an antigen-MHC class II complex. (Adapted from reference 106).

Antigen presenting cells (APCs) include macrophages, B-lymphocytes and dendritic cells which present on their surface MHC class II molecules. They function as a control of the T-helper cell activation, so that an antigen will only be recognised when in conjunction with the class II molecule. Antigens are typically complex peptides with small immunologically active regions, known as epitopes, which are bound by the receptor.

1.1.6. Major histocompatibility complex

The major histocompatibility complex (MHC) is a large genetic complex found on chromosome 6 which encodes the human leukocyte group A (HLA), the products of which play an important role in self/non-self discrimination in the immune response.[102-106] The genes exist in many different forms, known as alleles. The MHC has three regions which encode class I, class II and class III gene products. Class I and II molecules are membrane proteins, which are integral to the presentation of antigens. Class III are not membrane bound and do not take part in antigen presentation so will not be discussed further here.

MHC class I molecules are polymorphic transmembrane glycoproteins which are found on the surface of almost all nucleated cells. Structural information has been obtained by X-ray crystallographic analysis of the crystallised molecule.[103,106] It consists of an alpha chain (45 kDa) which is encoded by the A,

B and C regions of the HLA complex. This is non-covalently associated with a smaller protein β_2-microglobulin (12 kDa). The alpha chain is organised into three extracellular domains with the β_2m making up the fourth. It is attached to the membrane by a hydrophobic chain with a short hydrophilic chain linking it to the cytoplasm.[107] The structure for MHC class I is shown in Figure 1.1.6.1.

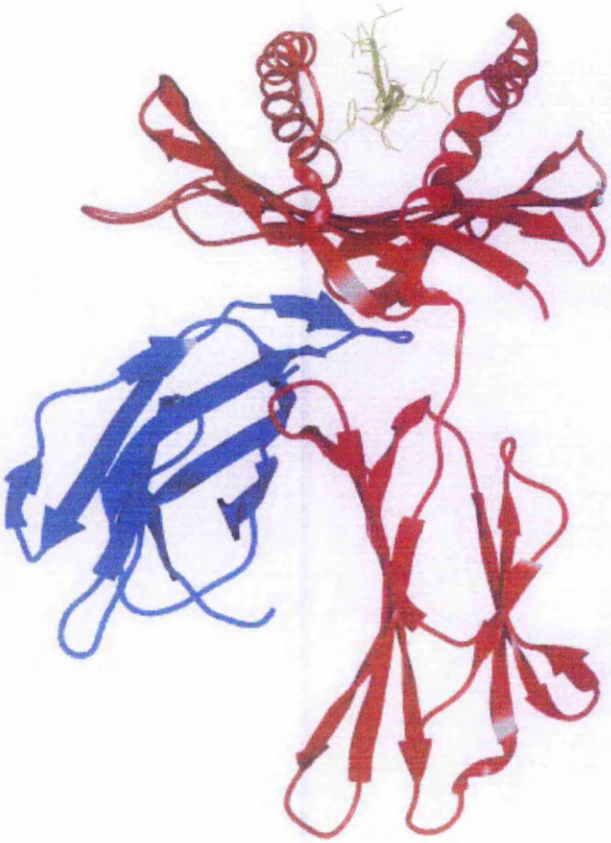


Figure 1.1.6.1. The structure of MHC class I. The alpha chain is shown in red and the β_2-microglobulin in blue. A peptide is shown in the binding cleft.

A peptide binding groove is formed by the alpha 1 and 2 binding domains and is highly polymorphic. Different allelic forms of the gene confer different

structures on the antigen binding cleft. These forms are inherited, one of A, B and C from each parent, and are known as haplotypes. There are almost 1000 haplotypes in the human population.[108] Each allelic variant binds a different set of peptides. Within the binding cleft are anchor residues, class I peptides are typically 8-11 amino acids in length, most commonly nonamers. The peptide amino acid residues bind to the anchor residue sites via their side chains which fit into pockets within the binding cleft. For class I there are two anchor positions, at position 9 on the carboxy terminus which is invariant and a further residue which is allele dependent and may reside at position 2, 3, 5 or 7. The MHC class I molecules presents peptides from endogenous proteins.[106, 109] It is shown schematically in Figure 1.1.6.2.

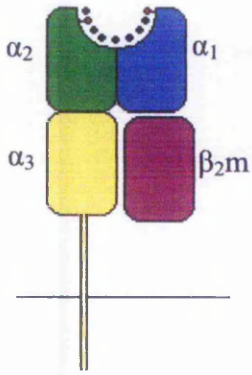


Figure 1.1.6.2. Representation of the transmembrane MHC class I molecule with nonamer peptide in the binding groove. Anchor residues, at the 2 and 9 position, are shown in red.

The MHC class II molecule is also a membrane bound glycoprotein molecule which shares extensive sequence homology with the class I molecule. It consists of two, non-covalently bound, polypeptide chains, the alpha chain is 34 kDa and the beta chain 28 kDa. Unlike the class I molecule, the class II molecule is found

only on the surface of specialist antigen presenting cells such as B-cells, dendritic cells and macrophages. It is encoded by the DP, DQ and DR region of the HLA complex. As with class I many alleles exist and are inherited, one from each parent. Structural information has been obtained by X-ray crystallographic analysis of the crystallised molecule.[109,110] It is made up of two external domains α1, α2 and β1, β2. A peptide binding groove is made from the α1 and β1 domains which, like class I α1 and 2 domains, are highly polymorphic. The structure of the class II molecule is shown in Figure 1.1.6.3.

Figure 1.1.6.3. The structure of MHC class II. The α1 and β1domains are seen to form the binding groove in which a peptide is nestled.

MHC class II peptides can vary in length and are usually between 12 and 25 residues long. This variability in peptide length is due to a lack of restriction in the binding groove. A core sequence of 9 amino acids sits in the binding groove with a further, indeterminate number of residues on either side. The MHC class II molecule presents peptides derived from exogenous proteins. It is shown schematically in Figure 1.1.6.4.

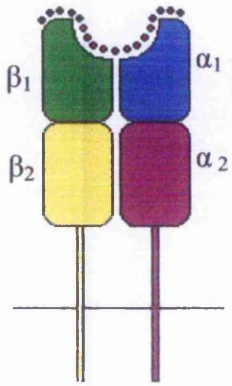


Figure 1.1.6.4. Representation of the transmembrane MHC class II molecule with a peptide core sequence is bound within the groove and shown in red, further amino acid residues are shown extending from the groove. The class II groove typically possesses 3 anchor positions, most importantly at position 1 but also 4, 6, 7 or 9.[111]

1.1.6.1. Processing pathways for class I and class II antigen presentation

Cytosolic proteins are made up of a variety of sequences involved in the routine cellular processes, including foreign structures such as viral proteins. These proteins are then degraded into peptides of optimal length, usually nonamers. The MHC class I complex is assembled in the endoplasmic reticulum with a calnexin chaperone for stability. Binding of a peptide within the groove then

occurs eliminating the need for the chaperone. The assembled complex is then transported through the golgi stack and to the cell surface.[112] This process is summarised in Figure 1.1.6.5.

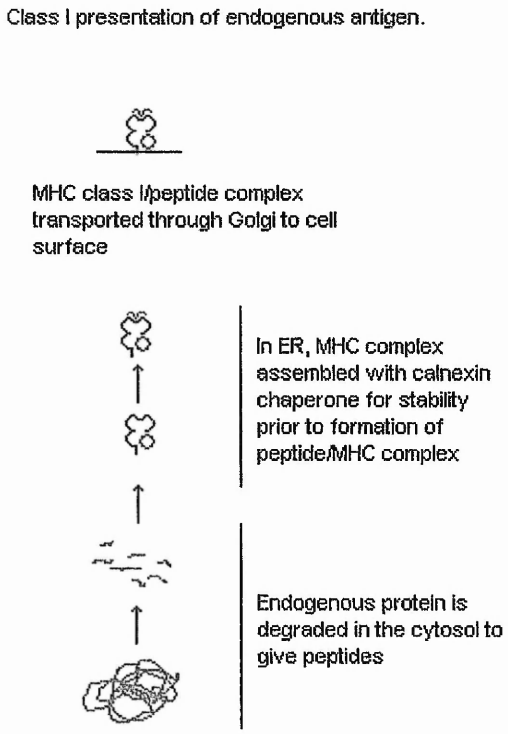


Figure 1.1.6.5. Class I endogenous antigen presentation pathway.

It is estimated that MHC class I complexed with a single peptide can vary in abundance from 1-10000 copy numbers per cell. Naturally, of these, the most abundant will be the self peptides.[113] The number therefore which can elicit a cytotoxic T-lymphocyte response would be considerably lower, in the region of 1-1000 copies per cell.[114,115] Table 1.1.6.1 lists a number of MHC class I peptides which have been identified and the levels at which they were estimated to be present on the surface of the cells.

Table 1.1.6.1. MHC class I peptide sequences and their estimated copy numbers per cell.[116]

Peptide	Definition	Copies/cell	Ref
YLDPAQQNL	MHN24 alloantigen	85-125	116
LSPFPFDL	p2Ca alloantigen	12000	117, 118
ALWGFFPVX	AH111 12-2 xenoantigen	1600	119
SLVELTSL	Bm1BZ194	10*	120
YIGEVLVSV	HA-2 minor HAg	260	121
SPSVDKARAEL	K-Y minor HAg	100	122
SLYNTVATL	HIV-1 gag	400	123
ILKEPVHGH	HIV-1 RT	12	118, 123
KYQAVTTTL	P198 tum-Ag	100	124
TYQRTRALV	Influenza ANP	100-540	125
ASNENMETM	Influenza ANP	100-220	124
SIINFEKL	OVA	88	118, 126
FLWGPRALV	MAGE-3	38-39	127
RRIYDLIEL	EBV	<1	128

*estimate 10 however assay prone to underestimate copy number due to competitive behaviour of other peptides in the extract

The peptides identified represent a range of originating proteins and a greater range of abundances from as low as 10-12 copies per cell up to 12000, although it may be difficult to really compare these copies/cell figures as they have been estimated using different protocols/techniques and assumptions. Hunt et al have identified a further fifteen peptides from HLA-B7 JY cells which were immunopurified and fractionated by RP-HPLC prior to analysis by HPLC ESI/MS/MS on a triple quadrupole mass spectrometer. These were identified with copy/cell numbers ranging from 90-1300.[129]

Although estimates as low as 10 copies per cell have been made these are in the

minority a copy/cell number between 50-100 would represent a sensible lower limit for which to aim.

The MHC class II molecule presents exogenous antigen on its surface. The exogenous protein is internalised by phagocytosis prior to degradation to peptide fragments in the endosome. The class II αβ dimer is assembled together with the invariant chain in the endoplasmic reticulum. The complex is then transported through the golgi stack and the invariant chain is proteolytically cleaved in a post-golgi compartment leaving class II invariant chain peptide (CLIP) in the peptide groove.[130-132] The molecule HLA-DM then facilitates the exchange of CLIP for an antigenic peptide prior to presentation on the cell surface. This presentation pathway can be seen in Figure 1.1.6.6.

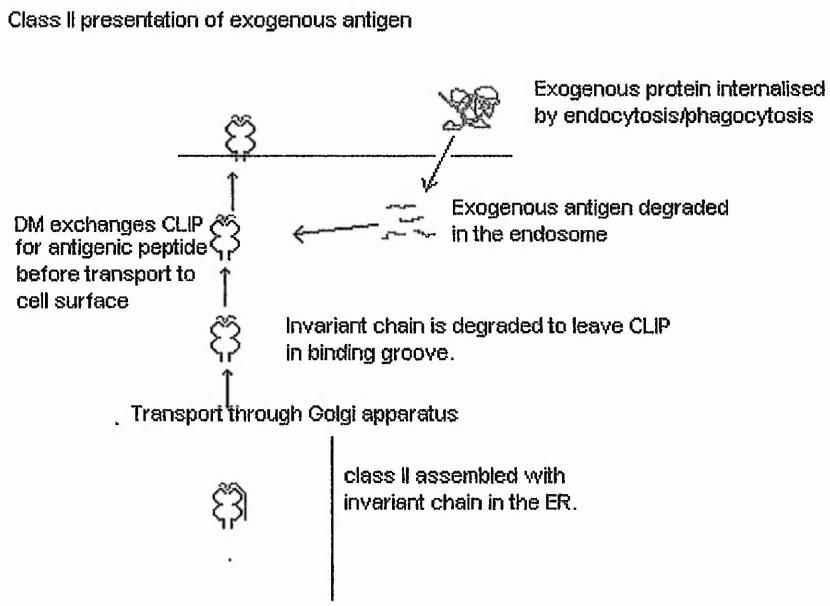


Figure 1.1.6.6. MHC class II exogenous antigen presentation pathway.

1.1.7. Cancer Immunotherapy

Cancer in general is extremely astute when it comes to resisting the advances of the immune system. Various types of the disease have developed immunologic escape mechanisms which may be exhibited either alone or in combination. They include.

- Down regulation of HLA class I expression, this makes the tumour a less attractive target for cytolytic T-cells.
- Surface expression of growth inhibiting/apoptosis inducing molecules (FasL, RCAS1) which react with the T-cells bearing their corresponding receptors.
- Internal defects in apoptosis, which favour the emergence of a tumour clone, also serve to make it harder to kill.

Mutation of the tumour suppressor protein p53 and its over expression is also common in human cancers.

The rationale behind research into cancer immunotherapy is that it should be possible to induce the immune system to reject cancerous cells. In order for this to happen it is necessary to identify antigenic protein or peptide targets in the patients' tumour and augment the existing weak anti-tumour response through therapeutic vaccination with tumour antigens by presenting them to the patient in such a way as to induce an immune response.[133, 134]

Tumour specific antigens (TSAs) are the ideal choice. These are antigens which are expressed solely by the malignancy, and are mutated proteins whose function is required to maintain the cancerous state, or embryonic proteins expressed due to a loss of the cellular control.[135-138] Tumour associated antigens (TAAs) are proteins which are not unique to cancer cells, but are over expressed compared to normal cells.

More than 100 genes encode antigens that are tumour associated or tumour specific and which are capable of acting as targets for cell mediated immune response. Tumour antigens can be divided into a number of different types and shown below are examples of each kind that have been identified.[103, 138-142]

- Viral antigens

  e.g. Epstein Barr virus (EBV), Burkitt's lymphoma; Human papilloma virus (HPV), cervical cancer; Human T-cell leukaemia virus.
- Over expressed and mutated antigen, these genes are expressed ubiquitously but are mutated in tumour cells.

  e.g. HER-2/neu, breast, lung; p53; MUC-1, colorectal, pancreatic; ras, pancreatic, colon, lung; c-myc
- Cancer testis, expressed in tumour and immune privileged tissue such as testis and placenta.

  e.g. MAGE, melanoma, lung, gastric; GAGE, melanoma; BAGE; NY-ESO-1; SSX
- Differentiation antigens, expressed in normal and tumour cells but restricted to the tissue type in question.

e.g. Melan-A/MART-1, melanoma; Tyrosinase, melanoma; gp100, melanoma; CEA, colorectal, lung, breast.

- Fusion Proteins, arise from chromosome translocation events and are found in tumour tissue.

  e.g. BCR/ABL (9:22), Leukaemia; ALL-1/AF-9 (9:11)

When attempting vaccination with tumour associated antigens the assumption is made that any autoimmune response is tolerable. One such side effect is the destruction of melanocytes, causing Vitiligo, seen in patients receiving immunotherapy targeting melanoma TAA.[143]

An approach pioneered by Boon and co-workers for the identification of tumour antigens has led to the identification of the cancer testis antigen MAGE and the melanoma antigens Melan-A/MART and tyrosinase.[144-148] This 'genetic approach' involves the transfection of cell lines with expression libraries of cDNA originating from diseased cells or tumours, into cells which expressed the desired MHC haplotypes. The transfection cells were then screened for expression of a tumour specific antigen by their ability to stimulate anti-tumour T-cells. The transcript which encodes for the MHC peptide was then isolated.

1.1.8. Identification of HLA peptides

The function of MHC class I and II molecules is to present antigens to CD8+ T-cells and CD4+ T-helper cells for immune surveillance. Characterisation of the

MHC antigen peptides is a continuing area of interest with a view to identifying potential vaccine candidates for immunotherapy. There are two major approaches which are used for the identification of HLA peptides.[140, 149-152] These approaches are described as follows.

1.1.8.1. Reverse Immunology

Under the reverse immunology approach, proteins associated with disease are searched for potential MHC binding epitopes using the consensus binding motifs of known MHC haplotypes. Prediction of allele specific sequence motifs are made by cataloguing the frequency of amino acids in different positions of peptides that are known MHC binders. Such motifs were first established employing Edman degradation on pooled MHC samples and confirmed and further redefined by tandem mass spectrometry.[108, 153, 154]

The establishment of these motifs has allowed the development of software tools that may be used in the prediction of corresponding sequences within a protein. Two such algorithms are SYFPEITHI (developed by Rammensee *et al.*[108]) and HLA_BIND (developed by K. C. Parker.[155]). Following software identification potential sequences can then be synthesised to establish their ability to stabilise the MHC molecule and stimulate an immunogenic CD8+ T-cell response. Most studies have identified HLA-A2 associated class I peptides as it is the most common human class I allele. A number of different allele specific peptides may be identified for each protein. The MAGE-1 protein is a good example of this, see Table 1.1.8.1.

Table 1.1.8.1. MAGE peptides identified using the reverse immunology approach.

Sequence	Source	HLA	Reference
EADPTGHSY	MAGE-1	A1	156
SLFRAVITK	MAGE-1	A3	157
EVYDGREHSA	MAGE-1	A28	
DPARYEFLW	MAGE-1	B53	
SAFPTTINF	MAGE-1	CW3	

However there are disadvantages to this approach, these are as follows:

i. The need for very well characterised binding motifs for the wide variety of haplotypes.

ii. It has been reported that these algorithms are correct only 30 % of the time. That is to say, less than one third of the sequences with the correct binding motif will actually bind to the MHC molecule.[158]

iii. Furthermore, the ability to bind and stabilise an MHC molecule does not guarantee its natural processing and surface expression.

iv. Potentially miss out on a wide range of peptides which have undergone post-translational modification such as glycosylation or phosphorylation.[159-163]

1.1.8.2. Direct Identification

The second approach involves the direct identification of MHC bound peptides from infected/transfected cells. This is generally done in one of three different ways.

Cell lysis

Cells are lysed with trifluouroacetic acid solution to give a complex mixture of intracellular and extracellular peptides. The addition of protease inhibitors to this solution is necessary to limit peptide degradation. The peptide mixture is fractionated by RP-HPLC and screened using a functional assay for the presence of particular T-cell epitopes. [166,167] However, only a small number of the peptides released would actually have been bound to MHC complex prior to cell lysis. As this peptide mixture contains the peptide content of the entire cell the subsequent identification of MHC class I peptides with any confidence is far from straightforward.

Immunoaffinity Purification

This approach, sometimes known as the direct biochemical approach, utilises the elution of HLA peptides from immunopurified MHC complexes.[168-170] Cells are lysed by employing a detergent, or by lowering the pH, allowing the release of the total cellular content. Similar to the cell lysis approach, the addition of protease inhibitors is necessary to limit peptide degradation. Monoclonal antibodies raised against particular class I haplotypes are then used to immunopurify class I complexes. An acid wash is then used to release the peptides from these immunopurified complexes. The released peptides are separated into fractions by RP-HPLC and each tested for their ability to stimulate T-cells.

This approach has been used extensively, in conjunction with mass spectrometric methods, by Hunt and coworkers and is described in detail later in this chapter.[171, 172] The major advantage of this approach is the ability to identify peptides which exhibit post translational modifications. The use of immunopurification also ensures a significantly cleaner extract than that observed by cell lysis alone. However, the possibility exists for MHC molecules to bind peptides which have been released from the cell interior by cytolysis. If these peptides are not normally expressed on the surface of cells and thus never presented for immune surveillance, this would lead to the incorrect identification of peptide vaccine targets.

Cell surface elution

An alternative method, which does not require cell lysis is the elution of peptides from the surface of cells using a mild acid wash. This elution occurs due to disruption of the MHC complex β_2-microglobulin, which destabilises the MHC complex and releases the peptide from the binding groove without affecting cell viability.[173, 174]

There are several advantages to this approach. The complexity of the elution mixture, compared to that obtained by cell lysis, is much reduced. After elution the MHC complexes will be reloaded with peptide allowing repeat elution of the same cells. This means that fewer cells may be required in order to generate sufficient peptide for analysis. Critically, unlike both of the previous methods, it can be said with a high degree of certainty that peptides identified after surface

elution were presented on the surface of cells and therefore subject to surveillance by the immune system.

The cell surface elution method has been applied to the identification of the MART-1/Melan-A peptide ILTVILGVL from the peripheral blood of melanoma patients. [174] The HLA-A3 bcr/abl peptide KQSSKALQR has also been identified from leukaemia patients and transfected cell lines.[175]

1.1.8.3. LC/MS methods for HLA peptide identification

Although the methodologies described in section 1.1.8.2 differ in their approach to peptide isolation, once this has been achieved the overall goal remains the same; to isolate and identify those peptides which are antigenic. There are three LC/MS based approaches commonly described in the literature for the identification of antigenic peptides following isolation. Examples of peptides isolated by these three methods can be seen in Tables 1.1.8.3a-1.1.8.3c. In most cases these examples have used immunopurification for the purpose of peptide isolation prior to analysis. However, these LC/MS methods may be applicable to other methods of isolation like those described in section 1.1.8.2. These LC/MS methods are described as follows.

- Predict/calibrate/detect

Epitopes of interest are predicted from protein sequences of known tumour or viral antigens. The corresponding synthetic peptides are then used to calibrate the LC/MS system prior to analysis of the extracted natural peptides using the same

LC/MS conditions. The peptides are identified by co-elution and tandem mass spectrometry. This method has been successfully applied to the identification of antigens from p53,[176] CEA[165] and MAGE.[177] Pascolo et al have used capillary LC/MS/MS with a split flow of ~300 nl/min and a 75 μm column and nanoflow ESI/MS/MS using a hybrid q-tof mass spectrometer for the identification of a MAGE-1 HLA-A2 epitope.163[99] Examples of peptides identified using this approach are given in Table 1.1.8.3a.

Table 1.1.8.3a. Examples of peptide sequences identified using the predict/calibrate/detect method.

Sequence	Source	HLA	Reference
KVLEYVIKV	MAGE-1	A1	177
PI/LDGEYFTI/L	p53	A2	176
GVI/LVGVAI/L/I/L	CEA	A2	165

- Differential or subtractive analysis

Hunt et al developed a method known as differential analysis which involves a comparison of the peptide products of infected/cancerous vs 'normal' cells. Two sets of MHC derived peptides, one from infected/cancerous tissue and one from healthy tissue are analysed by LC/MS. A comparison of the healthy and infected ion chromatograms is then performed over the entire scan range of the mass spectrometer using a window of 1 Da. This method is laborious and time consuming but will eventually reveal peptides unique to the diseased tissue.[171, 178] Employing differential/subtractive analysis the human B lymphoblastoid cell line CIR-A2.1 and its non-transfected CIR counterpart were immunopurified and

fractionated by capillary RP-HPLC/MS/MS on a triple quadrupole instrument. A broad band of material was observed in the CIR-A2.1 extracts but not the CIR extracts. A total of eight nonamer peptides were estimated to be present at a level of 100-3000 copies per cell. This study also investigated MHC binding motifs with all eight sharing a similar binding motif, a Leucine at position 2.[171] Similarly the viral MHC class I peptides KLWESPQEI was identified by differential analysis of EBV transformed cells and their measles virus infected counterparts. The MHC molecules were immunopurified and fractionated by RP-HPLC prior to ESI/MS/MS analysis on an ion trap mass spectrometer.[178] Differential analysis was used to identify two HLA-A2 peptides from the MAGE family, FLSLREAAL (MAGE-$1_{294-302}$) and FLWGPRALV (MAGE-$3_{271-279}$). Peptides were eluted from the surface of hepatocellular carcinoma tumour cells and healthy liver cells using a mild acid wash. Elution profiles were obtained by analysing peptides from both cell types by RP-HPLC-UV detection. Tumour specific peak fractions were identified by comparison of elution profiles and were analysed together with the non-tumour fraction by ESI-q-tof. Following this the MAGE-3 peptide was quantified and found to be present at 38-39 copies per tumour cell. CD8+ T-cell response was estimated by tetramer and IFN-γ cell release assay, this showed that the peptide produces an immune response *in-vitro*.[179] Combined with derivitisation, differential analysis was used to identify the HLA-A2 peptide RLASYLDRV expressed on Keratin-18 transfected cells. Normal and cancerous tissue were immunopurified and the peptides derivatised separately, one with a deuterated derivitisation agent. The two were then combined and fractioned by HPLC prior to ESI/MS analysis to quantify the

$^{2}D_{x}/^{1}H_{x}$ peak ratio and then MS/MS to perform sequence identification. Acetylation and guanidation/nicotinylation derivitisation were both utilised during this work.[180] Further examples of peptides identified using this approach are given in Table 1.1.8.3b.

Table 1.1.8.3b. Examples of peptides identified using differential/subtractive analysis.

Sequence	Source	HLA	Reference
GLDVLTAKV	Vinculin/HIV-1	A2	181
RILGAVAKV	Vinculin/HIV-1	A2	
KLWESPQEI	Measles/MV-C	A2	178
QLPEATFMV	Measles/MV-M	A2	
LMIDRYPVL	Measles/MV-H	A2	
RRYPDAVYL	Measles/MV-F	B27	
RLASYLDRV	Keratin-18	A2	180

- Ligand and T-cell split assay

Hunt's group also developed a comprehensive mass spectrometry based strategy to characterise antigens presented by MHC class I molecules.[40, 182] This protocol includes a multi-step reverse phase HPLC fractionation of peptides, CTL-based assay to identify active fractions and micro-LC/ESI/MS/MS to characterise the peptides. HPLC fractions are sampled and aliquots checked for biological activity, the active fractions are pooled and re-chromatographed and the procedure repeated, effectively reducing the peptide pool until it contains only peptide exhibiting immunological activity. In the third stage of chromatography

the immunologic assay is performed on-line with mass spectrometric identification by splitting the sample.[115, 172]

The general protocol of a ligand and T-cell split assay is described below:

1. The peptide-MHC complex is isolated by passing the supernatant of lysed cells through an allele specific monoclonal antibody column.
2. The liberated peptides are then separated from the high mass contaminants by ultrafiltration and fractionated by HPLC.
3. Each fraction is then analysed using a CTL-based assay.
4. Bioactive components are further fractionated with a second stage of RP-HPLC.
5. Each resulting fraction is then analysed by micro-LC/ESI/MS/MS in which the flow is split.

This approach has been used for the identification of a number of MHC bound peptides including a peptide epitope by high affinity CTL's from five melanoma patients. The peptide YLEPGPVTA was isolated from the melanoma cell line DM6 from HLA-A2 molecules. The peptide is isolated from 15 fmol of peptide extracted from 4.10^9 cells which implies only 2 peptide/HLA-A2 complexes per cell however low binding suggests disproportionate losses are experienced during immunoaffinity purification therefore as many as several hundred complexes per cell could be expected. A 12 % yield was estimated based on ~70 % recovery during class I purification and extraction and 25 % recovery during

the first two rounds of HPLC.[172] The ligand and T-cell approach has also been used to identify more than 30 novel HLA-presented peptides on renal clear cell carcinoma line. The cells were lysed and immunopurified prior to fractionation by microbore HPLC. These fractions were then rechromatographed and the eluent spotted onto a MALDI target. A novel peptide has been identified by this approach originating from the tumour associated gene NY-BR-16. Both on-line LC/MS and LC/MALDI/MS were used, a number of overlapping peptides were identified by both techniques and each identified a smaller group of different peptides, suggesting that the two techniques are complementary.[183] Further examples of peptides identified by the ligand and T-cell split assay are shown in Table 1.1.8.3c.

Table 1.1.8.3. Examples of peptides identified by the ligand and T-cell split assay.

Sequence	Source	HLA	Reference
YLEPGPVTA	Pmel-17/gp100	A2	172
SNFVFAGI	p68	H2-K	184
ALLAVGATK	Pmel-17/gp100	A3	185, 186
ALWGFFPVL	unknown	A2	119
YLDPAQQNL	zinc finger	A2	116
KTWGQYWQV	gp100	-	187
ITDQVPFSV	gp100	-	
VLYRYGSFSV	gp100	-	
AAGIGILTV	MART-1	-	
SQNFPGSQK	VMM18/VMM12 (melanoma)	A3	188
RLSNRLLLR	cancer testis antigen	A3	
SQNFPGSQ	VMM12 (melanoma)	A3	
RRIYDLIEL	EBV	B27	128
YMDGTMSQV	Tyrosinase	A2	185

It can be seen from the relative number of examples which have been found in the literature from each of these three methods that the Ligand and T-cell split assay approach is by far the most widely used and has been responsible for the identification of a number of MHC class I associated peptides. Although the approach requires multiple stages of separation, often leading to low peptide recoveries (~12 %, as described above) its ability to simultaneously identify the

peptide sequence whilst confirming its potential immunogenicity makes it a reliable tool for identifying MHC class I antigens.

Whatever the approach used, identification of tumour associated antigens is a considerable challenge as, at best they represent a tiny percentage of the cellular component and it may be difficult to ascertain their uniqueness to the cancer cells.

Several identified tumour antigens have been tested in clinical trials.[151]

- The cancer testis antigens MAGE-3; NY-ESO-1 and Melan-A/MART-1; tyrosinase and gp70 have all been tested and some promising results have been obtained.[146-148, 172, 189-190]
- Trials with peptide vaccines only,[191-192] with adjuvant and cytokines in melanoma and lymphoma patients.[190, 193-194]
- Use of antigen presenting cells (APC), dendritic cells (DC's) as a delivery system. The patients own dendritic cells are isolated and pulsed *in vitro* with peptides representing tumour antigens. The DC's are then reintroduced to the patient, they migrate to the lymph nodes and initiate the generation of an anti-tumour response.[193, 195-199]

It is clear, however, that in order for immunotherapy to be successful against cancer, the tumour load must first be reduced by chemotherapy, radiotherapy or surgery.

1.1.9. Chronic myeloid Leukaemia

Chronic myeloid leukaemia (CML) is a myeloproliferative disease which originates in the haemopoetic stem cell. It is characterised by an initial chronic phase of variable duration followed by an accelerated phase which progresses to a shorter terminal blastic phase. Once in blast crisis, patient survival is only 3-6 months.

CML is characterised by a chromosomal abnormality, a reciprocal translocation between chromosomes 9 and 22, t(9, 22) which results in a shortened chromosome 22 known as the Philadelphia chromosome and a lengthened chromosome 9, 9q+. The translocation gives rise to a chimeric gene, bcr/abl on the Philadelphia chromosome and the abl/bcr gene on the 9q+ chromosome, the fusion mRNA is translated into a chimeric fusion protein.[200-204] The breakpoint occurs on chromosome 9 in the c-abl gene between 1st and 2nd intron. This can be anywhere within a 300kb segment in the 5' end of the gene, upstream of 1st alternate exon Ib between Ia and Ib or downstream of Ia, this location has little impact. The breakpoint on chromosome 22 occurs in the bcr gene, at several locations. The major breakpoint region (M-bcr) in the bcr gene gives rise to two possibilities, between exon 2 (b2) and exon 3 (b3) or between exon 3 and exon 4 (b4). This breakpoint occurs in 95 % of all patients and results in hybrid bcr/abl mRNA molecules with a b2a2 and/or b3a2 junction encoding a p210$^{bcr/abl}$ fusion in each case.[205,206] In rare cases the bcr breakpoint occurs further upstream in the minor breakpoint region (m-bcr), the bcr/abl transcript contains an e1a2 junction

and is translated into a smaller 190 kDa fusion protein $p190^{bcr/abl}$. A representation of these locations is shown in Figure 1.1.9.1.

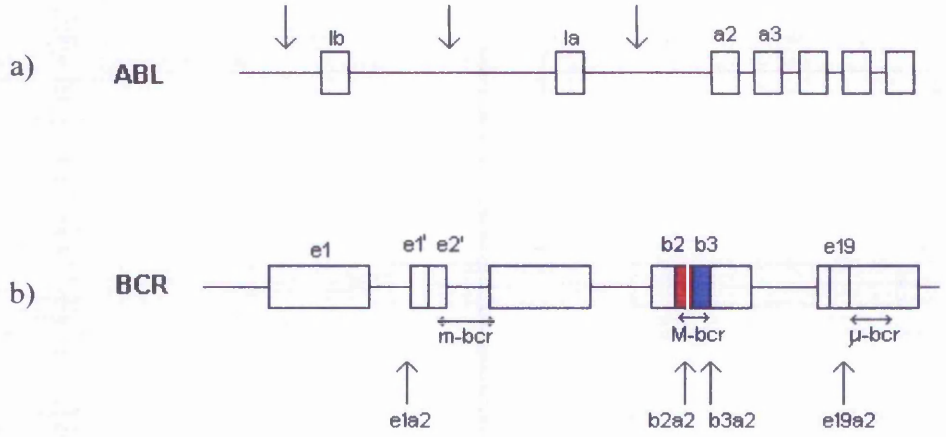


Figure 1.1.9.1. The breakpoint regions on a) the abl gene and b) the bcr gene which form the various bcr/abl fusion genes.

The protein which is encoded by the bcr/abl gene has elevated tyrosine kinase activity which is regarded as pivotal in the mechanism which underlies the chronic phase of CML.[207-209] The chimeric fusion protein is a tumour specific antigen as the junctional region contains areas of both the bcr and abl proteins in a unique association not present in other cells. The protein also contains a new amino acid at the breakpoint, a lysine residue in b3a2 and glutamic acid residue in b2a2. The chromosomal translocation which results in the formation of a Philadelphia chromosome with a b3a2 breakpoint is represented in Figure 1.1.9.2. The amino acid sequence of the resulting protein which spans the fusion region of the translocation is also shown. The unique amino acid residue which

has been spliced between the interrupted bcr and abl proteins, in the b3a2 case a lysine residue, is highlighted in yellow.

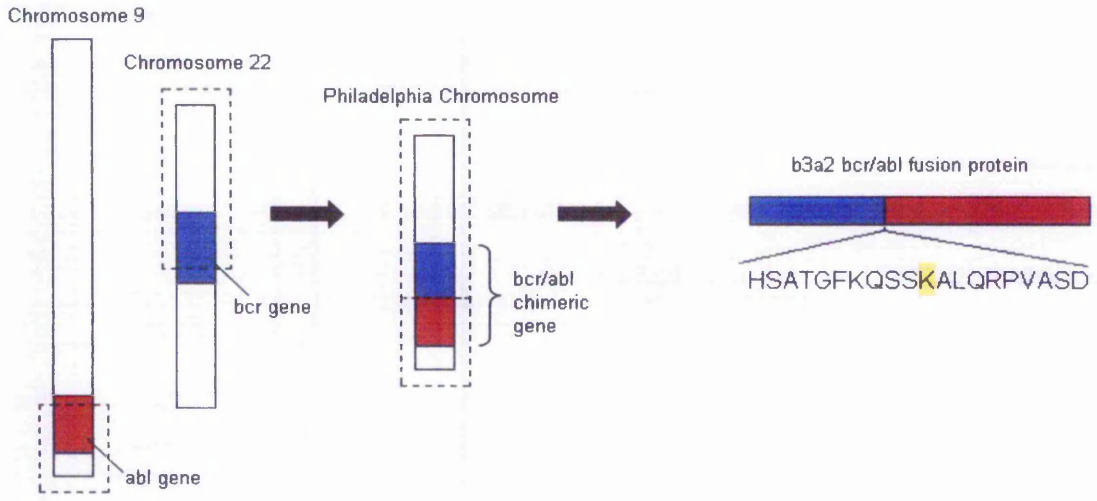


Figure 1.1.9.2 Representation of the reciprocal chromosomal translocation which generates the b3a2 bcr/abl fusion protein $p210^{bcr/abl}$. The uniquely produced lysine residue (K) is highlighted in yellow.

1.1.9.1. Treatment of CML

Chemotherapy

Treatment after initial diagnosis and during the chronic phase is through chemotherapeutic agents. Hydroxyurea and bisulfan were once the drugs of choice to reduce leukocyte numbers. However, these are no longer routinely used and have been superseded by interferon alpha which has shown great effectiveness during the chronic phase. Up to 80% of patients achieve haematological remission and a quarter will show a complete or major

cytogenetic remission.[210] This therapy shows prolonged survival and is very good during the early stages of disease although some patients can suffer badly from side effects.

Imatinib Mesylate (Gleevac)

A new drug has been highly successful, Imatinib mesylate or Gleevac is a rationally designed compound which blocks tyrosine kinase activity and so is an ideal weapon against the bcr/abl p210 protein as it is known that the tyrosine kinase activity of the protein is essential to its transforming ability, driving unchecked blood cell proliferation.[211, 212] It works by binding to the ATP binding site of ABL when the activation loop of kinase is closed and stabilises the protein in the inactive conformation.[213] It is a small molecule inhibitor which produces clinical remission in CML patients with minimal toxicity. It has also shown some success in treating patients in blast crisis.[214] The drug is now a frontline therapy but resistance to its effects is increasingly being encountered.[215]

Stem cell transplantation

The only treatment currently available which is considered a potential curative is allogenic stem cell transplantation with a HLA matched donor. Only a minority of patients are eligible for this treatment due to a lack of suitable donor or the patient's age, as the risks of this procedure become more pronounced in older patients. Autologous transplantation, where the patients own stem cells are harvested soon after treatment has begun and transplanted to the patient at a later stage, is sometimes explored. This can offer prolonged survival to patients who

are not responding well to interferon alpha treatment or who are not eligible for allogenic stem cell transplantation.

Vaccination with bcr/abl derived peptides

Bocchia *et al*[216] investigated whether peptides derived from the bcr/abl breakpoint are able to bind with high affinity to the groove of the HLA molecule. It was determined that four b3a2 peptides displayed intermediate to high strength binding to A3, A11, B8 and A3/A11 haplotypes. The ability of these peptides to elicit, *in vitro*, class I CTLs in healthy donors was tested.[216] A 25 amino acid peptide spanning the junctional region was also tested. Half of all the donors and all of the A3 positive ones mounted a specific T-cell cytolytic reaction to the peptides. The same peptides were also tested on patients with encouraging responses.[217] The first phase one dose escalation trial was performed on 12 patients.[218] The CML patients were all b3a2 and were dosed with 5 peptides with an immunologic adjuvant QS-21 with mixed responses. All but one patient had a high tumour burden. The single patient in cytogenetic remission experienced transient disappearance of b3a2 mRNA. Vaccination after treatment with Imatinib for 12 months or interferon alpha for 24 months has also been explored.[219] All but one of the patients saw an improved cytogenetic response with 2 patients achieving complete cytogenetic remission. Patients also experienced CD4 cell proliferation and IFNγ production. It is believed that the vaccination route is more successful when the tumour burden has been reduced as far as possible hence initial treatment with imatinib or interferon alpha improves the patients' chances.

If peptide vaccines are to be developed that are effective against CML and other cancers then it is necessary to show that specific antigens are associated with MHC class I and II molecules.

1.2.0. References

1. Thompson J.J., Rays of positive electricity and their application to chemical analysis, Longman Green and Co Ltd, London, **1913**
2. Belov M.E., Gorshkov M.V., Udseth H.R., Anderson G.A., Smith R.D., *Anal. Chem.* **2000**, 72, 2271-2279.
3. Barber M., Bardoli R.S., Sedgwick R.D. Tyler A.N., *J. Chem. Soc., Chem. Comm*, **1981,** 325-327.
4. de Hoffmann E., Stroobant V., *Mass Spectrometry Principles and Applications 2nd Ed,* **2002**, Wiley, UK.
5. Karas M., Hillenkamp F.H., *Anal. Chem.* **1988**, 60, 2229-3201.
6. Tanaka K., Waki H., Ido H., Akita S., Yoshida T., *Rapid Commun. Mass Spectrom.*, **1988,** 2, 151-153.
7. Dass C., *Principles and Practice of Biological Mass Spectrometry,* **2001,** Wiley, New York.
8. Dole M., Mach L.L., Hines R.L., Mobley R.C., Ferguson L.P., Alice M.B., *J. Chem. Phys.*, **1968,** 49, 2240-2249.
9. Whitehouse C.M., Dreyer R.N., Yamashita M., *Anal. Chem.*, **1985,** 5, 675-679.
10. Fenn J.B., Mann M., Meng C.K., Wong S.F., Whitehouse C.M., *Mass Spectrom. Rev.*, **1990**, 9, 37-70.
11. Fenn J.B., Mann M., Meng C.K., Wong S.F., Whitehouse C.M., *Science*, **1989**, 246, 64-71.
12. Smith R.D., Loo J.A., Edmonds C.G., Barinaga C.J., Udseth H.R., *Anal. Chem.*, **1990**, 62,882-899;
13. Smith R.D., Loo J.A., Loo R.R.O., Busman H., Udseth H.R., *Mass Spectrometr. Rev.*, **1991**, 10, 359-451.
14. Mehlis B., Kertscher U., *Analyt. Chim. Acta*, **1997**, 352, 71-83.
15. Cao P., Moini M., *Rapid Comm. Mass Spectrom.*, **1998**, 12, 864-870
16. Stults J.D., Marsters J.C., *Rapid Comm. Mass Spectrom.*, **1991**, 5, 359;
17. McCloskey J.A., Crain P.F., *Int. J. Mass Spectrom. And Ion Proc.*, **1992**, 118, 593-615.

18. Miketova P., Schram K.H., *Mol. Biotechnol.*, **1997**, 8, 249-253.

19. Muddiman D.C., Smith R.D., *Rev. Anal. Chem.*, **1998**, 17, 1-68.

20. Owens D.R., Bothner B., Phung Q., Harris K., Siuzdak G., *Bioorg. Med. Chem.*, **1998,** 6, 1547-1554.

21. Duffin K.L., Welply J.K., Huang E., Henion J.D., *Anal. Chem.*, **1992,** 64, 1440.

22. Shen X.D., Perreault H., *J. Chrom. A*, **1998**, 811, 47-59.

23. Herderich M., Richling E., Roscher R., Schneider C., Schwab W., Humph H. U., Schreier P., *Chromatographia,* **1997**, 45, 127-132.

24. Strege M. A., *J. Chrom. B.*, **1999**, 725, 67-78.

25. Nohmi T., Fenn J.B., *J. Am. Chem. Soc.*, **1992**, 114, 3241-3246

26. Katta V.,Chowdhury S.K., Chait B.T., *J. Am. Chem. Soc.*, **1990,** 112, 5348-5349;

27. Colton R., Traeger J. C., *Inorg. Chimica Acta*, **1992**, 201, 153.

28. Gaskell S.J., *J. Mass Spectrom.*, **1997**, 32, 677-688.

29. Kebarle P., Tang L., *Anal. Chem.*, **1993,** 65, 972A-985A.

30. Cole R.B., *Electrospray Ionisation Mass Spectrometry*, **1997,** Wiley, Chichester.

31. Gomez A., Tang K.Q., *Phys. Fluids,* **1994,** 6, 404-414.

32. Iribarne J.V., Thomson B.A., *J. Chem. Phys,* **1976,** 64, 2287-

33. Emmett M.R., Caprioli, R.M., *J.Am. Soc. Mass Spectrom.*, **1994,** 5, 605-613.

34. Wilm M., Mann M., *Anal. Chem.* **1996**, 68, 1-8.

35. Wilm M., Shevchenko A., Houthaeve T., Breit S., Schweigerer L., Fotsis T., Mann M., *Nature,* **1996,** 379, 466-469.

36. Abian J., Oosterkamp A. J., Gelpi E., *J. Mass Spectrom.* **1999,** 34, 244-254

37. McLuckey S.A., Vanberkel G. J., Glish G. L., Huang E. C., Henion J. D., *Anal. Chem.*, **1991,** 63, 375-383.

38. Miller P.E., Denton M.B., *J. Chemical Ed.*, **1986,** 63, 617-622.

39. Heinzle E., *J. Biotech.*, **1992,** 25, 81-114.

40. Yates J.R., *J. Mass Spectrom.*, **1998,** 33, 1-19.

41. Yost R.A., Enke C.G., *Anal. Chem.* **1979,** 51, 1251A-1264A.

42. March R.E., *J. Mass Spectrom.*, **1997,** 32, 351-369.

43. Dorey R.C., *Mass Spectrometry in the Biological Sciences,* **1992,** 79-92.

44. Stafford G.C., Kelley P.E., Syka J.E.P., Reynolds W.E., Todd J.F., *Int. J. Mass Spectrom. Ion Proc.*, **1984,** 60, 85.

45. Kinter M., Sherman N.E., *Protein Sequencing and Identification Using Tandem Mass Spectrometry,* **2000,** Wiley, New York.

46. Wiley W.C., McLaren J.B., *Rev. Sci. Instrum.*, **1955,** 16, 1150.

47. Cotter R.J., *Anal. Chem.*, **1992,** 64, 1027A-1039A.

48. Roepstorff P., Fohlman J., *J. Biomed. Mass Spectrom.*, **1984,** 11, 601.

49. Eng J.K., McCormack A.L., Yates J.R., *J. Am. Soc. Mass Spectrom.*, **1994,** 5, 976-989.

50. Biemann K., *Biomed. Environ. Mass Spectrom.*, **1988,** 16, 99-111.

51. Recommendations on nomenclature for chromatography, rules approved 1973, IUPAC Analytical Chemistry discussion commission on analytical nomenclature, *Pure Appl. Chem.*, **1974,** 37, 447.

52. Horvath C., Priess B., Lipsky S.R., *Anal. Chim. Acta*, **1967,** 38, 305.

53. Novotny M., *Anal. Chem.*, **1998,** 60, 502A-510A.

54. Ishii D., Asai K., Hibi K., Jonokuchi T., Nagaya M., *J. Chromatogr.* **1977,** 144, 157-168

55. Flurer C. L., Borra C., Andreolini F., Novotny M., *J. Chromatogr.* **1988,** 448, 73-86

56. Vissers J.P.C., Claessens M.A., Cramers C.A., *J. Chromatogr. A,* **1997,** 779, 1-28.

57. Tomer K.B., Moseley M. A., Deterding L. J., Parker C. E., *Mass Spectrom. Rev.*, **1994,** 13, 431-457.

58. Porath J., *Nature,* **1975,** 258, 598-599.

59. Ren D., Penner N. A., Selntz B. E., Mirzaei H., Regnier F., *J.Proteome Res.*, **2003,** 2, 321-329.

60. Chaga, G.S., *J. Biochem. Biophys. Methods,* **2001,** 49, 313-334.

61. Porath J., *Protein Expression and Purification,* **1992,** 3, 263-281.

62. Lindberg G., *Int. J. Peptide and Protein Purification,* **1991,** 38, 253-259.

63. Gaberc-Porekar V., Menart V., *J. Biochem. Biophys. Methods.* **2001,** 49, 355-360.

64. Zhang G., Fan H., Xu C., Bao H., Yang P; *Anal. Biochem.*, **2003**, 313, 327-330

65. Kagedal L., *Protein Purification,* New York, **1989**,

66. Chaga G. S., *J. Biochem. Biophys. Methods*, **2001**, 49, 313-334.

67. Wong J.W., Albright R.L., Wang N.L., *Sep. Purif. Methods* **1991**, 20, 49-106

68. Porath J, Belew M; *Affinity Chromatography and Biological recognition*, Academic Press, San Diego, **1983**.

69. Sulkowski E, *Trends in Biotechnology*, **1985**, 3, 1-7

70. Wang S., Zhang X., Regnier F. E., *J. Chrom. A*, **2002,** 949, 153-162.

71. Hochuli E., Dobeli H., Schacher A., *J. Chrom.*, **1987,** 411, 177-184.

72. Hochuli E., Bannwarth W., Dobeli H., Gentz R., Stuber D., *Biotechnology*, **1988,** 1321-1325.

73. Westra D.F., Welling G. W., Koedijk D. G. A. M., The T. H., Welling-Wester S., *J. Chrom. B*, **2001,** 760, 129-136.

74. Chaga G.S., Bochkariov D. E., Jokhadze G. G., Hopp J., Nelson P., *J.Chrom. A*, **1999,** 864, 247-256.

75. Wizemann H., von Brunn A., *J. Virol. Methods*, **1999**, 77, 189-197.

76. Casey J.L., Keep P. A., Chester K. A., Robson L., Hawkins R. E., Begent R. H. J., *J. Immunol. Methods*, **1995,** 179, 105-116.

77. Ueda E. K. M., Gout P. W., Morganti L., *J. Chrom. A*, **2003,** 988, 1-23.

78. Clausen M, Lamb C.J, Megnet R, Doerner P.W; *Anal Biochem.*, **1993**, 212, 537-539

79. Schon U, Schumann W, *Gene, 1994*, 147, 91-94

80. Witzgall R, O'Leary E, Bonvantre J.V, *Anal Biochem*, **1994**, 223, 291-298

81. Wu C.Y.E, Blaszczak L.C, Smith M.C, Skatrud P.L, *J Bacteriol.*, **1994**, 176, 1539-1541

82. Anderson L., Sulkowski E., *J. Chromatogr.*, **1992**, 604, 13-17

83. Hansen P., Lindberg G., *J. Chromatogr.*, **1995**, 690, 155-159.

84. Janson J.C; Ryden L., *Protein Purification 2nd edition*, **1998**, Wiley, New York

85. Ahuja S., *Chromatography and Separation Science*, **2003,** Academic Press, San Diego.

86. Niessen W. M. A., *Liquid Chromatography-Mass Spectrometry, 2nd Edition*, **1999**, Marcell Dekker, New York.

87. Mann, M, Hendrickson, RC, Pandey A., *Ann. Rev. Biochem.*, **2001**, 70, 437-473;

88. Shevchenko; Jensen O. N., Podtelejnikov L. V., Sagliocco F., Wilm M., Vorm O., Mortensen P., Shevchenko A., Boucherie H., Mann M., *Proc. Natl. Acad. Sci.* **1996,** 93, 14440–14445.

89. Peng J., Gygi S. P., *J. Mass Spectrom.*, **2001,** 36, 1083-1091.

90. Link A. J., Eng J., Schieltz D. M., Carmack E., Mize G. J., Morris D. R., Garvik B. M., Yates III J. R., *Nature Biotechmol.*, **1999**, 17, 676-682.

91. Harry J. L., Wilkins M. R., Herbert B. R., Packer N. H., Gooley A. A., Williams K. L., *Electrophoresis,* **2000,** 21, 1071-1081

92. Washburn M. P., Wolters D., Yates III J. R., *Nature biotechnol.* **2001,** 19, 242-247.

93. Wagner Y., Sickmann A., Meyer HE., Daum G., *J. Am. Soc. Mass Spectrom.*, **2003**, 14, 1003-1011

94. Heine G., Raida M., Forssmann W-G., *J. Chrom. A,* **1997**, 776, 117-124

95. Schrader M., Jurgens M., Hess R., Schulz-Knappe P., Raida M., Forssmann W-G., *J. Chrom. A,* **1997**, 776, 139-145.

96. Richter R., Schulz-Knappe P., Schrader M., Standker L., Jurgens M., Tammen H., Forsmann W-G., *J. Chrom. B,* **1999**, 726, 25-35.

97. Nagele E., Vollmer M., Horth P., *J. Chrom. A*, **2003,** 1009, 197-205

98. Zhang G., Fan H., Xu C., Bao H., Yang P., *Anal. Biochem.*, **2003**, 313, 327-330

99. Florens L., *Nature,* **2002,** 419, 520.

100. Tomlinson A.J., Jameson S., Naylor S., J. Chrom. A, 1996, 744, 273- 278

101. Opiteck G.J., Ramirez S.M., Jorgenson J.W., Moseley III M.A., *Anal. Biochem.,* **1998,** 258, 349-361.

102. Weir D.H., Stewart J., *Immunology 7th edition,* Churchill Livingstone, New York, **1993**.

103. Roitt I.M., Delves P.J., *Essential Immunology 9th Edition*, Blackwell, Oxford, **1997**.

104. Leffell M.S., Donnesberg A.D., Rose N.R., *Handbook of Human Immunology,* CRC Press, Boca Raton, **1997**.

105. Kuby J., *Immunology 3rd edition,* W.H. Freeman, New York, **1997**.

106. Goldsby R.A., Osborne B.A., Kindt T.J., Kuby J., *Immunology 5th Edition,* W.H. Freeman, New York, **2002**.

107. Bjorkmann P.J., Saper M.A., Samraoui B., Bennett W.S., Strominger J.L., Wiley D.C., *Nature,* **1987,** 329, 506-512.

108. Rammensee H.G., Bachman J., Emmerich N. P. N., Bachor O. A., Stevanovic S., *Immunogenetics,* **1999,** 50, 213-219.

109. Dessen A., Lawrence C. M., Cupo S., Zaller D. M., Wiley D. C., *Immunity,* **1997**, 7, 473-481.

110. Mullen M. M., Haan K. M., Longnecker R., Jardetzky T. S., *Molecular Cell,* **2002,** 9, 375-385.

111. Rammensee H.G., *Curr. Opin. in Immunol.,* **1995,** 7, 85-96.

112. Antoniou A.N., Powis S. J., Elliott T., *Curr. Opin. in Immunol,* **2003,** 15, 75-81.

113. Stevanovic S., Schild H., *Semin. Immunol.* **1999,** 11, 375-384

114. Wallny H. J., Deres K., Faath S., Jung G., Van Pel A., Boon T., Rammensee H. G., *Int. Immunol.,* **1992,** 4, 1085-1090.

115. Lemmel C., Stevanovic S., *Methods,* **2003,** 29, 248-259.

116. Wang W; Gulden P.H; Pierce R.A; Shabanowitz J.A; Man S.T; Hunt D.F; Engelhard V.H; *J. Immunol.* **1997**, 158, 5797-5804.

117. Udaka K; Tsomides T.J; Eisen H.N; *Cell,* **1992**, 69, 989.

118. Kageyama S; Tsomides T.J; Sykulev Y; Eisen H.N, *J. Immunol.*, **1995**, 154, 567-576.

119. Henderson R.A; Cox A.L; Sakaguchi E; Appella J; Shabanowitz J; Hunt DF; Engelhard V.H; *Proc. Natl. Acad. Sci.*, **1993**, 90, 10275.

120. Malarkannan S; Gonzalez F; NguyenV; Adair G; *J. Immunol.*, **1996**, 157, 4464.

121. de Haan J.M; Sherman N.E; Blokland E; Huczko E; Koning F; Drijfhout J.W; Skipper J; Shabanowitz J; Hunt D.F; Engelhard V.H; Goulmy E; *Science,* **1995**, 268, 1476.

122. Wang W; Meadows L.R; de Haan J.M; Sherman N.E; Chen Y; Blokland E; Shabanowitz J; Agulnik A; Hendrickson R.C; Bishop C.E; Hunt D.F; Goulmy E; Engelhard V.H; *Science,* **1995**, 269,1588.

123. Tsomides T.J; Aldovini A; Johnson R.D; Walker B.D; Young R.A; Eisen H.N; *J. Exp. Med.*, **1994**, 180, 1283.

124. Wallny H.J; Deres K; Faath F; Jung G; Van Pel A; Boon T; Rammensee H-G; *Int. Immunol.*, **1992,** 4, 1085.

125. Falk K; Rotzschke O; Deres K; Metzger J; Jung J; Rammensee H-G; *J. Exp. Med.*, **1991**, 174, 425

126. Rotzschke O; Falk K; Stevanovic S; Jung G; Walden P; Rammensee H-G; *Eur. J. Immunol.*, **1991,** 21, 2891.

127. Zhou M; Peng J-R; Zhang H-G; Wang H-X; Zhoung Z-H; Pan X-Y; Chen W-F; Leng X-S; *Immunology Letters,* **2005**, 99, 113-121.

128. Crotzer V.L; Christian R.E; Brooks J.M; Shabanowitz J; Settlage R.E; Marfo J.A; White F.M; Rickinson A.B; Hunt D.F; Engelhard V.H; *J. Immunol.*, **2000**, 164, 6120-6129.

129. Huczko E.L; Bodnar W.M; Benjamin D; Sakaguchi K; Zhu N.Z; Shabanowitz J; Henderson R.A; Appella E; Hunt D.F; Engelhard V.H; *J. Immunol.*, **1993**, 151, 2572-2587.

130. Newcomb J.R., Cresswell P., *J. Immunol.*, **1993,** 150, 499-507.

131. Machamer C.E., Cresswell P., *J. Immunol.*, **1982,** 129, 2564-2569.

132. Jones P.P., Murphy D.B., Hewgill D., McDevitt H.O., *Immunochemistry,* **1979,** 16, 51-60.

133. Admont A., Barnea E., Ziv T., *Molecular and Cellular Proteomics,* **2003,** 2, 388-398.

134. Salit R.B., Kast W.M., Velders W.M., *Frontiers in Bioscience,* **2002,** 7, e204-213.

135. Pardoll D.M., *Clin. Immunol.*, **2000,** 95, S44-S62.

136. Rosenberg S.A., *Nature,* **2001,** 411, 380-384.

137. Jager E., Jager D., Knuth A., *Curr. Opin. in Immunol.*, **2002,** 14, 178-182.

138. Boon T., Cerottini J. C., Vandeneynde B., Vanderbruggen P., Vanpel A., *Ann. Rev. Immunol.*, **1994,** 12, 337-365.

139. Melief C.J.M., Offringa R., Toes R.E.M., Kast W.M., *Curr. Opin. Immunol.*, **1996,** 8, 651-657.

140. Bonner P.L.R., Lill J.R., Hill S., Creaser C.S., Rees R.C, *J. Immunological Methods,* **2002,** 262, 5-19.

141. Fong L., Engleman E. G., A*nn. rev. immunol.* **2000,** 18, 245.

142. Scanlan M.J., Gure A.O., Jungbluth A.A., Old L.J., Chen Y.T., *Immunol. Rev.* **2002,** 188, 22-32.

143. Yee C., Thompson J. A., Roche P., Byrd D. R., Lee P. P., Piepkorn M., Kenyon K., Davis M. M., Riddell S. R., Greenberg P. D., *J. Exp. Med.*, **2000,** 192, 1637-1644.

144. Van der Bruggen P; Traversari P; Chouz P; Lurquin C; DePlaen E; Van den Eynde B; Knuth A; Boon T; *Science,* **1991,** 254, 1643.

145. Traversari C; Van der Bruggen P; Van Eynde B; Hainaut P; Lemoine C; Ohta N; Old L; Boon T; *Immunogenetics,* **1992,** 35, 145

146. Shoshan S. H., Admon A., *Pharmacogenetics,* **2004,** 5, 845-849.

147. Admon A., Barnea E., Ziv T., *Molecular and Cellular Proteomics,* **2003,** 2, 388-398.

148. Jager E., Jager D., Knuth A., *Int. J. Cancer,* **2003,** 106, 817-820.

149. Purcell A. W., Gorman J. J., *Molecular and Cellular Proteomics*, **2004**, 3, 193-208.

150. Falk K., Rotzcheke O., Rammensee H-G., *Nature,* **1990,** 348, 248-251.

151. Deininger M.W.N., Vieira S., Mendiola R., Schultheis B., Goldman J. M., Melo J. V., *Cancer Res.,* **2000,** 60, 2049-2055.

152. Parker K.C., Bednarek M.A., Coligan J.E., *J. Immunol.*, **1994,** 59, 1-14.

153. Traversari C; Vanderbruggen P; Luescher IF; Lurquin C; Chomez P; VanPel A; DePlaen D; Amar-Costesec A; Boon T; J. Exp. Med., 1992, 176, 1453-1457

154. Chaux P; Luiten R; Demotte N; Vantomme V; Stroobant V; Traversari C; Russo V; Shultz E; Cornelis G.R; Boon T; Vanderbruggen P; *J. Immunol.*, **1999**, 163, 2928-2936.

155. Gulukota K., Sidney J., Sette A., DeLisi C., *J. Mol. Biol.*, **1997**, 267, 1258-1267.

156. Hogan K.T., Eisinger D.P., Cupp S.B., Lekstrom K.J., Deacon D.D., Shabanowitz J., Hunt D.F., Engelhard V.H., Slingluff C.L., Ross M.M., *Cancer Res.*, **1998**, 58, 5144-5150.

157. Anderson M.H., Bonfill J.E., Neisig A., Arsequell G., Sondergaad I., Neefjes J., Zeuthan J., Elliott T., Haurum J.S., *J. Immunol.*, **1999**, 163, 3812-3818.

158. Harum J.S., Hoier I.B., Arsequell G., Neisig A., Valencia G., Zeuthen J., Neefjes J., Elliott T., *J. Exp. Med.*, **1999**, 190, 145-150.

159. Ogueta S., Rogado R., Marina A., Moreno F., Redondo J. M., Vazquez J., *J. Mass Spectrom.*, **2000**, 35, 556-565.

160. Zarling A.L., Ficarro S.B., White F.M., Shabanowitz J., Hunt D.F., Engelhard V.H., *J. Exp. Med.*, **2000**, 192, 1755-1762.

161. Barnea E., Beer I., Patoka R., Ziv T., Kessler O., Tzehoval E., Eisenbach L., Zavazava N., Admon A., *Eur. J. Immunol.*, **2001,** 31, 213-222.

162. Schirle M., Keilholz W., Weber B., Gouttefangeas C., Dumrese T., Becker H. D., Stevanovic S., Rammensee H. G., *Eur. J. Immunol.*, **2000,** 30, 2216-2225.

163. Traversari C., Vanderbruggen P., Luescher I.F., Lurquin C., Chomez P., Vanpel A., Deplaen E., Amarcostesec A., Boon T., *J. Exp. Med.*, **1992,** 176, 1453-1457.

164. Brichard V., Vanpel A., Wolfel T., Deplaen E., Lethe B., Coulie P., Boon T., *J. Exp. Med.*, **1993,** 178, 489-495.

165. Coulie P.G., Brichard V., Vanpel A., Wolfel T., Schneider J., Traversari C., Mattei S., Deplaen E., Lurquin C., Szikora J.P., Renauld J. C., Boon T., *J. Exp. Med.*, **1994,** 180, 35-42.

166. Rotzschke O., Falk K., Deres K., Schild H., Norda M., Jung J., Rammensee H.-G., *Nature*, **1990**, 348, 252-254

167. Falk K., Rotzschke O., Deres K., Metzger J., Jung G., Rammensee H.-G., *J. Exp. Med.*, **1991,** 174, 425-434.

168. Falk K., Rotzschke O., Stevanovic S., Jung G., Rammensee H., *Nature*, **1991,** 351, 290-296.

169. Rammensee H. G., Falk K., Rotzschke O., *Ann. Rev. Immunol.*, **1993,** 11, 213-244.

170. Rammensee H.-G., Friede T., Stevanovic S., *Immunogenetics*, **1995,** 41, 178-228.

171. Hunt D.F., Henderson R. A., Shabanowitz J., Sakaguchi K., Michel H., Sevilir N., Cox A. L., Apella E., Engelhard V. H., *Science*, **1992,** 255, 1261-1263.

172. Cox A.L., Skipper J., Chen Y., Henderson R. A., Darrow T. L., Shabanowitz J., Engelhard V. H., Hunt D. F., Slingluff C. L., *Science*, **1994,** 264, 716-719.

173. Storkus W. J., Zeh H. J., Salter R. D., Lotze M. T., *J. Immunother.* **1993,** 14, 94-103.

174. Castelli C; Storkus W.J; Maeurer M.J; Martin D.M; Huang E.C; Pramanik B.N; Nagabhushan T.L; Parmiani G; Lotze M.T; *J. Exp. Med.*, **1995**, 181, 363-368.

175. Clark R.E., Dodi I.A., Hill S.C., Lill J.R., Aubert G., Macintyre A.R., Rojas J., Bourdon A., Bonner P.L.R., Wang L., Christmas S.E., Travers P.J., Creaser C.S., Rees R.C., Madrigal J.A., *Blood,* **2001,** 98, 2887 - 2893.

176. Schirle M., Kielholz W., Weber B., Gouttefangeas C., Dumrese T., Becker H. D., Stevanovic S., Rammensee H-G., *Eur. J. Immunol.*, **2000**, 30, 2216-2225.

177. Pascolo S., Schirle M., Guckel B., Dumrese T., Stumm S., Kayser S., Moris A., Wallweiner D., Rammensee H-G., Stevanovic S., *Cancer Res.*, **2001,** 61, 4072-4077.

178. van der Heeft E., ten Hove G. J., Herberts C. A., Meiring H. D., van Els C. A. C. M., de Jong A. P. J. M., *Anal. Chem.*, **1998,** 70, 3742-3751.

179. Zhou M; Peng J-R; Zhang H-G; Wang H-X; Zhoung Z-H; Pan X-Y; Chen W-F; Leng X-S; *Immunology Letters*, **2005**, 99, 113-121.

180. Lemmel C; Weik S; Eberle U; Dengjel J; Kratt T; Becker H-D; Rammensee HG; Stevanovic S; *Nature Biotechnology*, **2004**, 22, 450.

181. DiMarzo Veronese F; Arnott D; Barnaba V; Loftus DJ; Sakaguchi K; Thompson C.B; Salemi S; Mastroianni C; Sette A; Shabanowitz J; Hunt D.F; Appella E; *J. Exp. Med.*, **1996,** 183, 2509.

182. Cox A.L; Huczko E.L; Engelhard V.H; Shabaonwitz J; Hunt D.F; *MHC Volume 1: A practical Approach*, **1997,** IRL Press, Oxford.

183. Hoffman S; Gluckmann M; Kausche S; Schmidt A; Corvey C; Lichtenfels R; Huber C; Abrecht C; Karas M; Herr W; *Molecular and Cellular Proteomics*, **2005**, 4, 1888-1897

184. Dubey P; Hendrickson R.C; Meredith S.C; Siegel C.T; Shabanowitz J; Skipper J.C.A; Engelhard V.H; Hunt D.F; Schreiber H; *J. Exp. Med.*, **1997**, 185, 695-705

185. Skipper J.C.A.; Hendrickson R.C; Gulden P.H; Brichard V; vanPel A; Chen Y; Shabanowitz J; Wolfel T; Slingluff C.L; Boon T; Hunt D.F; Engelhard V.H; *J. Exp. Med.*, **1996,** 183, 3527-3534;

186. Skipper J.C.A; Kittlesen D.J., Hendrickson R.C; Deacon D.D., Harthun N.L., Wagner S.N., Hunt D.F., Engelhard V.H., Slingluff C.L., *J. Immunol.*, **1996,** 157, 5027-5033.

187. Skipper J.C.A; Gulden P.H; Hendrickson R.C; Harthun N; Caldwell J.A; Shabanowitz J; Engelhard V.H; Hunt D.F; Slingluff C.L; *Int. J. Cancer*, **1999,** 82, 669-677.

188. Hogan K.T; Coppola M.A; Gatlin C.L; Thompson L.W; Shabanowitz J; Hunt D.F; Engelhard V.H; Slingluff C.L; Ross M.M; *Immunol. Letters*, **2003**, 90, 131-135.

189. Bakker A.B.H., Schreurs M.W.J., Deboer A.J., Kawakami Y., Rosenberg S.A., Adema G.J., Figdor C.G., *J. Exp. Med.*, **1994,** 179, 1005-1009.

190. Rosenberg S.A., Yang J.C., Schwartzentruber D.J., Hwu P., Marincola F.M., Topalian S.L., Restifo N.P., Dudley M.E., Schwarz S.L., Spiess P.J., Wunderlich J.R., Parkhurst M.R., Kawakami Y., Seipp C.A., Einhorn J.H., White D.E., *Nat. Med.*, **1998,** 4, 321-327.

191. Gjertsen M.K., Bakka A., Breivik J., Saeterdal I., Solheim B. G., Soreide O., Thorsby E., Gaudernack G., *Lancet,* **1995,** 346, 1399-1400.

192. Jaeger E., Bernhard H., Romero P., Ringhoffer M., Arand M., Karbach J., Ilsemann C., Hagedorn M., Knuth A., *Int. J. Cancer,* **1996,** 66, 162-169.

193. Nestle F.O., Alijagic S., Gilliet M., Sun Y. S., Grabbe S., Dummer R., Burg G., Schadendorf D., *Nat. Med.* **1998,** 4, 328-332.

194. Hsu F.J., Benike C., Fagnoni F., Liles T. M., Czerwinski D., Taidi B., Engleman E. G., Levy R., *Nat. Med.*, **1996,** 2, 52-58.

195. Nestle F.O., Banchereau J., Hart D., *Nat. Med.*, **2001,** 7, 761-765.

196. Lau R., Wang F., Jeffery G., Marty V., Kuniyoshi J., Bade E., Ryback M. E., Weber J., *J. Immunother.*, **2001,** 24, 66-78.

197. Sadanaga N., Nagashima H., Mashino K., Tahara K., Yamaguchi H., Ohta M., Fujie T., Tanaka F., Inoue H., Takesako K., Akiyoshi T., Mori M., *Clin. Cancer Res.*, **2001,** 7, 2277-2284.

198. Yu J.S., Wheeler C.J., Zeltzer P.M., Ying H., Finger D.N., Lee P.K., Yong W.H., Incardona F., Thompson R.C., Riedinger M.S., Zhang W.X., Prins R.M., Black K.L., *Cancer Res.*, **2001,** 61, 842-847.

199. Fong L., Brockstedt D., Benike C., Breen J. K., Strang G., Ruegg C. L., Engleman E. G., *J. Immunol.*, **2001,** 167, 7150-7156.

200. Rowley J.D., *Nature*, **1973**, 243, 290-293

201. de Klein A., van Kessel A.G., Grosveld G., Bartram C.R., Hagemeijer A., Bootsma D., Spurr N.K., Heisterkamp N., Groffen J., Stephenson J.R., *Nature*, **1982,** 300, 765-767

202. Heisterkamp N. Stephenson J.R., Groffen J., Hansen P.F., de Klein A., Bartram C.R., Grosveld G., *Nature*, **1983,** 306, 239-241

203. Shtivelman E., Lifshitz B., Gale R.P., Canaani E., *Nature*, **1985,** 315, 550-554

204. Shtivelman, E., Lifshitz B., Gale R.P., Roe B. A., Canaani E., *Cell*, **1986,** 47, 277-284

205. Melo J. V., *Blood*, **1996,** 88, 2375-2384.

206. Davis R. L., Konopka J. B., Witte O. N., *Mol. Cell. Biol.* **1985,** 5, 204-213

207. Daley G. Q., Vanetten R. A., Baltimore D., *Science*, **1990,** 247, 824-830.

208. Kelliher M. A., McLaughlin J., Witte O. N., Rosenberg N., *Proc. Natl. Acad. Sci* **1990,** 87, 6649-6653.

209. Chronic myeloid leukaemia-Leukaemia research fund booklet.

210. Oda T., Tamura S., Matsuguchi T., *Leukaemia*, **1995,** 9, 295-301

211. Lugo T.G., Pendergast A.M., Muller A.J., Witte O.N., *Science*, **1990,** 247, 1079-1081.

212. Schindler T., Bornmann W., Pellicena P., Miller W. T., Clarkson B., Kuriyan J., *Science*, **2000,** 289, 1938-1942.

213. Sawyers C.L., Hochhaus A., Feldman E., Goldman J.M., Miller C.B., Ottmann O. G., Schiffer C.A., Talpaz M., Guilhot F., Deininger M.W.N., Fischer T., O'Brien S.G., Stone R.M., Gambacorti-Passerini C.B., Russell N.H., Reiffers J.J., Shea T.C., Chapuis B., Coutre S., Tura S., Morra E., Larson R.A., Saven A., Peschel C., Gratwohl A., Mandelli F., Ben-Am M., Gathmann I., Capdeville R., Paquette R.L., Druker B. J., *Blood*, **2002**, 99, 10, 3530-3539.

214. Shah, N. P., Tran C., Lee F. Y., *Science*, **2004,** 305, 399-401.

215. Bocchia M., *Haematologica* **2002**, 87, 675-677.

216. Bocchia M., Korontsvit T., Xu Q., Mackinnon S., Yang S.Y., Sette A., Scheinberg D. A., *Blood*, **1996,** 87, 3587-3592.

217. Bosch G.J., Joosten A.M., Kessler J.H., Melief C.J., Leeksma O.C., *Blood,* **1996,** 88, 3522-3527.

218. Pinilla-Ibarz J., Cathcart K., Korontsvit T., Soignet S., Bocchia M., Caggiano J., Lai L., Jimenez J., Kolitz J., Scheinberg D. A., *Blood,* **2000,** 95, 1781-1787.

219. Bocchia M., Gentili S., Abruzzese E., Fanelli A., Iuliano F., Tabilo A., Amabile M., Forconi F., Gozzetti A., Raspadori D., Amadori S., Lauria F., *The Lancet,* **2005,** 365, 657-662.

CHAPTER 2

DEVELOPMENT OF MS AND CAPILLARY HPLC/MS FOR THE ANALYSIS OF PEPTIDES

Chapter 2 Aims Breakdown

In order to be able to successfully analyse MHC associated peptides, obtained via cell surface elution using the published method,[1] it was necessary to study the chromatographic and mass spectrometric conditions. The low levels of peptide which are anticipated for MHC peptides necessitate that the maximum possible sensitivity be obtained from the LC/MS system. It was necessary, therefore, to optimise chromatographic separation and mass spectrometric fragmentation to allow a successful isolation of the desired species followed by the maximum transmission of peptide ions which would be most useful in sequence determination.

In order to do this it was necessary to address the following points.

- The peptide fragmentation conditions.

  Firstly the charge state of the peptide ion ($[M+H]^+$, $[M+2H]^{2+}$ or $[M+3H]^{3+}$) would be studied to establish which would provide the most beneficial fragmentation pattern. Following this work, optimum conditions would then need to be established in order to generate a predominance of these ions by electrospray ionisation.

- The mass spectrometry ionisation source.

  The ionisation source was to be studied for two reasons. Firstly, as a chromatographic separation of low flow rate was desired, a

complementary electrospray ionisation source needed to be developed that could accept such flow rates. Secondly a low flow rate electrospray source is known to provide improved sensitivity based on their ability to generate smaller solvent droplets and better desolvation.[2]

- Instrumental sensitivity determined by chromatographic separation.

  As the sensitivity of the LC/MS instrument is dependent, to a certain extent, on the degree of chromatographic dilution work was carried out on the separation of analyte species employing low flow rates.[3]

- Degree of sample volume loading.

  It was necessary to configure a system which could accommodate the loading of large sample volumes (~1ml) which are expected from the previously developed MHC class I clean-up method.[1] A compatibility issue exists between the flow rate used in capillary LC and the injection volume requirements of the method. This conflict needed to be resolved.

The study, and compromise, of these four points would determine the optimum instrumental conditions for the analysis and identification of MHC class I and II peptides.

2.1.0. Introduction

The use of mass spectrometry for the analysis of peptides is a popular and ever growing field of interest. Whilst conditions for the analysis of tryptic peptides have been investigated, there have been few similar studies for MHC.[4] Different ionisation techniques are available to allow the study of peptide molecules, and these have been discussed in Chapter one, however this thesis will only be concerned with analysis using the technique of electrospray ionisation (ESI/MS).

Under ESI/MS conditions it is possible, and common, to encounter doubly protonated peptides ($[M+2H]^{2+}$) which undergo a charge separation reaction giving rise to an intense fragmentation series of usually singly charged b and y ions.[5,6] Protein identification is usually performed by characterisation of peptide products resulting from the tryptic digestion of the protein in question. Trypsin is favoured as the cleavage results in the presence of a basic amino acid residue at the C-terminus.[7] The presence of the basic residues encourages the formation of a doubly charged parent ion which fragments to form a series of y ions, making the spectra easier to interpret.[8]

A number of different factors are thought to influence the fragmentation of peptides. The use of collisionally induced dissociation, which is a relatively low energy technique, tends to favour b and y ion formation. Tabb *et al*[9] found that the amino acid residue adjacent to the amide bond influences its cleavage, for example, a proline residue enhances cleavage at the N-terminal amide bond and

fragments with loss of ammonia are common where Gly/Asn residues are present.[10] In small peptides, the number of basic sites determines the maximum charge state of the ion (Arg, Lys, His are all more basic than the peptides free N-terminus). The mobile proton model proposes that peptide dissociation results from charge directed cleavages initiated by intramolecular proton transfers. Fragmentation of most protonated peptides requires the involvement of a proton at the cleavage site, that cleavage is charge directed.[11,12] The ease of fragmentation is dependent on peptide composition and size, doubly charged peptides will fragment more easily than singly charged. Peptides with basic amino acid residues require more energy to dissociate, the mobile proton explanation for this is that a basic residue will have a strong affinity for the proton and therefore to move it from the side chain to the backbone to initiate charge directed cleavage and produce b and y ions more energy will be required.[13] The dissociation energy is greater for Arg containing peptides than Lys, mirroring the order of decreasing gas phase basicity.[11] Fragmentation of the highest possible charge state of the peptide can lead to the most informative fragmentation. However, if a single proton is fixed to a basic amino acid there remains no mobile proton to direct fragmentation which is therefore limited, even at severe low energy conditions. If the charge state of the peptide is greater than two then the location of additional non-terminal basic residues is important due to its effect on the migration of the mobile proton. The mobile proton will not reside close to another fixed proton therefore little fragmentation will be seen at these positions.[14] The mobile proton theory is illustrated in Figure 2.1.0.1 for the doubly charged peptide ATSFYK.

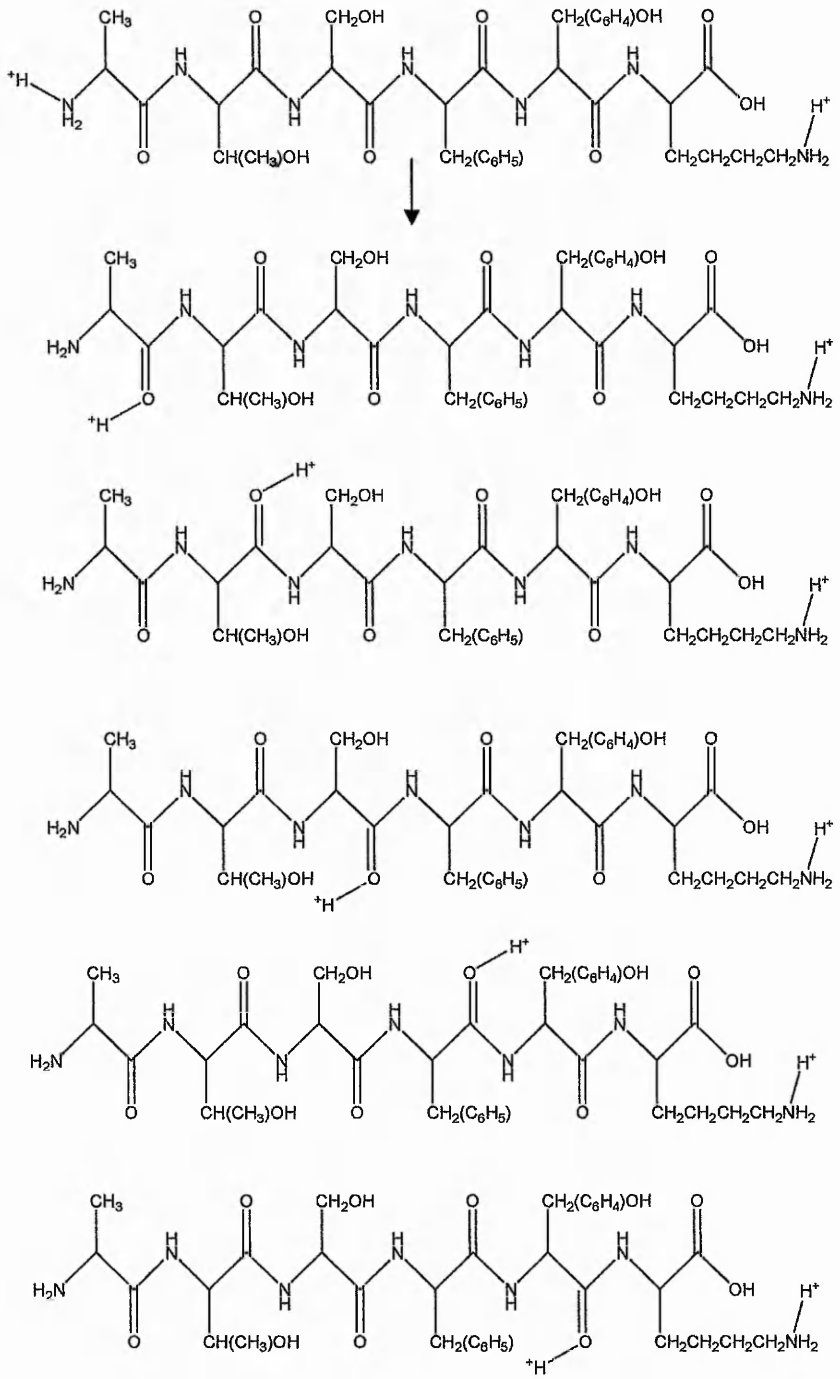


Figure 2.1.0.1. The peptide (ATSFYK) has two basic sites and so is doubly protonated when ionised via electrospray. The proton on the C-terminal lysine residue is fixed due to the basicity of that site. The proton attached to the N-terminal amine is not fixed and can migrate down the peptide chain by internal solvation resulting in a heterogeneous mixture of peptides with different sites of protonation along the same peptide sequence.

This also explains the simplicity of doubly protonated tryptic peptide fragmentation as, under low energy conditions, one proton remains associated with the most basic residue with the remaining proton free to initiate charge induced dissociation along the peptide backbone.[15]

The formation of the parent ion and the number of charges it will support is dependent on a number of factors during the electrospray process. Wang and Cole discussed composition and pH of the analyte solution. At low pH, <6.5, the calculated solution ratio of doubly : singly charged ions is much greater than that seen in the gas phase for the peptides bradykinin and gramicidin S.[16] Coulombic repulsion and attraction of counter ions were hypothesised to be responsible for this together with the protonation of non-aqueous solvent molecules. The application of the work was intended to be on-line LC/MS analysis therefore the solvent composition was not experimented with and was intended to mimic, as closely as possible, the post-column eluent. Instrument parameters were studied to determine the optimum conditions for the analysis of these peptides.

The temperature of the mass spectrometers heated capillary and the tube lens offset voltage were both studied in relation to the formation of peptide ions. The effect of these parameters on the charge state of the MHC class I and II peptide ions and the resulting fragmentation spectra obtained were both explored. Other variable such as capillary voltage and the voltages applied to the octapoles and lenses in the pre trap region can be collectively 'tuned' automatically by the software for the transmission of particular species. Figure 2.1.0.2 is a schematic

representation of the ion stack which illustrates the location of the heated capillary and the tube lens.

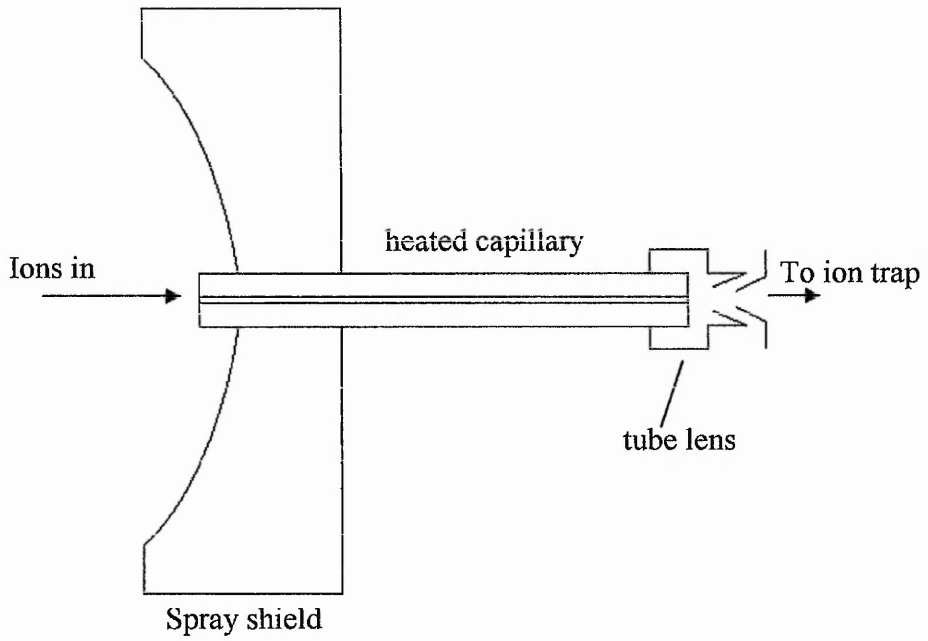


Figure 2.1.0.2 Schematic representation of the atmospheric pressure ionisation (API) stack demonstrating the location of the heated capillary and tube lens.

The potential advantages of micro/nano-spray ionisation over conventional electrospray have been discussed in Chapter 1 (section 1.1.1). Reduction in the spraying orifice size results in smaller droplet formation and better desolvation during the electrospray process thereby improving sensitivity at the detector. Enhancing sensitivity is desirable considering the low level of MHC peptides which is expected.

One of the aims of this work was to develop on-line LC/MS analysis to increase sample throughput and automation leading to a more robust system. The use of

capillary HPLC was therefore necessary in terms of compatibility of flow rate when coupled to a μ-spray ionisation source. It also benefits for an increase in sensitivity due to a reduction in chromatographic dilution as discussed in Chapter 1 (section 1.1.4).

One limitation of capillary HPLC involves the reduction in sample injection volume. It was known that the MHC class I method which had been previously developed resulted in samples in excess of 1 mL. It was therefore necessary to configure a system which would allow the loading of large sample volumes.

2.2.0. Materials and Methods

Investigations were performed into the optimum instrumental conditions for peptide identification by nano-spray ionisation mass spectrometry. Peptides were analysed using an in-house constructed static nano-spray ionisation source with unbroken nano-spray tips (Protana, Denmark), coupled to a Finnigan LCQ quadrupole ion trap mass spectrometer (QITMS) (ThermoFinnigan, Hemel Hempstead, UK). The tips were broken by gently touching them against the surface of the mass spectrometers heated capillary to one side of the inlet orifice using the xyz manipulator onto which the nanospray source was mounted.

Optimisation of mass spectrometric conditions was carried out by sustained spraying from a nanospray tip. This was done to ensure that optimisation was performed on a steady and continuous signal which would allow the relative intensities of ions to be obtained from averaged data of approximately 50 μscans. Trends in collision energy and fragmentation spectra remain consistent. Similarly the trends observed in peptide mass spectra with variation in capillary temperature and tube lens voltage remain applicable in the long term. As such replicate analyses were unnecessary, as was the repetition of the exercise on a different day.

Investigation of the effect of spray tip internal diameter was performed by infusion of peptide solutions at a flow rate of 3 μl/min. In each case peptide signal duration was up to 10 minutes which allowed the spray to stabilise and the

averaging of mass spectra over multiple μscans to compensate for scan to scan variations across the whole scan range of the mass spectrum. Once again this eliminated the need for replicate optimisation.

2.2.1. Effect of peptide charge state on collision energy and fragmentation spectra

The effect of charge state of the parent ion on collision energy required to fragment the ion and fragmentation pattern observed was explored for an 11mer peptide (GFKQSSKALQR, MWt 1248.7 $gmol^{-1}$, 100 pmol μl^{-1} in 50/50 % v/v acetonitrile/H_2O + 1 % acetic acid) and an 18mer peptide (ATGFKQSSKALQRPVASD, MWt 1890.2 $gmol^{-1}$ 100 pmol μl^{-1} in 50/50 % v/v acetonitrile/H_2O + 1 % acetic acid). Both of these peptides represent sequences from the bcr/abl protein junctional region. Analysis was performed using a static nanospray source with Au/Pd coated nano-spray tips (Nano-ES tip, Protana, Denmark).

The ion trap experimental conditions were as follows:

Sheath gas, 20 arb units; Auxillary gas, 0 arb units; Capillary temperature, 180 °C; Spray voltage, 1.1 kV; Capillary voltage, 46 V; Tube lens offset, -40 V; Collision energy, variable.

2.2.2. Optimisation of mass spectrometric tube lens voltage and heated capillary temperature

The effect of mass spectrometric conditions, specifically tube lens offset voltage and capillary temperature, on the distribution of charge state of peptide parent ions were investigated. Analysis was performed using a static nanospray source with Au/Pd coated nano-spray tips (Nano-ES tip, Protana, Denmark). Capillary temperature was allowed to equilibrate prior to analysis.

The ion trap experimental conditions were as follows:
Sheath gas, 20 arb units; Auxillary gas, 0 arb units; Capillary temperature, 180 °C / 250 °C; Spray voltage, 1.1 kV; Capillary voltage, 46 V; Tube lens offset, variable.

These conditions were then tested on a MHC class I predicted 9mer peptide (GFKQSSKAL, MWt 964.4 $gmol^{-1}$ 100 pmol μl^{-1} in 50/50 % v/v acetonitrile/H_2O + 1 % acetic acid) to ensure that these were the optimal conditions for the generation of useful mass spectra. The ion trap experimental conditions are as above with the exception of tube lens offset, -40 V; Capillary temperature, 180 °C.

2.2.3. Development and evaluation of a microspray ion source

An on-line miniature electrospray (microspray) source was constructed to accommodate the low flow required for HPLC columns with an internal diameter

of ≤300 μm. The source is shown schematically in Figure 2.2.3.1. An xyz manipulator with a PTFE mounting and a plastic tube holder was used to support and insulate a zero dead volume stainless steel union (SGE, Australia), which was used to connect a fused silica transfer line (100 μm i.d.) to the spraying tip. A voltage was then applied directly to the union.

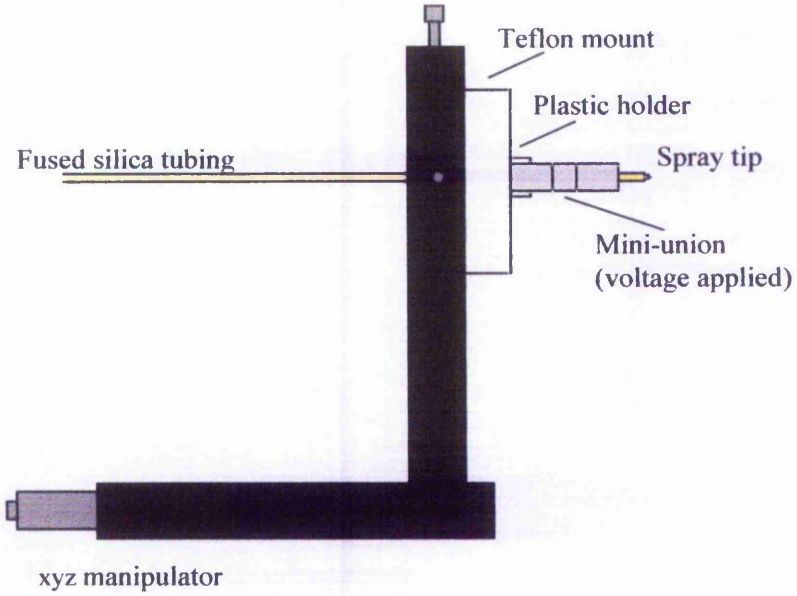


Figure 2.2.3.1 Schematic representation of the in-house constructed microspray source.

A comparison was made between the sensitivity of the microspray ion source and the normal LCQ ESI source, using a range of fused silica tip internal diameters. Flow injection analysis was used to introduce peptide solution (sequence ATGFKQSSKALQRPVASD, 1.0 pmol μl^{-1} in 50/50 % v/v acetonitrile/H_2O + 1 % acetic acid) from a 10 μL injection loop into a flow of 50/50 % v/v acetonitrile/H_2O + 1 % acetic acid. This flow was then pumped into the LCQ microspray ion source at a flow rate of 3 μl min^{-1} using the integral syringe pump. Full scan MS and MS/MS data were obtained for the Finnigan ESI source

and for the microspray source fitted with tips with internal diameters of 100 μm, 50 μm and 25 μm.

The ion trap experimental conditions were as follows: auxiliary gas, 0 arb units; capillary temperature, 180 °C; capillary voltage, 46 V; tube lens offset, -40 V; collision energy, 20 %. Finnigan ESI source: sheath gas, 60 arb units; spray voltage, 4.5 kV. Microspray source: sheath gas, off; spray voltage, 3.5 kV.

2.2.4. Nano-LC/μ-ESI/MS

An interface was constructed for use with a nano-HPLC system to allow on-line HPLC separation followed by micro-electrospray ionisation ion trap mass spectrometry. The microspray ESI source previously described (section 2.2.3) was modified to allow for sub μl min^{-1} flow rates by replacing the 100 μm i.d. inlet tubing with a 75 μm i.d. capillary LC column and 20 μm i.d. fused silica capillary tubing. The spray tip was replaced with a conductively coated tip of 20 μm i.d. pulled to a 5 μm i.d. spray orifice (New Objective, Massachusetts, USA).

Experimental conditions were as follows:

Sample solution: Peptide gp70 $SPSYVYHQF_{(454-462)}$ 18 fmol μl^{-1} in 0.01 % TFA(aq). Injection volume: 1 μl. Column: Pepmap 75 μm i.d. x 150 mm C18 nanobore column (Dionex, Surrey, UK). Mobile Phase A: H_2O + 0.09 % formic acid. Mobile Phase B: Acetonitrile + 0.08 % formic acid. Flow rate: 200 nl min^{-1}.

Gradient profile, 100 % A to 40 % B in 60 minutes. Pump: Ultimate™ nano LC pump (Dionex, Surry, UK).

Mass spectrometry was performed using on-line micro-electrospray ionisation mass spectrometry on a Finnigan LCQ ion trap (ThermoFinnigan, Hemel Hempstead, UK). Mass spectrometric conditions were as follows: spray voltage, 2.0 kV; capillary temperature, 180 °C; capillary voltage, 46 V; tube lens offset, -40 V; No sheath gas or auxiliary gas required.

2.2.5. Introduction of large sample volumes

A trap column, Pepmap 300 µm i.d. x 5 mm, (Dionex, Surrey, UK) was incorporated into the capillary LC system by modifying the Ultimate LC pump injection valve to allow the loading and preconcentration of large sample volumes. The modified injection valve configuration is shown schematically in Figure 2.2.5.1.

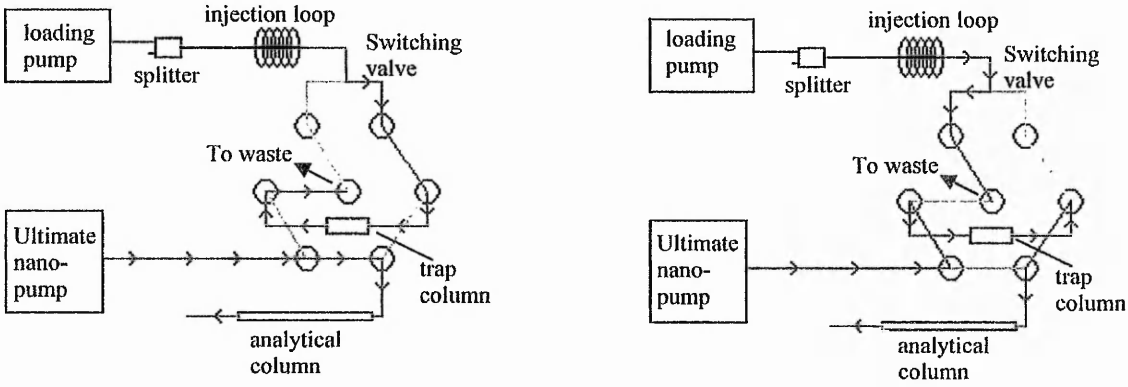


Figure 2.2.5.1 Schematic of the Ultimate switching valve configured for the loading of large sample volumes. (a) Shows the valve in the load configuration with the flow of

sample/solvent shown in red. (b) Shows the valve in the inject configuration with the flow of sample/solvent shown in blue.

Loading of the trap column was achieved by pumping the sample solution from the injection loop in 0.1 % TFA at a flow rate of 20 µl min^{-1}. Varying the loop volume allowed a sample volume of up to 1000 µl to be introduced in approximately 50 minutes. Using a Waters 501 HPLC pump operated at 0.6 ml min^{-1} which was reduced to 20 µL min^{-1} using a flow splitter (SGE, Australia). After sample loading was complete, the valve was switched and the gradient activated to allow elution of the analytes from the trap column onto the analytical column using the nano LC pump. The configuration was tested with the gp70 peptide $SPSYVYHQF_{(454-462)}$, 10 µl injections of peptide (1.8 fmol μl^{-1} in 0.1 % TFA) and 1000 µl injections of peptide (1.8 fmol total in 0.1 % TFA). All Ultimate™ and LCQ conditions were as described above (Section 2.2.5) except for the solvent gradient which was set to change from 90 - 40 % A in 90 minutes.

2.3.0. Results and Discussion

Investigations were performed into the optimum instrumental conditions for peptide identification by nano-spray ionisation mass spectrometry.

2.3.1. Effect of peptide charge state on collision energy and fragmentation spectra

The effect of peptide charge state on fragmentation has been investigated for tryptic peptides. Poor fragmentation efficiencies were observed in low energy collisionally induced dissociation of singly protonated peptides that contain strongly basic residues.[17]

Substantial differences in the extent of collisionally induced fragmentation are frequently observed for multiply protonated ions and their singly protonated counterparts. Investigations into peptide fragmentation have shown that doubly charged peptides fragment differently to their singly charged counterparts. The role of ionising protons and basic residues in the fragmentation process in terms of their location and interaction within the peptide ion has also been explored.[18-22]

The effect of the charge state of the peptide on collision energy required to fragment it and on the fragmentation spectra produced was investigated for three peptides composed of nine, eleven and eighteen amino acids, approximating

MHC class I and II length, originating from the junctional region of the bcr/abl fusion protein. A solution of an MHC class II 18mer peptide (ATGFKQSSKALQRPVASD, MWt 1890.2 $gmol^{-1}$ 100 pmol μl^{-1} in 50/50 % v/v acetonitrile/H_2O) was initially analysed using the nano-spray source. The doubly and triply charged ions were both subjected to collisionally induced dissociation to determine the optimum charge state for peptide fragmentation. The singly charged ion was not investigated as it has been shown that singly charged parent ions which contain one strongly basic residue fragment poorly due to the sequestering of the proton by the basic residue. In the case of the peptide RLICFSCFR, fragmentation of the singly charged parent ion (generated by FAB) gives poor fragmentation due to sequestering of the proton by one of the arginine residues.[23] The breakpoint region consists of a unique lysine (K) residue therefore all of the likely bcr/abl targets will fall into this category.[17] A solution of an MHC class I 11mer peptide (GFKQSSKALQR, MWt 1248.7 $gmol^{-1}$ 100 pmol μl^{-1} in 50/50 % v/v acetonitrile/H_2O) was subjected to the same analysis. Data was acquired continuously using nanospray tips loaded with the peptide solutions. The data was averaged over a period of *ca.* 50 μscans to account for slight variations in signal intensity with time. The MS and tandem MS data for the 18mer and 11mer peptides can be seen in Figures 2.3.1.1 and 2.3.1.2 respectively. For the 18 mer peptide a collision energy of 19 % was required to fragment the triply charged parent ion and a collision energy of 27 % for the doubly charged parent ion. Similarly for the CID of the 11mer peptide, a lower collision energy was required to fragment the triply charged ion compared with the doubly charged ion (15 % and 20 % respectively). Ions with greater charge

accelerated through the same potentials as ions of lower charges have greater energies so collision energy requirements may be reduced for multiply charged ions.[24] It has also been shown that triply charged ions experience 50 % higher collision activation than that for the doubly charged ions for a given potential difference due to their increased coulombic repulsion.[25] Figures 2.3.1.1 and 2.3.1.2, panels b and c, show the resulting tandem mass spectra for these two ions for the 18 mer peptide and 11 mer peptide respectively. It can clearly be seen that greater coverage of b and y ions is observed upon fragmentation of the doubly charged ion. Intensity of the product ions observed during CID of the $[M+2H]^{2+}$ in the high mass region, m/z 1000-2000, is greater than for the product ions of the triply charged, $[M+3H]^{3+}$, ions. For example in Figure 2.3.1.1 the b_{12}, b_{13} and y_{13} y_{14} (m/z 1247-1515) ions are clearly present in the fragmentation spectrum of the $[M+2H]^{2+}$ ion (c) but are absent in the fragmentation spectra of the $[M+3H]^{3+}$ ion (b). A similar pattern is seen for the 11mer peptide in Figure 2.3.1.2. The b_7 (m/z 789.4) and y_8 (m/z 917.5) ions are present in both MS/MS spectra, however the relative intensity of these ions in the product ion spectrum of the $[M+2H]^{2+}$ (c) ion is between 5 and 10 times greater than is observed in the product ion spectrum of the $[M+3H]^{3+}$ ion (b). The wider sequence coverage which can be obtained will allow greater confidence in the identification of unknown peptide sequences. It has been reported[26,27] that for the peptide RLCIFSCFR, fragmentation of the triply charged ions yields product ions with an efficiency two orders of magnitude greater than for the doubly charged ion. However, this peptide is unusual in that both Arginine residues are at the peptide terminal ends and therefore the mobile proton will not be hindered by the

presence of sequestered protons in the body of the peptide. In triply charged peptides the location of the basic residues is important as it will affect the migration of the mobile proton. The proton will not reside alongside another and therefore fragmentation at these sites will be limited.[14] This may explain the greater sequence ion coverage which was observed for the $[M+2H]^{2+}$ ion than the $[M+3H]^{3+}$ ion during these experiments.

To summarise, with both the 18mer and 11mer peptides, fragmentation of the doubly charged peptide ion yielded the most structural information about the peptide, so for both class I and class II peptides the $[M+2H]^{2+}$ ion was targeted in all subsequent experiments. As discussed in section 2.1.0, investigations of peptide dissociation have focused predominantly on tryptic peptides, these have been shown to generate predominantly doubly charged parent ions during ESI/MS due to the presence of C-terminal Lys/Arg residues.[8] The MHC processing pathway however does not guarantee the same cleavage therefore the fragmentation spectra which occur may be of greater complexity. In these examples extensive fragmentation is seen and as with the tryptic peptides the y series of ions predominates, b ions are also seen and these aid in identification. Coincidently, the 11mer peptide terminates in an Arg residue although the 18mer peptide does not. Multiply charged ions have higher kinetic energies than their singly charged counterparts, consequently when accelerated with the same potential they will retain more energy for electronic and vibrational excitation during collision.[13, 15, 28]

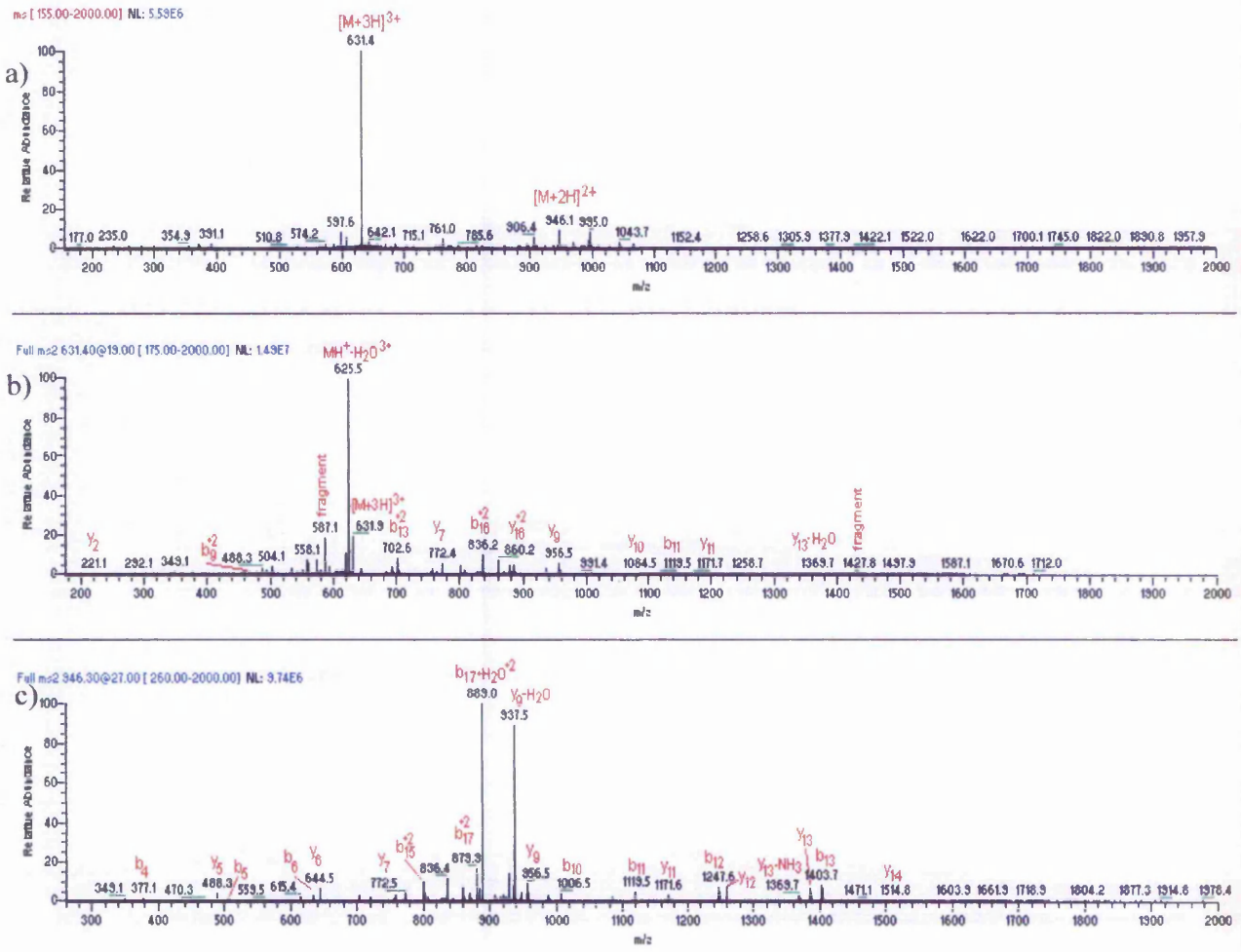


Figure 2.3.1.1 a) Full scan mass spectrum of peptide ATGFKQSSKALQRPVASD showing $[M+2H]^{2+}$ (m/z 946.1) and $[M+3H]^{3+}$ (m/z 631.4) ions. b) tandem mass spectrum of the $[M+3H]^{3+}$ ion. c) tandem mass spectrum of the $[M+2H]^{2+}$ ion. Mass spectra were averaged over *ca.* 50 µscans.

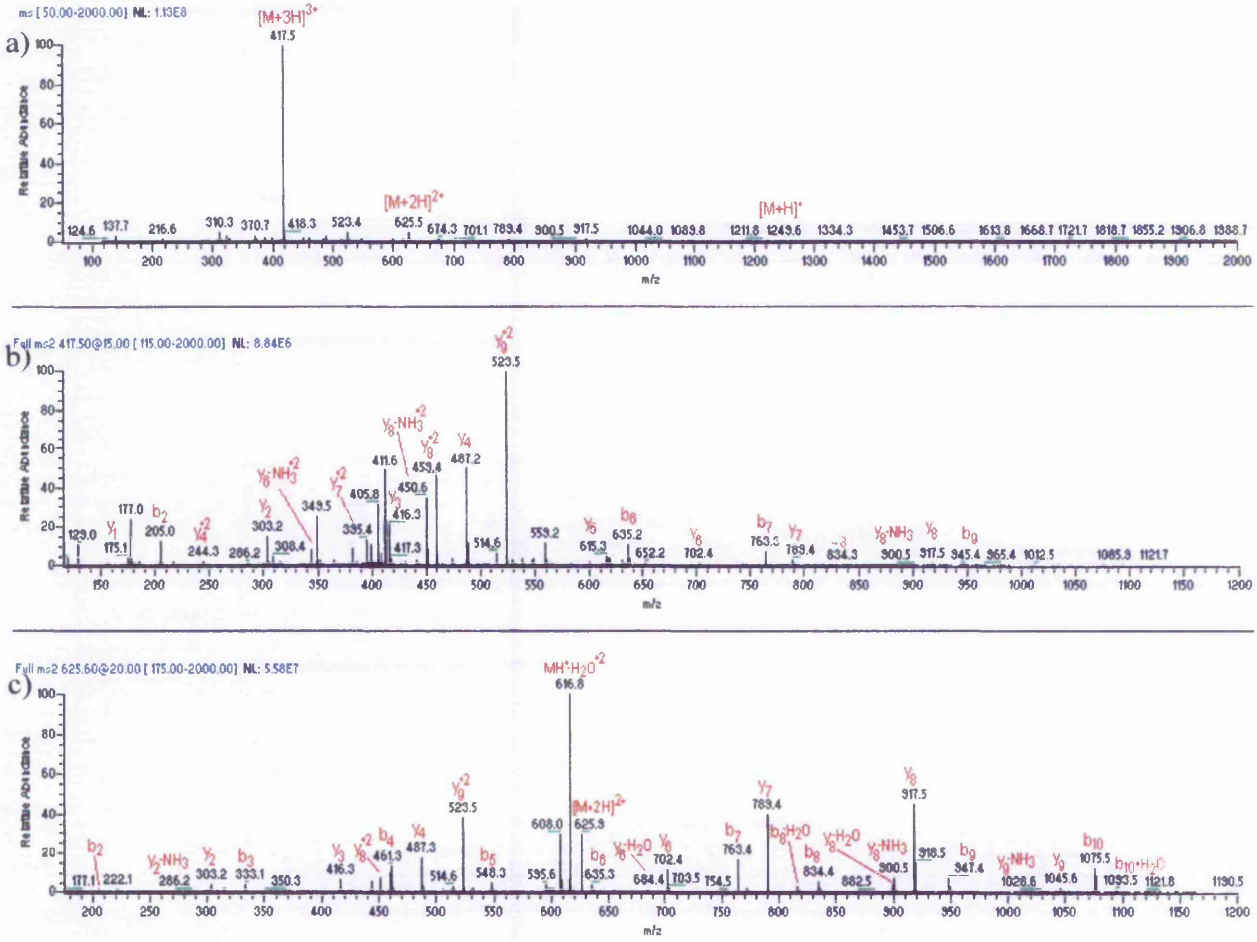


Figure 2.3.1.2 a) Full scan mass spectrum of peptide GFKQSSKALQR showing $[M+2H]^{2+}$ and $[M+3H]^{3+}$ b) Tandem mass spectrum of the $[M+3H]^{3+}$ ion. c) Tandem mass spectrum of the $[M+2H]^{2+}$ ion. Mass spectra were averaged over *ca.* 50 µscans.

It has been shown that doubly charged peptides give different fragmentation spectra to singly charged peptides, it follows that triply charged peptides will differ again. This is confirmed in Figures 2.3.1.1 and 2.3.1.2 which exhibit markedly different fragmentation spectra for the doubly and triply charged peptide ions, both in terms of intensity and identity of the ions produced. A greater abundance of y ions has been observed in Figures 2.3.1.1 and 2.3.1.2. It has been shown to be the case with the fragmentation of tryptic peptides that y-ions are more abundant than the complementary b-ions.[17,29] It has been suggested

that this is because y-ions are inherently more stable than the b-ions and therefore more resistant to further decomposition. [19]

Fragmentation of precursor ions to products of lower charge state clearly implies formation of complementary ions. Observation of complementary pairs facilitates the interpretation of product ion spectra. It has been noted that fragmentation of tryptic $[M+2H]^{2+}$ ions yielded product ion spectra in which b-ions (with charge retention at N-terminus) appeared at considerably greater relative abundance than y counterparts, due to higher stability of b-ions.[18] This is consistent with the idea of the b-ions site of charge being primarily on C-terminal amino acid residue side chain whereas y-ions have no strongly favoured site of charge.[23,27] On comparison of $[M+H]^{+}$ and $[M+2H]^{2+}$ ions it can be seen that high mass N-terminal (b) sequence ions have greater relative abundance in the fragmentation of doubly rather than singly charged ions.[24]

2.3.2. Optimisation of mass spectrometric tube lens voltage and heated capillary temperature

Analyte sensitivity and mass spectra depend not only on buffer composition and pH but also on ESI voltage and ion focusing parameters.[30] The relative proportion of singly vs multiply charged peptide ions depends on solvent composition and skimmer potential.[31]

Since the $[M+2H]^{2+}$ ion has been identified as the most useful charge state for the identification of peptide sequences (section 2.3.1) the optimum tube lens offset voltage (TLV) and capillary temperature for the transmission of these ions has been investigated. These parameters can have a marked effect on transmission of ion charge states and are relatively easy interface parameters to manipulate. 'Tuning' of the capillary and octapole voltages was additionally performed in conjunction with these determinations to establish maximum ion transmission. Two peptides, an 11mer and 18mer, were used for this optimisation. With the results confirmed by analysing a third peptide 9, amino acids in length, to confirm the findings. This set of peptides was chosen as they originate from the breakpoint region of the bcr/abl fusion protein which is the target for this work. The lengths represent MHC class I (9-11mer) and MHC class II (core sequence of ~9, indeterminate chain length).[32] Although only 3 peptides were studied during this optimisation they were sufficient to indicate a trend in the effect of conditions on mass spectra. Additionally they represented the breakpoint region which was ultimately of interest in this work. Furthermore the conditions optimised here were utilised in the development of nano-LC (section 2.3.4) using a 9mer peptide originating from the gp70 protein to target the $[M+2H]^{2+}$, hence further confirming the trends observed here.

2.3.2.1 Tube lens offset voltage

The tube lens offset voltage is the voltage offset from ground that is applied to the tube lens to focus ions towards the opening of the skimmer (a voltage is not applied to the skimmer on the LCQ ion trap). The tube lens offset voltage also

assists in the desolvation of adduct ions by accelerating the ions into the background gas. Low values of this parameter result in low energy collisions with the background gas which lead to desolvation, high values will lead to higher energy collisions which may result in collision induced dissociation.[33]

The 11mer and 18mer peptides were both analysed using the static nano-spray source. The tube lens offset voltage (as shown in Figure 2.1.0.2) was varied to ascertain the optimum condition for formation of the doubly charged ion for the two model peptides. Figures 2.3.2.1 and 2.3.2.2 show the effect on the % relative abundance of the $[M+H]^+$, $[M+2H]^{2+}$ and $[M+3H]^{3+}$ ions for both the 18mer and 11mer peptides observed when the tube lens offset voltage was varied, whilst maintaining a fixed capillary voltage (+46 V) and capillary temperature (180 °C). The data was averaged over a period of ca. 50 μscans to compensate for variability of signal intensity over time.

Figure 2.3.2.1 shows the trends in the relative abundance of the singly, doubly and triply charged ions for the 18mer peptide. At high negative tube lens voltages the $[M+3H]^{3+}$ ion dominates the spectra with the $[M+H]^+$ and $[M+2H]^{2+}$ ions present in only very low abundance. As the voltage increases there is a steady increase of $[M+H]^+$ and $[M+2H]^{2+}$ ion intensities whilst the $[M+3H]^{3+}$ ion maintains its dominance. Above a tube lens voltage of +40 V, the intensity of the $[M+3H]^{3+}$ ion begins to fall and the $[M+2H]^{2+}$ ion increases reaching 100 % relative abundance at a tube lens voltage of +50 V.

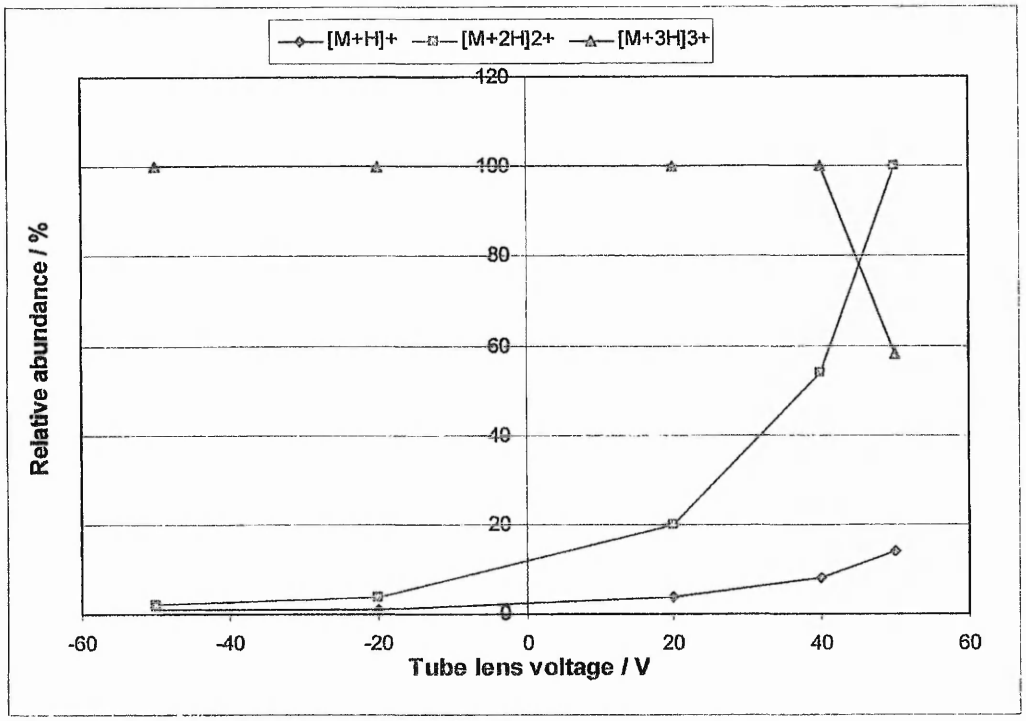


Figure 2.3.2.1. The variation of the relative abundance of the peptide ion against tube lens offset voltage for the singly, doubly and triply charged ions of the 18mer peptide ATGFKQSSKALQRPVASD (capillary temperature 180 °C).

A similar trend is observed in Figure 2.3.2.2 for the 11mer peptide. At negative tube lens voltages the $[M+3H]^{3+}$ ions predominate, although both the $[M+H]^{+}$ and $[M+2H]^{2+}$ have a greater relative abundance at negative tube lens voltages than that observed for the 18mer peptide ions. The $[M+2H]^{2+}$ ion becomes the dominant ion in the spectrum at a tube lens voltage of +20 V. The $[M+H]^{+}$ ion abundance continues to rise more gradually, and that of the $[M+3H]^{3+}$ falls, with the increased tube lens voltage. These observations suggest that higher positive tube lens voltages favour transmission of $[M+2H]^{2+}$ ions into the ion trap. It has been shown that a relatively high skimmer voltage optimises $[M+H]^{+}$ product whereas at a lower potential the $[M+2H]^{2+}$ ion is favoured.[31] Low electrospray

skimmer voltages favour survival of higher charge state ions.[23,34,35] A similar trend is observed in these experiments with the highest charge state ion (3+) favoured at low voltages and the increase in the lower charge state ions (1+ and 2+) as the voltage is increased. This effect is observed, for both peptides, until the $[M+2H]^+$ ion dominates the mass spectra. Had even higher voltages been explored it may have been the case that the $[M+H]^+$ ion would have eventually been dominant as described in the literature above.

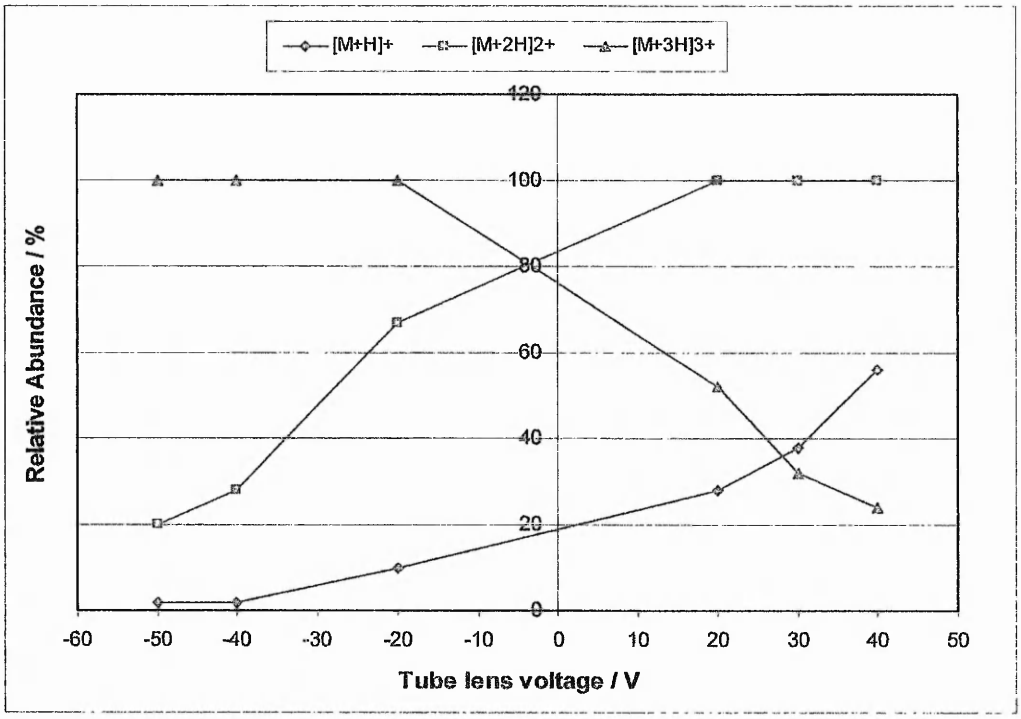


Figure 2.3.2.2. The variation of the relative abundance of the peptide ion against tube lens offset voltage for the singly, doubly and triply charged ion for the 11mer peptide GFKQSSKALQR (capillary temperature 180 °C).

2.3.2.2. Capillary temperature

A heated capillary aids in droplet desolvation. The use of high temperatures and occurrence of collisions within the capillary assist in the desolvation process and help to accomplish the declustering of ions from neutrals, thereby improving sensitivity.[36] Increasing the temperature will result in increased energy for the ions therefore more collisions and better desolvation and declustering however a greater incidence of in source fragmentation may also result. As such there must be a balance between increasing temperature to optimise desolvation but not so much that fragmentation occurs which will lead to a reduction in sensitivity.

The 11mer and 18mer peptides were also analysed at different heated capillary temperatures and tube lens offset voltages of -50 V and +40 V. Figures 2.3.2.3 and 2.3.2.4 show the effect of analysing these peptides at capillary temperatures of 180 °C and 250 °C. It has already been shown that effective ion desolvation occurs during nano-spray analysis when employing capillary temperatures of 150-200 °C.[37] The purpose of these experiments therefore was to examine the effect of increasing the capillary temperature to determine if there was a positive effect upon desolvation and transmission of a particular ion charge state. Figure 2.3.2.3 shows the experimental results for study of the 18mer peptide. It can be seen that at -50 V the heated capillary temperature has little effect on the charge generated on the peptide. At a tube lens voltage of +40 V, increasing the temperature from 180 °C to 250 °C has a significant effect on the level of doubly and triply charged ions. The $[M+2H]^{2+}$ ion increases from ~50 % up to 100 % relative abundance and the $[M+3H]^{3+}$ ion behaves in the opposite way with the

abundance falling to <40 % at the higher capillary temperature. Little effect is seen on the abundance of the $[M+H]^+$ ion.

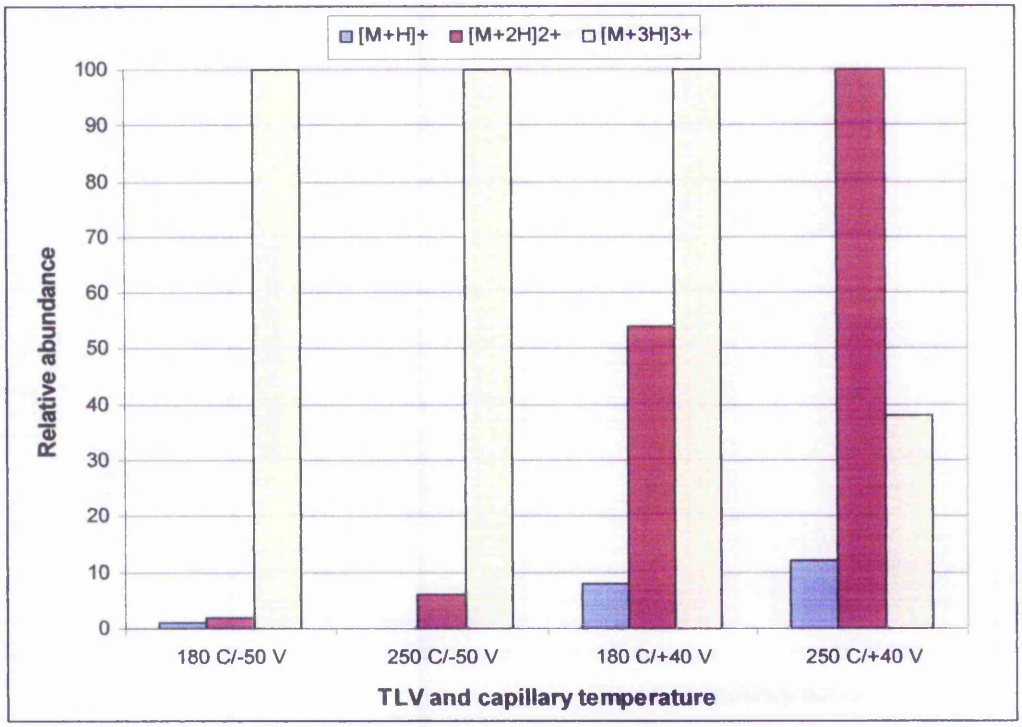


Figure 2.3.2.3. The variation of the relative abundance of the peptide ion charge state with capillary temperature for an 18mer peptide.

Figure 2.3.2.4 shows the experimental results for study of the 11mer peptide. At -50 V there was an increase in relative abundance of the doubly charged ion from ~20 % to ~70 % upon increasing the capillary temperature from 180 °C to 250 °C. For the same voltage, the singly and triply charged ion abundances remain relatively static over the same temperature range. At +40 V there was an increase in relative abundance of the $[M+H]^+$ ion up to 100 % from <60 % with an increase in temperature. Under the same variation in conditions the relative intensities of the $[M+2H]^{2+}$ and $[M+3H]^{3+}$ ions were both slightly reduced. The

increased transmission of the $[M+H]^+$ ion with increased capillary temperature may be explained by improved desolvation of these ions.

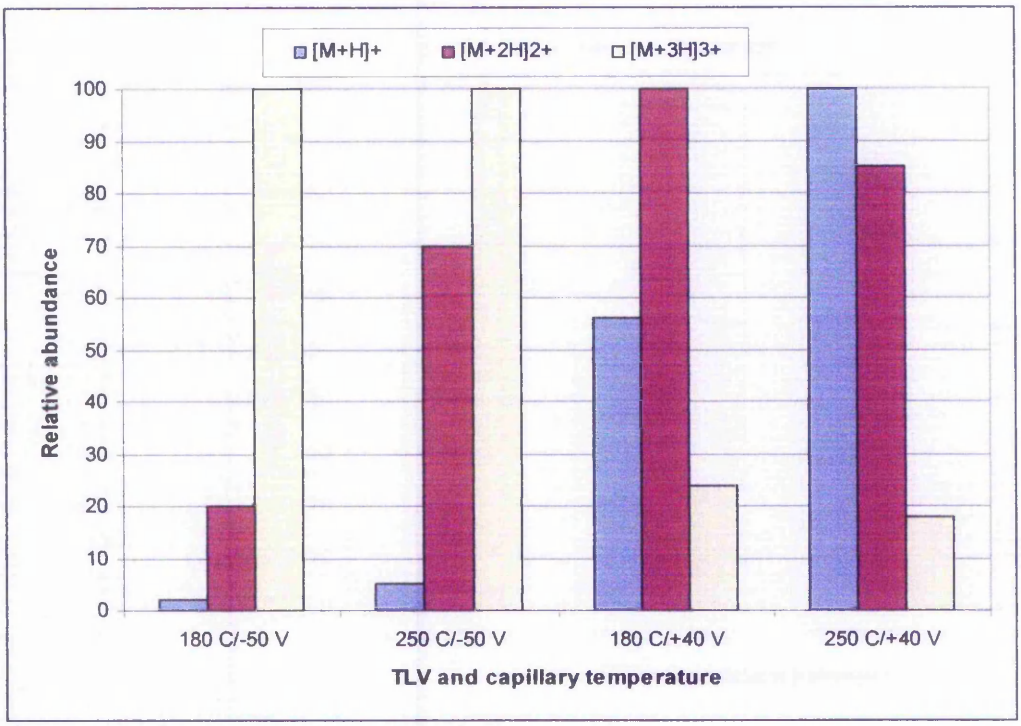


Figure 2.3.2.4. The variation of the relative abundance of each peptide ion charge state with capillary temperature for an 11mer peptide.

Examining the data from both the tube lens voltage and capillary temperature optimization it is apparent that in order to favour the transmission of doubly charged peptide ions, which have been shown to be the most desirable for sequence identification, a positive voltage and a higher capillary temperature are preferred. However, it is also important to look at the quality of the spectra obtained in these experiments, these may be seen in Figure 2.3.2.5. It should be noted that the peptide sequences on which these experiments were performed contain three basic residues (Arginine, R, pk_a ~12; Lysine, K, pk_a ~10.5)[38] which

make formation of a triply charged parent ion highly favourable. However, the concentration of basic residues is less common in the general peptide pool, particularly for shorter peptide lengths, due to the natural abundance of the residues in proteins. Both Arginine and Lysine occur at a level of ~5.7 %[38] in proteins, which is close to the 5% occurrence which would be true if all 20 naturally occurring amino acids had equal abundance. As the peptide chain shortens the probability of the appearance of one of these two diminishes. Therefore in a 9mer peptide the occurrence of a chain containing three of these amino acids would be low.

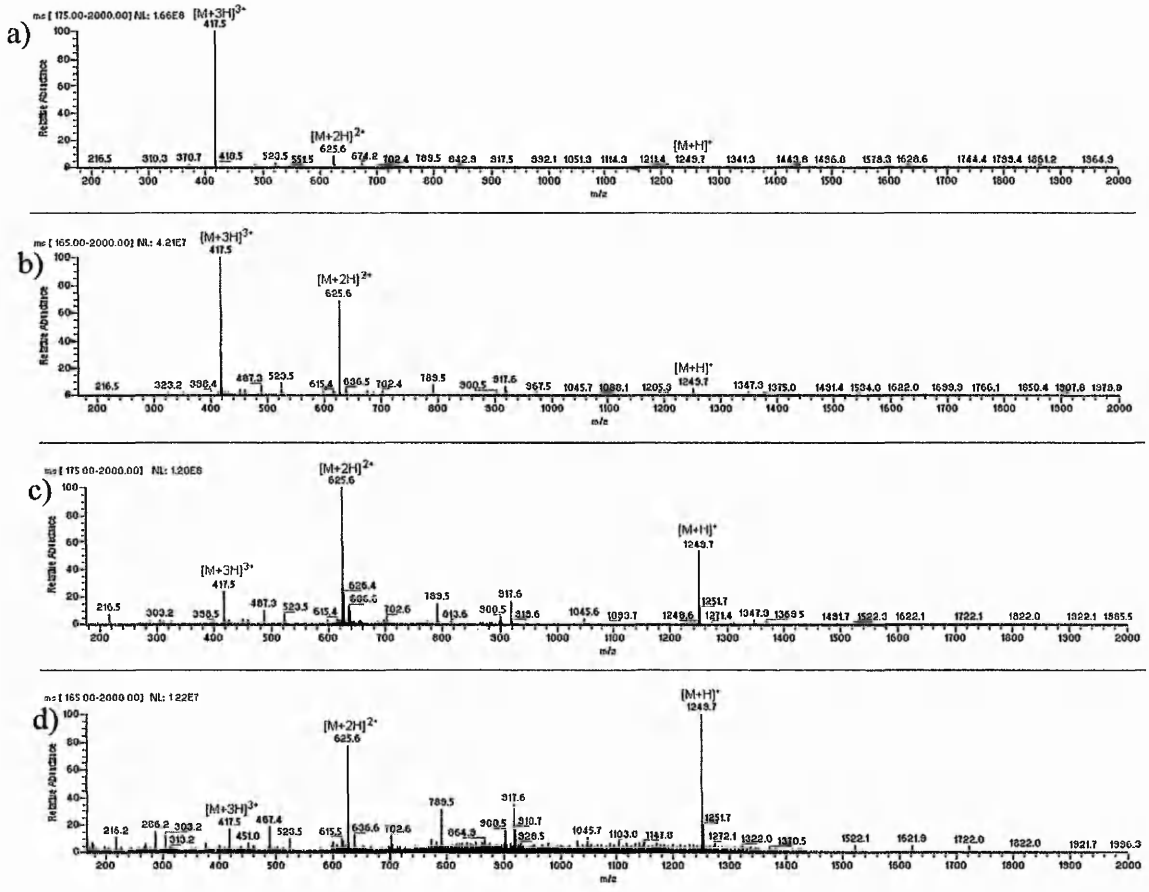


Figure 2.3.2.5. Full mass spectra of the 11mer peptide GFKQSSKALQR under different conditions. a) Tube lens offset, -50 V; Capillary temperature, 180 °C. b) Tube lens offset, -

50 V; Capillary temperature, 250 °C. c) Tube lens offset, +40 V; Capillary temperature, 180 °C. d) Tube lens offset, +40 V; Capillary temperature, 250 °C. Mass spectra were averaged over *ca.* 50 μscans.

It is clear from Figure 2.3.2.5 that the spectra shown in (a) and (b), recorded with a tube lens voltage of -50 V, are much cleaner than the spectra in (c) and (d) where the tube lens voltage was +40 V. The spectra recorded at a tube lens voltage of +40 V are characterised by the presence of peptide fragment ions derived from the doubly and triply charged ions (as seen in Figures 2.3.1.1 and 2.3.1.2). These arise from dissociation as a result of collisions between peptide ions and air in the mass spectrometer interface. This is commonly referred to as cone voltage or in-source fragmentation.[39] Optimum skimmer potential is unique for each peptide and represents a balance between generating ions of sufficient energy for ms/ms experiments and inducing the onset of other processes such as source fragmentation.[40] The extent of fragmentation in electrospray spectra is dependent on source conditions, particularly the potential applied to the skimmer.[25, 41-44] The magnitude of the skimmer potential determines the degree of collisional activation of ions accelerated through the intermediate pressure region in the source. The higher skimmer potentials generate ions with increased initial internal energies therefore more fragmentation occurs.[25, 31] This has been used to deliberately fragment the synthetic peptide HSIAGPPVPPR. By increasing the voltage between the end of the capillary and the first voltage lens, dissociation was induced leading to two stage fragmentation allowing the trap to be filled with ions of interest which were then further fragmented.[45] This has also been used for the doubly charged parent ions of Bradykinin and angiotensin II.[25]

In keeping with these observations the fragment ion abundances were significantly reduced when the tube lens voltage was set to -50 V presumably because the ion translational energy is lower at this voltage. A very small amount of fragmentation is seen in Figure 2.3.2.5 (a), but as the temperature was increased to 250 °C (Figure 2.3.2.5 (b)) the fragmentation was seen to increase slightly, particularly for the m/z 789.6 (y_7) and 917.5 (y_8) fragment ions.

These experiments suggest that a higher capillary temperature (250 °C) and a high tube lens voltage are the ideal conditions for the formation of doubly charged ions for the 18mer peptide. These conditions also generate a high abundance of the doubly charged ion for the 11mer peptide, although the singly charged ion dominates under these conditions (Figures 2.3.2.1-2.3.2.4). However it can also be seen in Figure 2.3.2.5 that at both tube lens voltages increasing the temperature from 180 °C to 250 °C results in an overall loss of sensitivity of approximately 5-10 fold. At -50 V, when increasing the capillary temperature from 180 °C to 250 °C (Figure 2.3.2.5, panels a and b) the signal intensity is reduced from 1.7 E8 to 4.2 E7. Similarly at +40 V, when increasing the capillary temperature from 180 °C to 250 °C (Figure 2.3.2.5, panels c and d) the signal intensity is reduced from 1.2 E8 to 1.2 E7. As stated earlier the use of high capillary temperature may improve sensitivity by assisting in the desolvation process, however, by increasing an ion's energy the risk of initiating unwanted fragmentation is also increased.[36] In these experiments it has been shown that increasing capillary temperature is detrimental to sensitivity. The application of a tube lens voltage of +40 V also leads to a decrease in signal to noise ratio and an

increase in cone voltage fragmentation resulting in a far less clean mass spectrum. It therefore must be concluded that the optimum conditions for doubly charged peptide ion formation are a lower capillary temperature and a lower tube lens voltage (180 °C and -50 V respectively). Although these conditions should favour the formation of a triply charged ion, this will only occur if the peptide contains multiple basic residues as with the 11mer and 18mer peptides used in this investigation. In practice this is uncommon and so the doubly charged ion will be formed preferentially. This can be seen in Figure 2.3.2.6 which shows the full scan mass spectrum of the 9mer MHC class I predicted peptide GFKQSSKAL,[46] a shortened version of the 11mer peptide which is without the terminal arginine residue and so does not contain 3 basic residues, this results in a spectrum in which the doubly charged ion predominates.

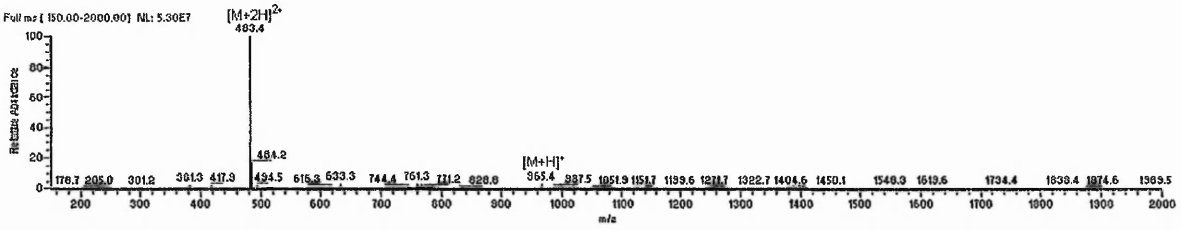


Figure 2.3.2.6. Full scan mass spectrum of 9mer peptide GFKQSSKAL generated under optimal doubly charging conditions, capillary temperature 180 °C and tube lens voltage -50 V.

2.3.3. Development and evaluation of a microspray ion source

An on-line microspray ion source was developed to assess the potential increase in sensitivity compared with conventional electrospray ionisation sources. Wilm and Mann proposed that the efficiency of analyte desorption increases with

decreasing droplet size due to its increased surface area to volume ratio. The transmission efficiency from needle to inlet is therefore increased, as much as 500 times, allowing detection limits as low as 50 amol.[2] A peptide solution (ATGFKQSSKALQRPVASD) was analysed using the LCQ ESI source and the on-line microspray source, operated at 3 μl min^{-1}, with 100 μm, 50 μm and 25 μm i.d. spraying tips. The peptide solution was infused through a 10 μl injection loop at a flow rate of 3 μl min^{-1} using an integral syringe pump resulting in a signal duration of up to *ca.* 10 minutes. Mass spectra were averaged over *ca.* 75 μscans. Compared to peptide study via ESI, the analytical conditions varied only in the sheath gas flow (60 arb units for the ESI but not used with the microspray ion source) and the spray voltage (4.5 kV for the ESI and 3.5 kV for the microspray ion source). The plots of spray tip size vs ion count (products of the $[M+3H]^{3+}$ ion) are shown in Figures 2.3.3.1 and 2.3.3.2.

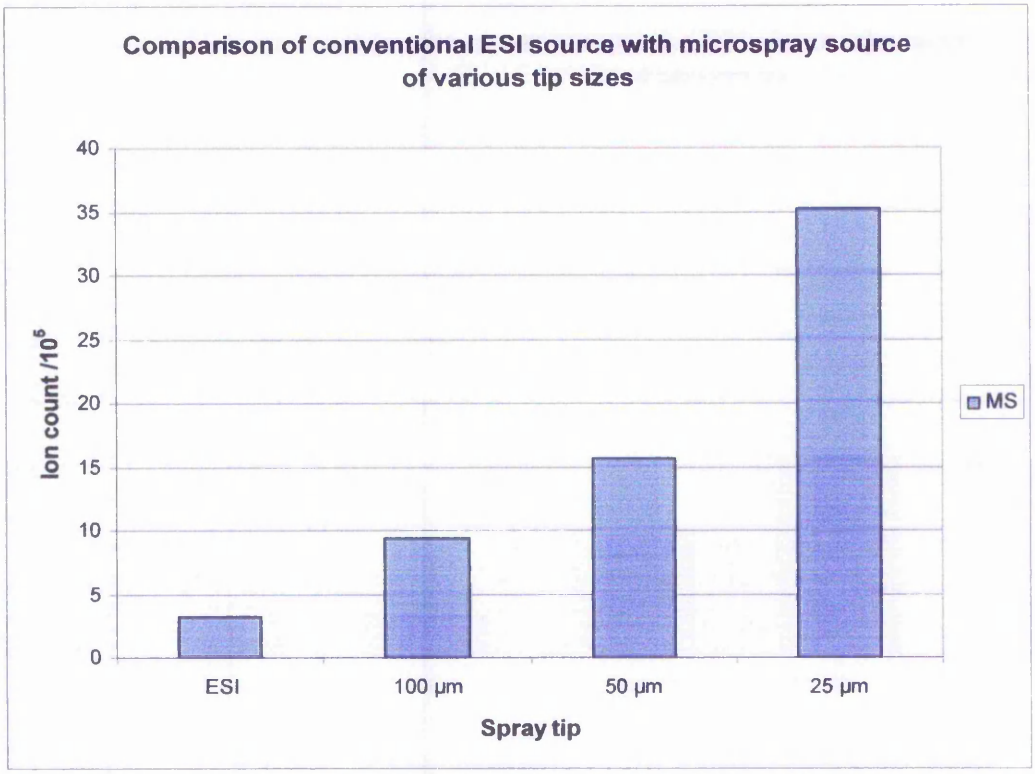


Figure 2.3.3.1. Ion count data for the peptide ATGFKQSSKALQRPVASD in full scan MS mode, $[M+3H]^{3+}$ (m/z 631.4) analysed by conventional ESI and microspray with various tip internal diameters.

Figure 2.3.3.1 shows a plot of the ion intensity for the $[M+3H]^{3+}$ ion of the peptide ATGFKQSSKALQRPVASD analysed by ESI/MS and microspray/MS with spray tip internal diameters of 100 µm, 50 µm and 25 µm. The graph shows a steady increase in the signal for the parent ion going from conventional ESI (Finnigan ESI source) to microspray ion source and again as the internal diameter of the spraying tip decreases. A steady increase is seen up to 50 µm and from 50 µm to 25 µm a sharp increase in signal is observed. The steady increase in sensitivity as tip id decreases can also been seen in Figure 2.3.3.2, which shows the signal intensity vs tip id for the main tandem fragment ion of the 18mer peptide, m/z 625.3.

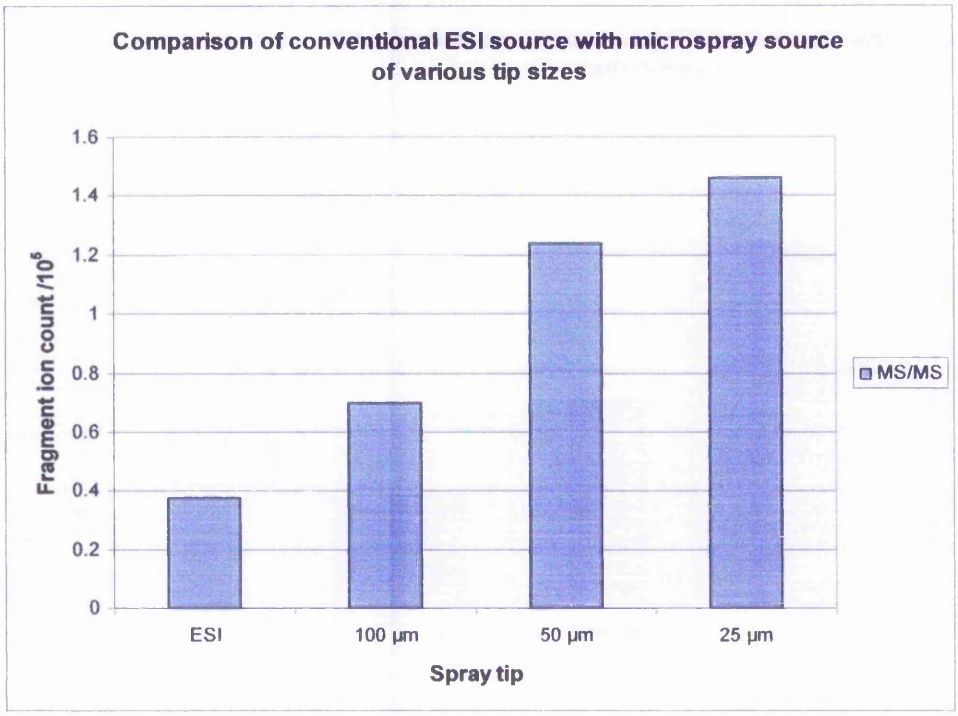


Figure 2.3.3.2. Fragment ion count data for the peptide ATGFKQSSKALQRPVASD in tandem mass spectra using the m/z 625.3 fragment ion to compare sensitivity of conventional ESI and microspray with various tip internal diameters.

This increase in sensitvity can also be observed in the mass spectra obtained from this analysis, which can be seen in Figure 2.3.3.3. Figure 2.3.3.3(a) shows the full scan mass spectrum and (b) the tandem mass spectrum for the 18mer peptide analysed using the conventional ESI source. Figure 2.3.3.3(c) shows the full scan mass spectrum and (d) the tandem mass spectrum for this peptide analysed using the 25 µm microspray source.

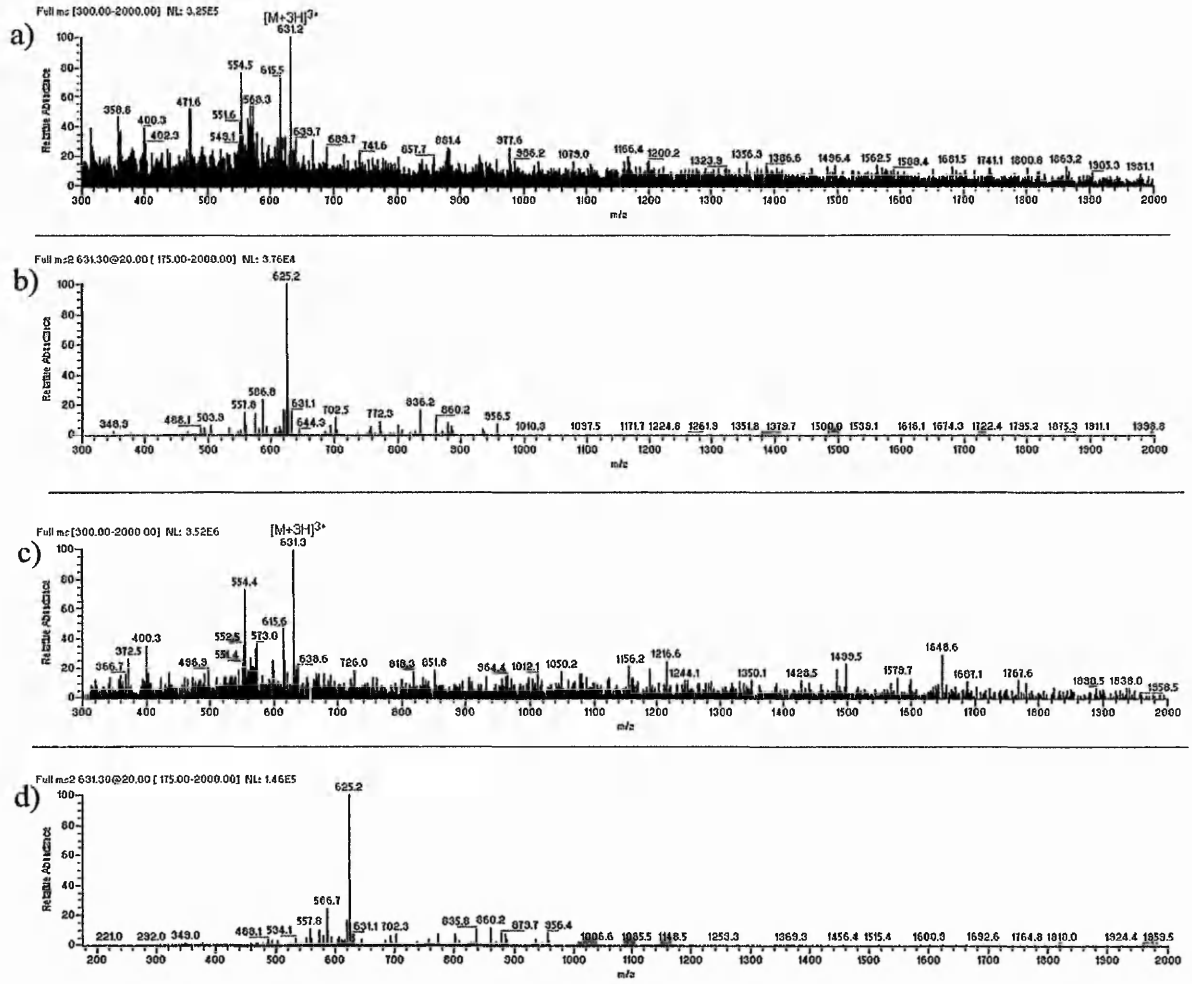


Figure 2.3.3.3. (a) Full scan mass spectrum of an ESI analysed 18mer peptide (b) Tandem mass spectrum of the 18mer peptide (m/z 631.2) from ESI analysis. (c) Full scan mass spectrum of a 25 µm microspray analysed 18mer peptide. (d) Tandem mass spectrum of the 18mer peptide (m/z 631.2) from 25 µm microspray analysis.

It can be seen in Figure 2.3.3.3 that the absolute signal intensity improved significantly from the conventional ESI source in Figure 2.3.3.3 (a) to the 25 µm microspray in 2.3.3.3 (c). The signal to noise ratio also improved for the 25 µm id spectrum compared with the conventional ESI spectrum. In both cases good tandem mass spectra are observed in Figure 2.3.3.3 (b) and (d) and again a much greater signal intensity was obtained using microspray analysis with the 25 µm

internal diameter spray tip. From this investigation we can infer that further increase in sensitivity may be observed as the spray diameter is decreased although this would require operating at sub μl min^{-1} flow rates as with the nano-electrospray source developed by Wilm and Mann which utilises a 1-2 μm spray orifice and flow rates of ~20 nl/min.[47]

2.3.4. Nano-LC/μ-ESI/MS

It has been shown in section 2.3.3 that significant improvement in instrument sensitivity can be obtained by reducing the electrospray spray tip diameter. The ability to couple the microspray device with capillary HPLC would further increase the sensitivity at the detector by a reduction in flow rate and therefore chromatographic dilution of the analytes as described in section 1.1.4.2. The coupling of nanoscale LC systems with micro-electrospray MS has resulted in detection limits in the attomole range.[48,49] Hunt's group have utilised 75 μm capillary columns in their investigation of MHC peptides from melanoma cells and EBV transformed cells.[50-52] The $[M+2H]^{2+}$ ion is targeted in all cases using the optimised conditions determined in Sections 2.3.1 and 2.3.2. The capillary LC/MS analysis of the peptide gp70 $SPSYVYHQF_{(454-462)}$ was performed successfully using a flow rate of 200 nl min^{-1} and a nano-ESI tip with a 5 μm internal diameter. The resulting chromatogram and mass spectrum are shown in Figure 2.3.4.1. A peak corresponding to the gp70 $[M+2H]^{2+}$ ion, m/z 564.3, can be seen eluting from the column after 28 minutes. The peak is sharp with a base width of ~1 minute for an 18 fmol injection of peptide. The LCQ was tuned to

optimise the abundance of the doubly charged ion using the conditions determined in section 2.3.2. However a response is also observed for the $[M+H]^+$ ion, m/z 1127.5 with a relative abundance of ~10 %.

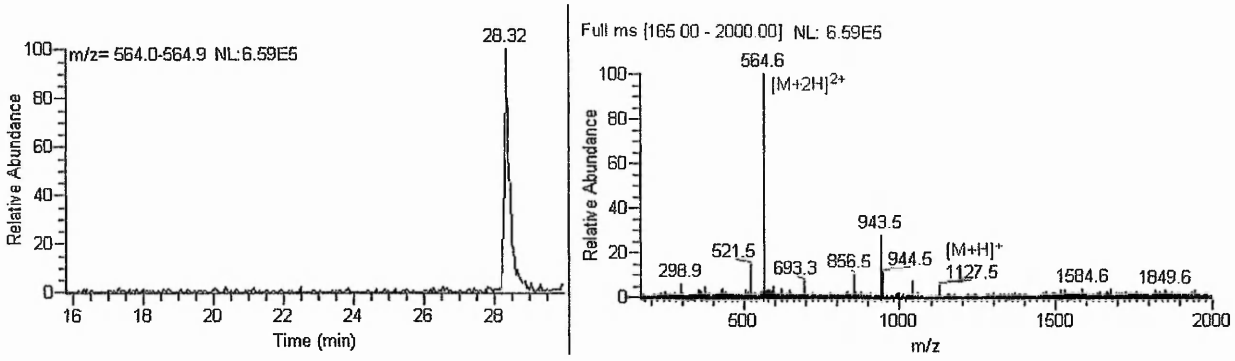


Figure 2.3.4.1. a) Single ion chromatogram of the peptide gp70 SPSYVYHQF$_{(454-462)}$ b) Mass spectrum of the gp70 peptide showing both the doubly charged ion, m/z 564.5 and the singly charged ion, m/z 1127.5.

The sensitivity of this system is excellent; the peptide peak is strong and the signal clearly detectable in full scan mode. Peak area reproducibility (calculated using the range error equation (equation 3.4.1) was 2 % for two replicate injections.

$$\left[\frac{X_n - X_1}{\bar{X}}\right] \times 100 \qquad \textbf{(Equation 3.4.1)}$$

X_n = highest result $\qquad$ $\bar{X}$ = mean result

X_1 = lowest result

2.3.5. Introduction of large sample volumes

The nano-LC column is suitable for the analysis of small sample injection volumes (≤1 µl), the maximum injection volume for a 75µm x 15cm microcolumn has been calculated to be not larger than 6.1 nl.[53] The low peptide concentration and high volume of sample anticipated after sample clean-up necessitate the injection of larger sample volumes than those commonly employed with capillary LC systems. It was therefore necessary to adapt the capillary LC/MS set-up to allow for injection volumes of up to 1000 µL. A trap column was therefore included in the system and evaluated for the loading of larger volumes of dilute solutions at higher flow rates (20 -30 µl min^{-1}) than those used for the analytical LC column (~200 nl min^{-1}). Microcapillary trap columns have been used in conjunction with narrowbore columns (75 µm - 300 µm i.d.) with injections volumes up to 5-8 µl.[54,55] A reverse phase teflon membrane inserted into a 10 port switching valve was used to inject sample volumes up to 100 µl for the investigation of peptide expression after measles virus infection.[56] However no record of such high volumes being loaded onto these capillary columns could be found in the literature. The valve configuration of the nano-LC pump injector was altered to accommodate the addition of a trap column.

Loading of large injection volume onto the trap column

A 10 µl injection was first loaded onto the trap column and then the valve was switched to allow the peptide to be transferred to the head of the analytical column and analysed using the conditions described in section 2.2.4. The

resulting single ion chromatogram and mass spectrum of the injected peptide can be seen in Figure 2.3.5.1. The trap and analytical column configuration was further tested by loading a sample from a 1000 µl injection loop of the peptide gp70 SPSYVYHQF$_{(454-462)}$ onto the trap column. The resulting chromatogram and mass spectrum can be seen in Figure 2.3.5.2.

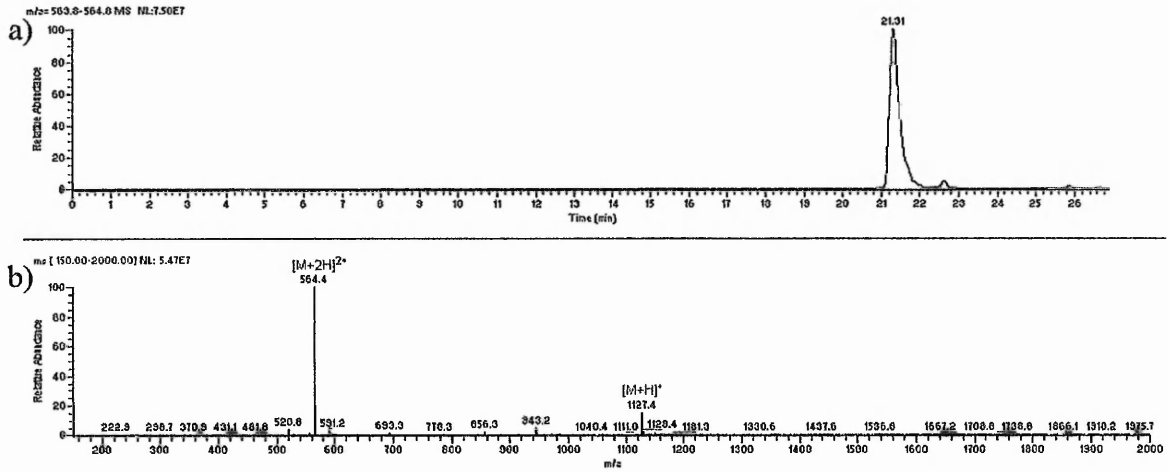


Figure 2.3.5.1. a) Single ion chromatogram of the peptide gp70 SPSYVYHQF$_{(454-462)}$ from a 10 µl injection using the trap column and analytical column configuration. b) Mass spectrum of the gp70 peptide showing both the doubly charged ion, m/z 564.5 and the singly charged ion, m/z 1127.5.

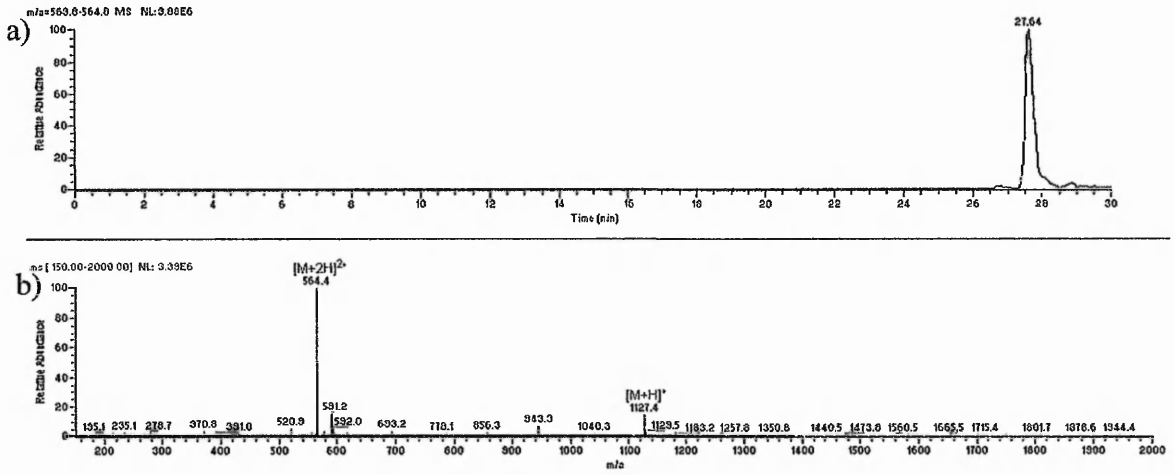


Figure 2.3.5.2. a) single ion chromatogram of the peptide gp70 SPSYVYHQF$_{(454-462)}$ from a 1000 µl injection using the trap column and analytical column configuration. b) mass spectrum of the gp70 peptide showing both the doubly charged ion, m/z 564.5 and the singly charged ion, m/z 1127.5.

Figure 2.3.5.2 shows the single ion chromatogram and mass spectrum of a 1000 µl loading of 1.8 fmol of the gp70 peptide using the newly configured trap and analytical column. The peptide is well retained on the trap column during the loading time required for both the 10 µl (0.5 minutes) and the 1000 µl (50 minutes) injection volumes. The peptide peak remains sharp, with base width of ~1 minute and good peak shape in each case. The signal intensity remains good for both the 10 µL and the 1000 µL injections (which is ten fold lower peptide concentration than for the 10 µL injection).

The use of a trapping column allows for the analysis of low concentration, large sample volume solutions without loss of chromatographic resolution and this configuration was used in all subsequent peptide analyses.

2.4.0. Conclusions

The effect of charge state on collision energy and fragmentation spectra of three MHC class I and II peptides, 9, 11 and 18 amino acids in length was investigated. It was observed that for larger peptides a greater collision energy was required for fragmentation to occur. The doubly charged peptides also require higher collision energies than their triply charged counterparts. It has been shown that the doubly charged peptides fragment to give greater coverage of b and y ions than the triply charged ions. The intensity of fragment ions in the higher mass region, m/z 1000-2000, is greater for the doubly charged ions. In some cases b and y ions were observed in this region only for the doubly charged ions or were up to 10 times greater intensity. As greater sequence coverage will improve confidence in the identification of unknown peptide sequences generation of the $[M+2H]^{2+}$ parent ion is preferred.

The tube lens offset voltage (TLV) and capillary temperature were investigated for their effect on the relative abundance of $[M+H]^{+}$, $[M+2H]^{2+}$ and $[M+3H]^{3+}$ parent ions. It was found that at high negative TLV the $[M+3H]^{3+}$ ion dominates the mass spectrum. The results seemed to suggest that the doubly charged ion is favoured with positive TLV and higher capillary temperature. However closer inspection of the mass spectra revealed that under these conditions a high degree of in source peptide fragmentation was observed and the signal to noise ratio was not very good. Increasing the capillary temperature from 180 °C to 250 °C also sees up to a ten fold loss in sensitivity. Therefore the optimum conditions for

formation of the $[M+2H]^{2+}$ ion were found to be a lower capillary temperature (180 °C) and a high negative tube lens offset voltage (-50 V).

A microspray ionisation source was assessed for the increase in sensitivity compared with conventional electrospray ionisation. Tip sizes of 100 μm, 50 μm and 25 μm were compared with the conventional ESI source and a steady increase in signal was observed as the spray tip diameter was decreased. An increase in signal to noise ratio was also observed.

A microspray source was constructed in-house and the spraying tip size was further reduced to ~5 μm in order to allow coupling with a capillary HPLC system. Excellent sensitivity was seen with the analysis of low femtomole quantities of peptides. This system was further adapted by incorporation of a trapping column and loading pump to allow the injection of much larger sample volumes than would normally be possible when operating at a flow rate of ≤200 nl min^{-1}. Injection volumes of up to 1000 μL of very dilute samples were successfully applied to the trapping column at a flow rate of ~30 μL min^{-1}, without loss of chromatographic resolution.

2.5.0. References

1. Clark R.E., Dodi I.A., Hill S.C., Lill J.R., Aubert G., Macintyre A.R., Rojas J., Bourdon A., Bonner P.L.R., Wang L., Christmas S.E., Travers P.J., Creaser C.S., Rees R.C., Madrigal J.A., *Blood,* **2001,** 98, 2887 - 2893.
2. Wilm M., Mann M., *Anal.Chem.*, **1996,** 68, 1-8.
3. Novotny M., *Anal Chem.*, **1998,** 60, 502A-510A.
4. Marina A. Garcia M.A., Albar J.P., Yague J., de Castro J.A.L., Vazquez J., *J. Mass Spectrom.*, **1999**, 34, 17-27.
5. Roepstorff P., Fohlman J., *J. Biomed. Mass Spectrom.* **1984,** 11, 601.
6. Sonsmann G., Romer A., Schomburg D., J. *Am. Soc. Mass Spectrom.*, **2002**, 13, 47-58.
7. Marie G., Serani L., Laprevote O., *Anal. Chem.*, **2000,** 72, 5423-5430.
8. Standing K.G., *Curr. Opin. Structural Biol.*, **2003**, 13, 595-601.
9. Tabb D.L., Smith L.L., Breci L.A., Wysocki V.H., Lin D., Yates J.R., *Anal. Chem.*, **2003**, 75, 1155-1163.
10. Wysocki V.H., Tsprailis G., Smith L.L., Breci L.A., *J. Mass Spectrom.*, **2000,** 35, 1399-1406.
11. Dongre A.R., Jones J.L., Somogyi A., Wysocki V.H., *J. Am. Chem. Soc.*, **1996**, 118, 8365-8374.
12. Vachet R.W., Glish G.L., *J. Am. Soc. Mass Spectrom.*, **1998,** 9, 175-177.
13. Downard K.M., Biemann K., *J. Am. Soc. Mass Spectrom.*, **1994,** 5, 966-975.
14. Kinter M., Sherman N.E., *Protein Sequencing and Identification Using Tandem Mass Spectrometry,* **2000,** Wiley, New York.
15. Tang X-J., Thibault P., Boyd R. K., *Anal. Chem.*, **1993**, 65, 2824-2834.
16. Wang G., Cole R.B., *Org. Mass Spectrom.*, **1994**, 29, 419-427.
17. Poulter L., Taylor L.C.E., *Int. J. Mass Spectrom. Ion Proc.*, **1989**, 91, 183-197
18. Tang X-J., Boyd R.K., *Rapid Comm. Mass Spectrom.*, **1992**, 6, 651-657.,
19. Tang X-J., Thibaiult P., Boyd R.K., *Anal. Chem.*, **1993,** 65, 2824-2834.

20. Burlet O., Orkiszewski R.S., Ballard K.D., Gaskell S.J., *Rapid Comm. Mass Spectrom.*, **1992**, 6, 658-662;

21. Burlet O., Yang C-J., Gaskell S.J., *J. Am. Soc. Mass Spectrom.*, **1992**, 3, 337-344;

22. McCormack A.L., Somogy A., Dongre A.R., Wysocki V.H., *Anal. Chem.*, **1993**, 65, 2859-2872.

23. Cox K.A., Gaskell S.J., Morris M., Whiting A., *J. Am. Soc. Mass Spectrom.*, **1996**, 7, 522-531

24. Tomer K.B., McGown S.R., Deterding L.J., *Org. Mass Spectrom.*, **1993**, 28, 534-541

25. Katta V., Chowdhury S.K., Chait B.T., *Anal. Chem.*, **1991**, 63, 174-178

26. Gaskell S.J., Bolger M.S., Riba I., Summerfield S.G., *Mass Spectrometry in Biomolecular Sciences* **1997** Ed Caproli R.M., Sidona G., Kluwer, Dordrecht.

27. Gaskell S., *J. Mass Spectrom.*, **1997**, 32, 677-688

28. Downard K.M., Biemann K., *Int. J. Mass Spectrom. Ion Processes*, **1995**, 148, 191-202.

29. Baringa C.J., Edmonds C.G., Udseth H.R., Smith R.D., *Rapid Comm. Mass Spectrom.*, **1989**, 3, 160-164

30. Smith K.D., Olivares J.A., Nguyen N.T., Udseth H.R., *Anal. Chem.*, **1988**, 60, 436-441

31. Kilby G.W., Sheil M.M., *Org. Mass Spectrom.*, **1993**, 28, 1417-1423

32. Roitt I.M., Delves P.J., *Essential Immunology 9th Edition*, Blackwell, Oxford, 1997

33. Xcalibur 1.3 software, ThermoFinnigan, Hemel Hempstead, UK

34. Ashton D.S., Beddell C.R., Cooper D.J., Green B.N., Oliver R.W.A., *Org. Mass Spectrom.*, **1993**, 28, 721-728;

35. Smith R.D., Loo J.A., Loo R.R., Busman M., Udseth H.R., *Mass Spectrom. Rev.*, **1991**, 10, 359-452.

36. Cech N.B., Enke C.G., *Mass Spectrom. Rev.*, **2001**, 20, 362-387.

37. Lill J. R., *PhD Thesis*, **2000**, Nottingham Trent University.

38. *Protein Purification 2nd edition* Ed Janson JC; Ryden L; **1998**, Wiley, New York

39. Harrison A.G., *Rapid Comm. Mass Spectrom.*, **1999**, 13, 1663-1670.

40. Alexander A.J., Boyd R.K., *Int. J. Mass Spectrom. Ion Proc.*, **1989**, 20, 211

41. Voyksner R., Pack T., *Rapid Comm. Mass Spectrom.*, **1991**, 5, 263;

42. Katta V., Chowdhury S.K., Chait B.T., *J. Am. Chem. Soc.*, **1990**, 112, 5348-5349;

43. Loo J.A., Edmonds C.G., Udseth H.R., Smith R.D., *Anal. Chimica Acta*, **1990**, 241, 167-173;

44. Hofstadler S.A., Bakhtiar R., Smith R.D., *J. Chem. Ed.*, **1996**, 73,A82-A88

45. Korner R., Wilm M., Morand K., Schubert M., Mann M., *J. Am. Soc. Mass Spectrom.*, **1996**, 7, 150-156

46. Norbury L.C., Clark R.E., Christmas S., *British Journal of Haematology*, **2000**, 109, 616-621

47. Shevchenko A., Wilm M., Vorm O., Mann M., *Anal. Chem.*, **1996**, 68, 850-858

48. Vanhoutte K., Hoes I., Lemiere F., van Dongen W., Esmans E. L., van Onckelen H., van den Eeckhout E., van Soest R.E.M., Hudon A.J., *Anal. Chem.*, **1997**, 69, 3161

49. Oosterkampe A., Gelpi E., Abian J., *J.Mass Spectrom.*, **1998**, 33, 976

50. Skipper J.C.A., Kittlesen D.J., Hendrickson R.C; Deacon D.D., Harthun N.L., Wagner S.N., Hunt D.F., Engelhard V.H., Slingluff C.L., *J. Immunol.*, **1996**, 157, 5027-5033.

51. Crotzer V.L., Christian R.E., Brooks J.M., Shabanowitz J., Settlage R.E., Marfo J.A., White F.M., Rickinson A.B., Hunt D.F., Engelhard V.H., *J. Immunol.*, **2000**, 164, 6120-6129.;

52. Hogan K.T., Coppola M.A., Gatlin C.L., Thompson L.W., Shabanowitz J., Hunt D.F., Engelhard V.H., Slingluff C.L., Ross M.M., *Immunol. Letters*, **2003**, 90, 131-135.

53. Chervet J.P., Ursem M., Salzmann J.P., *Anal. Chem.*, **1996**, 68, 1507-1512.

54. Vissers J.P.C., Hulst W.P., Chervet J-P., Snijders H.M.J., Cramers C.A., *J. Chrom. B.*, **1996**, 686, 119-128;

55. Yi E.C., Lee H., Aebersold R., Goodlett D.R., *Rapid Comm. Mass Spectrom.*, **2003**, 17, 2093-2098.

56. Ovsyannikova I.G., Johnson K.L., Naylor S., Poland G.A., *J. Immunol. Methods*, **2005**, 297, 153-167.

CHAPTER 3

DEVELOPMENT OF A METHOD FOR THE IDENTIFICATION OF MHC CLASS I AND CLASS II RESTRICTED PEPTIDES

Chapter 3 Aims Breakdown

A method for the identification of the MHC class I peptides developed at Nottingham Trent prior to the work described in this thesis, and used to confirm the presence of the HLA-A3 peptide KQSSKALQR, which originates from the bcr/abl fusion protein in patient and transfected cell lines. The method was as follows: [1]

Step 1 Surface elution of CML patient or transfected cells using Storkus[14] mild acid elution with citrate phosphate buffer at pH 3.3.

Step 2 Protein precipitation by addition of trichloroacetic acid (TCA), this acidified the solution to pH 0-1 denaturing the protein and causing precipitation, followed by centrifugation.

Step 3 Solid phase extraction of the supernatant using SCX phase cartridges to extract the peptides from the citrate phosphate buffer and other salts.

Step 4 Fractionation of the peptide SCX extract utilising RP-HPLC.

Step 5 Analysis of individual RP-HPLC fractions by off-line nano-ESI/MS/MS.

The method was tailored specifically towards the identification of the target KQSSKALQR peptide, which contains three basic amino acid residues (K, R) and as such has an abnormally high affinity for the SCX solid phase extraction cartridges which were used to remove the elution buffer salts. The presence of three basic residues is not representative of short chain peptides in general.

The objective of the work reported in this thesis was to develop this method in order to establish a generic procedure for the identification of MHC class I and II peptides. There were a number of ways in which it was hoped to improve on the original method:

- Increase the variety of peptides to which the method may be applicable.
- Establish improved peptide recovery through the clean-up procedure.
- Incorporate nano flow HPLC separation with on-line μESI/MS/MS.
- Incorporate peptide internal standards for a more robust method.

Specific modifications investigated for steps 2-5 are discussed below.

Step 2: Protein precipitation by addition of TCA

The addition of TCA for protein precipitation lowers the pH of the solution below the optimum operational condition of the SCX SPE cartridges utilised in the following step. This is not a problem for the strongly bound A3 peptide, but a consequence of the use of TCA is that peptides with lower affinity for that stationary phase may not be retained. The degree of losses which might be associated with the co-precipitation of peptides during protein precipitation is also not known. The investigation of alternative methods for the removal of proteins during this step was therefore a key objective of the work reported in this thesis.

Ultimately, it was found that a C18 solid phase extraction could be used for the removal of both extraneous proteins and citrate phosphate buffer salts in a single step, as well as concentrating the eluate. Solid phase extraction is a well established

technique, which has seen widespread use in the separation of small molecules of pharmaceutical interest from biological matrices such as urine and plasma.[2-5]

Step 3: SCX solid phase extraction

SCX is an orthogonal separation phase to C18 and therefore is an excellent way to improve separation efficiency in complex mixtures. However, a single step clean-up SPE does not fully utilize this separation. Incorporating an HPLC pump with SCX stationary phase column enables the stepwise elution of peptides in order of their increasing affinity for the stationary phase using salt fractions of increasing concentration. This resulted in 100 µl fractions of reduced complexity, which can be loaded directly onto a RP-HPLC trap column prior to capillary LC separation.

Step 4 and 5: HPLC fractionation and ESI/MS/MS analysis

The final improvement in this method was to combine HPLC and MS steps to allow on-line separation and identification of MHC peptides. The combination of these two steps reduces sample handling, increases sample throughput and is less labour intensive than off-line HPLC and static nano-ESI analysis. The use of µspray MS also lead to an improvement in sensitivity over conventional electrospray. In order to incorporate HPLC on-line with µspray it is necessary to utilize flow rates in the nL/min region, this means columns with internal diameters <100µm. These columns also serve to improve the sensitivity of the analysis by reducing chromatographic dilution.

3.1.0 Introduction

The work presented in this chapter concerns the development of methods for the identification of MHC restricted antigenic peptides presented on the surface of cells. The aim of the investigation was to establish a versatile mass spectrometry based approach that would allow both class I and class II restricted peptides to be identified, as the current view within the immunological community maintains both are necessary for the successful development of immunotherapeutic vaccines.[6]

A variety of procedures have been reported for the identification of MHC class I restricted peptides using ESI-MS based procedures. These have been recently reviewed[7] and were discussed in Chapter 1. These procedures included a method previously developed at Nottingham Trent University and used for the identification of the A3 class I bcr/abl peptide (KQSSKALQR) on the surface of leucophoresed white blood cells from patients diagnosed with chronic myeloid leukaemia.[1,8] This method involved the use of TCA protein precipitation, followed by strong cation exchange chromatography (SCX). The procedure relied on the very strong affinity of the stationary phase for the A3 peptide, which contains three basic residues and was therefore particularly well suited to this approach. Most peptides, however, do not have such a strong affinity for the SCX stationary phase so a new method was required which would be more generally applicable to the range of 9-11 mer MHC class I peptides predicted by the SYFPEITHI algorithm.[9]

A second objective of the investigation was to develop methods that minimised cell lysis by retaining the mild acid elution procedures reported previously. In the case of the class II peptides, the nature of the processing pathway means that the length of the

peptide antigen is highly variable and this makes the process of predicting epitopes considerably harder. Class II α and β subunits associate in the endoplasmic reticulum with the invariant chain.[10,11] The protein acts as a chaperone which ensures the correct assembly of the class II complex. Once bound, the invariant chain prevents binding of other peptides by the class II molecule during its transport from the endoplasmic reticulum to the Golgi.[12-14] Once the MHC molecule reaches the endoplasmic compartment, which is where the antigens can be found, proteolysis of the invariant chain occurs releasing the αβ dimer for binding of endocytically generated peptides prior to transport to the cell surface.

There are mutant human cell lines (TXB hybrid, T2) in which the class II antigen processing apparatus is defective. The cells cannot present peptides derived from intact protein antigen, but are able to present exogenously supplied peptides derived from the protein antigen. [14-16] Mutant αβ dimers are associated with peptides from the invariant chain between positions 81-104, these are known as the class II invariant chain peptide (CLIP). These have been shown to be present as part of the class II complex in T2DR3 cells.[17] An MHC related molecule known as DM catalyses the removal of CLIP and keeps the binding groove open to allow endosomally generated peptides to bind. The T2 cell line which has been used in this work is DM deficient, meaning they are unable to swap out the CLIP for an antigenic peptide. By using the mutant cell line which presents on its surface a nested set of the 'CLIP' peptides, that is a group of peptides that contain the same core sequence but vary in length at the amino or C-terminal ends, it is easier to predict possible CLIP peptide epitopes for which the length of the peptide is the only variable. This provided a convenient

system for the development and evaluation of a general method for the identification of MHC class II peptides.

3.2.0 Materials and methods

3.2.1 Elution of cell surface peptides

Cell surface peptides were eluted by mild acid washing with citrate phosphate buffer (0.131 M citric acid + 0.066 M sodium phosphate, adjusted to pH 3.3 using 10 M NaOH).[18] A minimum of 10^9 cells from leucapheresed CML patient blood or K562 transfected cell lines were washed in serum free RPMI media and centrifuged (1500 rpm, 5 minutes, 4 °C). The pelleted cells were resuspended in ~5 ml of citrate phosphate buffer for 5 minutes at room temperature and again centrifuged. The supernatant was collected, filtered through a 0.22 μm micropore filter and stored at -80 °C.

3.2.2 The recovery of standard peptides

A solution of citrate phosphate buffer (pH 3.3, 20 ml) was spiked with the hepatitis B peptide TPPAYRPPNAPIL (50 fmol). The solution was then subjected to the following clean-up procedure:

Trichloroacetic acid (72 % w/v, 2 ml) was added, on ice, to the buffer solution and the mixture was placed in an ultrasonic bath for 10 minutes. This was followed by 10 minutes centrifugation (15, 000 g, 4 °C) and a further 10 minutes ultrasonication on the supernatant. The supernatant was then subjected to strong cation exchange chromatography. The SCX column (1 ml, high S cartridge, Biorad, Hemel Hempstead, UK) was prepared by washing with NaOH (0.1 M, 10 ml), followed by HCl (0.1 M, 5 ml). The supernatant was applied to the column, washed with HCl (0.1

M, 10 ml) and eluted in NaOH (0.1 M, ~2.5 ml). Two 1 ml aliquots of the cation exchange eluate were then subjected to LC/MS analysis. A standard of hepatitis B (15.7 fmol/µl in 0.1 % TFA) equivalent to 100 % recovery was also analysed. The LC/MS set-up is shown schematically in Figure 3.2.2.1.

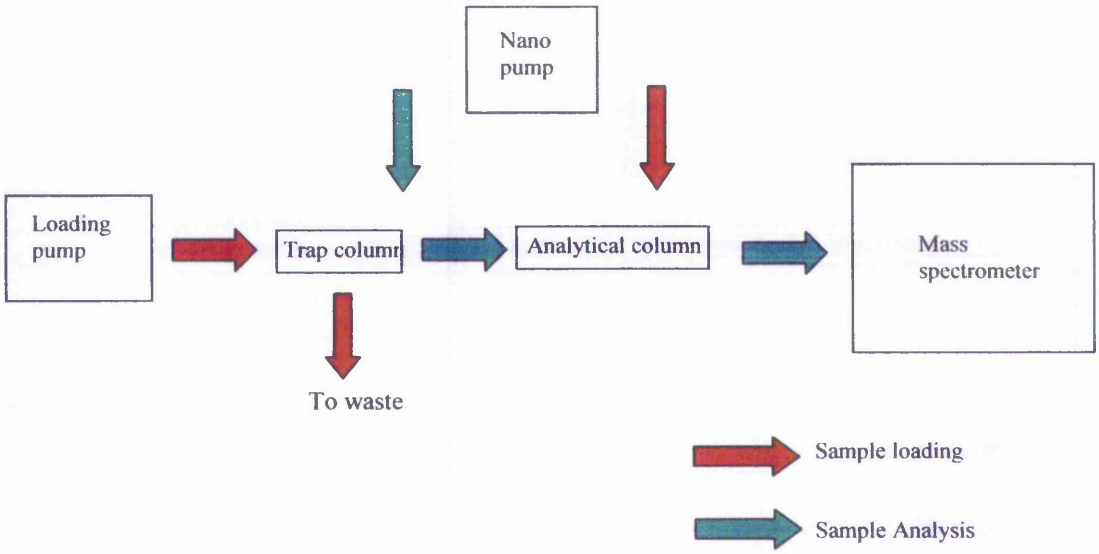


Figure 3.2.2.1 Schematic representation of the LC/MS set-up

The LC conditions were as follows:

Injection volume: 1000 µl (citrate phosphate buffer eluate); 1 µl (standard). Trap column: pepmap 300 µm i.d x 5 mm, C18 nanobore column (Dionex, Surrey, UK), loading buffer 0.1 % TFA at a flow rate of ~30 µl min^{-1}. Loading pump: Kontron 325 with microflow splitter (SGE, Milton Keynes, UK). Analytical column: pepmap 75 µm i.d x 150 mm, C18 nanobore column (Dionex, Surrey, UK). Mobile Phase A: H_2O + 0.09 % formic acid. Mobile Phase B: Acetonitrile + 0.09 % formic acid. Flow rate: 200 nl min^{-1}, Gradient profile: 10 % B – 60 % B in 90 minutes. Gradient pump: Ultimate™ nano-pump (Dionex, Surrey, UK).

Mass spectrometry was performed using on-line micro-electrospray ionisation mass spectrometry on a Finnigan LCQ quadrupole ion trap mass spectrometer (ThermoFinnigan, Hemel Hempstead, UK) using a home-made ion source. Mass spectrometric conditions were as follows:

Microspray voltage: 1.5-2.0 kV. Capillary Temperature: 180 °C. Capillary voltage: 20 V. Tube lens offset: -20 V. Collision energy: 30 %.

Aliquot 1 was analysed in data dependent mode with three scan events, full scan, zoom scan and MS/MS scan data acquisition on ions above a critical threshold (ion count 10^5). This was then repeated for the top three peaks with an exclusion time of 3.0 minutes.

Aliquot 2 was analysed to determine the recovery of the hep B peptide through the clean-up process. This was analysed in full scan mode and the same procedure was used for the standard solution.

Two 'test' patient eluates, provided by Professor Richard Clark, Royal Liverpool Hospital (UK), were subjected to the clean-up procedure described above. Both were then analysed in data dependent mass spectrometric acquisition mode following the method employed for fraction 1, above.

3.2.2.1 Peptide recovery through SCX

The recovery of peptides through the strong cation exchange method was tested using the following method. Citrate phosphate buffer (pH 3.3, 20 ml) was spiked with ~100 pmol of the following model peptides: p53 (RMPEAAPPV), β-gal (TPHPARIGL), gp70 (SPSYVYHQF), hep B (TPPAYRPPNAPIL), p53 (LLGRNSFEV). Protein precipitation was simulated using trichloroacetic acid (72 % w/v, 2 ml) and the supernatant was subjected to strong cation exchange chromatography procedure as described in section 3.3.2.2 above.

The eluted fractions were analysed, together with a standard peptide mixture that had not been subjected to TCA and SCX treatment, by LC/MS. All conditions were as previously described except for the injection volume, 1 µl. The mass spectrometric analysis method targeted for MS/MS the doubly charged ($[M+2H]^{2+}$) parent ion of each of the five peptides with a collision energy of 35 %.

3.2.2.2 Use of reduced bed volume SCX column

A UNO S polishing SCX column (160 µl volume; Biorad, Hemel Hempstead, UK) was investigated using the instrumental arrangement shown in Figure 3.2.2.2.

HPLC pump: Waters 501 HPLC pump with 455 LC spectrophotometer. Mobile phase: 0.1 M HCl, at a flow rate of 0.5 ml min^{-1}. Eluting solution: 0.1 M NaOH, 500 µl fractions were collected.

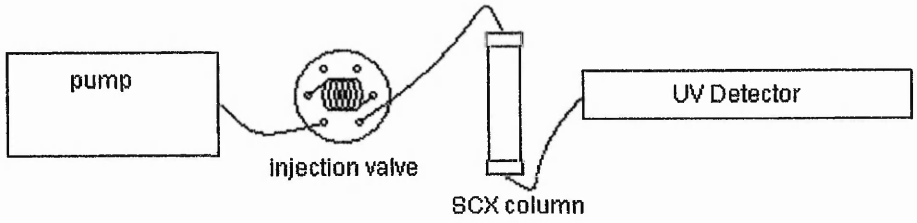


Figure 3.2.2.2 Schematic representation of the SCX instrumental set-up.

A test of this system was performed by spiking citrate phosphate buffer (pH 2.8, 2 ml) with 4 peptides; p53 (RMPEAAPPV), gp70 (SPSYVYHQF), hep B (TPPAYRPPNAPIL) and p53 (LLGRNSFEV) at ~100 pmol each. The fractions were then acidified with 0.1 % TFA and analysed by LC/MS, as has previously been described, and the recovery determined.

3.2.3 Investigation of protein removal methods

3.2.3.1 TCA precipitation and NaOH neutralisation

Two aliquots of citrate phosphate buffer (pH 3.3, 5 ml) were spiked with hep B at low level (~100 fmol) and high level (~100 pmol), TCA was added (72 % w/v, 2 mls) and pH adjusted with NaOH to ~pH 3 prior to SCX chromatography as described previously. The fractions were analysed by LC/MS as described in section 3.2.2.

3.2.3.2 Protein precipitation by centrifugation

The possibility of removing the protein in the eluate by centrifugation was explored. A test solution of bovine serum albumin (1 mg/ml in water) was centrifuged (Beckman Optima TLX ultracentrifuge, 100 000 rpm, 10 °C) for increasing lengths of time and the residual protein content measured by UV/vis spectrophotometer

(Beckman DU530 Life Science UV/vis spectrophotometer, λ= 278 nm) against a reference cell containing distilled water.

3.2.3.3 Removal of protein from patient eluates using ODS

The use of a C18 stationary phase was explored for the initial purification and protein removal from CML patient eluates.

The sample, citrate phosphate buffer (pH 3.3, 1 ml) spiked with 100-150 pmol of the peptides β-gal (TPHPARIGL), hep B (TPPAYRPPNAPIL), p53 (LLGRNSFEV). Column: Jupiter 30 x 1.0 mm C18 ODS guard column (Phenomenex, Cheshire, UK). Loading pump: Kontron 325 with microflow splitter (SGE, Milton Keynes, UK). Loading mobile phase 0.1 % TFA at a flow rate of ~250 μl min^{-1}. Elution was achieved with 200 μl injection of 60 % acetonitrile, five fractions of 50 μl each were collected.

The fractions were analysed by LC/MS as described in section 3.2.2 except for Injection volume, 1 μl. The mass spectrometric analysis method targeted for MS/MS the doubly charged ($[M+2H]^{2+}$) parent ion of each of the three peptides with a collision energy of 40 %.

3.2.4 Removal of proteins from patient eluates using solid phase extraction

3.2.4.1 Initial tests of solid phase extraction for recovery of peptides

The use of solid phase extraction (SPE) cartridges for the purification of the CML patient eluates was explored using the following procedure.

The sample citrate phosphate buffer (pH 3.3, 5 ml) spiked with ~100 pmol each of β-gal (TPHPARIGL), hep B (TPPAYRPPNAPIL), p53 (LLGRNSFEV). Each fraction was diluted from 100 µl to 500 µl in 0.1 % TFA. Solid phase cartridges, SEP PAK® (Water, UK), RP, C18, 360 mg sorbent, 55-105 µm particle size. Protocol: The cartridge was conditioned with 50 % acetonitrile (3 ml), and equilibrated with 0.1 % TFA (5 ml). The sample was applied to the cartridge, washed with 0.1 % TFA (2 ml) and eluted with 50 % acetonitrile, 6 x 500 µl fractions were collected. The fractions were analysed by LC/MS as described in section 3.2.3.3.

3.2.4.2 Refinement of solid phase extraction methodology

The sample, 5 ml of CML patient eluate was spiked with 50 pmol of p53 (STPPPGTRV) peptide. Protocol: The cartridge was conditioned with 70 % acetonitrile (3 ml) and equilibrated with 0.1 % TFA (5 ml). The sample was applied to the cartridge, washed with 0.1 % TFA (2 ml) and eluted with 70 % acetonitrile, 6 x 500 µl fractions were collected. These were diluted 5 µl to 50 µl with 0.1 % TFA.

The fractions were analysed by LC/MS as described in section 3.2.2 except Injection volume, 50 µl. The mass spectrometric analysis method targeted for MS/MS the doubly charged ($[M+2H]^{2+}$) parent ion with a collision energy of 40%.

3.2.5 Development of SCX methodology

The sample, 5 ml of 25 % acetonitrile was spiked with 25-30 pmol of hep B (TPPAYRPPNAPIL), β-gal (TPHPARIGL), A2 (SSKALQRPV), A3

(KQSSKALQR). A Polysulfoethyl A, 35 x 4.6 mm column (Hichrom, UK) was equilibrated with mobile phase; 0.1 % Formic acid + 5 % acetonitrile at a flow rate of 0.5 ml min^{-1}. Elution was performed with salt injections (1000 µl) of 60, 80, 100, 150, 200, 250, 300, 500 mM KCl + 0.1 % TFA + 5 % acetonitrile, two 500 µl fractions were collected for each salt fraction.

The fractions were analysed by LC/MS as described in section 3.2.2 except, injection volume, 1 µl. The mass spectrometric analysis method targeted for MS/MS the doubly charged ($[M+2H]^{2+}$) parent ion of each of the three peptides with a collision energy of 40 %.

3.2.5.1 SCX using buffering mobile phase

The procedure was repeated using a different mobile phase, 25 mM ammonium formate + 5 % acetonitrile adjusted to pH 3.0 using formic acid. The sample, 70 % acetonitrile with 250 fmol of hep B (TPPAYRPPNAPIL), β-gal (TPHPARIGL), A2 (SSKALQRPV), A3 (KQSSKALQR), diluted 1:2 with mobile phase. Elution was performed with salt injections (1000 µl) of 60, 80, 150, 250, 500 mM KCl + 0.1 % TFA + 5 % acetonitrile. Fractions were collected following the 150, 250, 500 mM salt injections (10 x 100 µl) and the 60 and 80 mM injections (1 x 1 ml).

The fractions were analysed by LC/MS as described in section 3.2.2.

Injection volume: 50 - 100 µl; The mass spectrometry analysis method targeted the doubly charged ($[M+2H]^{2+}$) parent ion of each of the three peptide ions for MS/MS at a collision energy of 40 %.

The procedure was repeated using the following salt fractions: 25, 75, 125, 175, 275, 500 mM KCl + 0.1 % TFA + 5 % acetonitrile, 8 x 100 µl fractions collected for each salt concentration and analysed by LC/MS/MS. The doubly charged ($[M+2H]^{2+}$) parent ion of hep B was targeted for MS/MS analysis. The sample, 70 % acetonitrile with 200 fmol of hep B (TPPAYRPPNAPIL) diluted 1:2 with mobile phase (25 mM ammonium formate + 5 % acetonitrile adjusted to pH 3.0 using formic acid).

The maximum loading volume of the high salt concentration fractions was determined. Samples 100 fmol hep B β-gal and A2 peptide in 50, 100 and 200 µl of 300 mM KCl + 0.1 % formic acid + 5 % acetonitrile. The fractions were analysed by LC/MS as described in section 3.2.2. The loading time included an additional 5 minutes of washing in 0.1 % TFA prior to gradient elution, injection volume, 100 µl. The mass spectrometry analysis method targeted the doubly charged ($[M+2H]^{2+}$) parent ion of each of the peptide ions for MS/MS with a collision energy of 40 %.

3.2.6 Evaluation of an integrated MHC class I clean-up method

Sample 1, 5 ml of a CML patient eluate spiked with 200 fmol each of A3 (KQSSKALQR), A2 (SSKALQRPV), β-gal (TPHPARIGL) and hep B (TPPAYRPPNAPIL) with 0.5 µl of TFA added. Samples 2 and 3 were as 1 except 100 fmol of each peptide was added prior to sample clean-up.

Solid phase extraction protocol.

Two solid phase cartridges (SEP PAK® (Water, UK), RP, C18, 360 mg sorbent, 55-105 µm particle size) connected in series were conditioned with 70 % acetonitrile (6

ml) and equilibrated with 0.1 % TFA (10 ml). The sample was applied and the cartridges washed with 0.1 % TFA (2 ml). The peptides were eluted with 70 % acetonitrile. The first 500 µl was discarded and the following 2-3 ml collected.

SCX protocol

A Polysulfoethyl A, 35 x 4.6 mm SCX column (Hichrom, UK) was equilibrated with mobile phase; 25 mM ammonium formate with 5 % acetonitrile at pH 3.0 with a flow rate of 0.5 ml/min. The sample was injected via a 750 µl injection loop and the column washed for 30-60 seconds. The peptides were eluted with injections of increasing salt concentration, 750 µl each of 25mM, 75mM, 125mM, 175mM, 275mM, 500mM mM KCl + 0.1 % formic acid + 5 % acetonitrile. Fractions (8 x 100µl) were collected starting immediately after each salt injection.

LC/MS Analysis

Injection volume: 100 µl. Trap column: pepmap 300 µm i.d x 5 mm, C18 nanobore column (Dionex, Surrey, UK), loading buffer 0.1 % TFA at a flow rate of ~30 µl min^{-1}. Loading pump (Dionex, Surrey, UK). Analytical column: pepmap 75 µm i.d x 150 mm, C18 nanobore column (Dionex, Surrey, UK). Mobile phase A: H_2O + 0.09 % formic acid. Mobile phase B: 80 % acetonitrile + 0.09 % formic acid. Flow rate: 180 nl min^{-1}, Gradient profile: 12 % B – 75 % B in 90 minutes. Gradient pump: Ultimate™ nano-pump (Dionex, Surrey, UK). Mass spectrometry was performed by on-line micro-electrospray ionisation mass spectrometry on a Finnigan LCQ quadrupole ion trap mass spectrometer (ThermoFinnigan, Hemel Hempstead, UK) using a home-made ion source. Mass spectrometric conditions were as follows: Microspray voltage, 1.5-2.0 kV. Capillary Temperature, 200 °C. Capillary voltage, 46

V. Tube lens offset, -40 V. Collision energy, 40 %. The mass spectrometric analysis method targeted for MS/MS the doubly charged ($[M+2H]^{2+}$) parent ion of each of the four peptides.

3.2.7 Analysis of k562 eluates and CML patient eluates for predicted peptides

K562 eluates were provided by Dr. Tony Dodi at the Antony Nolan Clinic, London, UK. The eluates (~ 5 ml) were spun where necessary (1500 rpm, 10 minutes) to remove particulates. The eluates were spiked with 100 fmol each of the internal recovery standard peptides β-gal (TPHPARIGL) and hep B (TPPAYRPPNAPIL) and 0.5 µl of TFA was added. They were then subjected to the solid phase extraction and SCX protocols which have been described in Section 3.2.5. The resulting fractions were analysed by capillary LC/MS/MS. The target peptides were bcr/abl derived, these are shown as Group 1 in Table 3.2.7.1. Half the fractions were analysed by targeting these peptide masses and the internal recovery standard peptides. The remaining fractions were analysed in data dependent mode using a full scan followed by zoom scan and tandem MS analysis targeting the three most intense ions (m/z 400-1500). Once an ion had been detected it was placed on an exclusion list for 3 minutes. An autosampler was integrated into the analytical system Famos™ (Dionex, Surrey, UK). A program was written to allow 100 µl injections and with a contact closure start to the mass spectrometer.

Some changes were made in subsequent patient and k562 eluates. The survey approach using the data dependent mode was discarded and a further set of bcr/abl derived target peptides was included, these are shown as Group 2 in Table 3.2.7.1.

Table 3.2.7.1 Peptide sequences and m/z targeted during LC/MS/MS analysis of eluates from CML patient cells and k562 transfected cells.

Group 1		Group 2	
Peptide sequence	$[M+2H]^{2+}$	Peptide sequence	$[M+2H]^{2+}$
KALQRPVASD	542.8	MLTNSCVKL	504.8
GFKQSSKAL	483.3	FLNVIVHSA	500.3
ATGFKQSSKA	512.8	FLSSINEEI	526.3
SSKALQRPV	493.3	FMVELVEGA	497.8
		QLLKDSFMV	540.8

3.2.8 MHC class II peptide analysis.

3.2.8.1 Identification of CLIP peptide from immunopurified class II molecules

The clean-up method developed for the purification of MHC class I surface presented peptides was tested for its application to the MHC class II peptides. This was initially applied to class II complexes which had been obtained by immunopurification of lysed T2 DR4 cells. Class II T2 DR4 cells were grown and immunopurified by Danielle Barry, Nottingham Trent University.[19] T2 DR4 cells were cultured in RPMI-1640 media with 10% FCS and 2mM L-glutamine. G418, 1mg/ml, was added to the cells with the media to maintain the transfection. The cells were split regularly when confluent to maintain maximum growth.

Immunopurification of peptides from T2 DR4 cells

Cells were lysed with 0.1 % Triton 20 in PBS with 5 mM EDTA. 1 μg/ml of the proteinase inhibitors leupetin and pepstatin was also added. The cells were subjected

to immunopurification using an ImmunoPure® rProtein A IgG plus orientation kit (Pierce, Rockford, US), which was used according to the manufacturer's instructions. Briefly, the column was equilibrated with ImmunoPure® antibody binding/wash buffer (5 ml), followed by the addition of 2-4 ml of antibody solution, which was suspended by inversion mixing (30 minutes, room temperature). The antibody, L243, was harvested from the mouse B-cell hybridoma cell line which was grown in RPMI-1640 serum free media with 2 mM L-glutamine and 10 % FCS prior to transfer to serum free HL-1 hybridoma media for harvesting of antibody. Excess antibody was washed from the column with wash buffer and the bound antibody cross linked by addition of cross-linking buffer solution and inversion mixing for one hour. The prepared column was then equilibrated in binding buffer (PBS, pH 7.2). The sample was diluted 1:2 with binding buffer and applied to the column and incubated for 1 hour at room temperature. The column was then washed with binding buffer prior to elution with ImmunPure® IgG elution buffer.

The target for the mass spectrometric analysis was the class II invariant chain peptide (CLIP) LPKPPKPVSKMRMATPLLMQALPM. Analysis was performed using capillary LC-nano/ESI/MS/MS using the following conditions. Injection volume: 100 µl. Trap column: pepmap 300 µm i.d x 5 mm, C18 nanobore column (Dionex, Surrey, UK), loading buffer 0.1 % TFA at a flow rate of ~30 µl min^{-1}. Loading pump: Kontron 325 with microflow splitter (SGE, Milton Keynes, UK). Analytical column: pepmap 75 µm i.d x 150 mm, C18 nanobore column (Dionex, Surrey, UK). Mobile phase A: H_2O + 0.09 % formic acid. Mobile phase B: 80 % acetonitrile + 0.09 % formic acid. Flow rate: 180 nl min^{-1}, Gradient profile: 12 % B – 75 % B in 90 minutes. Gradient pump: Ultimate™ nano-pump (Dionex, Surrey, UK).

Mass spectrometry was performed using on-line micro-electrospray ionisation mass spectrometry on a Finnigan LCQ quadrupole ion trap mass spectrometer (ThermoFinnigan, Hemel Hempstead, UK) using a home-made ion source. Mass spectrometric conditions were as follows: Microspray voltage, 1.5-2.0 kV. Capillary Temperature, 200 °C. Capillary voltage, 29 V. Tube lens offset, 0 V. Collision energy, 40 %. The mass spectrometric analysis method targeted for MS/MS the doubly charged ($[M+2H]^{2+}$, m/z 1138.5) and triply charged ($[M+2H]^{3+}$, m/z 892.5) parent ion of the CLIP peptide. A data dependent tandem scan on the strongest ion (m/z 400-1400) was also performed. The data was analysed using the TurboSEQUEST® (Thermofinnigan, Hemel Hempstead, UK) search algorithm against a database containing the entire Invariant chain protein sequence (from SWISPROT).[20]

3.2.8.2 Identification of CLIP from Class II surface eluted cells.

The mutant cell line T2 DR4 was cultured as described above and the class II cell surface peptides eluted using a novel elution buffer containing formic acid, sucrose and KCl.[19] The eluate was spiked with hep (100 fmol) and β-gal (100 fmol) and then subjected to solid phase extraction purification followed by strong cation exchange separation.

Solid phase extraction protocol

Two solid phase extraction cartridges: SEP PAK® (Waters, UK), RP, C18, 360 mg sorbent, 55-105 μm particle size were connected in series were conditioned with 70 % acetonitrile (6 ml) and equilibrated with 0.1 % TFA (10 ml). The sample was applied and the cartridges washed with 0.1 % TFA (2 ml). The peptides were eluted with 70

% acetonitrile, the first 500 µl fraction was discarded and the next 2-3 ml fraction was collected.

Strong cation exchange protocol

A Polysulfoethyl A, 35 x 4.6 mm column (Hichrom, UK) was equilibrated with mobile phase; 25 mM ammonium formate with 5 % acetonitrile at pH 3.0 with a flow rate of 0.5 ml min^{-1}. The sample was diluted 1:2 with mobile phase and injected via a 750 µl injection loop and the column washed for 30-60 seconds. The peptides were eluted with injections of increasing salt concentration, 750 µl each of 25mM, 75mM, 125mM, 175mM, 275mM, 500mM mM KCl + 0.1 % formic acid + 5 % acetonitrile. Fractions (8 x 100µl) were collected, starting immediately after each salt injection.

Each fraction was then subjected to analysis using capillary LC-nano/ESI/MS/MS. Injection volume: 100 µl. Trap column: pepmap 300 µm i.d x 5 mm, C18 nanobore column (Dionex, Surrey, UK), loading buffer 0.1 % TFA at a flow rate of ~30 µl min^{-1}. Loading pump: Kontron 325 with microflow splitter (SGE, Milton Keynes, UK). Analytical column: pepmap 75 µm i.d x 150 mm, C18 nanobore column (Dionex, Surrey, UK). Mobile phase A: H_2O + 0.09 % formic acid. Mobile phase B: 80 % acetonitrile + 0.09 % formic acid. Flow rate: 180 nl min^{-1}, Gradient profile: 12 % B – 75 % B in 90 minutes. Gradient pump: Ultimate™ nano-pump (Dionex, Surrey, UK).

Mass spectrometry was performed using on-line micro-electrospray ionisation mass spectrometry on a Finnigan LCQ quadrupole ion trap mass spectrometer (ThermoFinnigan, Hemel Hempstead, UK) using a home-made ion source. Mass spectrometric conditions were as follows: Microspray voltage, 1.5-2.0 kV. Capillary

Temperature, 200 °C. Capillary voltage, 29 V. Tube lens offset, 0 V. Collision energy, 40 %. The mass spectrometric analysis method targeted for MS/MS the doubly charged ($[M+2H]^{2+}$, m/z 1138.5 and m/z 1167.9) and triply charged ($[M+2H]^{3+}$, m/z 892.5 and 779.1) parent ion of two CLIP peptides ($CLIP_{81-101}$ and $CLIP_{81-104}$). A data dependent tandem scan on the strongest ion (m/z 300-2000) was also performed. The data was analysed using the TurboSEQUEST® (Thermofinnigan, Hemel Hempstead, UK) search algorithm against a database containing the entire Invariant chain protein sequence (from SWISPROT).[20] Potential matches were confirmed by manual matching of predicted fragment ion peaks with tandem mass spectra.

3.3.0 Results and Discussion

The nature of the following investigations into improving and developing a method for the identification of MHC restricted peptides was such that frequently experiments were performed only once as a means of determining degree of losses at various stages of the procedure to establish stages which required revision. As such the work was largely qualitative. Further detail is provided where applicable.

3.3.1 Elution of cell surface peptides

The identification of novel, immunogenic MHC class I and II peptides which are expressed at the surface of cells is important for the development of peptide based vaccines. The identification of MHC class I peptides by mass spectrometry is a growing area of research and a range of techniques have been employed to isolate and characterise antigenic peptides. These procedures were reviewed in Chapter 1. The identification of class II peptides has also been attempted using immunoaffinity purification following cell lysis. However, there is a need for the development of validated methods for the clean-up and identification of peptides, which may be successfully applied to both class I and class II antigens eluted from the surface of cells.

Previous studies in our laboratory have shown the potential of the use of a citrate phosphate buffer at pH 3.3 for the elution of class I peptides from the surface of leukaemic cells.[1,8] Tests were performed to determine cell viability following this elution.[21] Trypan blue dye exclusion and toxilight™ assays were performed on the supernatants of K562-A2 cells before and after mild acid elution. The k562 cell line

was derived from a patient suffering from chronic myeloid leukaemia (CML) who was in blast crisis. It is a highly undifferentiated cell line originating from a pluripotential haematopoietic stem cell. The cell line does not express any HLA molecules on its surface. It can, however, be transfected with genes for the surface expression of MHC class I molecules.[22,23] This enables transfection with a haplotype of choice for example A2.

The Trypan Blue dye exclusion assay is the standard approach for the determination of cell viability. The assay is based on the reactivity of the dye's chromophore with the cell. The chromophore is negatively charged and does not interact with the cell unless the membrane is damaged. Viable cells will therefore exclude the dye and can be seen under a microscope.[24]

The toxilight™ assay detects adenylate kinase which is released from cells when damage to the plasma membrane occurs. Following mild acid elution of K562-A2 cells 76 % of cells remained viable. The Trypan blue exclusion assay detected 14 % cell death. Storkus *et al*[18] reported 22 % cell lysis after a 5 minute incubation with buffer, which is comparable to the results obtained in this work. These data confirm that mild acid elution with citrate phosphate buffer results in minimal cell lysis and that peptides extracted into the buffer are more likely to have originated from the cells surface than from the intracellular compartment. Moreover, we can say that if peptides cannot be detected in the cell surface eluates, then they could not be associated with the MHC class I molecule on the cell surface and are therefore unlikely to be promising vaccine candidates. The disadvantage of the mild acid elution approach is that the peptide content of the elution buffer is much more

complex than lysates purified by immunoaffinity procedures, so a highly efficient clean-up is required. This chapter presents the development of a method for the identification of class I and II associated cell surface peptides without cell lysis, using a mild acid wash.

3.3.2 The recovery of standard peptides

The MHC class I method described in Section 3.2.2, and shown schematically in Figure 3.3.2.1, was developed from a previously reported procedure for the detection of the A3 peptide (KQSSKALQR).[1,7,8]

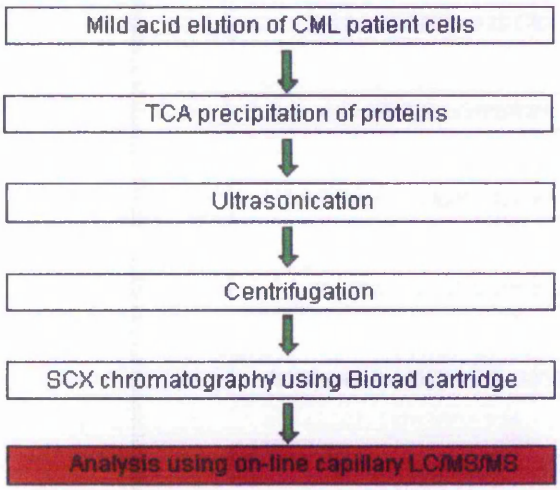


Figure 3.3.2.1 Schematic representation of the class I clean-up method. The method has been amended from the published method to incorporate on-line RP-HPLC/MS/MS analysis.

The method was evaluated for the purification of a solution of citrate phosphate buffer, pH 3.3, spiked with ~ 50 fmol of the hepatitis B peptide (TPPAYRPPNAPIL). The hep B peptide has previously been used as an internal recovery standard. As the peptide is a class II sequence of viral origin it would not be presented on the surface of class I molecules and would not be found in cells free of hepatitis B infection.

The purpose of this experiment was to test the method for the recovery of hep B in buffer, prior to the application of the method to the analysis of CML patient eluates. The buffer containing hep B was subjected to the clean-up procedure described in Section 3.2.2 and the SCX column eluate was split into two aliquots (1 ml each) for analysis. Aliquot 1 was analysed using a data dependent scan sequence to qualitatively confirm the presence of hep B peptide. The second aliquot was run in full scan mode and the peak area compared with a non-extracted standard to determine the recovery of the spiked peptide through the clean-up and loading procedure. The LC/MS data obtained following this eluate clean-up can be seen in Figure 3.3.2.2. The hep B spike was identified during data dependent analysis and tandem MS was performed on the $[M+2H]^{2+}$ ion revealing a fragmentation spectrum showing good coverage of b and y ions, as can be seen in Figure 3.3.2.2 (c), which corresponds to the MS/MS spectrum of the hep B standard.

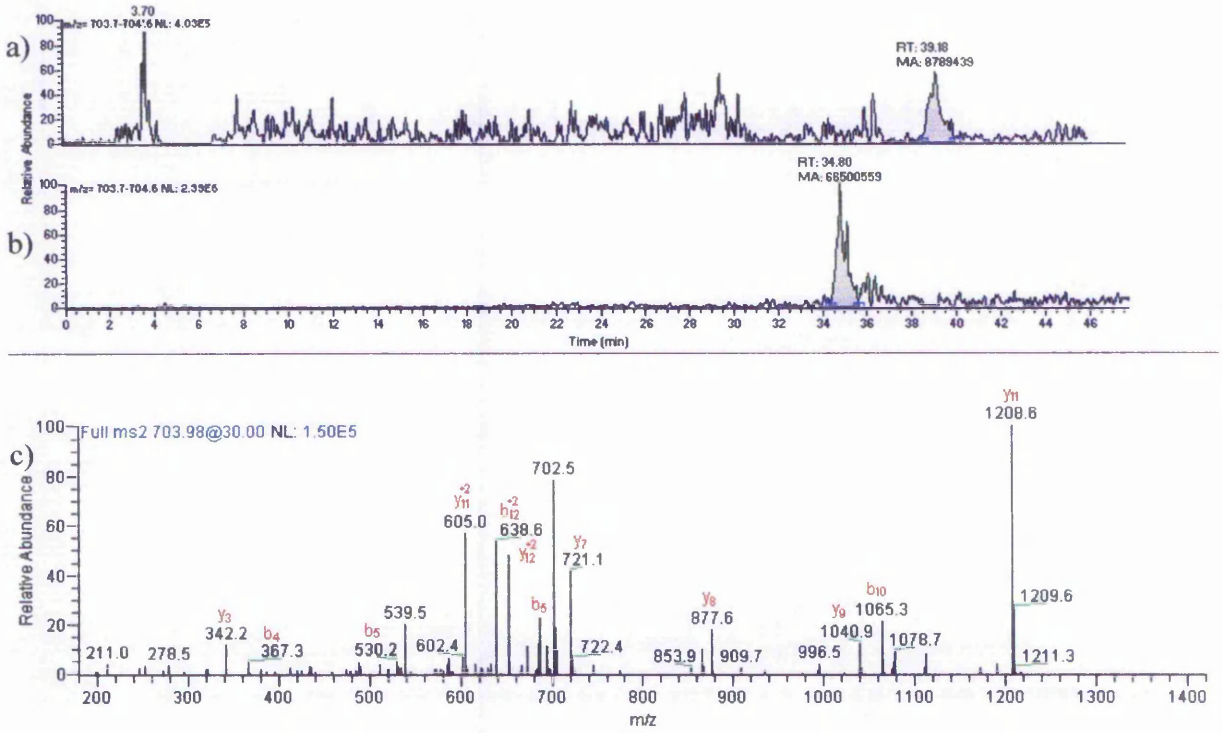


Figure 3.3.2.2 LC/MS analysis of citrate phosphate buffer spiked with the hep B peptide following clean-up procedure and a non-extracted standard, (a) Single ion chromatogram ($[M+2H]^{2+}$ m/z 704) of hepatitis B peptide in spiked buffer, (b) Single ion chromatogram ($[M+2H]^{2+}$ m/z 704) of a non-extracted standard containing hepatitis B peptide, (c) tandem mass spectrum of the $[M+2H]^{2+}$ ion (m/z 704.3) showing fragmentation ions and assigned to hepatitis B peptide.

The difference in the retention time of the hep B peptide between the two chromatograms is due to the difference in the sample injection volume. The extracted sample is 1000 µL and so the loading time compared with the 1 µL standard is much greater, therefore, back-flushing the sample from the trap column onto the analytical column takes longer.

Calculation of recovery

The recovery of the hep B peptide through the clean-up procedure was calculated using the peak areas obtained for the $[M+2H]^{2+}$ ion from the sample analysis (Figure 3.3.2.2 a) and the standard analysis (Figure 3.3.2.2 b) using equation 3.3.2.1.

$$\left[\frac{\text{Sample peak area}}{\text{Standard peak area}}\right] \times 100\ \% \qquad \left[\frac{8789439}{66500559}\right] \times 100\ \% = 13\ \% \qquad \textbf{Equation 3.3.2.1}$$

Clearly this level of recovery is very low, particularly when considering the simplicity of this spiked buffer sample compared to the complexity anticipated for patient cell eluate samples. The spiked peptide is being lost during one or all stages of the extraction procedure.

Two CML patient samples were surface eluted with citrate phosphate buffer. The eluates were spiked with hep B and were also subjected to the same clean-up and analysis procedure. In both cases the hepatitis B peptide spike was extremely weak in the full scan spectrum and in neither case was the parent ion selected for further data dependent analysis thus yielding no fragmentation spectra. It appears that the poor recovery observed when using this method on a citrate phosphate buffer spike has been reduced still further in the presence of a complex patient eluate matrix. It is clear that although this method was previously suitable for the identification of the A3 bcr/abl peptide, due to the unusually basic nature of the peptide, it is not suitable for use when searching a larger peptide pool. Therefore a more general method of clean-up was required.

3.3.2.1 Peptide recovery through SCX

The SCX method described in Section 3.2.2.1 was tested with a 20 mL sample of citrate phosphate buffer, pH3.3, spiked with five synthetic MHC peptides to determine the level of losses taking place during this step in the clean-up procedure. A non-extracted standard solution of equivalent concentration, approximately 40 fmol, was also analysed to allow a calculation of recovery of peptides from the SCX column, determined by LC/MS. The resulting chromatograms are shown in Figure 3.3.2.3. The recoveries for the synthetic peptides for a single experiment, calculated using the LC/MS peak areas are given in Table 3.3.2.1. These recoveries indicate that significant amounts of all the peptides are lost during this stage in the purification.

Table 3.3.2.1 Peptide recoveries through the SCX cartridge

Peptide	Sequence	$[M+2H]^{2+}$	% Recovery
$p53_{(65-73)}$	RMPEAPPV	484.3	8
β-gal	TPHPARIGL	481.3	22
gp70	SPSYVYHQF	564.5	15
Hep B	TPPAYRPPNAPIL	704.3	42
$p53_{(264-272)}$	LLGRNSFEV	517.8	24

The reason such poor recoveries are observed during this stage is not certain however the presence of TCA in the sample significantly lowers the pH of the sample solution below the optimum operating pH of the cation exchange column. This would adversely affect peptide binding.

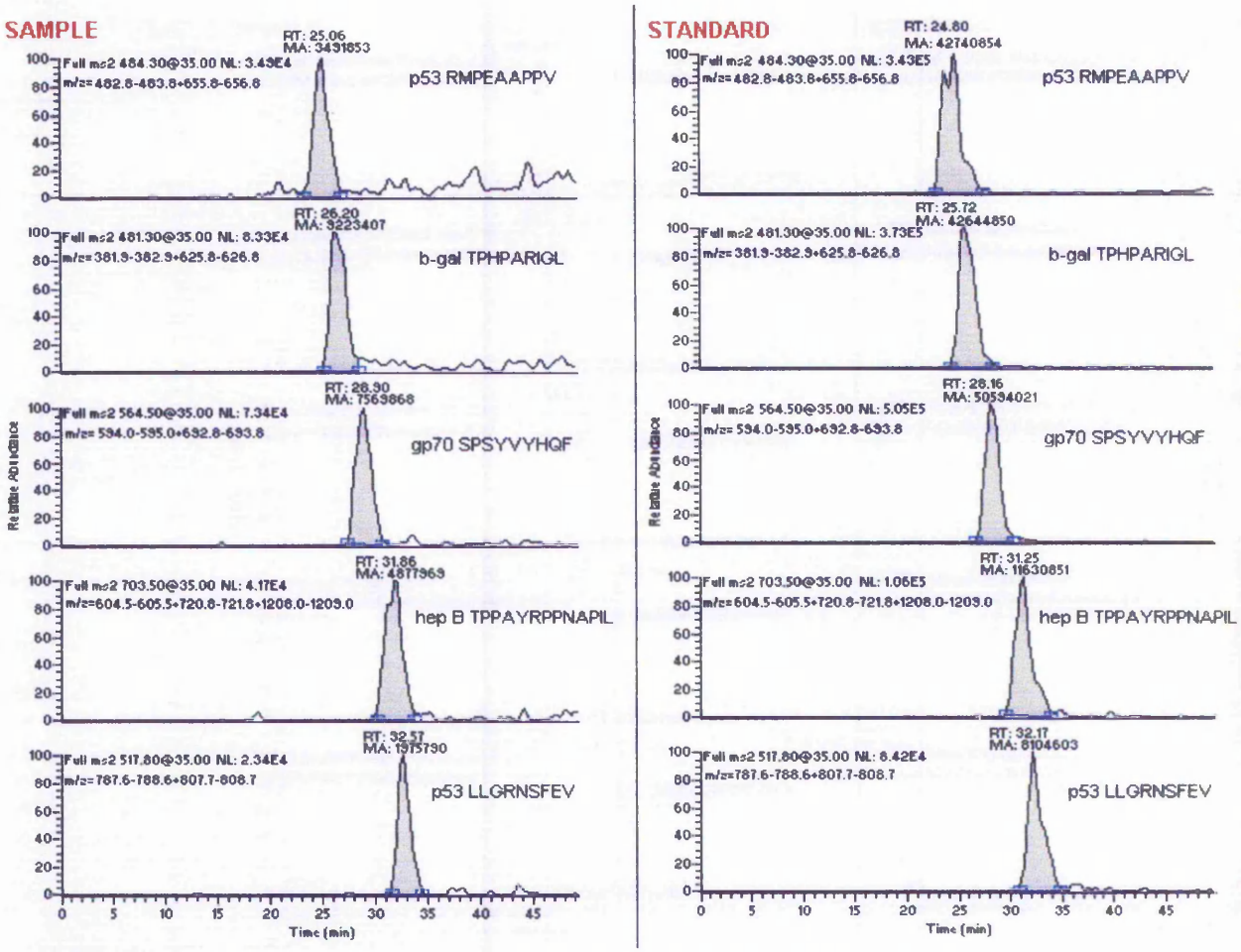


Figure 3.3.2.3 Selected product ion chromatograms of the five peptides used to test recovery through the SCX column (Biorad). Both the SCX sample and the corresponding standards are shown.

It was decided that repetition of this and similar experiments was unnecessary as the aim was to identify areas of the methodology in which significant losses were occurring rather than to quantify them absolutely.

The chromatograms shown in Figure 3.3.2.3 illustrate similar retention times for some of the peptides. Co-elution can lead to ion suppression, although it was not observed in this case. Any ion suppression would be mirrored in the standard analysis.

3.3.2.2 Use of reduced bed volume SCX column

The SCX polishing column has a smaller bed volume than the Biorad cartridge used previously and may be used in conjunction with a HPLC pump. The column was connected to an HPLC pump and UV detector as shown in Figure 3.2.2.1 and the loading buffer was pumped continuously through the column at 0.5 ml min^{-1}. The injector was fitted with a 500 µl injection loop which was used to introduce sample or eluent to the column. A test of the column was performed by loading a mixture of four peptides spiked into citrate phosphate buffer (pH 2.8) in 0.1 M HCl (pH 2) and eluted with 0.1 M NaOH (pH 13). This experiment was carried out only once in order to determine if a simple change to the original method, increasing the pH, together with incorporating an HPLC pumped system for a more consistent environment would have a beneficial effect on peptide retention. The recovery values show significant improvement compared to the Biorad cartridge and the previous method (see Table 3.3.2.2). Once again accurate quantification was not the aim of this experiment, rather, the aim was to ascertain the degree of loss which had occurred during the procedure. This was achieved by comparison of extracted and non-extracted peak areas. Additionally a qualitative comparison can be made between the results obtained here (Table 3.3.2.2) and in section 3.3.2.1 (Table 3.3.2.1). This indicates that the degree of recovery has been improved here.

Table 3.3.2.2 Peptide recoveries through the reduced bed volume SCX column

Peptide	Sequence	% Recovery
$p53_{(65-73)}$	RMPEAAPPV	38
gp70	SPSYVYHQF	93
Hep B	TPPAYRPPNAPIL	71
$p53_{(264-272)}$	LLGRNSFEV	55

The TCA precipitation step has not been included in this experiment therefore the sample pH of 2.8 is greater than after addition of TCA. A pH of 2.8 is not outside the optimum binding range of the SCX column. The new column coupled with the higher pH of the sample solution during loading has had a positive effect on the recovery of peptides in a standard solution. It was therefore of interest to investigate alternative procedures for protein removal required for a matrix of much greater complexity.

A major problem with the Biorad cartridge and polishing SCX column arises as a result of the TCA precipitation stage which immediately precedes it. TCA precipitation is used to precipitate the extraneous protein found in the patient eluates, particularly the β_2-microglobulin which is detached from the cells surface during the mild acid wash. Protein precipitation by pH adjustment has been used as a simple and inexpensive way to precipitate proteins.[25] This involves lowering the pH to a level (< pH 1) outside the optimum binding range of the cation exchange stationary phase. This is demonstrated above with the improved recoveries observed when the pH of the loading solution is raised. There are two potential solutions to this problem:

- Retain the TCA precipitation step and adjust the pH of the solution back up to an acceptable level prior to SCX clean-up.
- Remove the TCA precipitation step and find another way to remove protein from the eluate.

3.3.3 Investigation of protein removal methods

3.3.3.1 TCA precipitation and NaOH neutralisation

The possibility of continuing with the TCA precipitation step to remove the proteins, followed by raising the pH with NaOH to ~3 was explored using the hep B peptide (TPPAYRPPNAPIL) at both a low level (~100 fmol) and a high level (~100 pmol). The solutions were then subjected to SCX and LC/MS/MS analysis. In both cases the results were not encouraging with the hep B peptide showing very poor recovery, below the limit of detection of the system, therefore the data are not shown and this method was not pursued.

3.3.3.2 Protein precipitation by centrifugation

The possibility of removing the extraneous protein material from the eluates by centrifugation was investigated. A test solution of 1 mg ml^{-1} BSA in water was used to simulate the β_2-microglobulin and other proteins which may be present in cultured cells and patient eluates. The solution was centrifuged at 100 000 rpm and the absorbance at 278 nm (proteins in solution absorb UV light with maxima at ca. 200 and 280nm)[26] measured at various intervals to determine the length of time required for the protein to be removed from the solution. The resulting graph of absorbance, corresponding to protein content, against time is shown in Figure 3.3.3.1. The graph shows that the absorbance is steadily decreased as the centrifugation time increased, so the level of protein in solution was being reduced during spinning. However, the rate of decrease is quite slow with the reduction of protein concentration to approximately 70 % in 75 minutes. This appears to be a time consuming way of removing the protein in solution, which does not provide any other benefits, such as

reducing the volume or altering the salt content of the eluate. As this initial test did not yield any promising data it was not pursued any further.

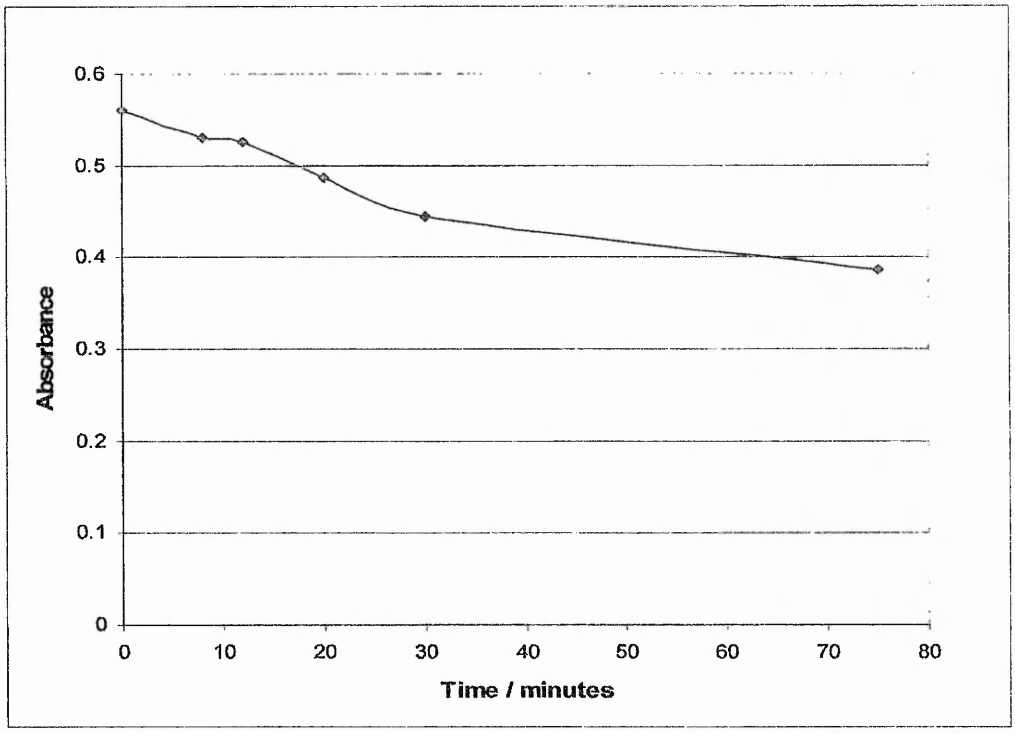


Figure 3.3.3.1 The variation of absorbance (278 nm) with centrifugation time for a BSA solution.

3.3.3.3 Removal of protein from patient eluates using ODS

The use of C18 columns was explored as a means of removing the proteins in patient eluates. This approach would have the added advantage of removing the citrate and phosphate salts present in these eluates and preconcentrating the samples to a more manageable volume. A Phenomenex ODS guard column was loaded with a sample of citrate phosphate buffer, spiked with three peptides hep B, β-gal and p53. The column was washed with 0.1 % TFA and eluted with 60 % acetonitrile at 200 μl min^{-1}. Fractions (50 μl) were collected at 15 s intervals. The resulting chromatograms are shown in Figure 3.3.3.2. Initial experiments were performed in the absence of protein to ascertain if the ODS column could retain peptides in the presence of large amounts

of buffer alone. The detection of the three peptides is observed in fractions eluted from the ODS column after loading in the presence of the citrate phosphate buffer. The volume in which the peptides were eluted has also been reduced approximately 10 fold compared to the volume of buffer in which they were loaded onto the column. This shows excellent promise in terms of removal of the buffer salts and a reduction in sample volume. It should also be possible to isolate the peptides from the protein using their relative affinities for the ODS stationary phase. The proteins have a greater affinity for the ODS than the peptides due to their much greater length and therefore the peptides can be selectively eluted from the ODS phase using a lower level of organic mobile phase modifier leaving proteins bound to the stationary phase. The ability of the C18 column to retain peptides in the presence of citrate phosphate buffer in this single experiment demonstrates a potential avenue of exploration, however, the capacity of guard column would not be sufficient for the kind of load which a patient eluate sample would represent and significant breakthrough would be expected. A solution to this problem would be to use a column of greater capacity with the same retention characteristics. A C18 solid phase extraction (SPE) cartridge fulfils these requirements, and it would be very easy to load the large (~5-20 ml) patient eluate volumes.

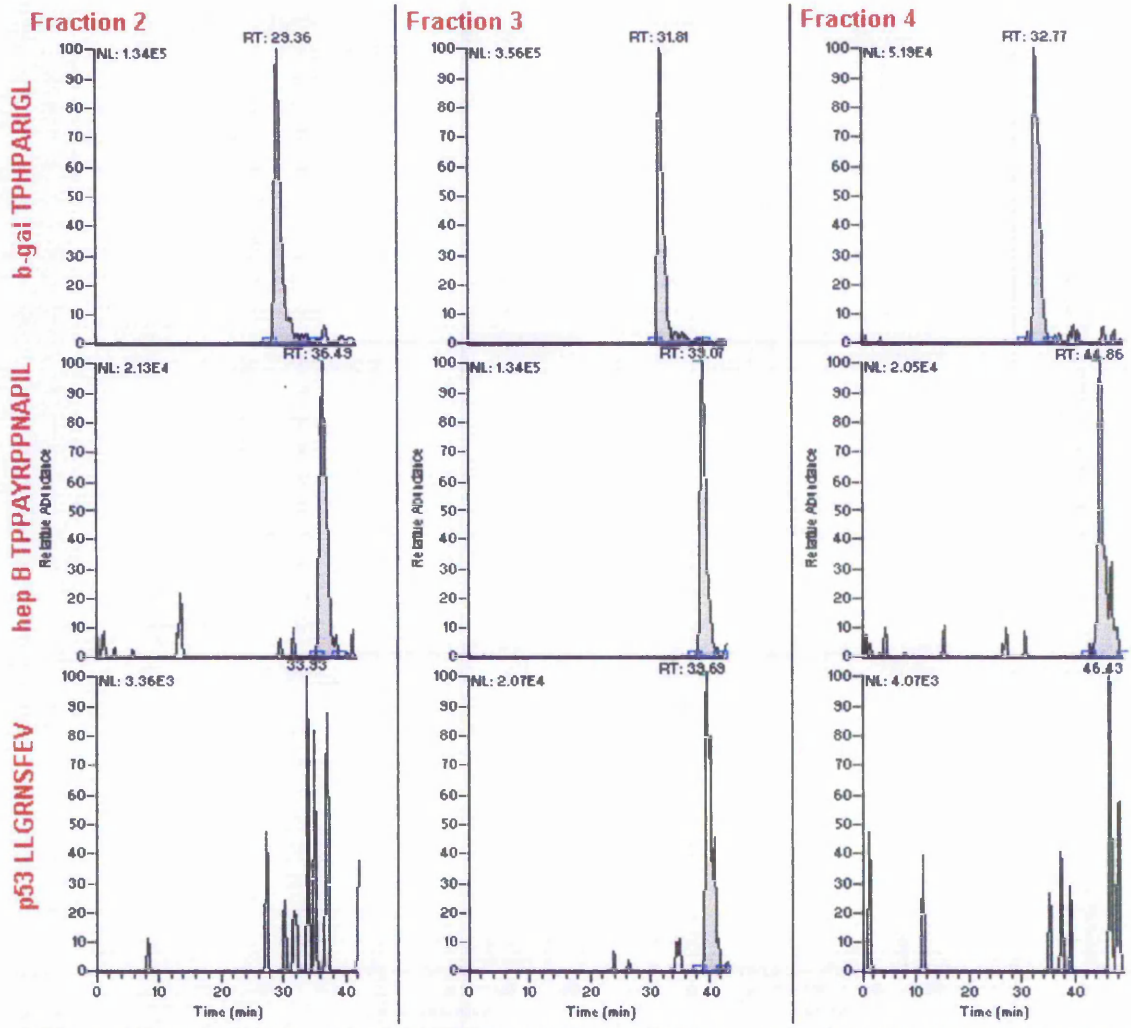


Figure 3.3.3.2 Selected product ion chromatograms of the three peptides (β-gal, hep B, p53) in three fractions eluted from the Jupiter HPLC ODS guard column.

3.3.4 Removal of proteins from patient eluates using solid phase extraction

3.3.4.1 Initial tests of solid phase extraction for recovery of peptides

A solution of the same mixture of standard peptides in citrate phosphate buffer (5 ml) was applied to a conditioned solid phase extraction cartridge in 0.1 % TFA and eluted with 50 % acetonitrile prior to the analysis of fractions by LC/MS. The peptide chromatograms shown in Figure 3.3.4.1 indicate that the peptides can be recovered from the solid phase extraction cartridges following loading in the presence of citrate

phosphate buffer. The ability of the cartridges to retain peptides in the presence of the buffer salts indicated that further tests of this methodology in the presence of more complex cellular matrices should be carried out.

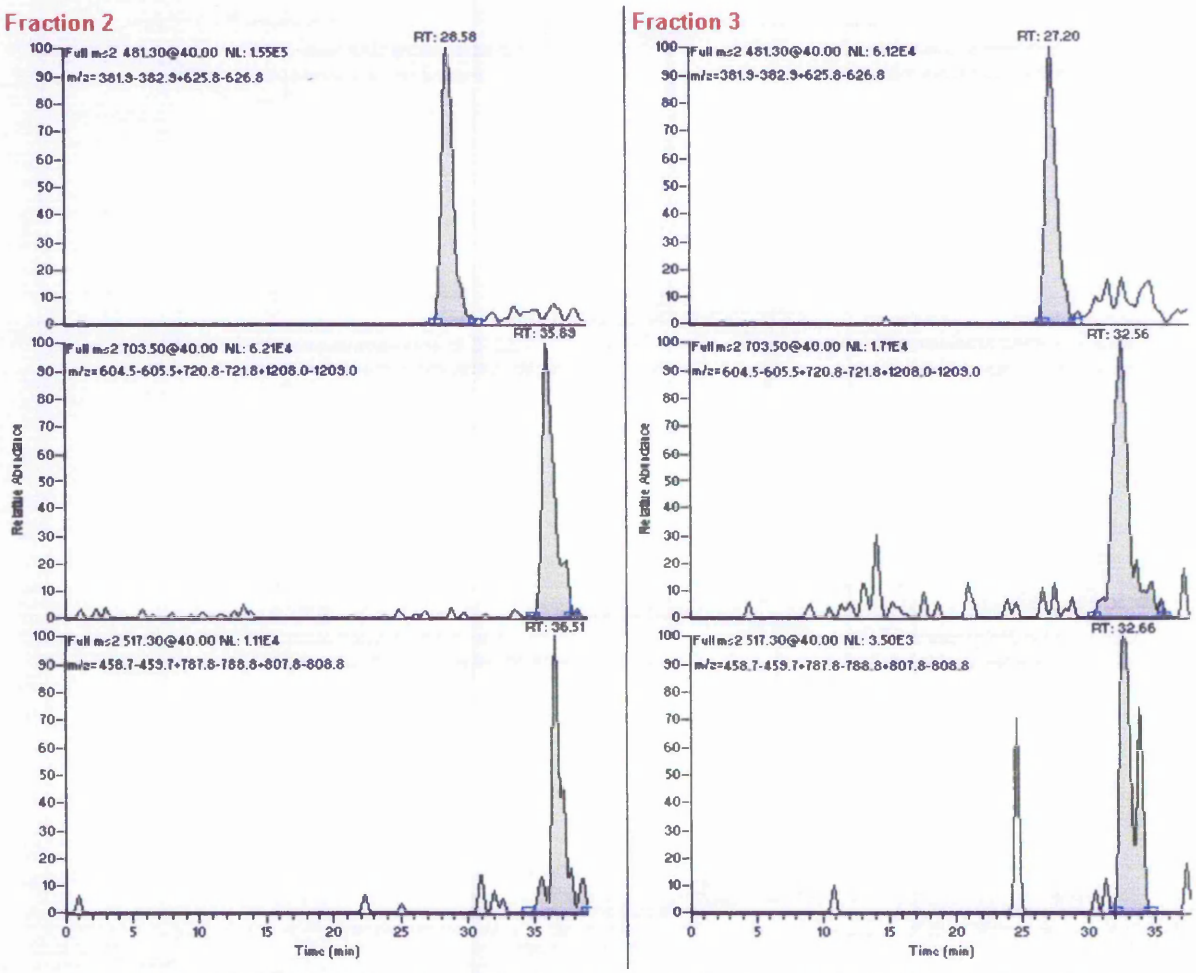


Figure 3.3.4.1 Selected product ion chromatograms of the three peptides (β-gal, hep B, p53) illustrating the recovery of peptides from the solid phase extraction cartridges over two fractions.

3.3.4.2 Refinement of solid phase extraction methodology

An initial attempt to use the SPE methodology described in section 3.3.4.1 for patient eluates spiked with peptides resulted in unacceptably poor peptide recoveries, <10 %. A possible explanation was that the capacity of the SPE cartridge was exceeded when used for the much more complex matrix of a CML patient eluate. This was confirmed

visually as the sample solution, which has a yellow colour, could be seen to travel through the cartridge stationary phase and saturate it. Two cartridges were therefore connected in series to double the capacity. This arrangement was tested by spiking a patient eluate with 50 pmol of the p53 peptide (STPPPGTRV). The elution solvent concentration was also increased from 50 % to 70 % acetonitrile to ensure the elution strength of the solution was sufficient to desorb the peptides. The eluate was collected in 500 µl fractions and the resulting chromatograms for fractions 3 and 4 are shown in Figure 3.3.4.2

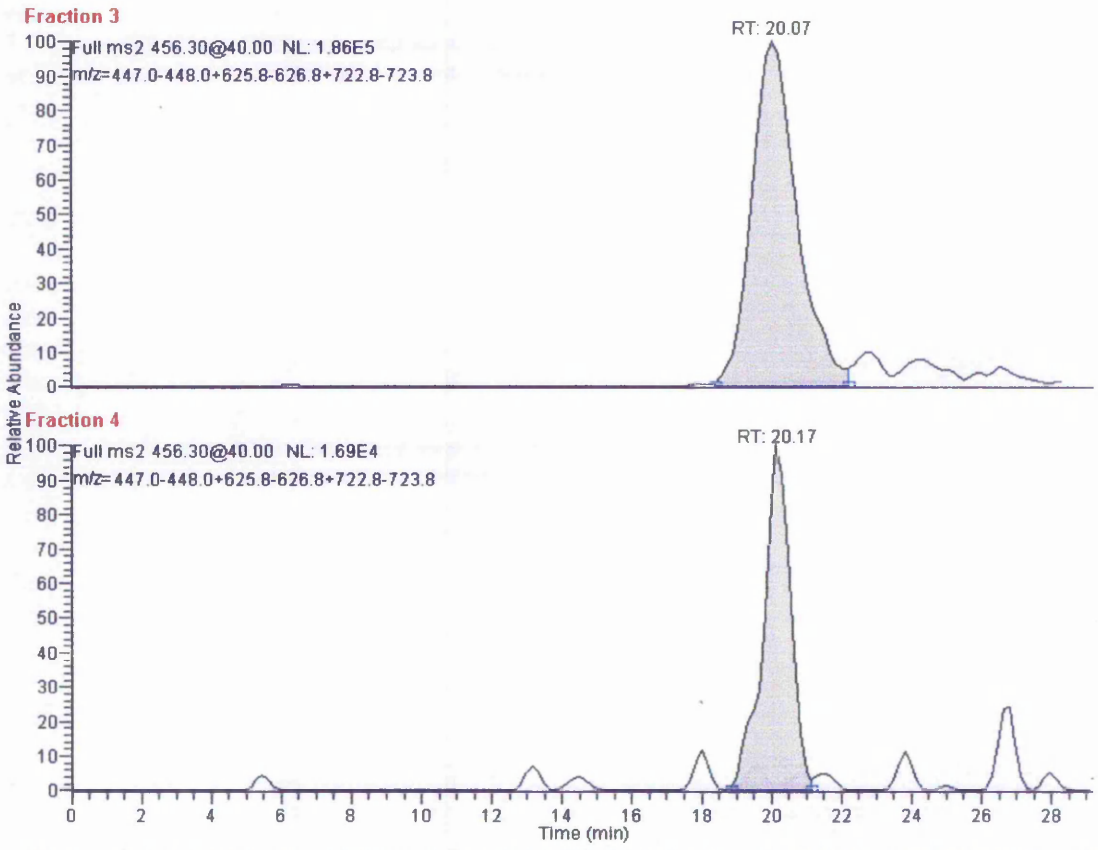


Figure 3.3.4.2 Selected product ion chromatograms (m/z 447 + 626 + 723) of the p53 (STPPPGTRV) peptide recovered from the SPE cartridge in the presence of a CML patient eluate.

Figure 3.3.4.2 shows that the addition of a second SPE cartridge, doubling the capacity of the column, has resulted in retention of the p53 peptide by the ODS stationary phase, in the presence of the CML patient matrix and the citrate phosphate buffer. Elution of the peptide was performed using 70 % acetonitrile and the main fractions in which the peptide has eluted have been identified as the third and fourth fraction after application of the eluting solvent. Recoveries were not calculated but overall recoveries are given in Section 3.3.6. Although performed with a relatively high level of peptide, the success of this test showed that SPE clean-up offers an alternative method to replace the TCA precipitation step which, in addition to removing extraneous proteins, also reduces the eluate volume, pre-concentrates the peptides, and removes the citrate phosphate buffer. Once again this initial test demonstrated a potential solution to the removal of proteins from surface eluates.

3.3.5 Development of SCX methodology

Attention was then directed towards the efficiency of the SCX stage of the clean-up procedure. A popular choice of column for peptide analysis is a polysulfoethyl strong cation exchange column (Hichrom, UK). These columns are suitable for incorporation into an on-line set-up utilising HPLC pumps and injection valves, unlike the Biorad cartridges which had been used previously. This allows for more efficient sample processing. The choice of an analytical SCX column is intended to alter the SCX methodology away from an SPE-type load on/wash off method and towards a true separation procedure. Elution is therefore performed using a salt solution which can be varied to allow stepwise elution of peptides. The instrumental configuration used for these experiments is shown in Figure 3.2.2.1. A 25 % acetonitrile solution was

spiked with 25-30 pmol of four peptides, hep B (TPPAYRPPNAPIL), β-gal (TPHPARIGL), A2 (SSKALQRPV) and A3 (KQSSKALQR). Elution from the column was achieved using potassium chloride salt solutions in a stepwise fashion which were collected and tested for the presence of peptide. Figure 3.3.5.1 shows the selected product ion chromatograms for each of the recovered peptides in the appropriate salt fraction.

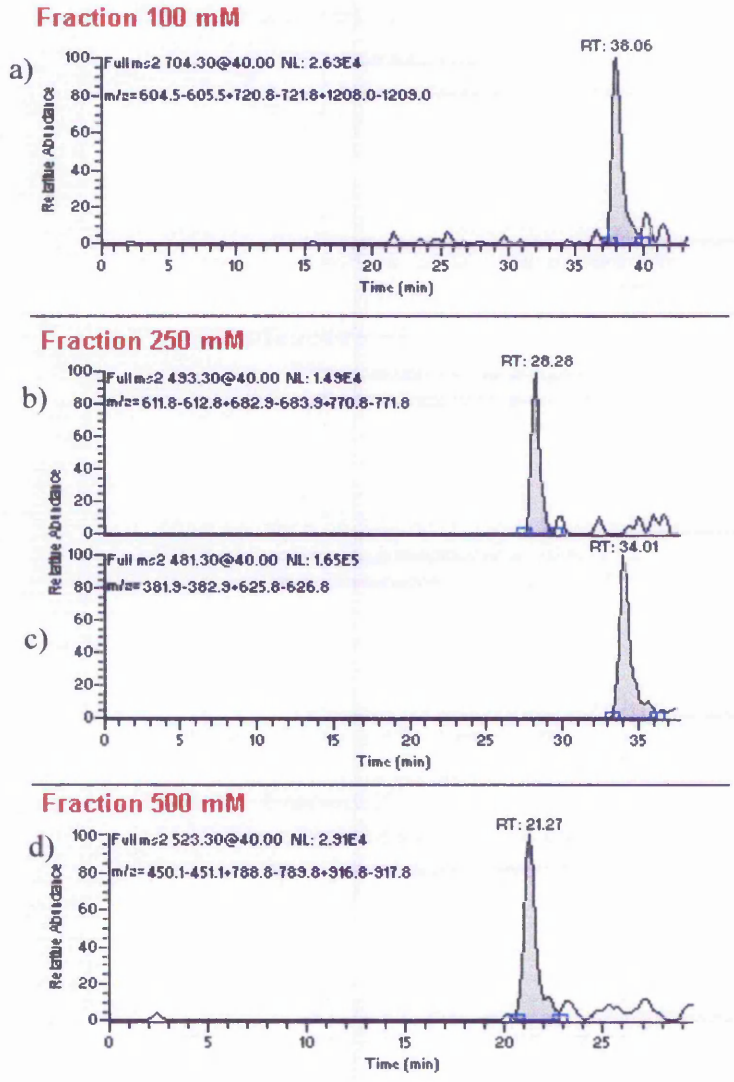


Figure 3.3.5.1 Selected product ion chromatograms illustrating the recovery of peptides in salt fractions eluted from the Polysulfoethyl SCX column. (a) Fraction 100 mM contains the hep B

(TPPAYRPPNAPIL) peptide. (b) Fraction 250 mM contains both the A2 (SSKALQRPV) and β-gal (TPHPARIGL) peptides. (c) Fraction 500 mM contains the A3 (KQSSKALQR) peptide.

The chromatograms shown in Figure 3.3.5.1 illustrate the initial success of tests using the new SCX column. The hep B is the first eluting peptide in the 100 mM fraction, this tallies with previous experiments which have shown the hep b peptide to have the lowest affinity for SCX stationary phase. Conversely, the A3 peptide has the highest affinity of those tested, for the SCX stationary phase and therefore requires a much stronger salt concentration, 500 mM in order to achieve elution from the column. These are the two extremes of the situation, it is expected that the majority of peptides encountered will experience a medium affinity for the SCX stationary phase as it is likely that the majority of class I peptides will contain at least on basic residue, based on their natural abundance (~5.7 % for both K and R)[25] as is the case for both the A2 and β-gal peptides.

This was a successful test which was performed only once as it was decided to incorporate a buffering mobile phase in order to regulate the peptide on-column environment for a more robust final method.

3.3.5.1 SCX using buffering mobile phase

Further to these preliminary tests, alternative SCX mobile phases were investigated which would create a more buffered environment for the column bed and so improve reliability of this column loading stage of the clean-up. The 0.1 % TFA loading buffer was therefore replaced with 25 mM ammonium formate with 5 % acetonitrile. The number of salt injections was reduced and the number of fractions collected per salt concentration was increased. The initial salt fractions of 60 and 80 mM were collected

as a 1 ml aliquot, whilst higher salt fractions were collected in 100 µl fractions. On LC/MS analysis of the SCX fractions, it became apparent that the peptide elution had shifted to lower salt concentrations. The hep B, A2 and β-gal peptides were all observed to elute using salt concentrations in the range 60-150 mM. The corresponding selected product ion chromatograms of these peptides in their main fraction are shown in Figure 3.3.5.2

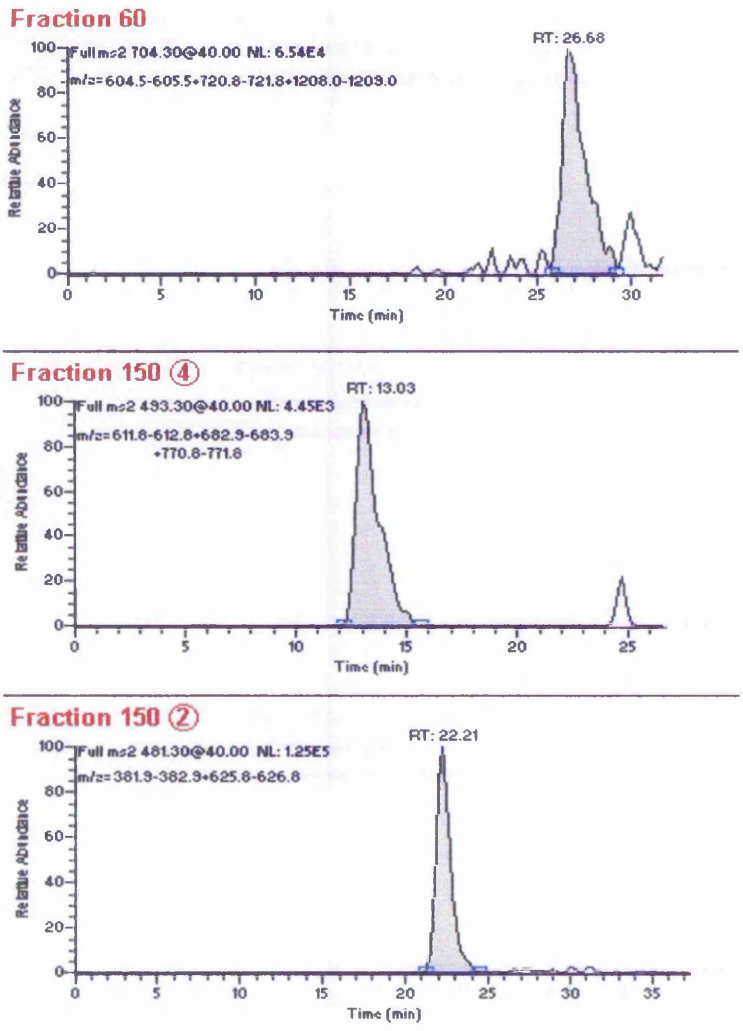


Figure 3.3.5.2 Selected product ion chromatograms illustrating the recovery of peptides from the SCX column following loading in 25 mM ammonium formate + 5 % acetonitrile mobile phase. (a) hep peptide recovered from fraction 60 mM. (b) A2 peptide recovered from fraction 150 mM, 4. (c) β-gal peptide recovered from fraction 150 mM, 2.

The use of a buffering mobile phase during loading was a successful modification to the SCX protocol that would hopefully provide a greater degree of reproducibility and enhance the robustness of the method due to the increased stability of the column environment. The salt concentrations used to elute the peptides from the SCX column were modified following the change in elution properties observed for the buffered mobile phase to: 25, 75, 125, 175, 275, 500 mM KCl + 0.1 % formic acid + 5% acetonitrile. The increase in concentration is steeper after 175 mM to reflect the reduced likelihood of peptides eluting at higher salt concentrations, the majority are expected to elute with salt injections below 275 mM. Of those peptides used in the method development only the A3 peptide (KQSSKALQR) with its three basic residues and strong affinity for the SCX stationary phase required a stronger salt concentration to elute. A test of the SCX methodology incorporating these changes was performed using the hep B peptide with the eluent from the column collected in 8 x 100 µl aliquots for each 750 µl KCl salt injection. The results of this test are shown in Figure 3.3.5.3. Each experiment was performed singularly as the SCX method was being continually refined and developed, here it was necessary to observe the elution trends of the standard peptides with the altered elution salt concentrations.

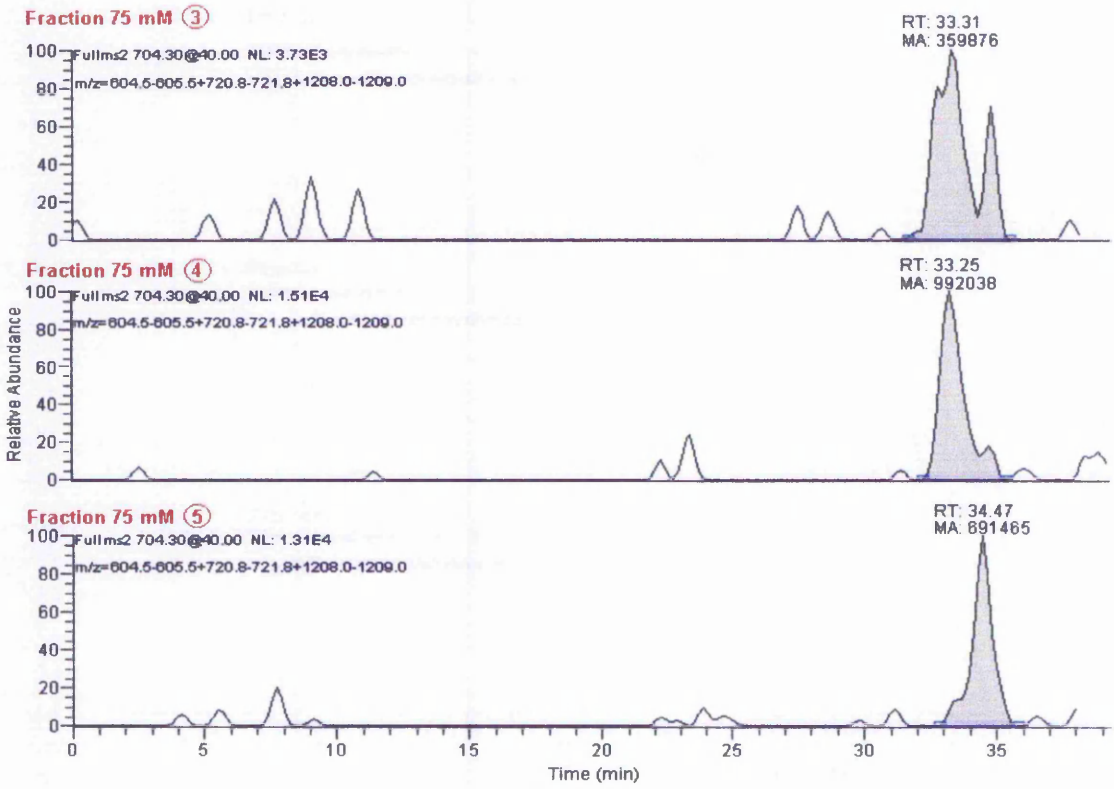


Figure 3.3.5.3 Selected product ion chromatograms showing the recovery of hep B peptide (TPPAYRPPNAPIL) from the SCX column in the 75 mM salt fractions (a) Fraction number 3. (b) Fraction number 4. (c) Fraction number 5

The hep B peptide eluted in the 75 mM salt fraction, over three of the eight 100 µl fractions collected. The distribution over these fractions is shown graphically in Figure 3.3.5.4. These data suggest that ~80 % of the total hep B elutes in fractions 4 and 5, so analysis of these fractions alone will result in little loss in sensitivity.

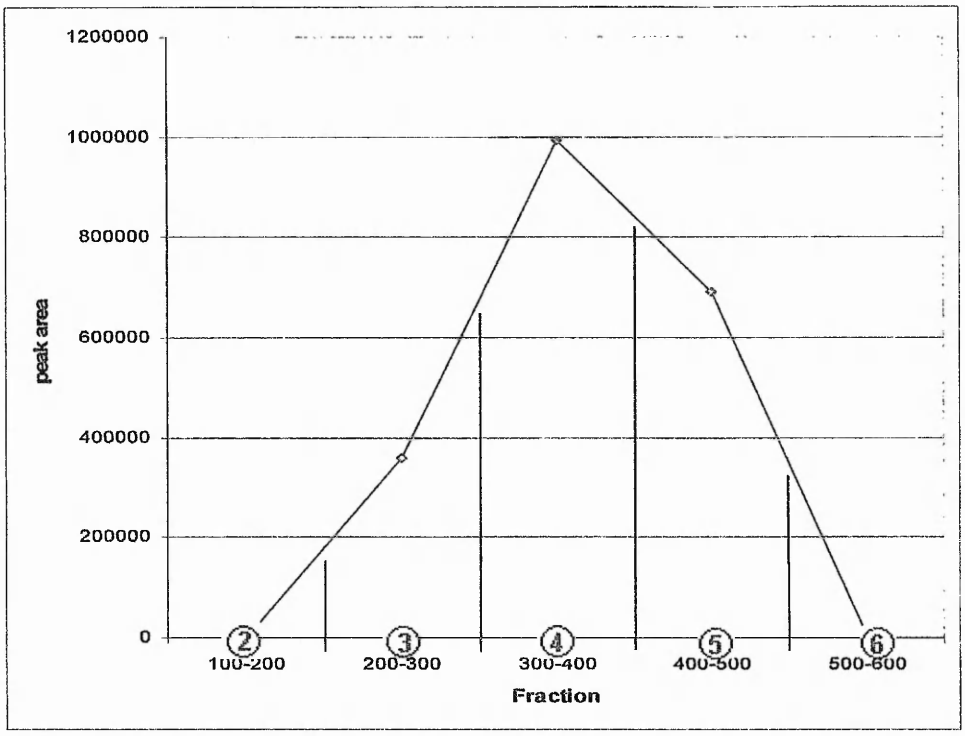


Figure 3.3.5.4. The distribution of the hep B peptide over the fractions of the 75 mM salt injections.

The presence of high concentrations of potassium chloride in the SCX fractions results in limitations on the volume which can be loaded onto the 300 μm trap column during LC/MS/MS analysis without breakthrough. It has been shown in Section 3.3.4 that peptides typically elute from the SCX column in a volume 300-400 μl. The maximum volume which could be loaded onto the trap column was explored. A qualitative comparison of loading 50, 100 and 200 μl fractions of the hep B, β-gal and A2 peptides in 300 mM KCl was made, as this represents the upper end of the potential salt concentrations which will be routinely analysed. The 50 and 100 μl peptide injections resulted in detection of comparable levels of all three peptides. However attempts to load 200 μl of salt and peptide resulted in losses, particularly for the A2 peptide. The maximum loading volume was found to be 100 μl for a procedure including loading the salt fraction and a 5 minute wash of the trap column with 0.1 %

TFA loading buffer prior to gradient elution. Figure 3.3.5.5 shows the selected product ion chromatograms and MS/MS spectra obtained for the A2, β-gal and hep B peptides loaded in 100 µl KCl solution (300mM).

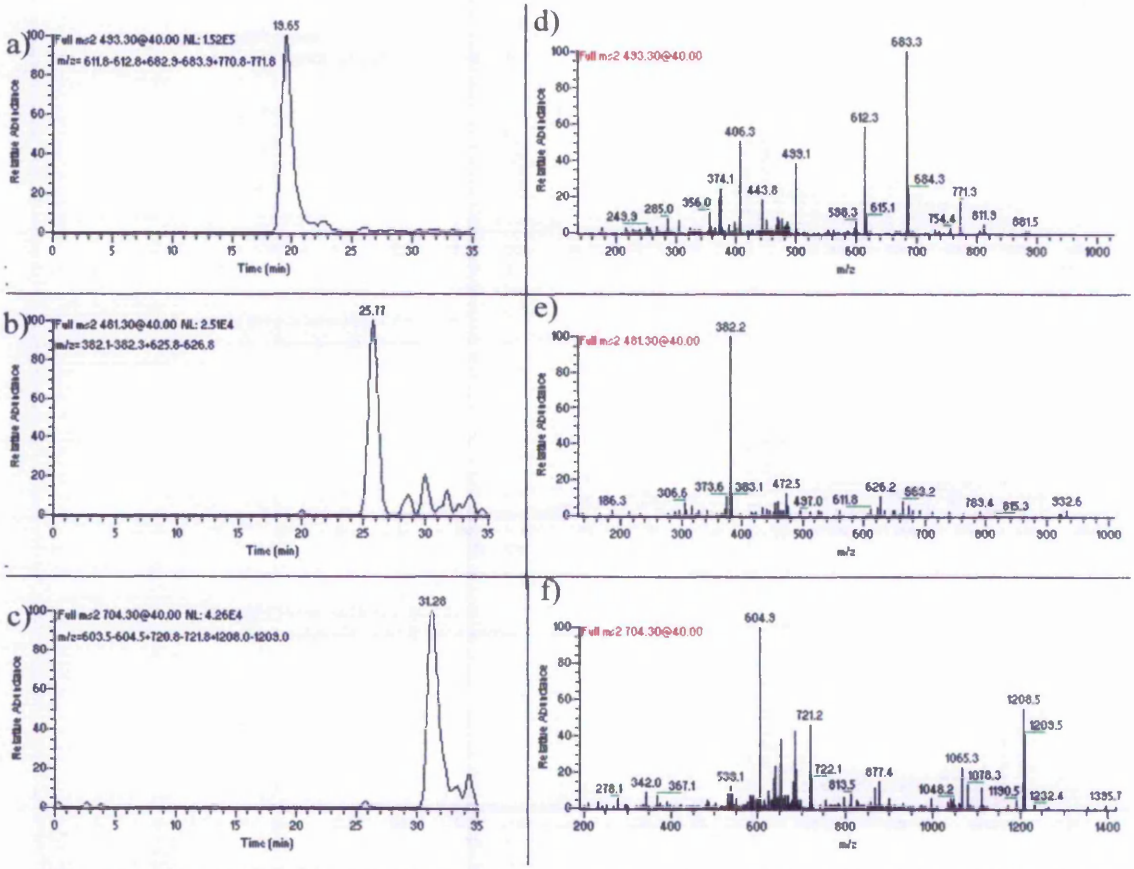


Figure 3.3.5.5 (a) Selected product ion chromatograms of the three peptides (a) A2 (b) β-gal (c) hep B (d) Tandem mass spectrum of the A2 peptide (m/z 493.3) (e) Tandem mass spectrum of the β-gal peptide (m/z 481.3) (f) Tandem mass spectrum of the hep B peptide (m/z 704.3).

An injection volume of 100 µl containing 300 mM potassium chloride is tolerated by the LC loading column therefore the 100 µl fractions collected from the SCX column may be loaded in their entirety onto the LC/MS system. This can be anticipated, despite an absence of recovery data, as a strong signal is seen in each case although only 100 fmol of peptide was loaded. This does not preclude the possibility of losses however they must be minimal or the signal observed would be considerably reduced.

3.3.6 Evaluation of an integrated MHC class I clean-up method

The various stages of eluate clean-up described in the previous sections were combined into an integrated method.

The method included the following optimised steps:

1) Surface elution of CML patient cells using Storkus[18] mild acid elution with citrate phosphate buffer pH3.3 to minimise cell lysis.
2) Solid phase extraction of the surface eluate using C18 cartridges to remove protein and buffer salts and achieve a reduction in sample volume.
3) Separation of peptides from the SPE extract using strong cation exchange chromatography. Peptides are eluted with fractions of salt of increasing concentration.
4) Analysis using on-line capillary RP-HPLC/micro-electrospray tandem mass spectrometry for peptide identification.

This method was evaluated using a CML patient eluate. The eluate, obtained from leucophoresed CML patient cells by the Richard Clarke group at the Royal Liverpool Hospital, was spiked at a low level (~200 fmol each) with the synthetic peptides hep B (TPPAYRPPNAPIL), β-gal (TPHPARIGL), A2 (SSKALQRPV) and A3 (KQSSKALQR). The hep b and β-gal peptides were added as internal standard recovery peptides to this eluate and all subsequent samples. The hep B is a class II peptide of viral origin and the β-gal a class I peptide of murine origin, so neither would be native to any human CML patient samples. The A2 and A3 peptides represent bcr/abl derived peptides predicted to be bound to MHC class I molecules in the literature[27] and by the SYPEITHI algorithm,[9] which showed poor recoveries in

preliminary experiments. The eluate was then subjected to the SPE and SCX clean-up methodology and analysed by capillary LC/μ-ESI/MS/MS.

The hep B, A2 and β-gal peptides were all recovered from this spiked eluate through the SPE and pre column loading. The selected product ion chromatograms showing the recovery of the peptides through the clean-up procedure are presented in Figures 3.3.6.1 to 3.3.6.3. The A3 peptide was not detected using this protocol because the highly basic nature of this peptide results in very poor retention on ODS where typically it shows extensive breakthrough.[8] Since the ODS SPE clean-up and loading onto the ODS trap column precedes the analytical column, loss of the A3 peptide is expected to be very significant at both stages.

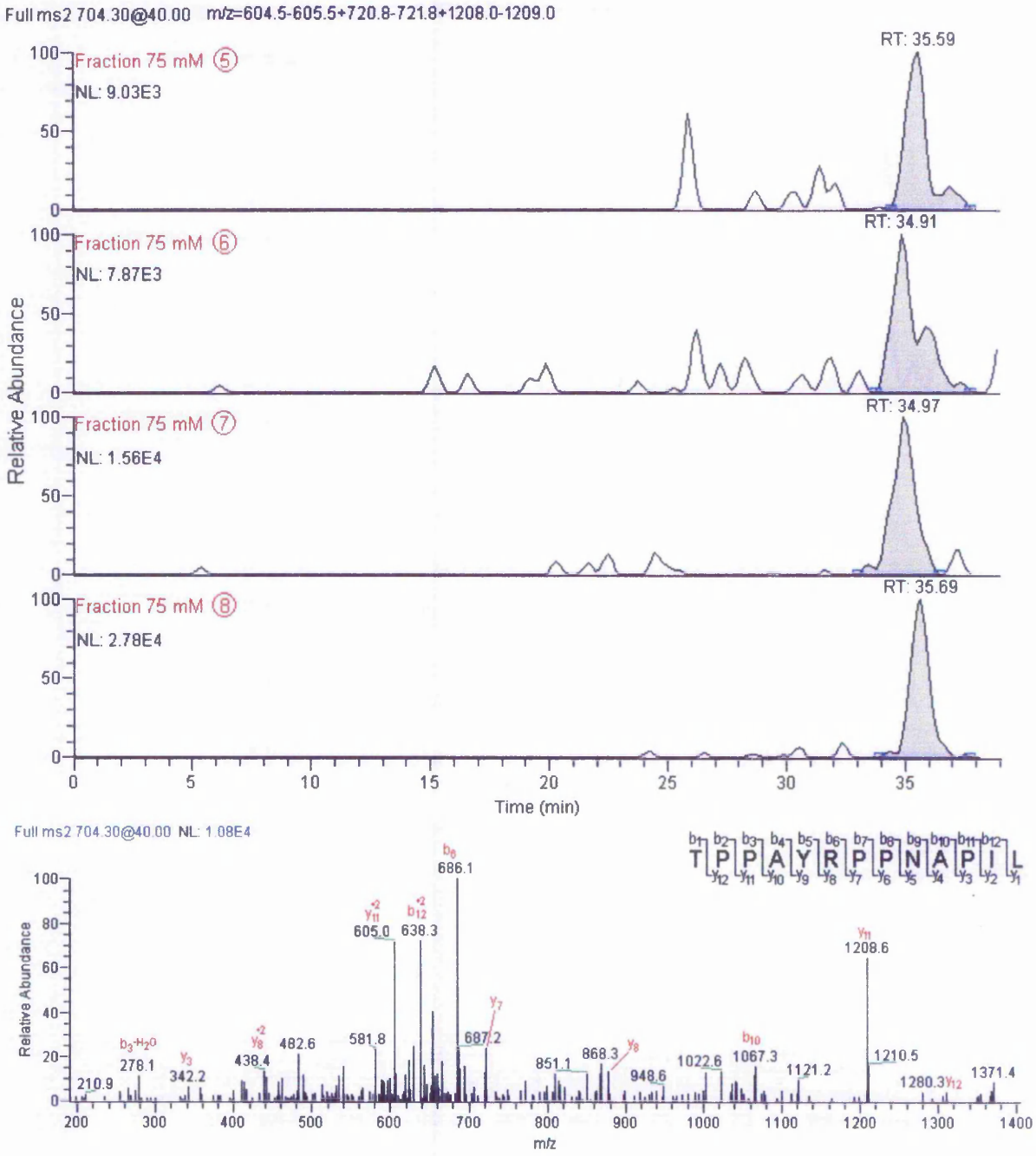


Figure 3.3.6.1 Selected product ion chromatograms showing the recovery of the hep B peptide (TPPAYRPPNAPIL) through the integrated MHC class I clean-up procedure. The peptide was found in the 75 mM salt fractions (5-8) as expected.

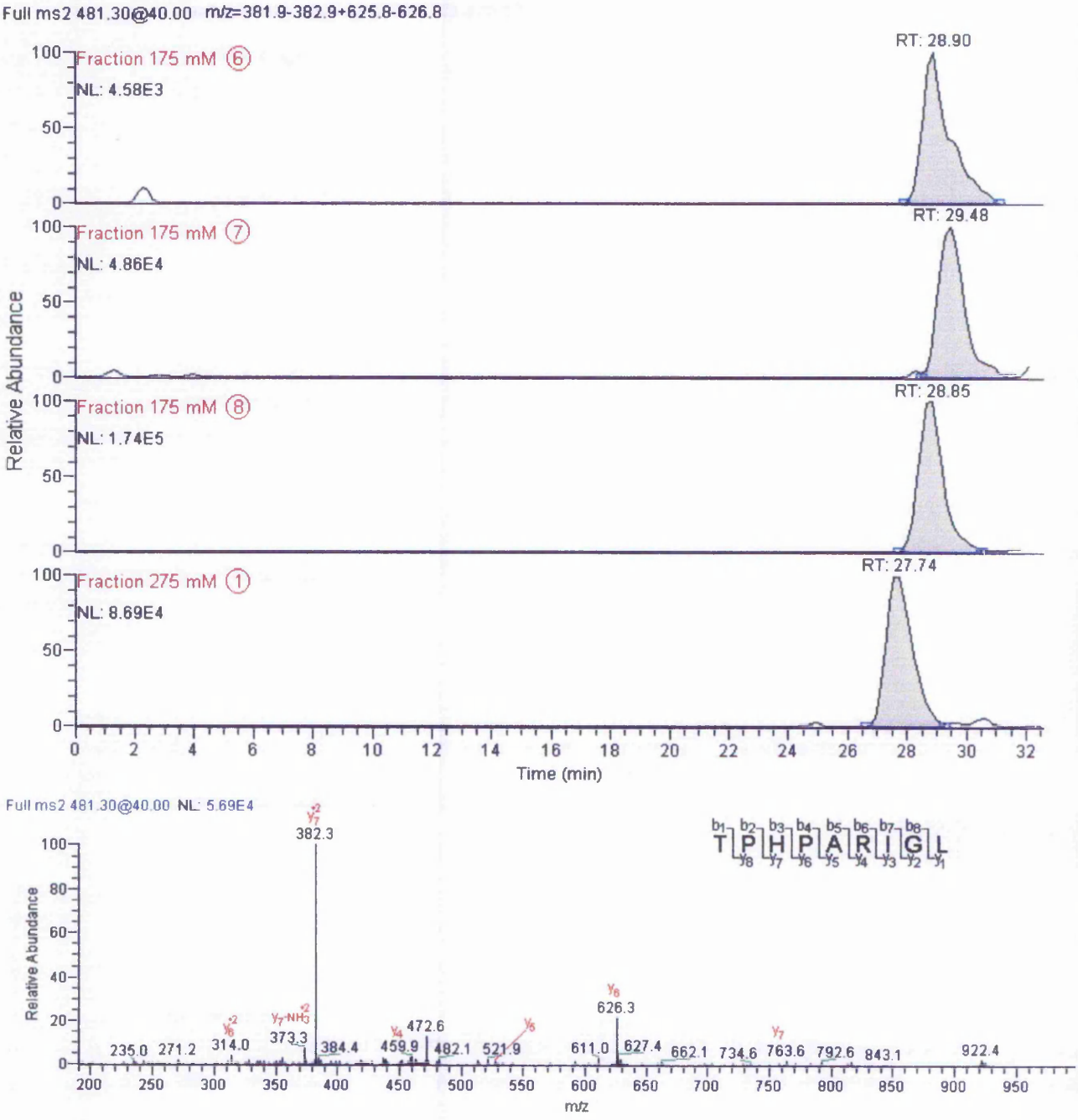


Figure 3.3.6.2 Selected product ion chromatograms showing the recovery of the β-gal peptide (TPHPARIGL) through the integrated MHC class I clean-up procedure. The peptide was found in the 175 mM salt fractions (6-8) and in the 275 mM fraction 1.

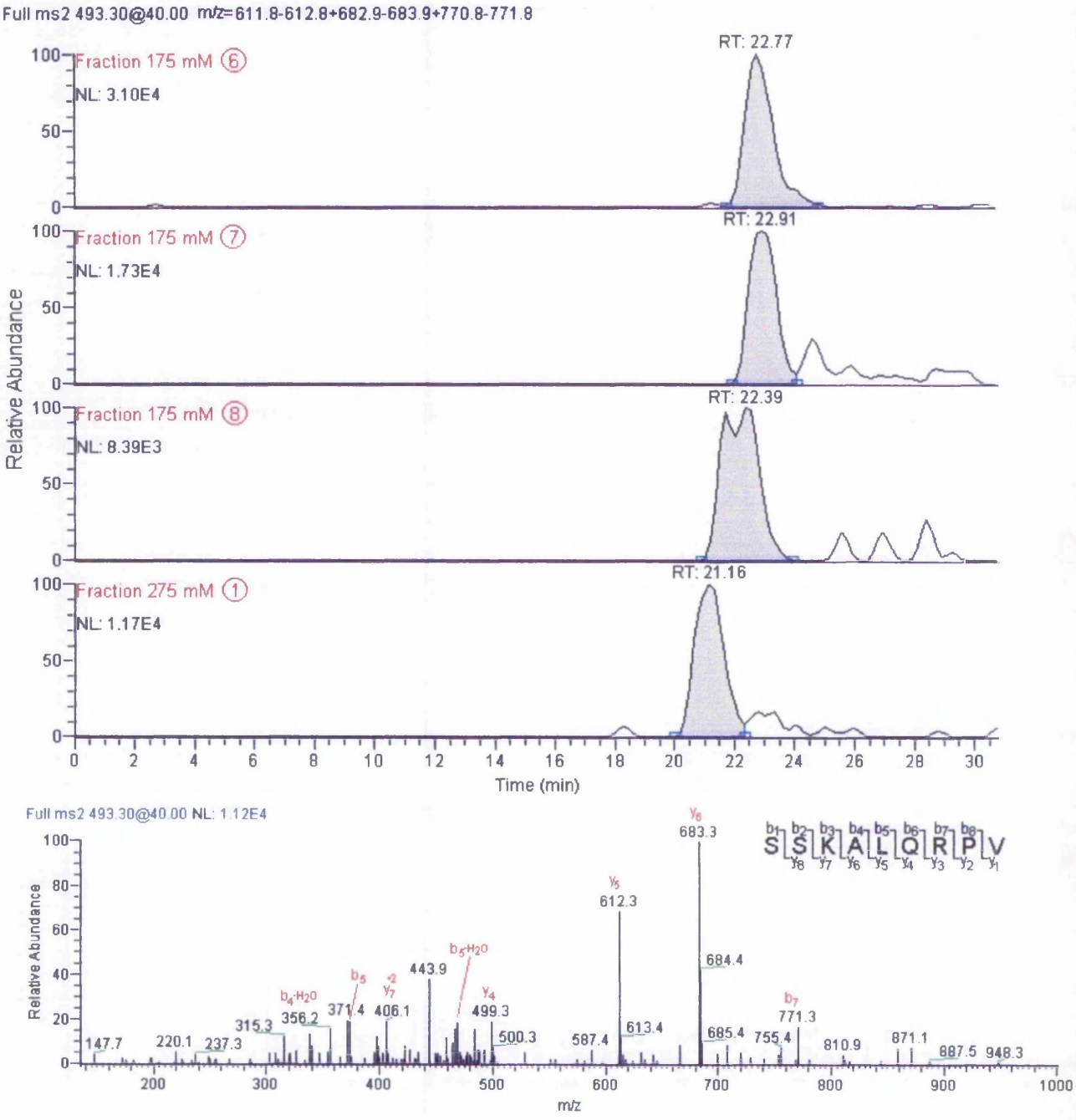


Figure 3.3.6.3 Selected product ion chromatograms showing the recovery of the A2 peptide (SSKALQRPV) through the integrated MHC class I clean-up procedure. The peptide was found in the 175 mM salt fractions (6-8) and in the 275 mM fraction (1).

Following this successful test of the integrated procedure on the CML patient 1 eluate a repeat test was performed in duplicate (patients 2 and 3) with the eluates spiked at a lower level of only 100 fmol of each peptide. Although it is difficult to predict the level at which MHC class I peptides will be found *in vivo*, detection at a level of 100 fmol or better is the aim. All other conditions were as described before. In each case the recovery was determined by comparing the peak area of each fraction to that of a standard equivalent to 100 % recovery. The recoveries determined for the three patient eluates can be seen in Table 3.3.6.1.

Table 3.3.6.1 Recovery of internal standard peptides (β-gal and hep B) and the bcr/abl derived A2 peptide from patient cell eluates.

	% Recovery		
Sample	**Hep B**	**β-gal**	**A2**
Patient 1	73	52	11
Patient 2	72	78	10
Patient 3	40	81	16
Mean	62	70	12

The calculated recoveries show some patient to patient variability but are > 60 % for the hep B and β-gal peptides. The variation in recovery between patient samples was not unexpected. Visual comparison of the individual eluates showed a lack of consistency in terms of colour and levels of precipitated proteins which could explain the disparity in recoveries for the model peptides. However, the magnitudes of peptide recoveries are broadly comparable. They are a significant improvement on the ~13 % recovery achieved for hep B using the MHC class I A3 method described in section

3.3.2. The bcr/abl A2 fusion peptide ($SSKALQRPV_{35-43}$) was included as a test of MHC class I peptide recovery. The recoveries for this peptide are lower than those for the hep B and β-gal peptides, probably because of the number of basic residues in this peptide. Detection of the A2 peptide by tandem mass spectrometry was still achieved with a recovery of ~12 % in the presence of eluate constituents. This corresponds to a conservative detection limit for the A2 peptide 25-50 copies per cell assuming 2.10^9 cells initially (assume ~ 5-10 fmol limit of detection and 12 % recovery).

3.3.7 Analysis of k562 eluate and CML patient eluates for predicted peptides

The aim of this work was to establish robust methods for the identification of antigenic peptide sequences in patients and transfected cell lines by targeting specific bcr/abl protein sequences within and without the junctional region predicted to bind by the SYFPEITHI algorithm[9] and in the literature[27-31] or to exclude sequences as vaccine candidates because they are not present on the cell surface at immunologically significant levels. The HLA.A2 peptides were preferentially targeted because they were predicted by the SYFPEITHI algorithm to show strong binding to the MHC and are derived from the bcr/abl junction region, these have been designated as group 1. Other peptides were selected because of association with the bcr or abl proteins but are not derived from the junctional region, these have been designated as group 2.

The validated method was applied to the analysis of k562 cultured cells and CML patient samples for predicted target peptides derived from the bcr/abl protein. The bcr/abl derived peptides were targeted specifically by tandem MS and in some cases a

survey scan was also performed using the ion trap software data dependent feature to select the most abundant peak(s) for tandem mass spectrometry. The number of peptide masses that could be targeted in a single run was limited to six (including the standard peptide) because of the cycle time of the experiment.

Table 3.3.7.1 Peptide sequences and masses targeted during MS/MS analysis

Group 1		Group 2	
Peptide sequence	**$[M+2H]^{2+}$**	**Peptide sequence**	**$[M+2H]^{2+}$**
KALQRPVASD	542.8	MLTNSCVKL	504.8
GFKQSSKAL	483.3	FLNVIVHSA	500.3
ATGFKQSSKA	512.8	FLSSINEEI	526.3
SSKALQRPV	493.3	FMVELVEGA	497.8
		QLLKDSFMV	540.8

In each analysis, 100 fmol each of the internal recovery standard peptides, hep B and β-gal, was added to the eluate prior to sample clean-up. All of the samples were positive for the bcr/abl protein with the b3a2 breakpoint. The sample information and peptide targets can be found in Table 3.3.7.2. These analyses represent actual patient samples and therefore there was no scope for repetition.

Table 3.3.7.2 Summary of bcr/abl peptide targets and peptides identified from k562 cell lines and CML patient samples.

Sample	Group 1 target	Found?	Group 2 target	Found?	Internal Standards recovered
K562 A2 (1)	Yes	No	No	-	Yes
K562 B8	Yes	No	No	-	Yes
K562 A2 (2)	Yes	No	Yes	No	Yes
Patient 180.01 (HLA.A2 positive)	Yes	No	No	-	Yes
Patient 98.01 (HLA.A2 positive)	Yes	No	Yes	No	Yes

As can be seen in Table 3.3.7.2, none of the targeted peptides were identified in either the CML patient samples or k562 transfected cell line eluates. One of the primary aims of this work was to determine whether the bcr/abl 'A2' peptide (SSKALQRPV) was actually expressed on the surface of HLA-A2 bcr/abl positive cells. In order to make this determination a method was developed for the analysis of both MHC class I and class II restricted peptides eluted from the surface of cells. The A2 peptide was searched for in both CML patient eluates and k562 cells and was not found despite being predicted to be present.[27] The results reported therefore indicate that this peptide is not presented *in vivo* above the limit of the detection for the method.

A target cell displays on its surface up to several hundred thousand MHC class I molecules complexed with an array of peptides representing diverse intracellular proteins, including products encoded by pathogens such as viruses and products that

arise from aberrant protein expression in tumour cells. The expression 'copies per cell' is used to describe the number of a particular MHC-peptide complex which are expressed on the surface of each individual cell. There have been various levels of copies per cell for immunogenic peptides expressed on the surface of cells reported in the literature. These range from as low as <1 up to 12000, see Table 1.6.1.1. Although this is not an exhaustive list, it is illustrative of the difficulty in assigning a level of copies per cell which are immunologically significant. The most common level is *ca.* 100 copies per cell and it has been suggested that 100-400 pep-MHC complexes per cell are needed for T-cell activation.[32-35]

Reported studies suggest that the copy number can vary over a range of many thousand fold. It is possible that few peptide complexes are required when the affinity of the TCR that recognises them is high and that a large number of complexes are needed when they and the corresponding TCR interact with low affinity.[36] Hunt's group[37] have isolated MHC class I peptides by immunopurification of EBV transformed B-LCL cell line. Four epitopes were identified in the context of 3 HLA B27 subtypes. Three of the four peptides were found to induce a CTL response in only one of the subtypes, although all four were found to be present and frequently at greater abundance in that subtype than the immunogenic peptide. In addition the immunodominant sequence, RRIYDLIEL, was found to be the least abundant, being present at < 1 copy number per cell. It has been concluded that the affinity between peptide and MHC molecule is not directly related to cell surface density. [37-39] Peptides which have been identified by motif scanning of proteins to have high affinity, however, may not be presented endogenously.[37-40] Similarly low affinity peptides have been found to be important antigens.[41-43]

The method reported here has a limit of detection for the so called bcr/abl A2 peptide (SSKALQRPV), which has shown to have a lower recovery through the clean-up than others tested, of approximately 25-50 copies per cell, assuming 2 x 10^9 cells were eluted, a limit of detection of 5-10 fmol and a recovery of 12 % through the procedure. This is a level at which all but the very rare peptide epitopes would expect to exceed. Recoveries of other peptides would be expected to be significantly higher (cf >60 % for the internal recovery standards) corresponding to < 5 copies per cell. These copy numbers correspond to the lower end of the range for possible MHC class I peptide sequences reported in the literature, suggesting that the targeted sequences are not present at immunologically significant levels.

Yotnda et al tested a range of bcr/abl peptides for their ability to bind to HLA-A2 T2 cells. This is a cell line which presents on its surface empty, and therefore unstable, MHC class I molecules. Binding ability was assessed using a thermostabilisation assay and the SSKALQRPV peptide was found to have significant binding avidity. Furthermore they found that it stimulates in-vitro specific CTL responses using PBL from healthy donors and CML patients.

However, several other groups[28,30,44-46] have investigated the bcr/abl protein for potential HLA-A2 binding peptide sequences, but none of the groups were able to find any strongly binding epitopes. In addition, a reverse immunology investigation of the bcr/abl protein found that the putative binding protein KQSSKALQR cannot be generated *in vitro* via cleavage by either the constitutive or the immunoproteosome.[47] Attempts by the Rees group to generate strong HLA-A2 restricted CTL in HLA-A2

transgenic mice for this peptide and others (ATGFKQSSKA, KALQRPVASD) were also unsuccessful.[48]

Clearly these findings are contradictory with the evidence presented here supporting the view that the A2 peptide, KQSSKALQR, is not presented on the surface of CML patient and k562 cells expressing the bcr/abl protein.

Other sequences from the bcr/abl breakpoint region were also investigated and were similarly not found in the samples analysed. There appears to be a general consensus emerging from the groups mentioned above suggesting there is no 'A2' peptide associated with the bcr/abl breakpoint region. The breakpoint region is therefore unlikely to contain a promising peptide vaccine candidate for HLA-A2 positive patients.

A further conclusion of this work is the necessity for complementary techniques to be brought together for the identification of antigenic peptides in the selection of candidate peptides for clinical trials. Computer modelling, in-vitro testing and mass spectrometry represent three interdependent techniques required for confident identification of a surface expressed antigenic peptide, see Figure 3.3.7.1.

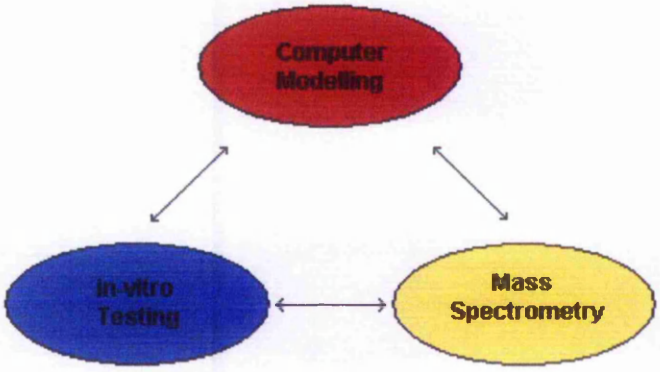


Figure 3.3.7.1 The triangle of techniques required for confident identification of MHC peptides.

Computer modelling is useful in predicting potential epitopes from known antigenic proteins of interest and so narrowing down the searching field for mass spectrometric identification. However, a predicition of binding ability does not necessitate expression. For example, Hunt *et al*[48,49] found during the isolation of MHC class I peptides from melanoma cells by immunopurification and RP-HPLC fractionation that the peptide presented is different to the one deduced from the gene. The epitope YMNGTMSQV was predicted, but instead YMDGTMSQV was expressed. They found that the modified peptide is recognised more effectively by CTL than the predicted sequence. Both peptides have a similar binding affinity and would be predicted to bind to MHC class I molecules. The conclusion was that spontaneous deamidation was unlikely due to the abundance of the modified peptide and that the epitope arises via enzymatically posttranslational modification resulting in the conversion of asparagines to aspartic acid. Similarly the presence of a peptide in an MHC class I molecule does not assure its antigenicity and so in-vitro testing of candidate peptides is important. This can also aid in identifying potential mass spectrometric targets as antigenicity is also no guarantee of surface expression.

K562 B8 eluate

The analysis of the K562 B8 transfected eluate did not result in the identification of any of the target bcr/abl peptides, but the data dependent survey analysis revealed the presence of peptides associated with the cancer testis antigen protein family GAGE. An 18 mer (EVEPATPEEGEPATQRQD) and a 16 mer (PPNPEEVKTPE) peptide were identified by the SEQUEST algorithm as originating from the Human GAGE-1 protein. SEQUEST works by establishing the similarity of a reconstructed mass spectrum to one which has been determined experimentally. The magnitude of the xcorr value indicates the quality of the match between the sequence and the mass spectrum. Generally the cross correlation score cut off is around 2.0.[50,51] The assigned full scan and tandem spectra are shown in Figures 3.3.7.2 and 3.3.7.3. Both peptides yielded strong product ion spectra, shown in Figures 3.3.7.2(b) and 3.3.7.3 (b) with good b and y ion sequence coverage. The SEQUEST algorithm identified these sequences with xCorr values of 4.06 (18 mer) and 3.84 (16 mer) and Scores of 669 (18 mer) and 889 (16 mer).

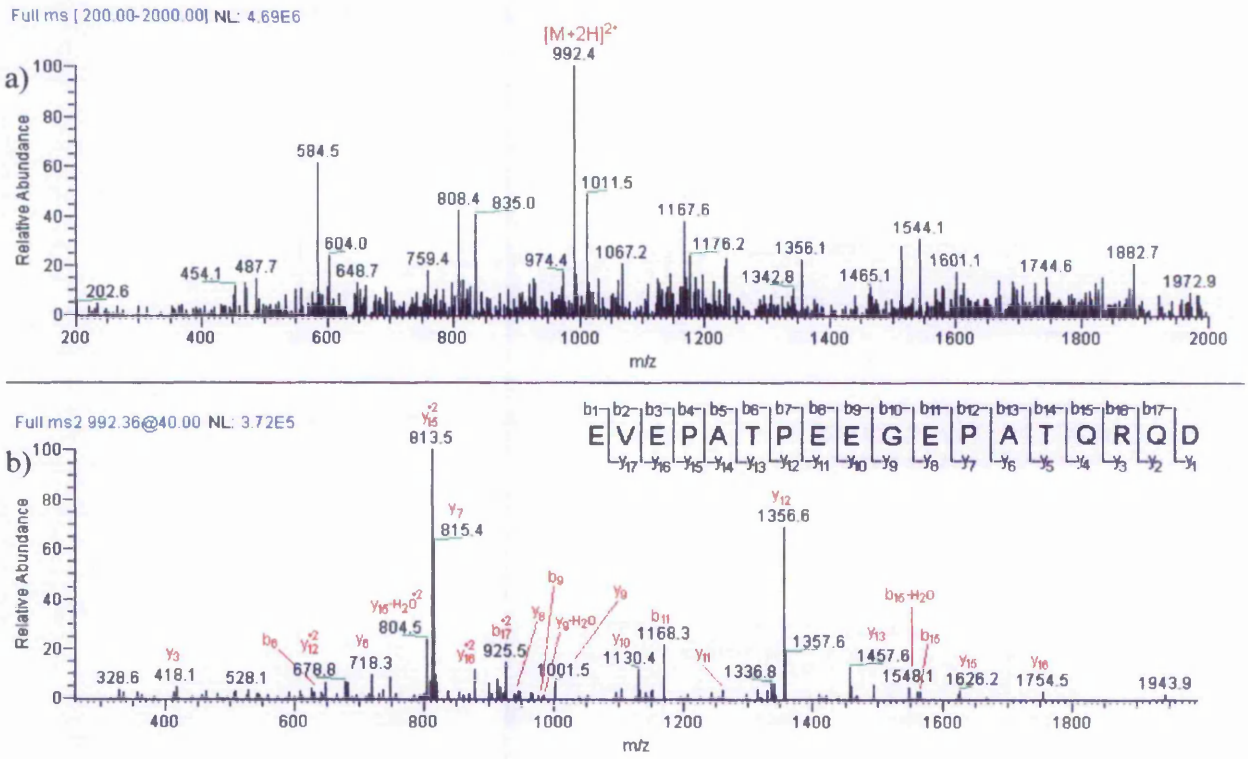


Figure 3.3.7.2 (a) Full scan mass spectrum of the cancer testis antigen GAGE family 18 mer peptide EVEPATPEEGEPATQRQD $[M+2H]^{2+}$ ion (m/z 992.4). (b) Assigned product ion spectrum from the tandem MS analysis of the doubly charged parent ion (m/z 992.4).

The identification of two peptides from the same protein increases confidence that the assignment is correct, as does the excellent match of the individual spectra to the assigned sequence, although, the size of these peptides precludes their being bound to the MHC class I molecule. It is not possible to say with certainty whether these peptides migrated from inside the cells to the cell surface. However, they are unlikely to be surface bound peptides associated with the class II molecule because the class I elution method is not expected to be effective in eluting these class II peptides. It was shown in section 3.3.1 that a small amount (~22 %) of cell lysis occurs during the surface elution and the possibility exists that these peptides originate from within the cells. The GAGE family of genes which express the proteins belong to the same

family as the MAGE genes, present only in male germ line cells, they are silent in most normal adult tissues. They have been discussed in the literature and are associated with a variety of cancers including melanoma, leukaemia, lymphoma and various carcinomas.[52-55]

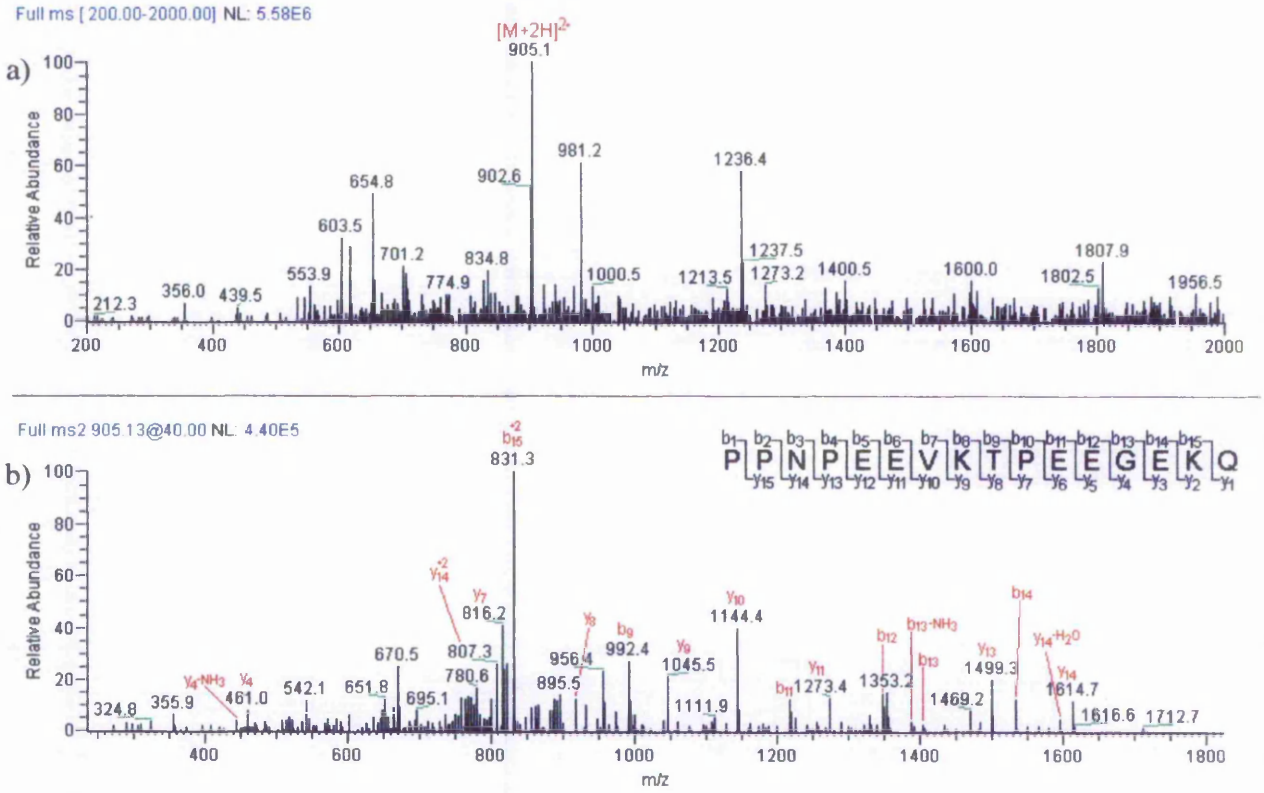


Figure 3.3.7.3 (a) Full scan mass spectrum of the cancer testis antigen GAGE family 16 mer peptide PPNPEEVKTPEEGEKQ $[M+2H]^{2+}$ ion (m/z 905.1). (b) Assigned product ion spectrum from the tandem MS analysis of the doubly charged parent ion (m/z 905.1).

In each eluate analysis where a data dependent MS/MS scan was acquired using the survey method, whether from k562 cell lines or patient samples, a number of native housekeeping peptides were identified after searching the raw data with the SEQUEST algorithm against a database of human proteins. For example, the peptides and proteins found in the patient sample 180.01 and identified in this manner are shown in Table 3.3.7.3

Table 3.3.7.3 Peptide sequences and the proteins from which they originated found in the patient sample 180.01 and identified by the SEQUEST algorithm, with correlation factor (xCorr) and score results.

Sequence	xCorr	Score	Protein
IMRSGASGPENFQVG	3.74	1027	Talin
IMRSGASGPENF	3.47	1417	Talin
GEYKFQNAL	3.07	844	Albumin precursor
GEYKFQNALL	3.06	780	Albumin precursor
TGYGMPRQIL	2.84	755	Transgelin 2
LVVYPWTQRF	2.83	848	Haemoglobin gamma G
GTNRGASQAGMTG	2.47	458	Transgelin 2
GTNRGASQAGMTGYG	4.16	1480	Transgelin 2
LVVPWTQRFFES	2.65	670	Haemoglobin beta
AQPPSSLVIDKESEVYKML	3.22	630	PDZ & LIM domain protein
KYPSPFFVF	3.08	850	Leukophysin
KYPENFFLL	2.95	1284	Protein phosphatase 1
EGGPEAPIPPPPQPREKVS	2.56	355	Zyxin
AAAAIEAAAKKLE	3.94	1904	Talin
NPLPSKETIEQEKQAGES	3.47	576	Prothymosin beta
TQPRGPPASSPAPAPKF	3.38	684	Zyxin
VNPFRPGDSEPPPAPGAQRAQ	3.21	1080	Zyxin
TVPVEAVTSKTSNIR	2.99	254	Cortactin oncogene
WQVKSGTIFDNF	2.78	1302	Calreticulin precursor
IYQGRLWAF	2.71	823	Cortactin oncogene
SPVTPKFTPVASK	2.74	599	Zyxin
ECLHPTKFLNNGTCTAEAEGKFS	2.35	769	Human complement factor 1
EEGAGAE	2.05	544	Zondahesin
GPEAAKSDETAAK	2.50	641	Heat shock protein
VPVEAVTSKTSNIR	2.09	447	Cortactin oncogene
APPKPPLPEGEVPPPRPPPPEEK	4.28	844	Vinculin isoform VCL
APVISAEKAYHEL	4.24	1150	Tubulin alpha 3
LSALEEYTKKLNTQ	3.66	890	Apolipoprotein A-I precursor

SGPKPFSAPKPQ	3.15	1373	Adenyl cyclase associated protein
LASVSTVLTSKYR	3.08	1160	Haemoglobin alpha 2
SLDRNLPSDSQDLGQHGLEED	3.07	993	Proteoglycan 1 secretory granule
SGPKPFSAPKPQT	2.95	800	Adenyl cyclase associated protein
HVQPQPQPKPQVQ	2.94	605	Zyxin
TQEKNPLPSKE	2.60	817	Prothymosin beta 4
NLAKGKEESLDS	2.52	689	Pro platelet basic protein
AEPAVQRTLLEK	3.28	565	Antigen identified by monoclonal antibodies
ESEEKGDPNKPSGF	2.05	301	PDZ & LIM domain protein
KFIDTTSKF	3.75	1027	Ribosomal protein L3
PEPAKSPAPKKGS	2.81	364	H2B Histone family
VVSLGSPSGEVSHPRKT	3.04	723	Alpha 2 HS glycoprotein
SAKPTKPAASDLPVPAEGVRNIK	3.23	520	Caldesmon 1 Isoform

These peptides represent a range of housekeeping proteins, cytoskeletal proteins such as tubulin, and actin; DNA-binding histone proteins; Actin binding proteins such as talin and transgelin. The score is a preliminary scoring system, which is based on the number of ions in the MS/MS spectrum which match the experimental data. A value of >200 indicates a good match although a larger sequence will score higher than a shorter sequence. The Xcorr factor is the cross correlation value from the search and is indicative of the best match in the search. An Xcorr value in the region of 2.00 would suggest a good correlation.

3.3.8 MHC class II peptide analysis.

The class II invariant chain peptide, CLIP, is involved in the assembly of the MHC class II molecule in the endoplasmic reticulum. The mutant cell line T2 DR4 cannot

present peptides derived from intact protein antigen, therefore are transported to the cell surface with the CLIP peptide still in the binding groove. The class II peptide chain length is variable which leads to increased difficulty in class II prediction. The T2 mutant cell line is therefore ideally suited for use in the investigation of the transferability of the class I clean-up method to class II peptide identification as only a single family of peptides is expected to be present in the binding groove. Due to the time consuming nature of cell growth to the level required for these experiments each was limited to a single performance.

3.3.8.1 Identification of CLIP peptide from immunopurified class II molecules

T2DR4 cells which are DM deficient were incubated and the class II molecules isolated by immunopurification with an L243 antibody column using an ImmunoPure® kit and the peptides eluted from the molecule using an IgG elution buffer. The peptides were then analysed by capillary LC/MS/MS for the presence of the class II invariant chain peptide (CLIP). The mass spectrometric method used targeted the full CLIP sequence LPKPPKPVSKMRMATPLLMQALPM. The mass spectrometry procedure was also set up to perform a data dependent tandem mass spectrometry scan on the strongest ion in the preceding full scan.

The targeted CLIP sequence was not found in the T2DR4 cell extract and the data dependent scan data was therefore analysed by searching with the TurboSEQUEST algorithm. The search identified a peptide sequence originating from the class II invariant chain, with a doubly charged parent ion at m/z 1167.9, eluting at 38 minutes which yielded a tandem mass spectrum with a SEQUEST score (Sp) of 188 and a correlation factor (xCorr) of 2.60. The cross correlation factor indicates a good match,

although the score is less than 200. However taken together with the correlation factor, the SEQUEST search indicated a match worth further investigation.

The sequence identified (LPKPPKPVSKMRMATPLLMQA) was from the CLIP peptide with the last three C-terminal residues missing. There is no restriction on the length of peptide which protrudes from either side of the MHC class II peptide groove so a shorted sequence can readily bind to the MHC with high affinity. Closer investigation of the data revealed that the triply charged ion had also been selected for tandem mass spectrometry in the data dependent scan. The combined chromatogram for the doubly and triply charged ions is shown in Figure 3.3.8.1(a) The resulting full scan mass spectrum which shows both $[M+2H]^{2+}$ and $[M+3H]^{3+}$ ions is shown in Figure 3.3.8.1 (b).

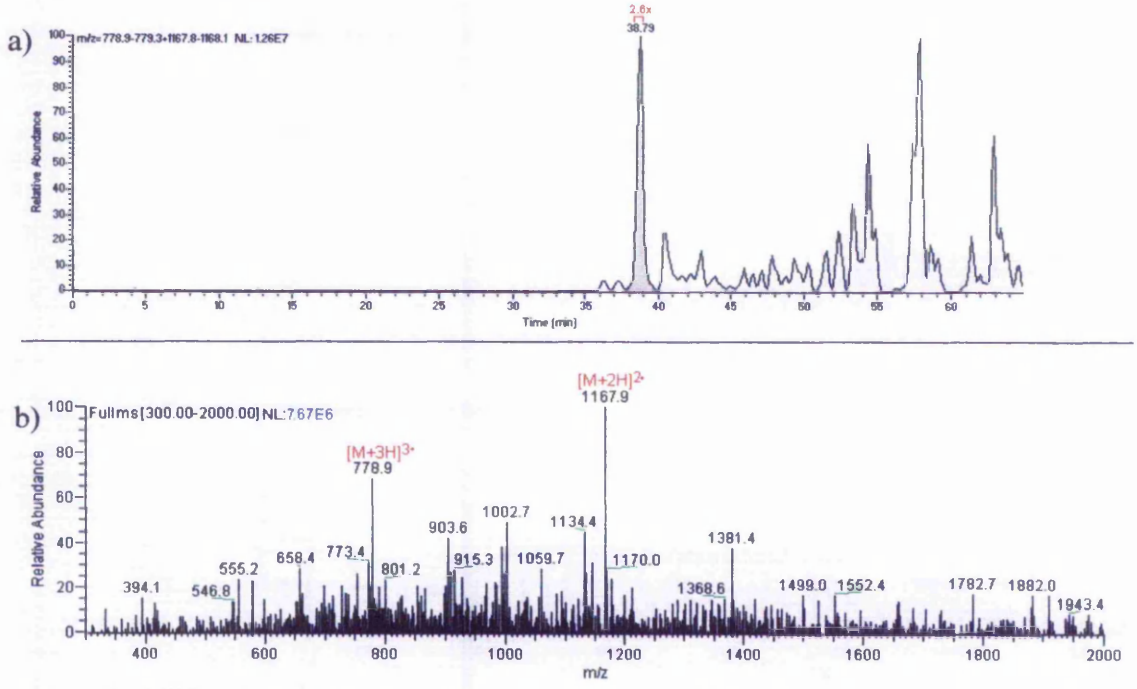


Figure 3.3.8.1 (a) Selected ion chromatogram (m/z 1168 +779) showing the combined peak for the doubly and triply charged ions eluting at 38.8 minutes (x 2.6 magnification). (b) Full scan mass spectrum showing the $[M+3H]^{3+}$ (m/z 778.9) and the $[M+2H]^{2+}$ (m/z 1167.9) ions.

The peak in the chromatogram is clear and discrete, although the scale has been expanded to allow better visualisation. The strong peaks later in the chromatogram are unrelated to the peptide ions. The sequence identified using the TurboSEQUEST algorithm, was subjected to manual confirmation of the fragment ion assignment by comparison of the spectra with fragments predicted using the algorithm MS-Tag on the UCSF website.[56] The manually assigned tandem mass spectra are shown in Figure 3.3.8.2.

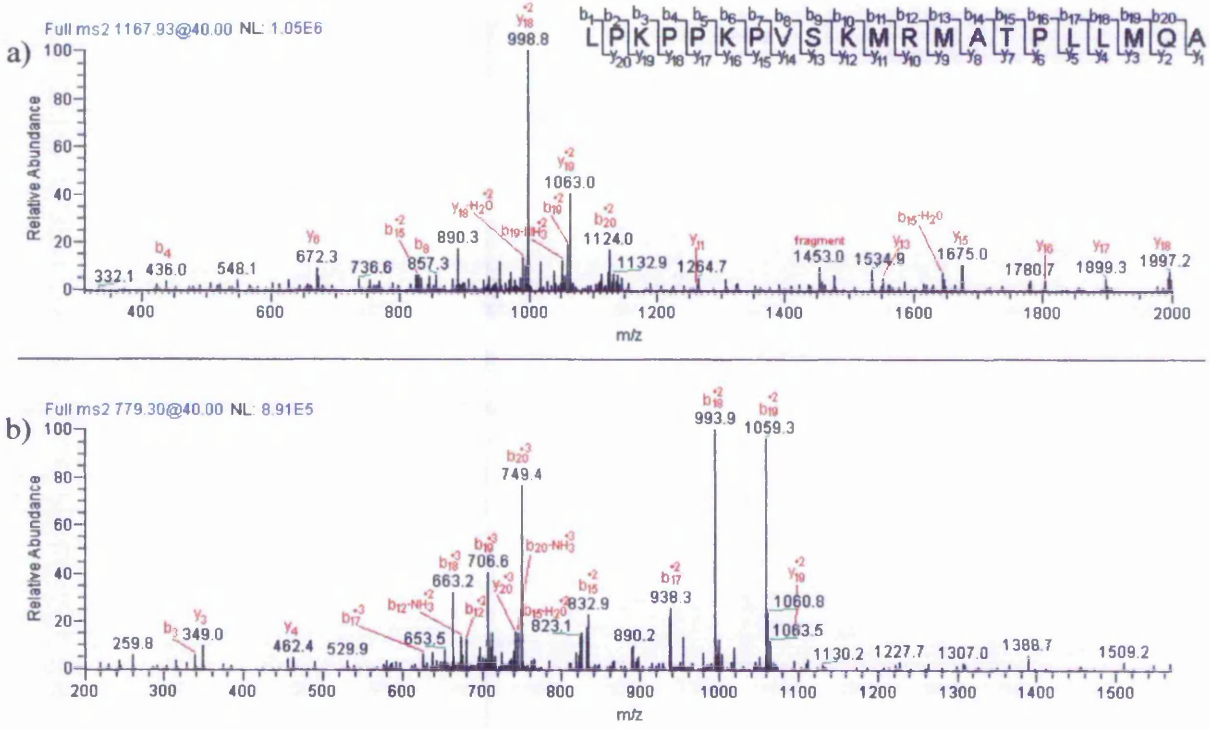


Figure 3.3.8.2 Assigned tandem mass spectra for the CLIP peptide LPKPPKPVSKMRMATPLLMQA. The ladder sequence of b and y ions is shown. (a) tandem mass spectrum of $[M+2H]^{2+}$ ion (m/z 1167.9) (b) tandem mass spectrum of $[M+3H]^{3+}$ ion (m/z 779.3).

It is clear from the tandem mass spectra in Figure 3.3.8.2, that the assignment of b and y ions shows good sequence coverage, particularly when the doubly and triply charged ion fragments are taken together. It should be noted that the scan range during

MS/MS of the triply charged ion stops at m/z 1500, this is because the acquisition software has incorrectly assumed the parent ion to be doubly charged. However, this does not appear to impact on identification as the majority of the strong fragment peaks are in the region below m/z 1000. From this analysis we can conclude that the CLIP peptide is expressed by T2DR4 MHC class II molecules and following immunopurification, can be identified, even without specific targeting of the correct sequence length, by the LC/MS methods developed in this work

3.3.8.2 Identification of CLIP from Class II surface eluted cells.

The aim of these experiments was to determine whether the method developed for the clean-up and identification of MHC class I peptides, and shown to be applicable to immunopurified class II peptide from cell lysates, can also be applied to surface eluted MHC class II molecules. A novel class II surface elution protocol has been developed in parallel with the class I protocol, which allows the surface presented MHC class II bound peptides to be eluted in a similar way to that of the MHC class I peptides with minimal cell lysis.[18] The T2DR4 cells were cultured and eluted using the class II buffer. The eluate was then spiked with the internal recovery peptides hep b and β-gal and subjected to the solid phase extraction and strong cation exchange protocols prior to capillary LC/μESI/MS/MS analysis of the resulting fractions. Taking into consideration the results from the analysis of the immunopurified DR4 cells, two additional scan events were added to the method to specifically target the doubly and triply charged ions of the CLIP peptide that was identified. The presence of the same CLIP peptide which had been discovered in the immunopurified cells, $CLIP_{81\text{-}101}$, LPKPPKPVSKMRMATPLLMQA, was confirmed from the tandem mass spectra in the class II buffer eluate without immunopurification. The chromatograms

and mass spectrum, which shows the presence of both the doubly and triply charged peptide ions are shown in Figure 3.3.8.3.

Due to the targeting nature of the data acquisition method, the chromatographic peaks which are shown in Figure 3.3.8.3 (a) and (b) are strong, sharp and discrete. The mass spectrum of the two ions was more complex and has been presented in the range m/z 300-1200 in Figure 3.3.8.3 to allow a clearer view of the two parent ions. Both ions were targeted for tandem mass spectrometry and were assigned their b and y fragment ions by manual comparison of the fragments predicted by the UCSF MS-Tag function.[56] The assigned tandem mass spectra can be seen in Figure 3.3.8.4.

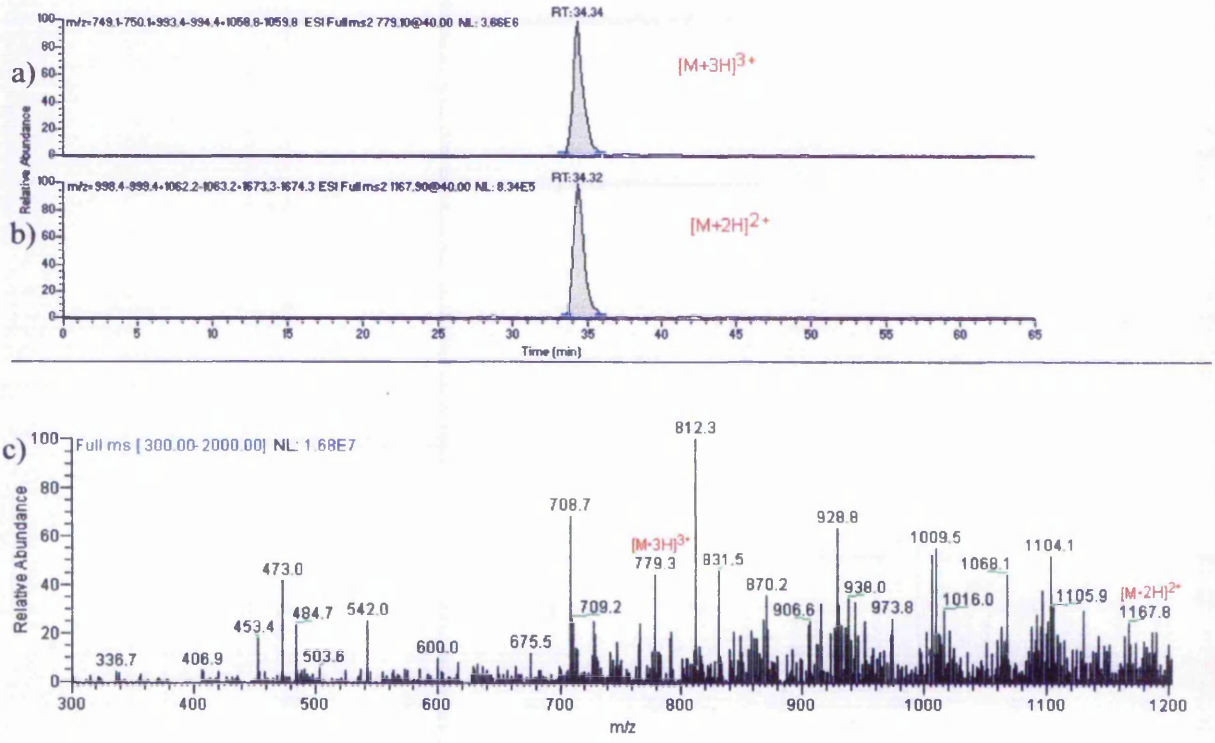


Figure 3.3.8.3 Chromatograms and mass spectrum of the CLIP peptide LPKPPKPVSKMRMATPLLMQA identified from surface eluted T2DR4 MHC class II complexes. (a) Product ion chromatogram (m/z 750 + 994 + 1059) of the $[M+3H]^{3+}$ CLIP ion. (b)

Product ion chromatogram (m/z 999 + 1063 + 1674) of the $[M+2H]^{2+}$ CLIP ion. (c) Full scan mass spectrum showing both parent ions $[M+3H]^{3+}$ (m/z 779.3) and $[M+2H]^{2+}$ (m/z 1167.8).

The b and y assignment for both parent ions is very comprehensive, particularly when taken together, and the assignments confirm the identification of these two peptide ions as originating from the CLIP peptide. The MS/MS spectra shown in Figure 3.3.8.4 also match well with the corresponding spectra shown in Figure 3.3.8.2. It can be seen in Figure 3.3.8.4 (b) that there are few prominent ions above m/z 1200, which confirms the earlier assumption that the loss of this portion of the spectrum during the immunopurification analysis is of little consequence

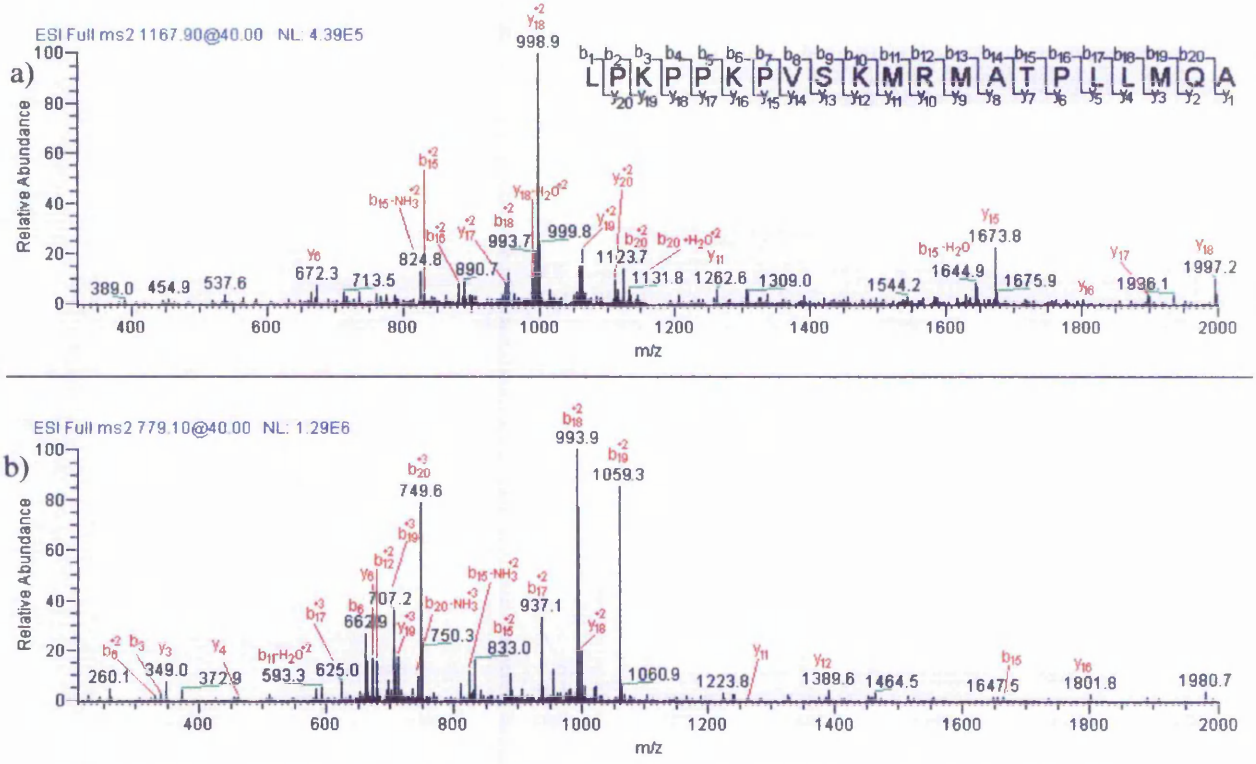


Figure 3.3.8.4 b and y fragment assigned tandem mass spectra of the CLIP peptide LPKPPKPVSKMRMATPLLMQA. The peptide ladder is also shown. (a) tandem mass spectrum of $[M+2H]^{2+}$ ion, (m/z 1167.8) and (b) tandem mass spectrum of $[M+3H]^{3+}$ ion, (m/z 779.3).

In conclusion, the method developed for the identification of MHC class I peptides has been successfully extended to the analysis of immunopurified and surface eluted MHC class II peptides. Using a mutated cell line which cannot display endocytotically digested peptides on the cell surface, it was possible to greatly reduce the number of possible bound peptides. These were a nested set originating from the class II invariant chain peptide (CLIP) which is associated with the class II molecule during its assembly in the endoplasmic reticulum. Using conventional immunopurification techniques we were able to confirm the presentation of a particular CLIP peptide by the T2DR4 cells and verify the suitability of the LC/MS analytical method for the detection of the peptide. The MHC class I clean-up and analysis methods were then performed on a MHC class II surface eluate and the CLIP peptide was identified as being presented on the surface of the T2DR4 cells.

3.4.0 Conclusions

The method which presented here allows MHC class I and class II peptides eluted from the surface of infected and transfected cells to be targeted and, if present, identified. The prevailing method for the identification of MHC class I peptide antigens was pioneered by Hunt *et al*[57] involving the antibody purification of MHC class I molecules and complexed peptide following cell lysis and acid elution of the peptides from the complex. The eluted peptides are fractionated by HPLC fractionation, often coupled with assays to establish immunological activity of peptides within fractions, prior to sequence identification by tandem mass spectrometry. Although the Hunt technique is well established and has been used to identify a number of MHC associated peptides it suffers from the underlying disadvantage that the peptides identified cannot be said, with confidence, to have originated from the cell surface. A comparison of the two methods is presented in Table 3.4.1

The method involves mild acid elution, to minimise cell lysis, followed by SPE clean-up, orthogonal SCX and capillary RP-LC fractionation and tandem mass spectrometry. The mild acid elution approach, ~80% or more of the cells remain intact and viable after treatment with citrate phosphate buffer. The possibility of the MHC complexes picking up peptides released from the cells interior is considerably reduced and confidence of a surface expressed identification is correspondingly higher.

Table 3.4.1. A comparison between the MHC cell surface method presented in this thesis and the Hunt immunoaffinity purification method

Parameter	Cell Surface Elution Method	Hunt Method [42,57]
Analytical Methodology	Utilises an orthogonal mode of separation, SCX and RP-HPLC to maximise separation of species.	Multiple stages of RP-HPLC, the same mode of separation as the cell surface elution method.
Efficiency of Analysis	SCX and single RP-HPLC processes should provide separation of species with minimal analyte losses.	Multiple stages of RP-HPLC will increase the risk of analyte losses, particularly highly polar peptides.
Specificity	This method has the potential to incorporate stages to provide information on immunogenicity.	Involves the incorporation of a T-cell assay to provide information on potential immunogenicity.
Peptide Yield	12 % (For the bcr/abl peptide SSKALQRPV). >60 % (For other model peptides).	12 % (Based on the melanoma peptide YLEPGPVTA).
Number of Cells Studied	2.0×10^9	4.0×10^9
Certainty of Analysis	Surface elution maximises the likelihood of the analysis of surface expressed peptides.	Immunoaffinity purification of MHC molecules from lysed cells provides no assurance of the analysis of surface expressed peptides.
Limit of Detection	5 to 10 fmol	~15 fmol

The method presented here has been shown to have a limit of detection of a 5-10 femtomoles which is comparable with, or slightly better than that achieved using the immunoaffinity purification method.[42,58]

The use of a mild acid wash for the surface elution of MHC peptides results in a more complex sample than is obtained by immunopurification. However, the use of multiple stages of separation simplifies the mass spectrometers task. This method

utilises orthogonal modes of separation, fractionation by strong cation exchange chromatography followed by RP-HPLC in contrast to the Hunt method which uses multiple stages of RP-HPLC. The multidimensional approach has been used by other groups for the analysis of complex mixture.[59-61] This approach is well suited to the analysis of the surface elution as, in addition to reducing the complexity of the sample, it also reduces the sample volume with each stage of separation, which is useful when the initial sample volume is 5-20 ml. A benefit of carrying out the Hunt method would lie in the incorporation of the T-cell assay that would maximise the information that could be obtained from an analysis. However, such an assay could also be incorporated into the cell surface elution method.

The method is more widely applicable to MHC class I peptides than the previous A3 clean-up method which has been developed in this group for the analysis of surface eluted peptides. The protein precipitation step using trichloroacetic acid was eliminated in favour of a solid phase extraction protocol, which served the dual purpose of removing the extraneous protein and concentrating and desalting the cell eluate.

Strong cation exchange chromatography was then employed for the fractionation of the peptides. Several columns were investigated under a variety of binding and elution conditions to ascertain the optimum system for peptide separation and maximum recovery. Analysis was then performed using the nano-LC/μ-ESI/MS/MS system, of which the development and optimisation was described in Chapter 2.

The integrated method was satisfactorily tested with recovery for the peptides hep B and β-gal of 60-70 %. The A2 predicted peptide suffered significantly lower recovery, ~12 %, however there is sufficient sensitivity to allow detection of an immunologically significant level of 25-50 copies of peptide per cell. This method was applied to the analysis of HLA.A2 CML patient eluates and k562 A2 transfected cell lines in search of the A2 predicted peptide SSKALQRPV and other targeted bcr/abl peptides. None of the predicted peptides were observed in these eluates at detectable levels. These data provide powerful evidence that MHC binding peptides presented *in silico* and by binding assays and proteosomal digestion studies may not be presented to the immune system via MHC class I molecules. It is therefore necessary to undertake the direct analysis of patient samples for predicted peptides, in addition to algorithm prediction and *in vitro* assays, when peptide vaccine candidates are selected for clinical trials.

Analysis of the HLA. B8 positive CML patient eluate revealed the presence of peptides associated with the cancer testis antigen family GAGE. The size of peptide excludes them as MHC class I bound and it is unlikely that they were class II bound peptides as the mild acid elution method has not been shown to be effective in releasing these peptides. The remaining possibility is that these peptides originate from within the cell and are released during the small amount of cell lysis which has been shown to occur. A number of native housekeeping proteins were also identified with high confidence from the CML patient eluates using the SEQUEST algorithm.

The integrated clean-up method was further tested for its application to the identification of MHC class II peptides. DM deficient T2DR4 cells were utilised for

their simplified class II repertoire as their mutation ensures they present only the class II invariant chain peptide (CLIP). A sequence from the CLIP peptide was identified using the integrated clean-up method from cells which had been immunopurified. This initial test confirmed the potential application of the clean-up method to the identification of class II peptides. The immunopurification method gives rise to a very clean sample although it does not allow confidence in surface expression. A novel class II surface elution buffer was applied to the same T2DR4 cells and the eluate subjected to the integrated clean-up method. The same CLIP sequence was identified from the surface eluate.

3.5.0 References

1. Clark R.E., Dodi I.A., Hill S.C., Lill, J.R., Aubert G., Macintyre A.R., Rojas J., Bourdon A., Bonner P.L.R., Christmas S.E., Travers P.J., Creaser C.S., Rees R.C., Madrigal J.A., *Blood,* **2001,** 98 2887-2893.
2. Barberiheyob M., Merlin J.L., Weber B., *J. Chromatogr. Biomed. Appl.,* **1992,** 573, 247-252.
3. Azuma R., Urakawa A., *J. Chromatogr. B;* **1997**, 691, 179-185.
4. Al-Rawithi S., Hussein R., Raines D.A., AlShowaier I., Kurdi W., *J. Pharma. Biomed. Anal.*, **2000,** 22, 281-286.
5. Abu-Qare A.W., Abou-Donia M.B., *J. Chromatogr. B,* **2001**, 754, 503-509.
6. Pardoll D.M., *Nat. Med. Vaccine Suppl.*, **1998,** 525-531.
7. Bonner P.L.R., Lill J.R., Hill S.C., Creaser C.S., Rees R.C., *J. Immunol. Methods,* **2002,** 262, 5-19.
8. Lill J.R., *PhD Thesis*, **2000,** Nottingham Trent University.
9. Rammensee H.G., Bachman J., Emmerich N.P.N., Bachor O.A., Stevanovic S., *Immunogenetics,* **1999,** 50, 213-219.
10. Machamer C.E., Cresswell P., *J. Immunol.*, **1982,** 129, 2564-2569.
11. Jones P.P., Murphy D.B., Hewgill D., McDevitt H.O., *Immunochemistry,* **1979,** 16, 51-60.
12. Roche P.A., Cresswell P., *Nature,* **1990,** 345, 615-618.
13. Newcomb J.R., Cresswell P., *J. Immunol.*, **1993,** 151, 4153-4163.
14. Riberdy J.M., Newcomb J.R., Surman M.J., Barbosa J.A., Cresswell P., *Nature,* **1992,** 360, 474-477.
15. Ceman S., Rudersdorf R., Long E.O., Demars R., *J. Immunol.*, **1992,** 149, 754-761.
16. Mellins E., Smith L., Arp B., Cotner T., Celis E., Pious D., *Nature,* **1990,** 343, 71-74.
17. Riberdy J.M., Avva R.R., Geuze H.J., Cresswell, P., *J. Cell Biol.*, **1994,** 125, 1225-1237.
18. Storkus W.J., Zeh H.J., Salter R.D., Lotze M.T., *J. Immunother.*, **1993**, 14, 94-103.
19. Personal Communication, Danielle Barry., Interdisciplinary Biomedical Research Centre, Nottingham Trent University, UK.
20. Expasy website (http://www.expasy.org).
21. Tests were performed by Dr. Amanda Miles, Interdisciplinary Biomedical Research Centre, Nottingham Trent University, UK.

22. Lozzio C.B., Lozzio B.B., *Blood,* **1975**, 45, 321-324.

23. Lozzio B. B., Lozzio C. B., Bamberger E. G., Feliu A. S., *Proc. Soc. Exp. Biol. Med.,* **1981,** 166, 546-550

24. Freshney R., *Culture of Animal Cells: A manual of Basic Techniques,* **1987,** Liss A. R., New York.

25. Ed Janson J-C., Ryden L, *Protein Purification Principles, High-Resolution Methods, and Applications, Second Edition,* **1998**, Wiley-VCH, New York.

26. Stoscheck C.M., *Methods in Enzmology,* **1990,** 182, 50-69

27. Yotnda P., Firat H., Garcia-Pons F., Garcia Z., Gourru G., Vernant J-P., Lemmonier F. A., Leblond V., Langlade-Demoyen P., *J. Clin. Investigation*, **1998,** 101, 2290-2296.

28. Buzyn A., Ostankovitch M., Zerbib A., Kemula M., Connan F., Varet B., Guillet J-G., Choppin J., *Eur. J. Immunol.,* **1997**, 27, 2066-2072.

29. Norbury L.C., Clark R.E., Christmas S.E., *British J. Haematology.,* **2000,** 109, 616-621.

30. Bocchia M., Wentworth P.A., Southwood S., Sidney J., McGraw K., Scheinberg D.A., Sette A., *Blood,* **1995**, 85, 2680-2684.

31. Bocchia M., Korontsvit T., Xu Q., Mackinnon S., Yang S.Y., Sette A., Scheinberg D. A., *Blood,* **1996,** 87, 3587-3592.

32. Christinck E.R., Luscher M.A., Barber B.H., Williams D.B., *Nature*, **1991,** 352, 67

33. Vitiello A., Potter T.A., Sherman L.A., *Science,* **1990,** 250, 1423.

34. Harding C.V., Unanue E.R., *Nature,* **1990,** 346, 574

35. Demotz S., Grey H.M., Sette A., *Science,* **1990,** 249, 1028

36. Kageyama S; Tsomides T.J; Sykulev Y; Eisen H.N, *J. Immunol.,* **1995,** 154, 567-576.

37. Crotzer V.L., Christian R.E., Brooks J.M., Shabanowitz J., Settlage R.E. Marto J.A., White F.M., Rickinson A.B., Hunt D.F., Engelhard V.H., *J. Immunol.,* **2000,** 164, 6120-6129;

38. Brickner A.G., Warren E.H., Caldwell J.A., Akatsuka Y., Shabanowitz J., Eisenlohr L.C., Hunt D.F., Engelhard V.H., Riddell S.R., *J. Exp. Med.,* **2001**, 193, 195-206;

39. Pierce R.A., Field E.D., Mutis T., Golovina T.N., Kap-Herr C., Wilke M., Pool J., Shabanowitz J., Pettenati M.J., Eisenlohr L.C., Hunt D.F., Goulmy E., Engelhard V.H., *J. Immunol.*, **2001**, 167, 3223-3230.

40. Chen Y., Sdney J., Southwood S., Cox A.L., Sakaguchi K., Henderson R.A., Appella E., Hunt D.F., Sette A., Engelhard V.H., *J. Immunol.*, **1994,** 152, 2874-2881.

41. Skipper J.C.A., Hendrickson R.C., Gulden P.H., Brichard V., Van Pel A., Chen Y., Shabanowitz J., Wolfel T., Slingluff C.L., Boon T., Hunt D.F., Engelhard V.H., *J. Exp. Med.*, **1996**, 183, 537-534

42. Cox A.L., Skipper J., Chen Y., Henderson R.A., Darrow T.L., Shabanowitz J., Engelhard V.H., Hunt D.F., Slingluff C.L., *Science,* **1994,** 264, 716-719

43. Parkhurst M.R., Salgaller M.L., Southwood S. Robbins P.F., Sette A., Rosenberg S.A., Kawakami Y., *J. Immunol.*, **1996,** 157, 2539-2548.

44. Cullis J.O., Barnet A.J., Goldman J.M., Lechler R.I., *Leukemia,* **1994**, 8, 165-170.

45. Papadopoulos K.P., Suciu-Foca N., Hesdorffer C.S., Tugulea S., Maffei A., Harris P.E., *Blood,* **1997,** 90, 4938-4946.

46. Berke Z., Anderson M.H., Pedersen M., Fugger L., Zeuthen J., Haurum J.S., *Leukemia,* **2000,** 14, 419-426.

47. Schwarz M., *PhD Thesis*, **2005**, Freie Universitat, Berlin, Germany.

48. Riley C., *MSc Thesis,* **2005**, Nottingham Trent University, UK

49. Mosse C.A., Meadows L., Luckey C.J., Kittlesen D.J., Huczko E.L., Slingluff C.L., Shabanowitz J., Hunt D.F., Engelhard V.H., *J. Exp. Med.*, **1998**, 187, 37-48.

50. Moore R.E., Young M.K., Lee T.D., *J. Am. Soc. Mass Spectrom.*, **2000**, 11, 422-426.

51. Ducret A., Van Oostveen I., Eng J.K., Yates J.R. III, Aebersold R., *Protein Sci.* 1998, 7, 706-719.

52. Van den Eynde B., Peeters O., De Backer O., Gaugler B., Lucas S., Boon T., *J. Exp. Med.*, **1995,** 182, 689-698

53. De Backer O., Arden K.C., Boretti M., Vantomme V., De Smet C., Czekay S., Viars C.S., De Plaen E., Brasseur F., Chomez P., Van den Eynde B., Boon T., van der Bruggen P., *Cancer Res.*, **1999,** 59, 3157-3165.

54. Chang H-K., Park J., Kim W., Kim K., Lee M., Park U., Choi B., *Gynecologic Oncology*, **2005**, 97, 342-347.

55. Kobayishi Y., Higashi T., Nouso K., Nakatsukasa H., Ishizaki M., Kaneyoshi T., Toshikuni N., Kariama K., Nakayama E., Tsuji T., J. *Hepatology*, **2000**, 32, 612-617.

56. MS-tag website (http://prospector.ucsf.edu).

57. Cox A.L., Huczko E.L., Engelhard V.H., Shabanowitz J., Hunt D.F., *MHC volume 1: A Practical Approach,* **1997,** IRL Press, Oxford.

58. Hunt D.F., Henderson R.A., Shabanowitz J., Sakaguchi K., Michel H., Sevilier N., Cox A.L., Appella E., Engelhard V.H., *Science,* **1992,** 255, 1261-1263.

59. Heine G., Raida M., Forssmann W-G., *J. Chrom. A,* **1997**,776, 117-124.

60. Nagele E., Vollmer M., Horth P., *J. Chrom. A,* **2003**, 1009, 197-205.

61. Zhang G., Fan H., Xu C., Bao H., Yang P., *Anal. Biochem.,* **2003,** 313, 327-330.

CHAPTER 4

IMMOBILISED METAL ION AFFINITY CHROMATOGRAPHY (IMAC) FOR THE SELECTIVE ISOLATION OF PEPTIDES

Chapter 4 Aims Breakdown

The purpose of the work presented in this chapter was to explore alternative methods for peptide purification by immobilised metal ion affinity chromatography (IMAC).

The following points were explored.

- Investigation of conditions for a simple IMAC method for the selective purification of histidine containing peptides using immobilised copper ions.

- To investigate conditions for an IMAC method that is selective for all peptides via interaction with the N-terminus.

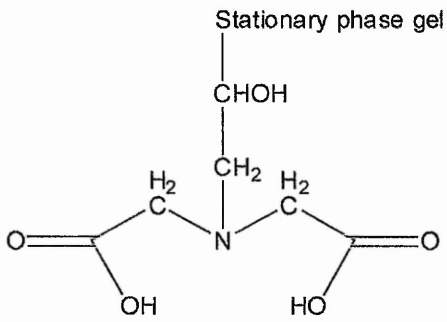
- Apply the IMAC method to the purification of biological samples including protein digests and cell surface eluates.

4.1.0 Introduction.

Immobilised metal ion affinity chromatography (IMAC), also known as metal chelate chromatography, was first introduced in 1975.[1] In recent years IMAC has proved to be a powerful tool for the isolation of species in proteomics. Only a small number of peptides in a protein digest are required to identify the originating protein. By reducing the number of peptides analysed the complexity of a sample is reduced and it becomes easier to usefully analyse. By selecting low abundance peptides, for example those containing histidine residues whose natural abundance in proteins is ~2.2 %, this can be achieved.[2-4] According to genomic analysis 82.5 % of human proteins will contain a histidine residue.[3] The separation of peptides and proteins is possible through co ordination between immobilised metal ions and electron donor groups present on amino acid side chains at the protein surface.[5-8] The most commonly used metal ions are Cu^{2+}, Ni^{2+}, Co^{2+} and Zn^{2+}.[9] Although many residues such as Glu, Asp, Tyr, Cys, His, Arg, Lys, Met can participate in binding, the actual protein retention in IMAC is based primarily on the availability of histidine residues. Cysteine and Tryptophan can also contribute to binding if they are in the vicinity of accessible histidine residues.[6] It has also been shown that above pH8, the α-amino group of the peptide amino-terminus has an affinity for Cu^{2+}-IDA.[10,11] Copper is the most commonly used metal ion due to its high affinity for the histidine residue. Other metal ions have also been used for histidine isolation, particularly nickel when chelated with nitrotriacetic acid (NTA). This approach is most successful when applied to adjacent histidine residues, which are uncommon in nature, or engineered histidine affinity tags.[12-18] In comparison with the widespread use of

IMAC for protein fractionation, the technique has been little used for the selective retention of peptides.[11,18-24]

A metal ion acceptor or chelating group such as iminodiacetic acid (IDA) (Figure 4.3.0.1) immobilised on agarose forms a column stationary phase to which metal ions, for example Cu (II), can be bound and immobilised. These metal ions act as an electron acceptor or Lewis acid forming a bond with the chelating ligands on the analyte. The analyte is captured by the metal complex and binds to the metal ion. This is illustrated by the binding of imidazole nitrogens of histidine to immobilised copper in Figure 4.3.0.2.



IDA

Figure 4.3.0.1 The chelating agent iminodiacetic acid (IDA) bound to agarose prior to metal ion binding forms the basis of an IMAC column.

Analyte molecule elution is achieved by lowering the pH of the mobile phase resulting in protonation of the nitrogen and disruption of the co-ordination bonds between the amino acid and immobilised metal ion.

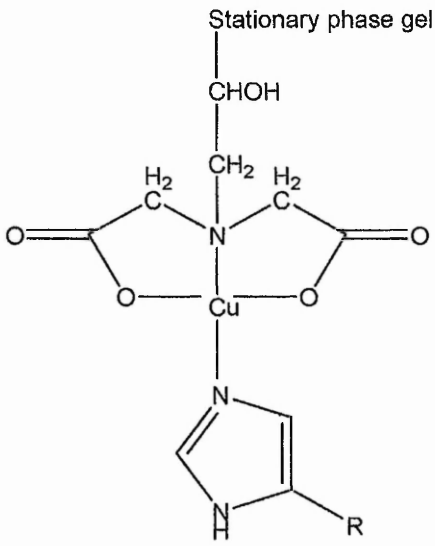


Figure 4.3.0.2 Stationary phase with bound IDA and Cu ion complex bound to histidine residue.

The technique of IMAC has also been used for more widespread applications. Cu, Zn, Co and Ni have been sucessfully applied to the isolation of human monoclonal immunoglobulin fragments following pepsin digestion.[25] A new polyacrylamide cryogen based support in conjunction with Cu-IDA has been compared with conventional Cu-IDA sepharose columns for the capture of histidine rich Urokinase using 20 mM HEPES with 0.2 M NaCl at pH7.0. It was found that the cryogen based column has better selectivity and operational efficiency.[26] IMAC purification of antibodies for immonpharmacotherapy on Hi-tagged cocaine binding scFv antibody.[27] IMAC has also been applied to non-biological samples for the purification of copper complexing ligands (Proteins and peptides and compounds with N-heterocyclic aromatic structures, e.g. humic and fulvic acids) from soils using a buffer of 2 mM MOPSO with 0.5M NaCl at pH 6.9.[26]

The mechanism of IMAC has been utilised in conjunction with BIACORE which is a biosensor instrument employing surface plasmon resonance detection. The ligand was immobilised on a sensor chip to which the analyte (in solution) was bound. The interaction was monitored and a sensorgram, a plot of interaction over time, was compiled to reveal binding characteristics. This was carried out using Ni^{2+}-NTA on the proteins GroEL and GroES using Mops-NaOH at pH7.2. It was found that proteins with only one histidine tag were bound very weakly at neutral pH.[29]

This chapter describes an investigation of the potential of IMAC using immobilised copper ions for the selective retention of histidine containing peptides and the application of the technique to the clean-up of cellular lysates.

4.2.0 Materials and Methods

Immobilised metal ion affinity chromatography (IMAC)

Investigations were performed on the use of IMAC for the purification of model peptides. The synthetic peptides used during this investigation were the histidine containing and non-histidine containing 9 mer MHC class I peptides, whose sequences are given below. For each sample approximately 100 pmol of each peptide were dissolved in 3-5 mls of buffer.

gp70 SPSYVY**H**QF$_{(423-431)}$ β-galactosidase TP**H**PARIGL$_{(876-884)}$

p53 LLGRNSFEV$_{(264-272)}$ p53 RMPEAAPPV $_{(65-73)}$

A Hi-Trap 1 ml metal chelating column (Amersham Biosciences, Buckinghamshire, UK) was used for IMAC purification. The column was loaded with 750 µl of 0.15 M $CuSO_4$, this left the column incompletely saturated with copper which was clear to see as only approximately three quarters of the column length turned blue. This incomplete saturation was carried out to prevent metal leaching. It was then washed with 2 column volumes of distilled deionised water to remove excess unbound copper. Equilibration was followed with 10 column volumes of buffer. The peptide sample was then applied and the column washed with 5 column volumes of buffer and elution was performed with 0.1 % TFA. The first 1 ml fraction of eluate was discarded and the following 3 ml were collected and combined. Analysis of the collected fractions was performed using capillary LC-nano/ESI/MS/MS. A schematic of the analysis process can be seen in Figure 4.2.0.1.

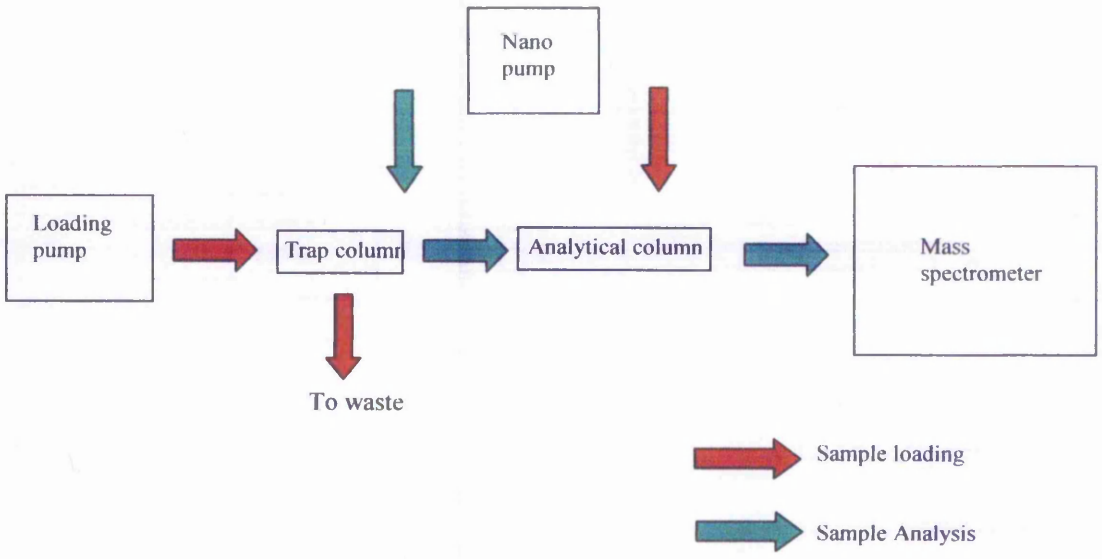


Figure 4.2.0.1. A schematic representation of the LC/MS system set-up.

The LC experimental conditions employed during this study were as follows: Injection volume: 1 µl. Trap column: pepmap 300 µm i.d x 5 mm, C18 nanobore column (Dionex, Surrey, UK), loading buffer: 0.1 % TFA at a flow rate of ~30 µl min^{-1}. Loading pump: Kontron 325 with microflow splitter (SGE, Milton Keynes, UK). Analytical column: pepmap 75 µm i.d x 150 mm, C18 nanobore column (Dionex, Surrey, UK). Mobile Phase A: H_2O + 0.09 % formic acid. Mobile Phase B: 80 % Acetonitrile + 0.09 % formic acid. Flow rate: 180 nl min^{-1}, Gradient profile; 12 % B – 75 % B in 60 minutes. Gradient pump; Ultimate™ nano-pump (Dionex, Surrey, UK).

Mass spectrometry was performed using on-line micro-electrospray on a Finnigan LCQ quadrupole ion trap mass spectrometer (ThermoFinnigan, Hemel Hempstead, UK). The typical mass spectrometric conditions were as follows:

Microspray voltage, 1.5-2.0 kV. Capillary Temperature, 200 °C. Capillary voltage, 46 V. Tube lens offset, -40 V. Collision energy, 35 %.

4.2.1 Selective binding of histidine containing peptides.

Investigations were performed into the optimum buffer conditions for the selective binding of histidine containing peptides using immobilised metal ion affinity chromatography (IMAC) methodology. Comparisons were made between the following loading buffers.

(i) The HEPES (4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid) buffer, 25 mM adjusted to pH 7.2 with 0.1 M NaOH.

(ii) The triethanolamine buffer, 5 % adjusted to pH 7.2 with conc. HCl.

4.2.2 Non-selective binding of peptides

Investigations were performed on the potential of immobilised metal ion affinity chromatography for the non-selective binding of peptides. The buffers explored during these experiments were as follows.

(i) CAPS buffer, 50 mM adjusted to pH 11.0 with 1.0 M NaOH.

(ii) Diethanolamine buffer, 50 mM adjusted to pH 10.0 with conc. HCl.

Experiments were then performed on the non-selective binding of the four peptide mixture using a diethanolamine loading buffer, 50 mM adjusted to pH 10.0 with conc. HCl. Once loading was complete, the non-histidine containing peptides were then eluted using the HEPES pH 7.2 buffer followed by the release of the histidine containing peptides by applying a 0.1 % TFA acid elution.

4.2.3 Application of IMAC method to cellular extracts

The histidine selective IMAC method, following binding of peptides at pH7.2 using the HEPES buffer, was used to purify and simplify peptides extracted from a cultured cell line. The cell line CT26 CL25 (mouse colon cell line transfected with the β-galactosidase gene) was grown in DMEM with 2 mM glutamine, 10 % FCS and 500 μg ml^{-1} G418 antibiotic to select for the transfected cells. They were grown at 37 °C with 5 % CO_2 atmosphere. When 80-90 % confluent the cells were split by treatment with trypsin/versene to release the cells which were then washed in fresh media and transferred to new flasks. The cells were centrifuged to pellet and washed in serum free culture media (RPMI-1640) to remove any serum proteins from the cell surface. The cells were centrifuged and all media drawn off then they were re-suspended in phosphate buffered saline (PBS).

Extraction of intracellular peptides

A mixture of chloroform : methanol : HCl (2000 : 1000 : 2) was added to the suspended cells (360 μl per 100 μl of solution, 2.4 ml added in total). The sample

was then vortex mixed (30 s), a further 800 µl of chloroform was then added and the sample vortex mixed again (30 s). The sample mixture was then centrifuged (5 minutes at 2500 rpm) to allow separation into layers. The lower fraction contained only waste products and solvent, whilst the upper layer was the methanolic fraction which contained free amino acids and peptides. These liquid phases were separated by a solid protein plug, shown schematically in Figure 4.2.3.1. The methanolic fraction was then collected and retained for separation by immobilised metal ion affinity chromatography.

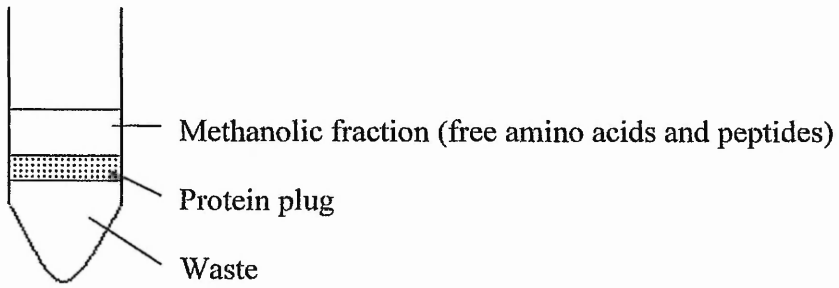


Figure 4.2.3.1 A schematic representation of chloroform : methanol extraction of peptides and proteins from cultured cells.

The protein plug was removed and washed twice by addition of 1 ml of ethylacetate followed by centrifugation (5 minutes at 2500 rpm). The plug (dry weight 96.1 mg) was subjected to tryptic digestion as follows. The protein was solubilised in diethanolamine (0.1 M, pH 8.4, 1 ml). To this was added a solution of trypsin (10 mg ml^{-1}, 30 µl), the solution was then vortex mixed and incubated (37 °C for 36 hours). The digest was quenched and undigested protein precipitated by the addition of TCA (72 % w/v, 100 µl), prior to centrifugation of

the remaining digest (10 minutes, 10000 g, 4 °C). The supernatant was collected and adjusted to ~pH 8 by the addition of 50 % NaOH (20 µl).

IMAC purification of intracellular peptides

The methanolic fraction and the protein digest were diluted 1 : 2 with HEPES buffer (25 mM pH 7.2) and independently subjected to the following IMAC procedure on a Hi-Trap 1 ml metal chelating column (Amersham Biosciences, Buckinghamshire, UK). The column was charged with $CuSO_4$ (0.15 M, 750 µl) and washed with 2 column volumes of double distilled H_2O to remove excess unbound copper. The column was equilibrated with 10 column volumes of HEPES buffer (25 mM, pH 7.2). The sample was then applied and the column washed with 3 column volumes of HEPES buffer and analyte molecules were then eluted with 0.3 % TFA. The first 1 ml fraction was discarded and the following 3 mls of eluate collected as a single fraction.

Analysis by nano-LC/µ-spray MS/MS

Analysis was performed using capillary LC-µ/ESI/MS/MS. The LC experimental conditions were as follows: Injection volume, 1000 µl. Trap column, pepmap 300 µm i.d x 5 mm, C18 nanobore column (Dionex, Surrey, UK), loading buffer, 0.1 % TFA at a flow rate of ~30 µl min^{-1}. Loading pump: Kontron 325 with microflow splitter (SGE, Milton Keynes, UK). Analytical column, pepmap 75 µm i.d x 150 mm, C18 nanobore column (Dionex, Surrey, UK). Mobile Phase A: H_2O + 0.09 % formic acid. Mobile Phase B: 80 % Acetonitrile + 0.09 % formic

acid. Flow rate: 180 nl min^{-1}, Gradient profile; 12 % B – 75 % B in 60 minutes. Gradient pump; Ultimate™ nano-pump (Dionex, Surrey, UK).

Mass spectrometry was performed using on-line micro-electrospray on a Finnigan LCQ quadrupole ion trap mass spectrometer (ThermoFinnigan, Hemel Hempstead, UK). The typical mass spectrometric conditions were as follows:

Microspray voltage, 1.5-2.0 kV. Capillary Temperature, 200 °C. Capillary voltage, 46 V. Tube lens offset, -40 V. Collision energy, 35 %. Data was acquired using Xcalibure 1.3 software (ThermoFinnigan, Hemel Hempstead, UK) operating in data dependent mode and acquiring tandem mass spectra for the top 3 parent ions. Data analysis was performed using the TurboSEQUEST® search engine (ThermoFinnigan, Hemel Hempstead, UK).

4.2.4 IMAC purification of CML patient eluates

The IMAC purification method was applied to the purification of cell surface eluates obtained from patients suffering from chronic myeloid leukaemia (CML). Two patient eluates designated patient 355 and 256 were provided by Professor Richard Clark of the Royal Liverpool Hospital (UK).

Surface elution of CML cells.

Peripheral blood mononuclear cells were harvested from patients with CML by leucophoresis. Surface elution was achieved using an acid elution technique

adapted from Storkus *et al.*[30] Glassware was silanised prior to use using Sigmacote (Sigma, UK). A minimum of 2.10^9 cells were washed in RPMI-1640 serum free media and centrifuged (1500 rpm, 4 °C, 5 minutes). Pelleted cells were then resuspended in 20 ml of citrate phosphate buffer (0.131 M citric acid + 0.066 M sodium phosphate adjusted to pH 3.3 with 10 M NaOH) for 5 minutes and centrifuged again. The supernatant was collected and filtered through a 0.22 µm micropore filter and stored at -80 °C.

Solid phase extraction and IMAC purification of CML surface eluates.

The CML eluates were cleaned-up using solid phase extraction cartridges SEP PAK® RP, C18, 360 mg sorbent, 55-105 µm particle size (Waters, UK). Two cartridges were used in series. The cartridges were conditioned with acetonitrile (70 %, 6 ml) and equilibrated with TFA (0.1 % v/v, 10 ml). The sample was then loaded and washed with TFA (0.1 % v/v, 2 ml). The cartridges were then eluted with 70 % acetonitrile (the first 500 µl of eluate were discarded). The resulting fractions were diluted 1 : 7 with 50 mM HEPES buffer to reduce the acetonitrile content to less than 10 % prior to loading onto the IMAC column. IMAC was then performed using the HEPES loading buffer (25 mM adjusted to pH 7.2 with 0.1 M NaOH) as described in section 4.2.0.

Analysis by nano-LC/µ-spray MS/MS

LC/MS/MS analysis was performed largely as described in section 4.2.3, except that the gradient was extended to allow improved separation. Injection volume: 20 µl, Gradient 12 % B-75 % B in 90 minutes. Data acquisition was

accomplished using the Xcalibur 1.3 data dependent 'triple play' program. Briefly, a full scan was performed followed by a zoom scan (increased resolution scan for charge state determination) and finally a tandem MS scan (collision energy 40 %). Individual ion masses were selected a maximum of 2 times before being put on an exclusion list for a duration of 3 minutes, once 3 minutes had passed the ions were once again free to be picked up. The zoom scan and tandem mass spectrometry scans were repeated for the top 6 parent ions observed in the full scan spectrum. This sequence of scans was repeated throughout the duration of the analysis. The mass range for the surface eluted peptide analysis was restricted to m/z 400-850. Data analysis was performed on the raw data using the TurboSEQUEST® searching algorithm (Thermofinnigan, Hemel Hempstead, UK).

4.2.5 IMAC purification of CML patient cells

Cell surface peptides extracted by surface washing of cells and intracellular peptides extracted from whole cellular lysates by tryptic digestion, as detailed in Section 4.2.3, were subjected to a further stage of SPE and IMAC clean-up. Two patient cell eluates, designated patient 301 and patient 286, and corresponding cell pellets were obtained from Professor Richard Clark of the Royal Liverpool Hospital (UK). A summary of the procedures these cells were subjected to can be seen in Figure 4.2.5.1.

Surface elution of CML cells

The cell surface peptides were obtained by mild acid elution as described in section 4.2.4 except the washing of the cells with RPMI-1640 serum free media was repeated three times prior to re-suspension in citrate phosphate buffer to ensure removal of serum proteins.

Extraction of intracellular peptides

The free intracellular peptides were extracted by treatment with chloroform/methanol and the resulting protein plug was digested tryptically, both of these procedures were as previously described in section 4.2.3.

Solid phase extraction and IMAC treatment of CML surface eluates and tryptic digests

The CML surface eluates and tryptic digest samples were subjected to an initial solid phase extraction clean-up followed by IMAC purification using the methods described in section 4.2.4.

Analysis by nano-LC/ESI/MS/MS

LC/MS/MS analysis was performed as described in section 4.2.4, except the LC mobile phase gradient was 12 % B - 75 % B in 60 minutes. The mass range scanned during the analysis of the cell surface eluates was m/z 400-1400 and this was increased to m/z 400-2000 to accommodate higher molecular weight peptides for the analysis of the methanolic fractions and intracellular tryptic digest.

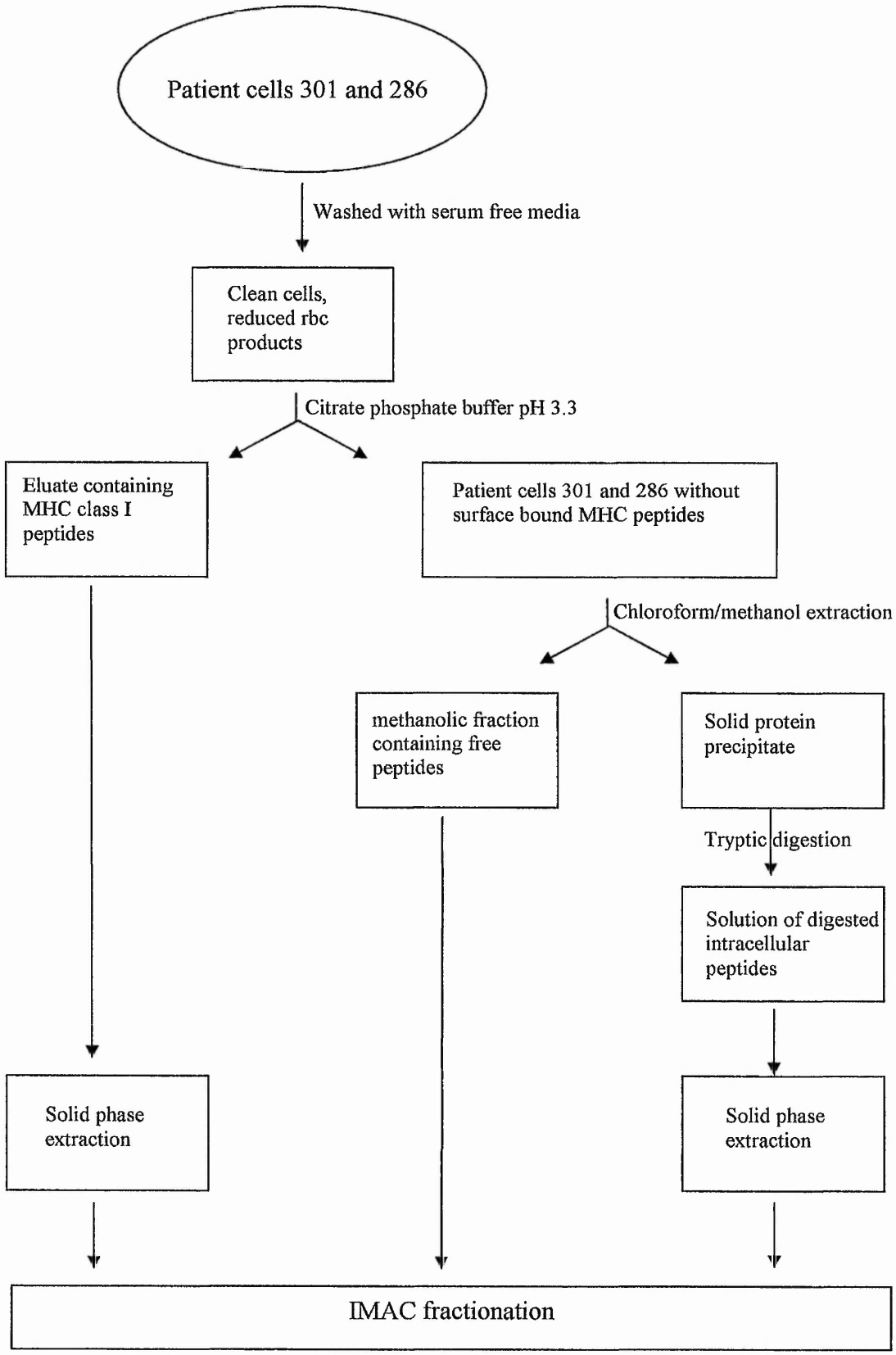


Figure 4.2.5.1 Flow chart summary of the treatment of patient cells 301 and 286.

4.3.0 Results and Discussion

Copper has been shown to be an effective metal for the retention of histidine containing peptides using the technique of immobilised metal ion affinity chromatography (IMAC).[3,11,20,21,24] It has been reported that histidine, although of low natural abundance, is present in 82.5 % of all human proteins. As such it represents an ideal target for the selective purification of complex biological samples.[3] Described below are the results of investigations carried out to determine the potential use of immobilised copper ions for the purification of histidine containing peptides.

4.3.1 Selective binding of histidine containing peptides.

The model synthetic 9 mer MHC class I histidine containing peptides (gp70 SPSYVY**H**QF$_{(423-431)}$, β-galactosidase TP**H**PARIGL$_{(876-884)}$) and non-histidine containing peptides (p53 LLGRNSFEV$_{(264-272)}$, p53 RMPEAAPPV $_{(65-73)}$) were applied to an Hi-Trap metal chelating column, which had been previously loaded with copper, to evaluate the selective binding of histidine containing peptides. Ren *et al* investigated three IMAC resins against histidine containing peptides to test selectivity and specificity. They found that the Hi Trap column used here performed most effectively with the least non-specific binding of the three resins.[23] Two different buffers were investigated, HEPES (4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid) buffer and triethanolamine buffer, both at pH 7.2. At this pH the imidazole nitrogens in the histidyl residues will be in their non-protonated state and therefore encourage binding to the metal complex.

The LC/MS/MS selected product ion chromatograms obtained following elution of the IMAC column with 0.1 % TFA are shown in Figures 4.3.1.1 and 4.3.1.2. Both buffers effected the selective binding of the two histidine containing peptides. Similarly for both buffers the two non-histidine containing peptides were not retained.

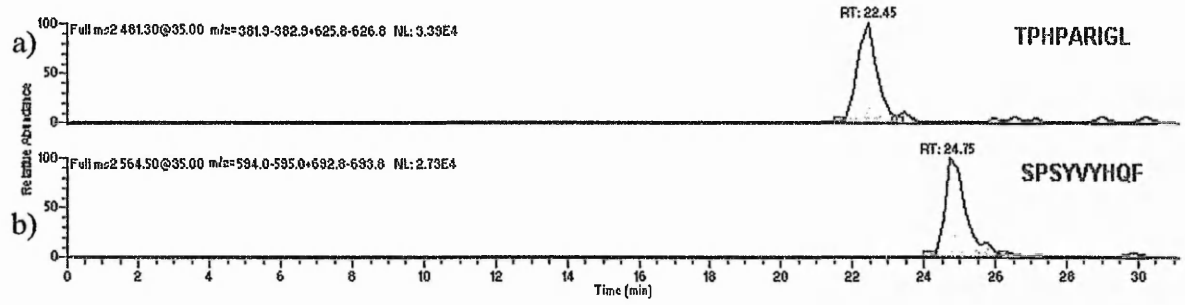


Figure 4.3.1.1 Selected product ion chromatograms for the IMAC separation using the triethanolamine buffer, pH 7.2. (a) β-galactosidase chromatogram (b) gp 70 chromatogram.

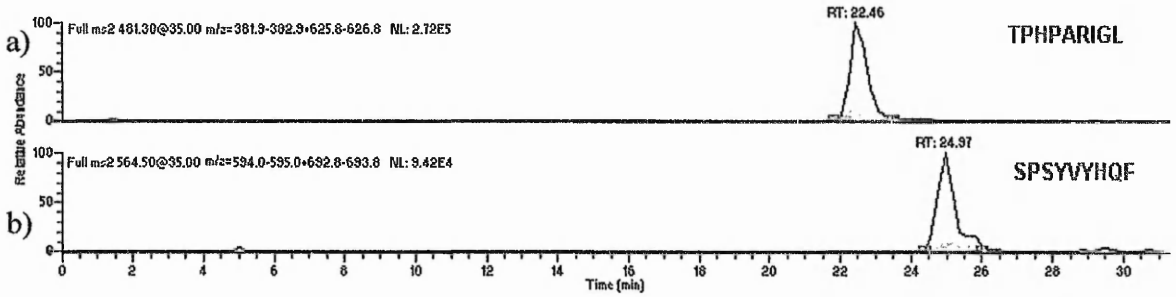


Figure 4.3.1.2 Selected product ion chromatograms for the IMAC separation using the HEPES buffer, pH 7.2. (a) β-galactosidase chromatogram (b) gp 70 chromatogram.

Recoveries of the histidine-containing peptides were determined by LC/MS/MS analysis of the peptides with and without IMAC fractionation at the same

concentration and under the same analytical conditions. These recoveries are presented in Table 4.3.1.1.

Table 4.3.1.1 Recoveries of the histidine containing peptides SPSYVYQF and TPHPARIGL from the IMAC column using the HEPES (mean data) and triethanolamine buffers.

	% Recovery	
Peptide	**TPHPARIGL**	**SPSYVYHQF**
HEPES buffer	103	102
Triethanolamine buffer	12	44

Following a promising initial test of the HEPES buffer the experiment was repeated and the data presented in Table 4.3.1.1 is the mean recovery for these two experiments (TPH, 120 and 86 %, range error = 33 %; SPS, 83 and 121 %, range error = 37 %). Initial testing of the triethanolamine buffer did not look promising and therefore no further tests were performed with this system. It is clear from the recovery data in Table 4.3.1.1 that, although both buffers successfully discriminated in favour of the histidine containing peptides, the HEPES buffer facilitates better retention of the histidine containing peptides on the column (average recovery was ~100 % for both peptides). In comparison the triethanolamine buffered experiment shows a considerably reduced retention of <50 % for both peptides and as low as ~12 % for the β-galactosidase peptide. Although the range error for the HEPES buffer data is high a reproducible trend of high recovery was observed. Such a trend was ultimately the desired outcome, in contrast to the triethanolamine buffer whose initial recoveries were judged too low to be worth repeating. A possible explanation for these observations,

regarding each buffer, is that HEPES with a pk_a of 7.48 (and a buffering pH range of 6.8-8.2) is better able to buffer at pH 7.2 than triethanolamine buffer which has a higher pk_a of 7.76 (and a more limited pH range of 7.0-8.3). An operating pH of 7.2 lies closer to the midrange for the HEPES buffer. This was a successful test of the method which was shown to be selective for histidine containing peptide standards. This method was explored for the purification of more complex samples, this work is discussed later in the chapter (section 4.3.3 onwards).

4.3.2 Non-selective binding of peptides

Following the success of the histidine selective method described in section 4.3.1, the use of the IMAC methodology for the non-selective binding of peptides was explored using buffers at pH 10.0 – 11.0. Since at this pH the α-amino group of the N-terminus will not be protonated it will be susceptible to association with the immobilised metal complex, the mechanism of which is shown in Figure 4.3.2.1.[10,11] Hansen *et al* have investigated the importance of the α-amino group in the purification of histidine containing peptides at neutral pH. Using immobilised Cu^{2+} metal ions, peptides were strongly retained even with a blocked α-amino group. With immobilised Ni^{2+} metal ions, blocking the α-amino group or removing Imidazole moiety reduced binding by up to a factor of 10.[11]

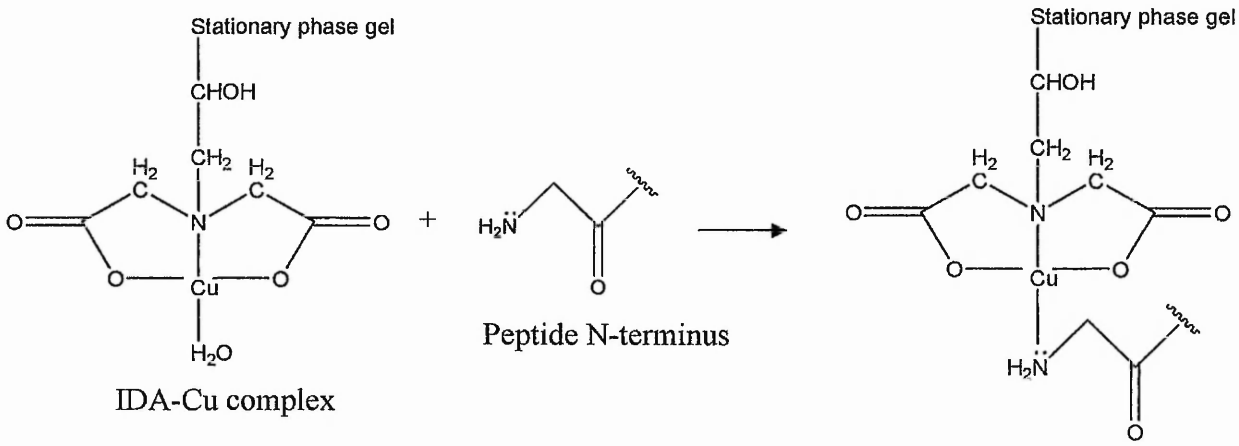


Figure 4.3.2.1 The mechanism of the non-selective binding of peptides to Cu-IMAC stationary phase via the N-terminus.

A solution containing the four model histidine and non-histidine containing 9 mer peptides, each ~100 pmol in buffer, was loaded onto the copper charged IMAC column. Two different buffers were used: CAPS buffer at pH 11.0 and a diethanolamine buffer at pH 10.0. The LC/MS/MS selected product ion chromatograms of these peptides following loading in the diethanolamine buffer and the CAPS buffer followed by elution with 0.1 % TFA are shown in Figures 4.3.2.2 and 4.3.2.3 respectively.

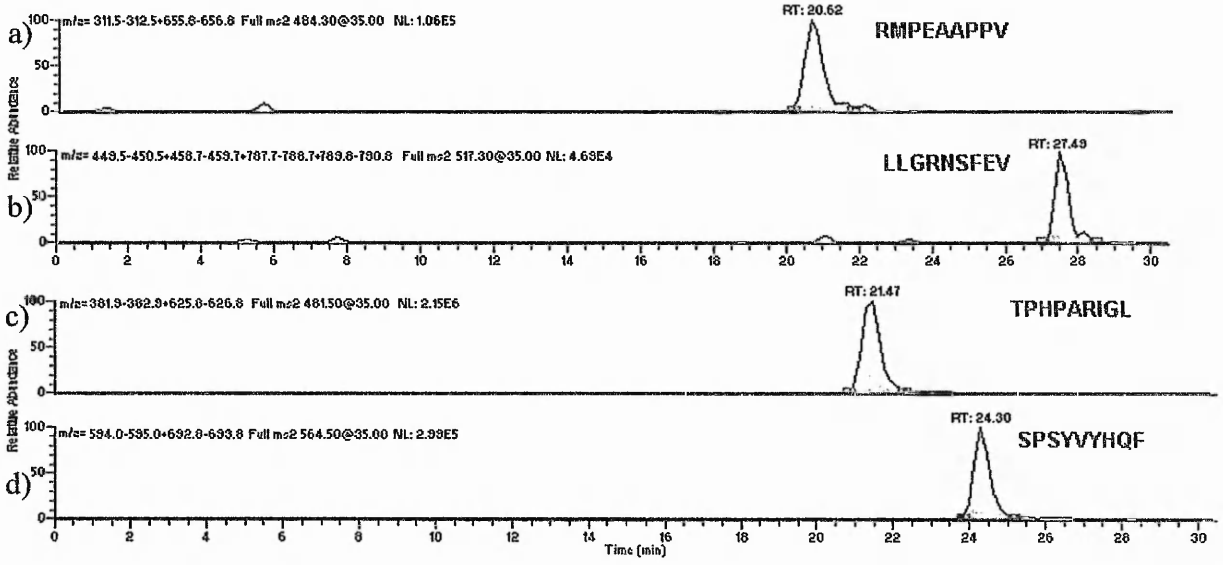


Figure 4.3.2.2 Selected product ion chromatograms of IMAC purified four peptide mixture using the diethanolamine buffer pH 10.0. (a) p53 RMPEAAPPV (b) p53 LLGRNSFEV (c) β-galactosidase TPHPARIGL (d) gp70 SPSYVYHQF.

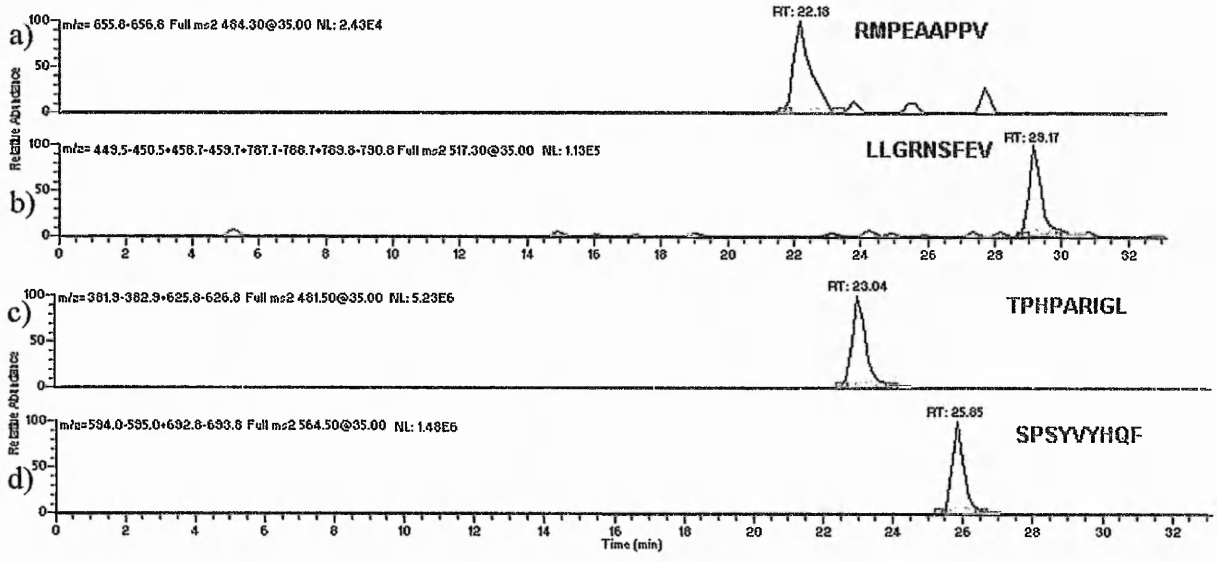


Figure 4.3.2.3 Selected product ion chromatograms of IMAC purified four peptide mixture using the CAPS buffer pH 11.0. (a) p53 RMPEAAPPV (b) p53 LLGRNSFEV (c) β-galactosidase TPHPARIGL (d) gp70 SPSYVYHQF.

It can be seen in Figures 4.3.2.2 and 4.3.2.3 that the presence of all four peptides is observed in the selected product ion chromatograms, indicating the non-selective binding of all four standard peptides to the Cu/IMAC column. An estimate of recovery was made in each case to allow more informative comparison between the two buffers. These data are presented in Table 4.3.2.1.

Table 4.3.2.1: Recoveries for the non-selective binding of peptides using the diethanolamine and CAPS buffers.

	% Recovery			
Peptide	**RMPEAAPPV**	**LLGRNSFEV**	**TPHPARGL**	**SPSYVYHQF**
diethanolamine buffer	37	33	120	25
	12	72	134	22
Mean	25	53	127	24
CAPS buffer	10	47	106	68
	0	28	97	4
Mean	5	38	102	36

The data in Table 4.3.2.1 clearly shows that although both of these buffers allow the binding of both histidine and non-histidine containing peptides to the IMAC column the recoveries of the non-histidine containing peptides are quite low. Less than 50 % recovery was observed in all cases, significantly lower than this was seen for the RMPEAAPPV peptide using the CAPS buffer which is not present at all during the second analysis. Similarly the recovery of the histidine containing SPSYVYHQF peptide has also decreased significantly when

compared with the HEPES buffer at pH 7.2 (Section 4.3.1). Both the histidine containing peptides show reproducible recovery with the diethanolamine buffer (range error; TPH 11%; SPS 13%) however the range errors for the non-histidine containing peptides are in excess of 70%. Only the β-gal peptide performs consistently with the CAPS buffer (range error 9%), range error for the remaining peptides are all in excess of 50%. From these data the diethanolamine buffer was concluded to be the most consistent, although some refinement is clearly required. This method has shown some promise but has highlighted the requirement for further development to improve consistency although it can be said from this data that it should be the diethanolamine buffer which is pursued rather than the CAPS.

Following the tentative success of the previous two methods, elements of both were combined to explore the possibility of a two step elution approach. Both histidine and non-histidine containing peptides were loaded onto the column using the diethanolamine buffer. The aim behind such a process was to attempt to elute the non-histidine containing peptides (by lowering the pH to 7.2 with a HEPES buffer wash) prior to elution of the histidine containing peptides (with the 0.1 % TFA acid elution). Each elution was collected and analysed independently by LC/MS/MS. The resulting chromatograms can be seen in Figure 4.3.2.4. Figures 4.3.2.4 (a) and (b) show the non-histidine peptides eluted by washing with the HEPES buffer and Figures 4.3.2.4 (c) and (d) show the histidine containing peptides eluted by the application of 0.1 % TFA. The peptide recoveries showed a similar pattern to that observed above with the

diethanolamine buffer alone. Once again the recovery of the β-galactosidase peptide was excellent. The other histidine containing peptide from gp70 was not well recovered ~32 % which was slightly better than was seen previously. Both the p53 non-histidine containing peptides however showed slightly poorer recoveries ~20 %.

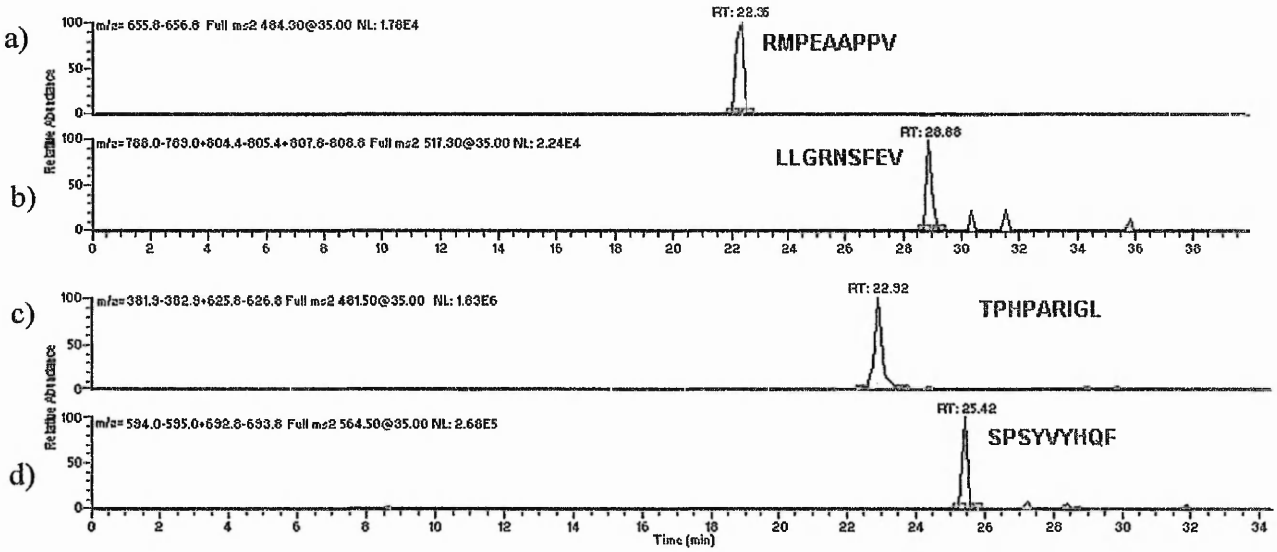


Figure 4.3.2.4 Selected product ion chromatograms of IMAC purified peptide mixture loaded using the diethanolamine buffer at pH 10.0 and eluted using a two step procedure with HEPES buffer (pH 7.2) followed by 0.1 % TFA. (a) p53 RMPEAAPPV eluted with HEPES buffer (b) p53 LLGRNSFEV eluted with HEPES buffer (c) β-galactosidase TPHPARIGL eluted with 0.1 % TFA (d) gp70 SPSYVYHQF eluted with 0.1 % TFA.

Repeated attempts to improve these recoveries were unsuccessful and several attempts made to replicate this non-selective binding using new Hi-trap columns also proved unsuccessful with only the histidine containing peptides being captured by the Cu/IMAC column during binding with the pH 10 and 11 buffers. It was proposed that a high degree of between column variability was affecting

the weaker interactions between the peptide N-terminus and the Cu ion. The selective binding of the histidine containing peptides was unaffected by changing columns. Consequently attempts to combine non-selective and selective histidine binding in a single procedure were not pursued further.

4.3.3 Application of IMAC method to complex cellular samples

One approach in proteomics is to decrease the complexity of tryptic peptide mixtures and reduce the number of peptides that must be examined to identify proteins Digestion of a protein results in a series of peptides of variable size, above 6 residues the peptide become increasingly unique to the point where only one peptide may be required to identify the parent protein.[22] Such signature peptides can be used to reduce sample complexity.[31] Cu(II)-IMAC has been used in this way for the simplification of protein tryptic digests by the isolation of peptides containing histidine and cysteine residues.[3,23,24]

The selectivity of the Cu/IMAC method developed in Section 4.3.1 for model histidine-containing peptides suggests that the histidine containing peptides in complex proteomic samples, such as cell lysates and eluates, might be fractionated using this approach. The method was therefore evaluated for the purification and simplification of proteins from cultured cell lines. The murine cell line designated CT26.CL25, a colon carcinoma cell line with high expression of the tumour associated antigen β-galactosidase,[32] was selected for analysis. Cells were cultured and the population lysed using acidified

chloroform/methanol mixture (Section 4.2.3) to solubilise and disrupt membranes.[33] The protein is precipitated and partitioned between the two solvent phases and the lipid fraction extracted into the chloroform layer which was discarded. The peptides and free amino acids were extracted into the methanolic fraction of the liquid/liquid extraction. The protein pellet was washed, resolubilised and digested. The proteins were digested with trypsin prior to clean-up by IMAC. Both the peptides and protein tryptic digests were then applied to the Cu/IMAC column using the histidine selective methodology to purify and reduce the complexity of the samples. The retained fractions were then analysed by nano-LC/MS/MS for peptide sequence and protein identification.

The base peak chromatogram for the protein digest is shown in Figure 4.3.3.1. A number of prominent peaks can be seen along the chromatogram, these are particularly intense between 20 and 45 minutes. The raw tandem mass spectral data was searched against a murine database using the TurboSEQUEST algorithm. The highest matching peptide sequences are shown in Table 4.3.3.1, together with the SEQUEST score and xCorr values.

The SEQUEST score is a preliminary scoring system, which is based on the number of ions observed in the MS/MS spectrum which match the predicted data for a particular peptide sequence. A value of >200 indicates a good match. It should be noted that a larger sequence will generally score higher than shorter sequences.

The Xcorr factor is the cross correlation value from the search and is indicative of the best match in the search. An Xcorr value in the region of 2.00 suggests a good correlation. In all cases these results provide a good first stage in identifying peptides, although manual confirmation of assignment is also required particularly if the assigned values are borderline.

In all cases the correlation factor (Xcorr) was greater than 1.8 and considerably higher for most peptide sequences, a SEQUEST score greater than 200 is also observed in all cases. The parent proteins are listed in the table and represent a range of cellular housekeeping proteins. Of the peptide sequences identified 90 % are histidine containing confirming that the IMAC method remains highly selective during the purification of a complex mixture such as a whole cell digest. The majority of proteins which have been identified by the SEQUEST algorithm are single hit results. This was expected as the histidine residue has such a low abundance in naturally occurring proteins. Genomic analysis performed by Wang *et al* have predicted that, in humans 82.5 % of proteins will be histidine containing and that *in silico* tryptic digestion of these proteins will yield 16.8 % histidine containing peptides.[3] The liklehood therefore of multiple peptides originating from the same protein containing histidine residues is slim.

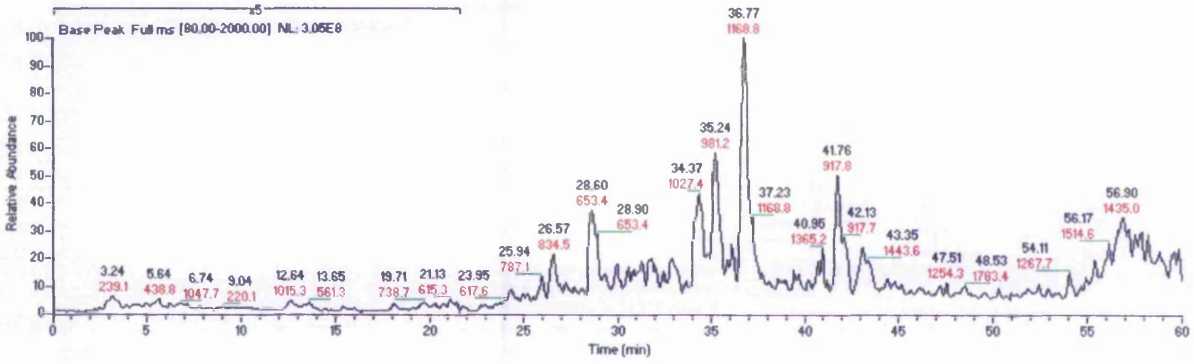


Figure 4.3.3.1 Full scan base peak chromatogram of the LC/MS analysis of the CT26.CL25 murine cell digest following IMAC purification. The base peak m/z is annotated in red below the retention time. The early segments of the chromatograms have been expanded to visualise the baseline.

The CT26.CL25 cell line over expresses the β-galactosidase tumour antigen. The LC/MS data was searched against a database containing this protein but the presence of peptides originating from this antigen were not detected in the presence of the total cell proteome. However a selection of peptides originating from various cellular proteins was identified using the SEQUEST algorithm with a selection of the strongest matches give in Table 4.3.3.1.

Table 4.3.3.1 Peptide sequences identified by TurboSEQUEST search algorithm from analysis of IMAC purified CT26.CL25 cellular protein tryptic digest. TurboSEQUEST score and X correlation factor and parent protein are shown.

Sequence	XCorr	Score	Protein
HISQISVADDDDESLLGH	4.88	2095	Protein kinase C inhibitor
HLEINPDHSIIETLR	3.69	955	Heat shock protein 86
YPIEHGIITNWDDMEK	3.60	892	Alpha actin

SLTNDWEDHLVAK	3.56	1128	Heat shock protein 86
SLTNDWEEHLVAK	3.40	1314	Heat shock protein 84
ISSVQSIVPALEIANAHR	3.32	263	Chaperonin
LLGGVTIAQGGVLPNIQAVLLPK	3.27	584	Histone 5
HLSNIPPSVSEDDLK	2.98	402	Polypyrimidine tract binding protein
DRPFFPGLVK	2.97	725	Nucleoside diphosphate kinase
HLAGLGLTEAIDK	2.36	430	Collagen binding protein
VIHDNFGIVEGL	2.50	562	Glyceraldehyde 3-phosphate dihydrogen
HQGVMVGMGQK	2.54	590	Alpha-2-actin
ALPGHLKPFETL	2.32	966	Glutathione S-transferase
VVSALPIQH	1.82	487	Ig gamma 2b chain precursor
KPLPDHVSIVEPK	2.80	683	Ribososmal protein S3
HAVSEGTK	2.75	1015	Histone H2BF
FIQHPKNFGLASYLERKSVPDCVL	2.09	423	Nuclear receptor co repressor
HLKIVSNKLESVEEQVSTVMKTEEMEAKR	2.02	410	Cell proliferation antigen
SRLQVRSIQQVVAKQSHQK	1.96	221	Cytochrome C oxidase subunit
QKIQGPTKWNKAKRHGESIKGKTESSK	1.93	319	Coagulation factor VIII
GVSQAVEHINK	2.12	363	Enolase 1 alpha non neuron

The proteins which have been identified represent a range of housekeeping proteins found in cells. For example heat shock proteins are present in all cells and become elevated when cells are exposed to stress. Chaperonins assist in the folding of proteins and actin which is a cytoskeletal protein.

The LC/MS/MS analysis of the IMAC purified methanolic fraction did not yield

particularly strong matches when searched using the SEQUEST algorithm. It was concluded that the methanolically extracted histidine containing peptides were not present in sufficient numbers or of sufficient concentration to be readily detected by the mass spectrometric process.

Overall the evaluation of the IMAC method using immobilised copper ions showed that histidine-containing peptides could be isolated successfully allowing for the simplification of an extremely complex cellular proteome digest and the subsequent identification of a number of housekeeping proteins by LC/MS/MS data dependent acquisition and TurboSEQUEST data searching algorithm.

4.3.4 IMAC purification of CML patient eluates

Following the success of the IMAC purification method with the murine cells it was then applied to the clean-up of cell surface eluates obtained from patients suffering from chronic myeloid leukaemia. A mild acid wash was used to elute peptides from the class I molecules on the cell surface using the procedure described in detail in chapter 3. Citrate phosphate buffer (pH 3.3) was used to destabilise the MHC β_2-microglobulin chain, thus releasing the MHC associated peptide from the binding groove and into the eluate. The eluate was then applied to a solid phase extraction cartridge to concentrate and desalt the eluted peptides prior to IMAC purification. This analysis was performed on two separate patient eluates designated Patient 355 and Patient 256. A full scan base peak chromatogram can be seen in Figure 4.3.4.1 for both patient 355 (a) and 256 (b).

The full scan base peak chromatograms shown in Figure 4.3.4.1 show many similarities between the two patients. Analysis of the TurboSEQUEST search results reveals an abundance of haemoglobin-related and serum peptides. These suggest that the extracts are contaminated by the products of red blood cells. The cells were obtained by leucophoresis of whole blood at the University of Liverpool and the red blood cells should be separated out by this procedure. However, the presence of red blood cell products in the extract can occur. This was clearly the case for the patient samples 355 and 256, where the haemoglobin fragments dominate the chromatogram.

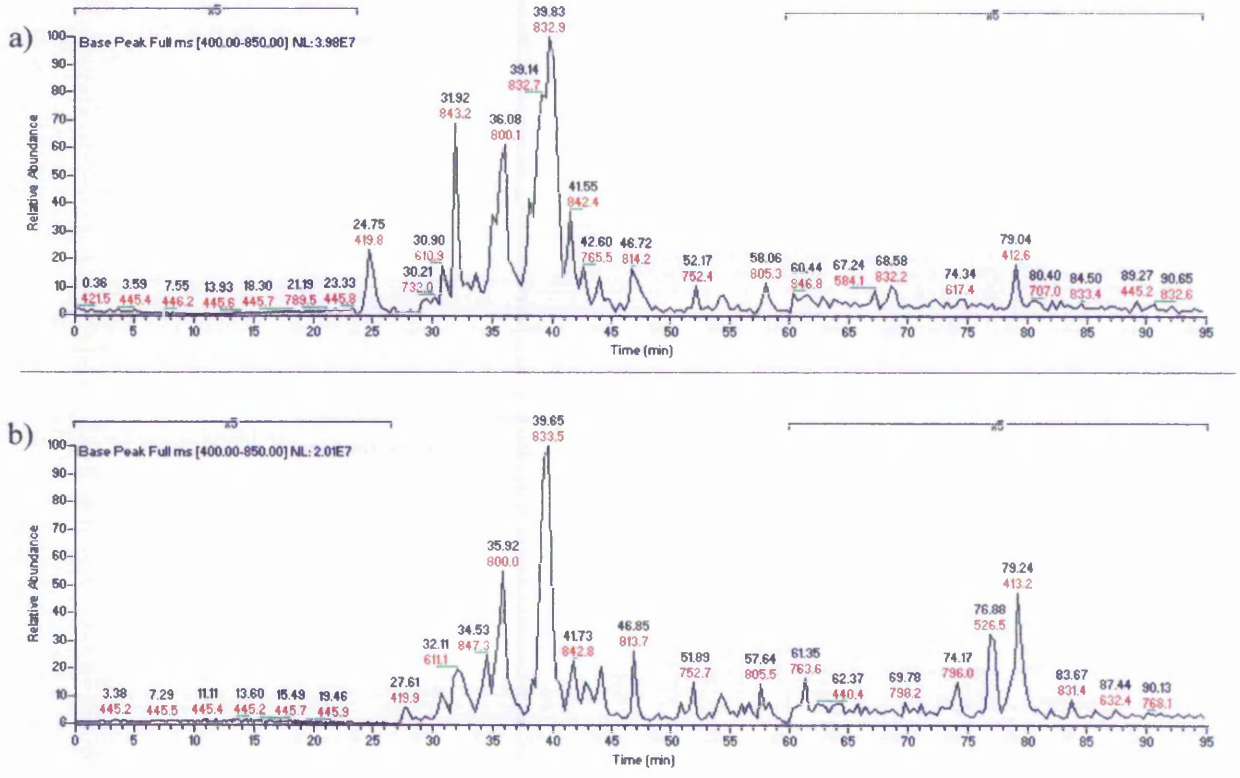


Figure 4.3.4.1 The LC/MS base peak chromatograms for patient 355 (a) and patient 256 (b). The base peak m/z is annotated in red below the retention time. Segments of the chromatogram have been expanded to visualise the baseline.

A selection of matches for both patient samples are presented in Table 4.3.4.1, which shows the Xcorr factor and score assigned to each match during the search.

Table 4.3.4.1 Peptide sequences identified by TurboSEQUEST search algorithm from analysis of IMAC purified CML cell surface eluates for patients 256 (blue) and 355 (yellow). TurboSEQUEST score and X correlation factor and parent protein are shown.

Sequence	Xcorr	Score	Protein
YQKVVAGVANALAHKYH	3.65	996	Haemoglobin beta
TLAAHLPAEFTPAVHASLDKF	3.37	1408	Haemoglobin alpha
AAHLPAEFTPAVHASLDKF	3.14	1021	Haemoglobin beta
FVSNHAY	1.93	363	Aldolase A
VLSPADKTNVKAAWGKVGAHAGEYG	4.02	1515	Haemoglobin alpha
AAHLPAEFTPAVHASLDK	3.93	1140	Haemoglobin alpha
YQKVVAGVANALAHKYH	3.88	2012	Haemoglobin beta
PVNFKLLSH	3.39	1121	Haemoglobin alpha
LSFPTTKTYFPHF	2.97	865	Haemoglobin alpha
LLVRYTKKVPQVSTPTL	2.84	645	Albumin precursor
YTKKVPQVSTPTL	2.55	753	Albumin precursor

The full scan single ion chromatograms for the top three SEQUEST hits from the analysis of patient eluate 355 are shown in Figure 4.3.4.2. All three originate from the haemoglobin alpha or beta chain. It is clear from the chromatograms that these ions are strong, with total ion counts in the 10^7 region.

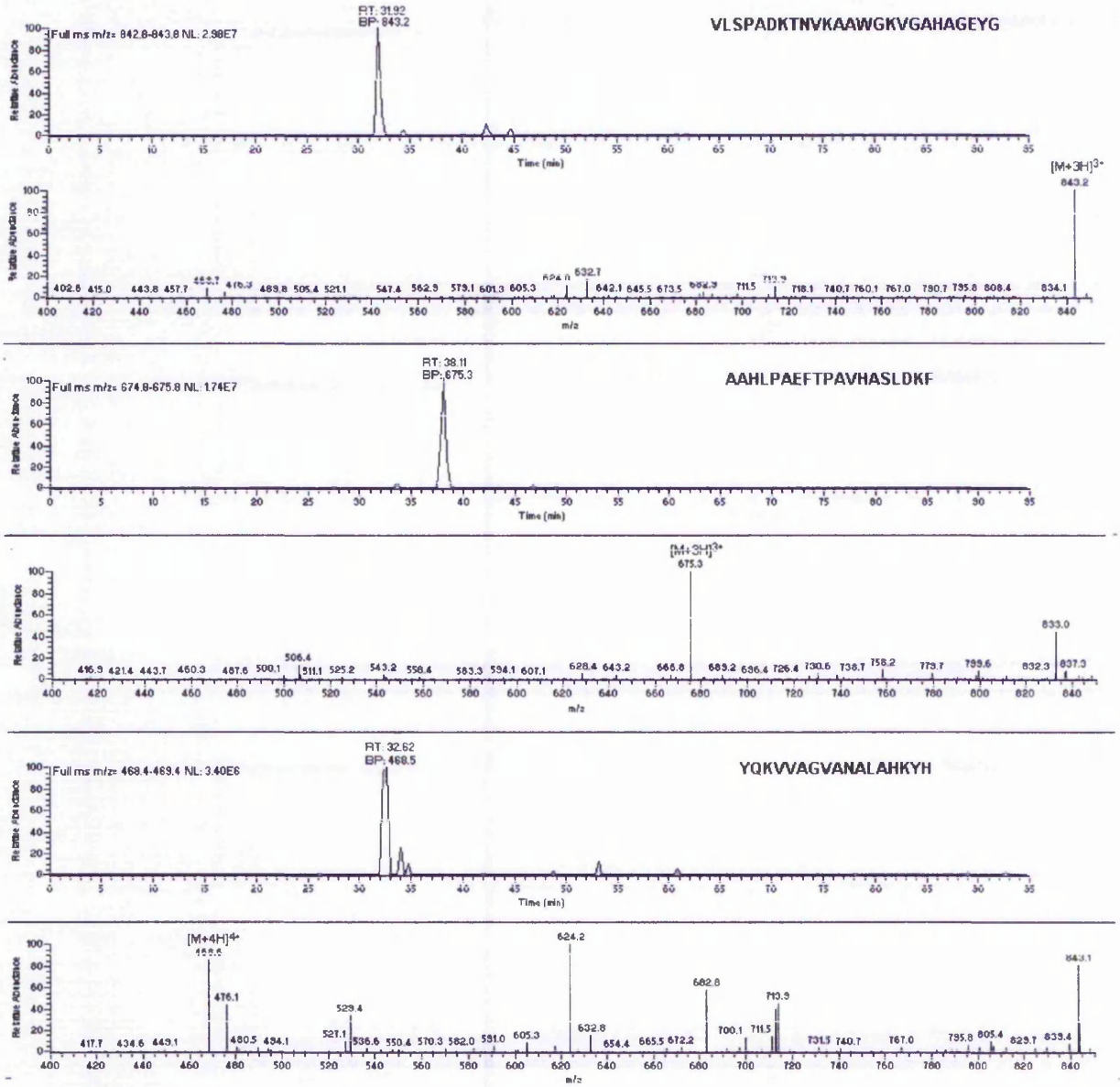


Figure 4.3.4.2 Full scan single ion chromatograms showing the top 3 SEQUEST hits for patient 355.

Most of the dominant peptides recovered are haemoglobin and albumin derived and it can be seen that all of these haemoglobin peptides are histidine containing. Surprisingly the two albumin precursor peptides are both non-histidine containing. However, human serum albumin is by its nature a 'sticky' protein, which may explain why it was recovered as a result of non-selective binding either to the IMAC support or potentially to other bound peptides during the

IMAC procedure. It was noted above that although the IMAC procedure is selective, during purification of very complex samples it is not 100 % selective for histidine-containing peptides and some non-histidine containing peptides may also be retained, particularly when they are present at high levels. The recovery efficiency of the IMAC column may be very low in these cases, but the high initial concentration of peptides means traces of peptides remain and are detected by LC/MS. Ren *et al* have reported that Cu^{2+}-IMAC is highly selective for histidine containing peptides from β-galactosidase protein although a low degree of non-selective selection observed.[22]

One of the peptides identified in Table 4.3.4.1 does not originate from the haemoglobin or albumin proteins. The chromatogram and tandem mass spectrum of the peptide, FVSNHAY, which originates from the Aldolase muscle protein can be seen in Figure 4.3.4.2. This was the fourth strongest matching assignment from the SEQUEST search for patient 256. The corresponding chromatogram, mass spectrum and assigned tandem mass spectrum for the $[M+2H]^{2+}$ ion (m/z 420) are shown in Figure 4.3.4.2.

Clearly the presence of haemoglobin and serum albumin in the eluates is hindering their analysis by swamping the signal of other native peptides. A change to the sample handling after the leucophoresis stage is required in order to eliminate this problem. These data show that native as well as tryptically digested histidine-containing peptides may be fractionated successfully using the

Cu/IMAC procedure to simplify complex biological samples for LC/MS analysis.

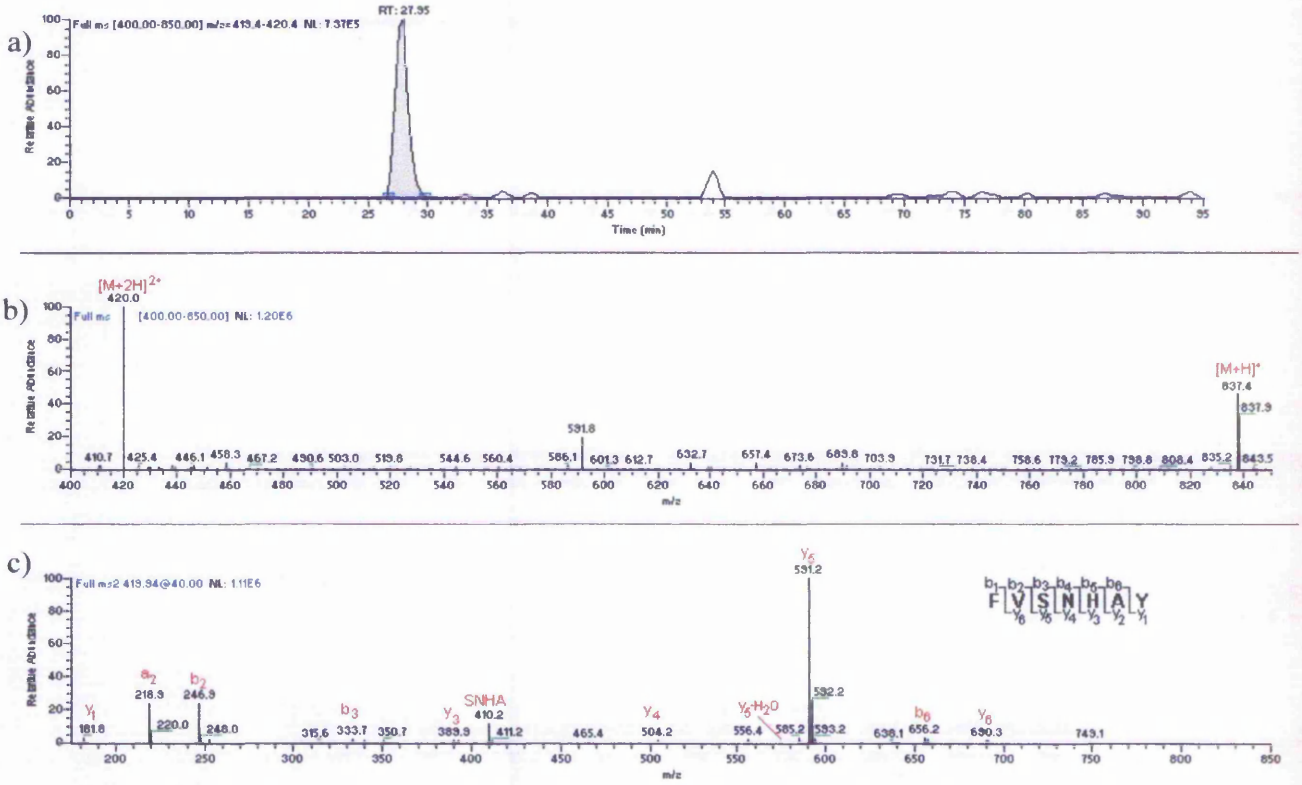


Figure 4.3.4.2 Aldolase A peptide FVSNHAY a) Chromatogram of doubly charged parent ion m/z 420.0. b) Full scan mass spectrum showing the doubly charged ion m/z 420.0 and the singly charged ion m/z 837.4. c) Tandem mass spectrum of the doubly charged ion (m\z 420) with b and y assignment.

4.3.5 IMAC purification of CML patient cells

The analysis of patient eluates performed in section 4.3.4 was hindered by the presence of excessive amounts of haemoglobin and serum albumin fragments. In an effort to eliminate this problem a further set of patient samples, designated patient 301 and patient 286, were obtained. During the cell washing processes

these cells were subjected to an additional repeat washing stage using serum free media in an attempt to eliminate haemoglobin and serum derived contaminants. The washed cells were eluted with citrate phosphate buffer as described in section 4.3.3, to elute cell surface peptides. Furthermore the 301 and 286 cells were subjected to the chloroform/methanol procedure to extract the free intracellular peptides and the protein precipitated as a solid mass which was then resolubilised and digested with trypsin. The CML surface eluates and tryptic digests were then subjected to SPE and the eluates from this and the methanolic fraction were each applied to the IMAC purification step. A schematic of the preparation stages for each extract is displayed in Figure 4.2.5.1. Analysis of IMAC fractions was carried out by chromatography and tandem mass spectrometry and the data was analysed using the TurboSEQUEST search algorithm to identify peptide sequences and the protein of origin in each case.

Surface elution of CML cells

An example of a base peak chromatogram for both patient 301 and 286 samples can be seen in Figure 4.3.5.1. This figure shows both of the patient sample chromatograms have strong signals with a total ion count in the region of 10^9. In both cases the overlapping of individual peaks has resulted in chromatograms which have an unresolved appearance, however data dependent tandem mass spectrometry analysis ensures that useful data is obtained in spite of this overlap of signals.

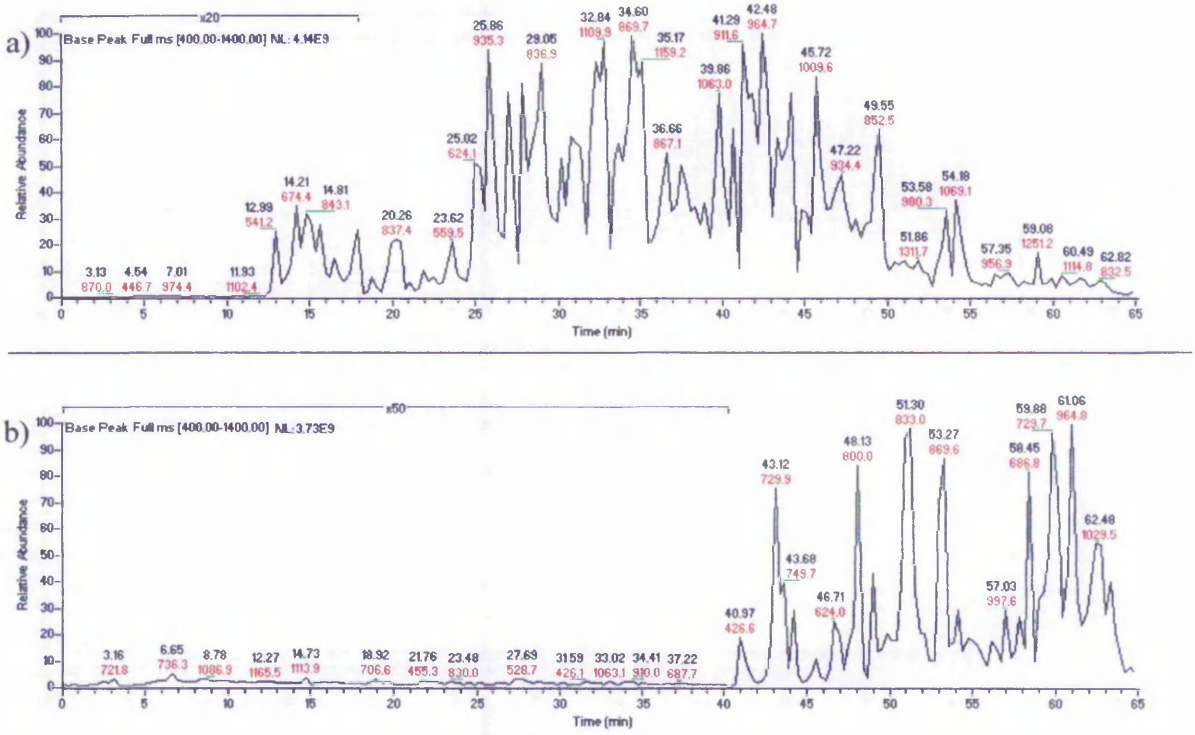


Figure 4.3.5.1 Base peak chromatograms for IMAC purified cellular eluates of CML patient cells. The base peak m/z is annotated in red below the retention time. Segments of the chromatogram have been expanded to visualise the baseline. (a) Patient 301. (b) Patient 286.

The results of the SEQUEST searches performed on the patient eluates 301 and 286 can be seen in Table 4.3.5.1. The peptide sequence is shown with the histidine residues highlighted in red. This table also shows the Xcorr and Score assigned by the SEQUEST search to each match. The sequences are all histidine containing peptides and represent a range of cellular proteins. Although a few peptides originate from the haemoglobin family these no longer swamp the signal as was seen with the eluates for patients 256 and 355 (Section 4.3.4) which had not undergone the extra washing stage with serum free media prior to elution.

Table 4.3.5.1 Peptide sequences identified by TurboSEQUEST search algorithm from analysis of IMAC purified CML cell surface eluates for patients 301 (blue) and 286 (yellow). TurboSEQUEST score and X correlation factor and parent protein are shown.

Sequence	Xcorr	Score	Protein
MAHMASKE	2.52	777	Glyceraldehyde 3-phosphate
ECQSHKLTVE	2.35	596	Proteosome subunit
YQKVVAGVANALAHKYH	4.98	2534	Haemoglobin
HDSFLKAYPSQKRT	3.76	1165	S100 Ca binding protein Calgizzarin
FGPKGFGRGGAESHTFK	3.19	1066	Cysteine rich protein
FVSNHAY	2.26	1168	Aldolase
NALAHKYH	2.66	1133	Haemoglobin
YQKVVAGVANALAHKYH	5.41	2738	Haemoglobin
HDSFLKAVPSQKRT	3.56	1145	S100 Ca binding protein Calgizzarin
KAKKPAAAAGAK	2.99	623	Histone H1-4
HKKSHEESHKE	2.82	1201	S100 Ca binding cystic fibrosis antigen
PPQYPVVPVHLDRII	2.60	690	Ribonuclease RNase A family

The peptide sequence HDSFLKAVPSQKRT which originates from the protein S100 Ca binding, Calgizzarin, and was found in both patient samples is shown in Figure 4.3.5.2. The parent ion is very strong with a total ion count in the 10^8 region and both the doubly charged and triply charged ions are present and were automatically identified for tandem mass spectrometry by the data dependent scanning function. The b and y assignment of the fragmentation spectrum is shown in Figure 4.3.5.2c and shows excellent correlation with the predicted ladder sequence ions.

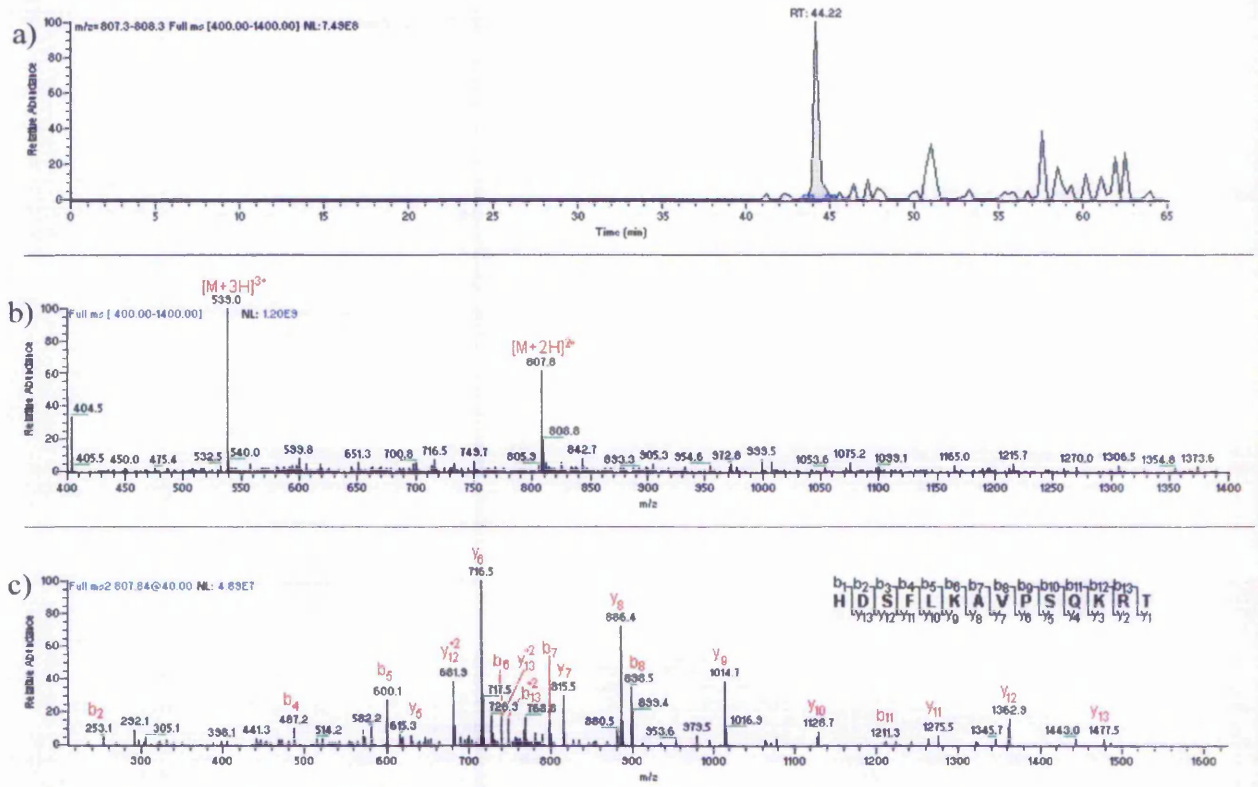


Figure 4.3.5.2 An example peptide identification by the SEQUEST algorithm from CML patient eluate following IMAC purification. Sequence HDSFLKAVPSQKRT from protein S100 Ca binding protein Calgizzarin. (a) Chromatogram of the doubly charged parent ion (m/z 807.8). (b) Full scan mass spectrum of the base peak showing the doubly charged, m/z 807.8 and triply charged, m/z 539.0 peptide ions. (c) Tandem mass spectrum of the doubly charged parent ion m/z 807.8 with assignment of b and y ions.

Extraction of intracellular peptides

The 301 and 286 patient cells were subjected to cell lysis and peptide extraction. Addition of acidified chloroform/methanol allowed the free intracellular peptides to be extracted into the methanolic fraction. The resulting LC/MS analysis is illustrated in Figure 4.3.5.3, which shows a base peak chromatogram for the methanolic extract of samples from both patient 301 and patient 286.

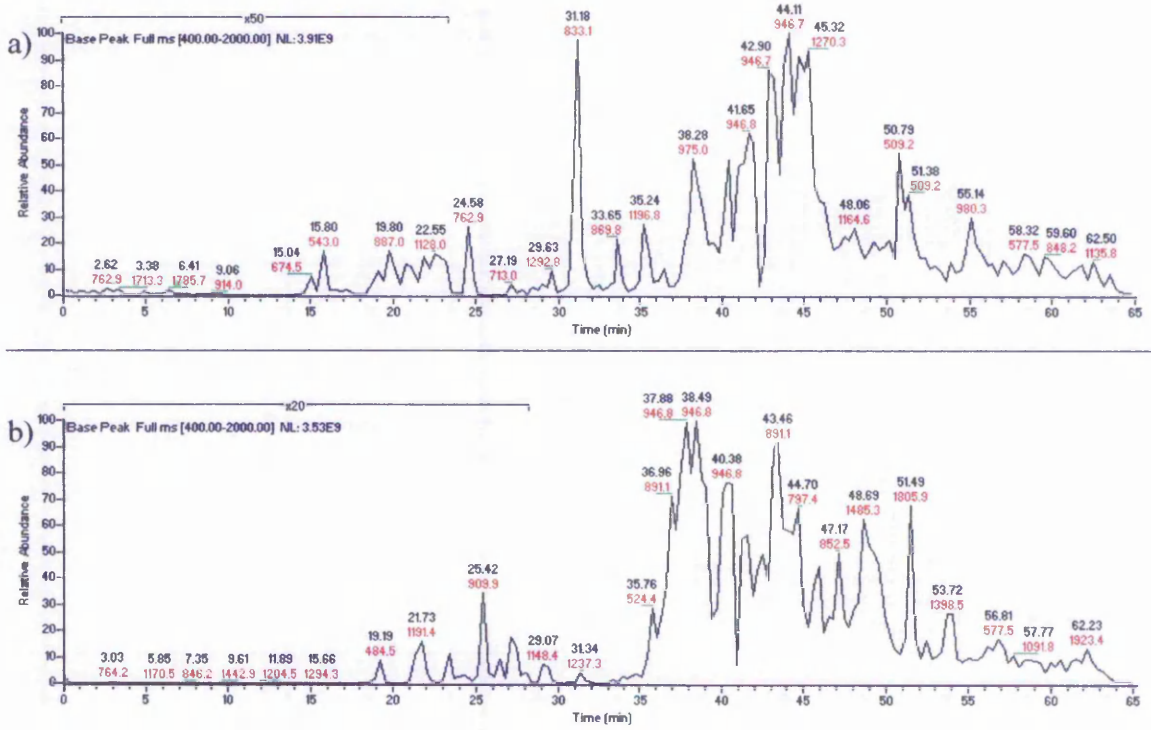


Figure 4.3.5.3 Example base peak chromatograms of IMAC purified methanolic fraction from CML patient cells. The base peak m/z is annotated in red below the retention time. Segments of the chromatogram have been expanded to visualise the baseline. (a) Patient 301. (b) Patient 286.

A selection of the peptides most clearly identified from the SEQUEST searches performed on the methanolic extracts of patient cells 301 and 286 can be seen in Table 4.3.5.2. The peptide sequence is shown with the histidine residues highlighted in red. The table also shows the Xcorr and Score assigned by the SEQUEST search to each match. All but one of the peptides identified were histidine-containing and represent a range of housekeeping proteins, again showing a high degree of selectivity for this residue.

Table 4.3.5.2 Peptide sequences identified by TurboSEQUEST search algorithm from analysis of IMAC purified CML patient cells from the methanolic fraction, patient 301 (blue) and patient 286 (yellow). TurboSEQUEST score and X correlation factor and parent protein are shown.

Sequence	Xcorr	Score	Protein
LEGQTQVAENPHSEY	4.36	1168	dpy-30 like protein
HDSFLKAVPSQKRT	3.57	991	S100 Ca binding protein Calgizzarin
DDEEAEEKAPVKK	3.73	1468	Nucleophosmin
KENKNEKVIEH	3.46	1570	S100 Ca binding protein Calgizzarin
RFPTDQLTPDQER	2.89	940	Myeloperoxidase
HEGDEGPGHHHKPGLGEG	2.93	1580	Calgranulin
VHSAGNIKKEK	2.00	448	Zinc finger protein 33

An example base peak chromatogram of the tryptic digest performed on protein extracted from the CML patient cells is shown in Figure 4.3.5.4. The overall total ion count is a massive 10^9 and some very strong sharp peaks can be observed within the chromatogram. These raw data files were then subjected to peptide identification using the SEQUEST searching algorithm. The strongest matches resulting from these searches can be seen in Table 4.3.5.3.

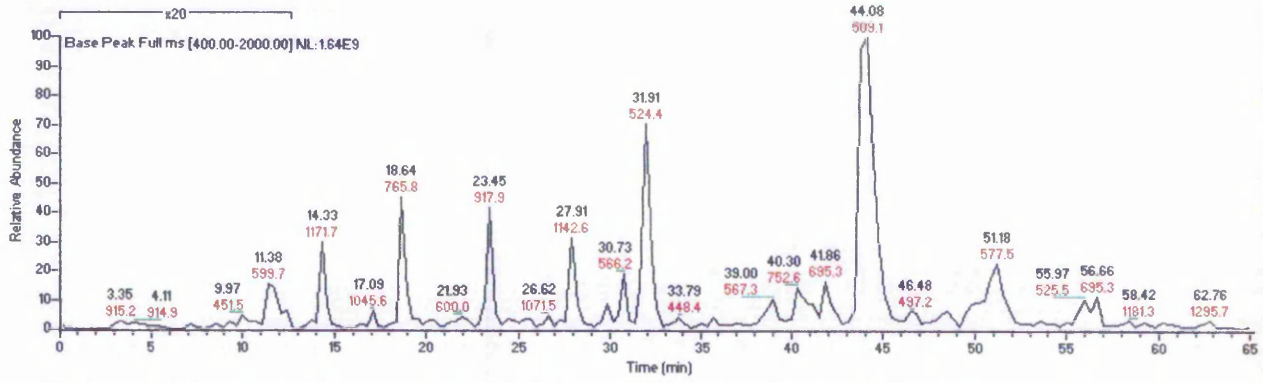


Figure 4.3.5.4 Example base peak chromatogram of IMAC purified tryptic digest of protein plug extracted from CML patient cells. The base peak m/z is annotated in red below the

retention time. Segments of the chromatogram have been expanded to visualise the baseline.

Table 4.3.5.3 Peptide sequences identified by TurboSEQUEST search algorithm from analysis of IMAC purified CML patient cellular protein digest for patient 301 TurboSEQUEST score and X correlation factor and parent protein are shown.

Sequence	Xcorr	Score	Protein
VLGAFSDGLAHLDNLK	5.04	233	Haemoglobin
LYPIANGNNQSPVDIK	4.95	1016	Carbonic anhydrase
LISWYDNEFGYSNR	4.59	1979	Glyceraldehyde 3-phosphate dehydrogenase
TPAVHASLDKFLASVSTVLTSK	4.57	1407	Haemoglobin
VTIAQGGVLPNIQAVLLPK	4.35	1160	H2A Histone
VAPEEHPVLLTEAPLNPK	4.35	1018	Beta actin
VVAGVANALAHKYH	4.34	1829	Haemoglobin
VGAHAGEYGAEALER	4.30	1768	Haemoglobin
NIETIINTFHQYSVK	3.95	876	S100 Ca binding proein S9
TYFPHFDLSHGSAQVK	3.85	965	Haemoglobin
ETNLDSLPLVDTHSKR	3.85	561	Vimentin
VLKQVHPDTGISSK	3.49	1084	H2B Histone
VVDLMAHMASK	3.43	1570	Glyceraldehyde 3 phosphate dehydrogenase
IVGGHEAQPHSRPY	3.36	963	Human proteinase 3
VLDFEHFLPMLQTVAK	3.29	645	Myosin
GGPGSAVSPYPTFNPSSDVVAALHK	3.08	431	Annexin
SETGAGKHVPR	3.01	954	Tubulin alpha 3
HAVSEGTKAVTKYTSSK	5.12	2553	H2B Histone
AVEHINKTIAPALVSK	4.14	1546	Anolase
HQGVMVGMGQKD	3.32	1671	Alpha actin
HAVSEGTK	2.71	956	H2B Histone
TADGIVSHLKK	2.33	691	Protein disulphide isomerase
LLPKKTESHHK	2.18	517	H2A Histone

AVEHINKTIAPALVSK	3.81	982	Enolase 1
AVGRLEAVSHTSDMHR	2.95	1751	Adenyl cyclase associated protein
AIRHPQYNQR	2.14	512	Cathepsin G

It is clear that the majority of peptide sequences that are identified in Table 4.3.5.3 contain one or more histidine residue which are shown highlighted in red. Again the number of non-histidine containing peptides is <10 % of the identified peptides. This represents high selectivity when considering the relative abundance of the histidine and/or cysteine residue is approximately 10-18 %.[14] Although there are still a number of haemoglobin originating peptides identified their presence is again much reduced. The peptides are derived from a range of proteins cytoskeletal proteins (vimentin, myosin, tubulin, actin), DNA binding proteins (histones), and calcium binding proteins (annexin and calgranulin).

4.4.0 Conclusions

The use of immobilised metal ion affinity chromatography (IMAC) has been investigated for the purification of histidine containing peptides. A selective method for the binding of histidine containing peptides to a Cu/IMAC column was developed which uses a HEPES buffer at pH 7.2 and allows selective binding of histidine containing model peptides with efficiency close to 100 %.

A non-selective IMAC binding method applicable to the purification of both histidine and non-histidine containing peptides was developed using buffers at pH 10/11. The peptide recoveries exhibited were somewhat lower with this non-selective method due to the weaker interaction between the N-terminus and the metal ion. Attempts were made to establish a two stage washing procedure with non-selective binding followed by elution of non-histidine and then histidine containing peptides. However reproducibility proved problematic, this was attributed in part to high column to column variability which combined with the weaker binding interaction resulted in poor robustness and this method was not pursued further. Perhaps these experiments would have been more successful if a different metal system was utilised, Hanson *et al* have shown that the effect of the α-amino group on peptide retention can be more pronounced when using a nickel system.[11]

Only a small number of peptides in a protein digest may be required for identification of the parent protein. A technique which reduces the number of

peptides analysed will result in more useful analysis. The histidine selective IMAC method was successfully applied to the purification of extremely complex cellular extracts. Both cell eluates and entire cell content from lysed cells, subjected to protein digestion, from patients with chronic myeloid leukaemia were purified using the IMAC column prior to analysis by nano-LC/μ-ESI/MS/MS. The SEQUEST algorithm was then used to identify the peptide sequences and their originating proteins. The method was shown to be 90 % selective when applied to these greatly complex samples and a large variety of housekeeping proteins were identified.

A Cu/IMAC method was successfully developed for the simplification of complex cellular extracts allowing a greater ease in the identification of proteins. The method which has been developed is selective for histidine containing peptides. It utilises a simple HEPES loading buffer, at physiological pH and elution is achieved in one stage by lowering the pH, resulting in a sample which is well suited for a further dimension of separation such as HPLC. It is robust, maintaining a high level of selectivity in the presence of complex cellular digests and cell surface eluates.

This method has proven a successful avenue for the purification and simplification of complex cellular samples. The assumption could be made that it would be useful in achieving the thesis main aim of the identification of a bcr/abl tumour antigen. However, the absence of a histidine residue in the breakpoint region of the bcr/abl protein determines that this is not the case. This does not

exclude the methods applicability to other such determinations but merely restricts the target sequences to those containing a histidine residue.

4.5.0 References

1. Porath J., *Nature,* **1975,** 258, 598-599.
2. Protein Purification 2nd edition Ed Janson J.C., Ryden L. 1998, Wiley, New York
3. Wang S., Zhang X., Regnier F.E., J. *Chrom. A.,* **2002,** 949, 153-162
4. Geng *J. Chrom B.*, **2001**, 752, 293
5. Belew H., Yip T., Anderson L., Porath J., *J. Chrom. A,* **1987,** 403, 197,
6. Gaberc-Porekar V., Menart V., *J. Biochem. Biophys. Methods,* **2001,** 49, 335-360.
7. Chaga G.S., *J. Biochem. Biophys. Methods,* **2001**, 49, 313-334.
8. Ueda E.K.M., Gout P.W., Morganti L., *J. Chrom. A,* **2003,** 988, 1-23.
9. Wong J.W., Albright R.L., Wang N-H.L., *Separation and Purification methods*, **1991**, 29, 49-106.
10. Andersson L, Sulkowski E., *J. Chrom.*, **1992**, 604, 13-17
11. Hansen P., Lindberg G., *J. Chrom. A,* **1995,** 690, 155-159.
12. Hochuli E., Dobeli H., Schacher A., *J. Chrom.*, **1987,** 411, 177-184
13. Hochuli E., Bannwarth W., Dobeli H., Gentz R., Stuber D., *Biotechnology,* **1988,** 1321-1325.
14. Hochuli E., Piesecki S., *Methods,* **1992**, 4, 68-72
15. Westra D.F., Welling G. W., Koedijk D. G. A. M., The T. H., Welling-Wester S., *J. Chrom. B,* **2001,** 760, 129-136;
16. Chaga G.S., Bochkariov D.E., Jokhadze G.G., Hopp J., Nelson P., *J.Chrom. A,* **1999,** 864, 247-256;
17. Wizemann H., von Brunn A., J. Virol. Methods, **1999**, 77, 189-197;
18. Casey J.L., Keep P.A., Chester K. A., Robson L., Hawkins R.E., Begent R.H. J., *J. Immunol. Methods,* **1995,** 179, 105-116.
19. Yip T., Nakagawa Y., Porath J., *Anal. Biochem.* **1989,** 183, 159.
20. Kronina V.V., Worth H-J., Hearn M..T.W., *J. Chrom.* **1999,** 852, 261-272.
21. Lindberg G., Bennich H., Engström Å., *Int. J. Peptide Protein Res.* **1991,** 38, 253.

22. Ren D., Penner N.A., Slentz B.E., Mirzaei H., Regnier F., *J. Proteome Res.*, **2003,** 2, 321-329.

23. Ren D., Penner N.A., Slentz B.E., Inerowicz H.D., Rybalko M., Regnier F.E., *J. Chrom. A.*, **2004,** 1031, 87-92.

24. Amini A., Chakraborty A., Regnier F. E., *J. Chrom. B*, **2002,** 772, 35-44.

25. Todorva-Balvay D., Pitiot O., Bourchim M., Srikrishnan T., Vijayalakshmi M., *J. Chrom. B.*, **2004**, 808, 57-62.

26. Kumar A., Bansal V., Andersson J., Roychoudhury P.K., Mattiason B., *J. Chrom. A,* **2006,** 1103, 35-42.

27. Moss J.A., Coyle A.R., Ahn J-M., Meijler M.M., Ofter J., Janda K. D., *J. Immunol. Methods*, **2003**, 281, 143-148.

28. Paunovic I., Shulin R., Nowack B., *J. Chrom. A*, **2005**, 1100, 176-184

29. Nieba L., Nieba-Axmann S.E., Persson A., Hamalainen M., Edebratt F., Hansson A., Lidholm J. Magnusson K., Frostell Karlsson A., Pluckthun A., *Anal. Biochem.*, **1997**, 252, 217-228.

30. Storkus W.J., Zeh H. J., Salter R. D., Lotze M. T., *J. Immunother.* **1993**, 14, 94-103.

31. Geng M., Ji J., Regnier F.E., *J. Chrom. A.*, **2000,** 870, 295-313

32. Wang M., Bronte V., Chen P. W., Gritz L., Panicali D., Rosenberg S. A., Restifo N. P., *J. Immunol.* **1995**, 154, 4685-4692.

33. Jackowski S., Rock C.O., *Arch. Biochem. Biophys.*, **1989**, 268, 516-524.

CHAPTER 5

CONCLUSIONS AND FURTHER WORK

Chapter 5: Conclusions and Further work

5.1.0 Chapter 2: Development of MS and capillary HPLC/MS for the analysis of peptides.

During the analysis of a series of MHC class I and II peptides, the effect of molecular charge state on fragmentation collision energy and the spectra generated was successfully studied.

It was observed that peptides composed of a greater number of amino acids or a higher charge state require higher collision energies. When identifying unknown peptides, the formation of a doubly charged parent ion is preferred as such ions yield a large coverage of b and y ions with a high intensity of fragment ions in the higher mass region.

The abundance of singly, doubly and triply charged parent ions was investigated with regard to the tube lens voltage (TLV) and capillary temperature parameters. To avoid signal to noise ratio and sensitivity loss problems, optimum conditions for the formation of the doubly charged parent ion were a low capillary temperature (180°C) and a highly negative TLV of -50V. It should be noted, however, that these conditions favour the high intensity formation of triply charged parent ions during the analysis of peptides containing three basic residues.

During the development of a microspray ionisation source, an increasing sensitivity and signal to noise ratio was observed as tip diameter was decreased

(100 μm to 25 μm). For a 5μm tip diameter microspray ionisation source, constructed in-house, incorporating a trapping column and loading pump, 1000 μl of very dilute samples could be analysed successfully at high sensitivity and signal : noise ratio and no loss of chromatographic resolution.

5.2.0 Chapter 3: Development of a method for the identification of MHC class I and II restricted peptides.

The effectiveness of the Storkus method was confirmed, employing a toxilight™ assay, when MHC class I bound peptides were eluted from the surface of cells whilst maintaining cellular integrity. The data obtained was comparable to that reported by Storkus as 76% of k562 cells remained viable after exposure to the method conditions. The A3 clean-up method, performed on CML patient eluates spiked with synthetic MHC class I and II peptides proved to be particularly poor (≈ 13%) as the conditions are specific to the A3 peptide. A more widely applicable clean-up method was developed, employing solid phase extraction and strong cation exchange chromatography, which was successfully used in the recovery of a series of peptides. The application of this method has revealed the following.

Although the recovery of the A2 predicted peptide was low (≈ 12%), the method was sufficiently sensitive to allow detection of immunologically significant levels of peptide per cell. Consequently this method was applied to the analysis of HLA.A2 positive CML patient eluates and k562 A2 transfected cell lines in search of several A2 predicted peptides. None of these peptides were present at

detectable levels in the eluates studied so it was concluded that they are not expressed on the cell surface at immunological levels. These peptides are therefore not suitable vaccine candidates.

The method was also applied to the clean-up and analysis of immunopurified DM deficient T2DR4 cells with the class II invariant chain peptide (CLIP) identified. A novel class II surface elution buffer was used to elute the same T2DR4 cell lines. The eluate was then subjected to the clean-up protocol and the same CLIP peptide was identified confirming the potential application of the MHC method to the identification of surface eluted MHC class II peptides.

5.3.0 Chapter 4: Immobilised metal ion affinity chromatography (IMAC) for the selective isolation of peptides.

Immobilised metal ion affinity chromatography (IMAC) was investigated as a process for the purification of histidine containing peptides. Two approaches to this purification were studied.

A selective method for binding histidine containing peptides to a copper/IMAC column using a HEPES buffer (pH 7.2) was developed, this allowed highly efficient and selective binding of such peptides. A non-selective IMAC binding method, incorporating a pH 10/11 buffer, was also developed that exhibited binding between the column and both histidine and non-histidine containing peptides. Peptide recoveries employing this method were considerably lower than that in the histidine selective method due to the weaker interactions between the N-terminus and the copper ion. A two stage wash procedure, to remove non-

histidine containing and then histidine containing peptides bound to the column during the non-selective binding process, was investigated but reproducibility problems resulted in this part of the study being abandoned.

The histidine selective IMAC method was successfully employed in the purification of complex cellular extracts. Cell eluates and entire cellular content, of cells from patients suffering from CML, were purified using this method prior to analysis by nano-LC/μ-ESI/MS/MS. The SEQUEST algorithm was then used in the identification of peptide sequences and their originating proteins.

5.4.0 Further Work

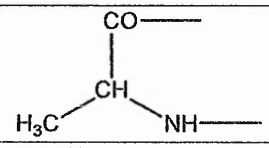
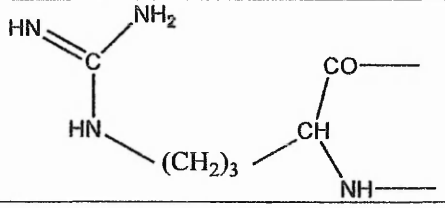
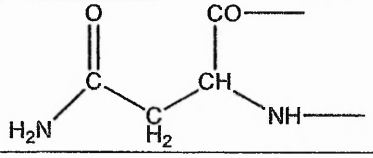
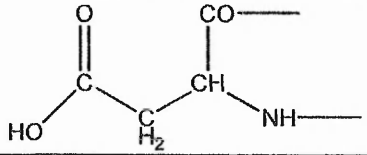
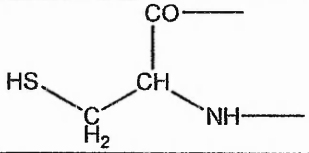
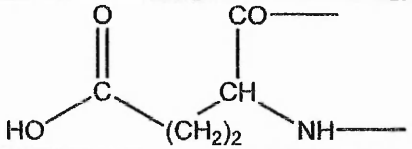
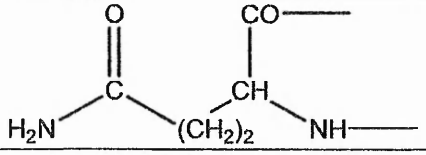
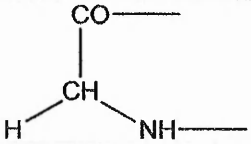
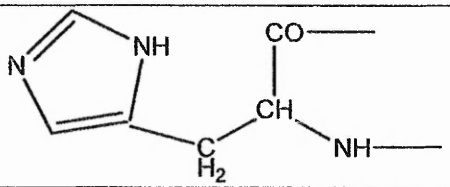
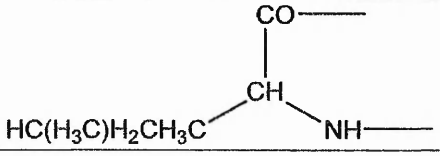
The MHC class I and II clean-up and analysis method has so far been applied to only a small number of CML patient samples and has for the most part been restricted to a single haplotype. Further application of this method to a wider range of CML patient and k562 transfected cells would prove useful in confirming the presence or absence of other MHC class I predicted peptide antigens for a variety of haplotypes. Furthermore, the application to eluates obtained from other types of malignancies may prove useful in identifying antigenic peptides.

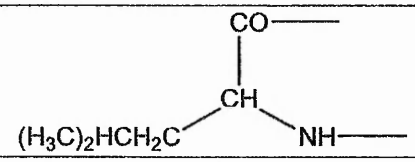
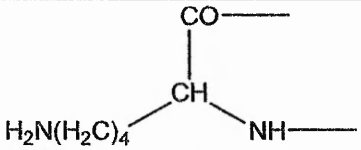
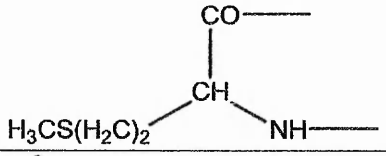
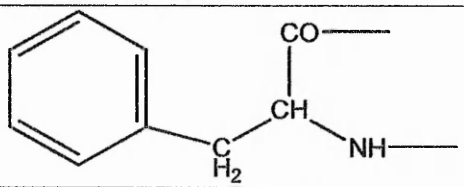
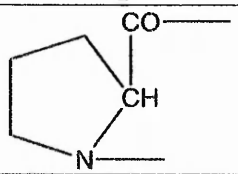
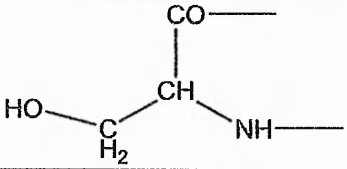
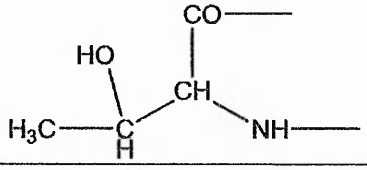
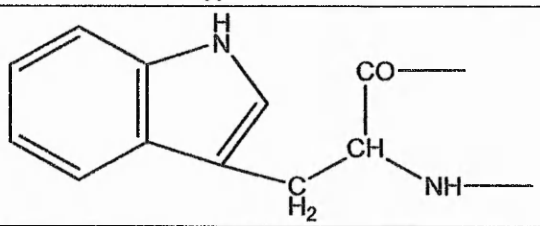
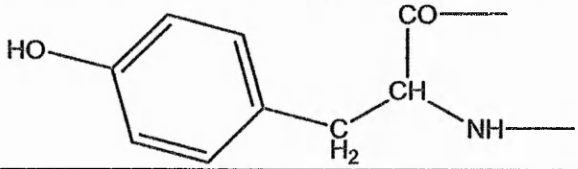
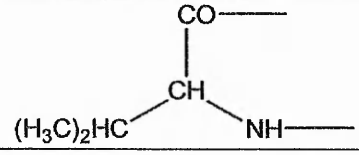
Application to the analysis of class II surface eluates may also be performed. This may prove a more challenging proposition than with class I due to the uncertain length of the class II peptides and hence the uncertainties associated with predicting a target. However the nature of the analytical method has been

designed and tested to incorporate searching for both targeted sequences and unknown sequences.

The immobilised metal affinity chromatography method which was developed in parallel with the MHC class I and II clean-up method also has potential for more widespread use.

Appendix

Name (Abbrev.)	Symbol	Residue Structure
Alanine (Ala)	A	
Arginine (Asn)	R	
Asparagine (Asn)	N	
Aspartic Acid (Asp)	D	
Cysteine (Cys)	C	
Glutamic Acid (Glu)	E	
Glutamine (Gln)	Q	
Glycine (Gly)	G	
Histidine (His)	H	
Isoleucine (Ile)	I	

Leucine (Leu)	L	
Lysine (Lys)	K	
Methionine (Met)	M	
Phenylalanine (Phe)	F	
Proline (Pro)	P	
Serine (Ser)	S	
Threonine (Thr)	T	
Tryptophan (Trp)	W	
Tyrosine (Tyr)	Y	
Valine (Val)	V	

Peptide abbreviation	Sequence	Rationale
A3 peptide	KQSSKALQR	Originates from the breakpoint region of the bcr/abl fusion protein. Binds to HLA-A3 molecules
A2 peptide	SSKALQRPV	Originates from the breakpoint region of the bcr/abl fusion protein. Predicted to bind to HLA-A2 molecules
Hep B peptide	TPPAYRPPNAPIL	Originates from the hepatitis B protein
β-gal peptide	TPHPARIGL	Originates from the β-galactosidase protein
$p53_{(65-73)}$ peptide	RMPEAAPPV	Originates from the p53 protein. (65-73) indicates region of the protein sequence.
$p53_{(264-272)}$ peptide	LLGRNSFEV	Originates from the p53 protein (264-272) indicates region of the protein sequence.