Capillary Electrophoresis with Multiple Readout Techniques for Protein Analysis

Miriam Sarah Goyder

Department of Chemistry, Imperial College London

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy of Imperial College London and the Diploma of Imperial College

September 2011
Declaration

I hereby certify that the material presented in this thesis, which I now submit for the award of Doctor of Philosophy of Imperial College London, is my own work unless otherwise cited or acknowledged within the body of the text.

Miriam Sarah Goyder

September 2011
Abstract

In the era of proteomics, new technologies in separation and identification are required. Separation methods, such as capillary electrophoresis or liquid chromatography, are a crucial part of high throughput proteomic workflows. In this thesis, novel approaches to proteomics using capillary electrophoresis are presented.

A platform of technologies based on capillary electrophoresis with continuous deposition of separated proteins onto metallic substrates enables subsequent analyses and identification. Since sample deposition and identification are decoupled, multiple readout techniques can be explored. Readout techniques used include matrix assisted laser desorption/ionisation mass spectrometry (MALDI-MS), electron-vibration-vibration two dimensional infrared spectroscopy (EVV 2DIR) and fluorescence microscopy. This technology was used without the deposition interface, to achieve advances in ribosomal separations or with the deposition interface, to develop new proteomic strategies of separation and readout.

The eukaryotic ribosomal proteins were separated using capillary electrophoresis for the first time. Over 26 peaks were resolved in less than 10 minutes. An outstanding RSD migration time of < 0.5% was achieved, demonstrating that the readout could provide a ribosomal ‘fingerprint’.

Separations of proteins were successfully analysed using a standard MALDI-MS instrument. This work was advanced by the offline coupling of CE to MALDI-imaging and applied to the ribosomal proteins to demonstrate a novel workflow from cell culture to protein identification.

Quantitative analysis of protein levels is an important part of proteomics, but is difficult to achieve using mainstream technologies with high throughput and accuracy. EVV 2DIR is a non-linear spectroscopy which is able to achieve absolute quantification of proteins.[1] Coupling of EVV 2DIR to CE (CE-2DIR) was demonstrated through the deposition and analysis of peptide and proteins. CE-2DIR offers great promise as a new proteomic tool.
I would like to dedicate this thesis to Dr Judit Nagy (*in memorium*). It was a pleasure and an honour to work with her and she has inspired me.

**Acknowledgements**

First of all, I would like to thank my supervisors, Andrew deMello, Oscar Ces and David Klug - Andrew, for giving me a sound grounding in capillary electrophoresis and for his wisdom, Oscar for his continuing encouragement all along the way, and David for being full of great ideas and for helping me to develop and learn to discuss as a scientist. I am extremely grateful to Keith Willison for the supervision he has given me on the biological aspect of this work. I have been amazed by his wealth of knowledge, and his excitement in protein complexes has inspired me in my work with the ribosome.

A big thank you goes to Liz McCormack and Heather Rada at The Institute of Cancer Research (ICR) for the invaluable expertise that they shared with me during my ribosome purifications, and for making ICR such a nice place to come and work at. I would also like to thank Isaac Garcia-Murillas at the ICR for the initial assistance he gave me on working with yeast.

I would like to thank Paul Hitchen at the Mass Spectrometry Core Facility in Centre for Integrative Systems Biology for performing LC-MS/MS analyses of ribosome samples, and for his patience with my numerous requests!

I am very thankful to Judit Nagy as it was through her guidance and encouragement that I pursued the CE-MALDI work. I thank her for the kindness she showed me, for the opportunities she opened to me and for building my confidence. I am very grateful to Ian Edwards, Andrew Rogers and Emmanuelle Claude at Waters Corporation, Manchester for working on the CE-MALDI-imaging with Judit and me. They were a pleasure to work with and I am grateful for the kindness they showed me on my visit and the insight they gave me into state of the art mass spectrometry.

There are many people whom I have been working alongside and would
like to thank. In particular, I would like to thank Elizabeth Gardner with whom I collaborated on the CE-2DIR work, but who has also been a good friend and source of support, and Chris Barnett, who’s knowledge knows no bounds and has been an invaluable collaborator and source of wisdom throughout this work. Thank you Rafal for your support as my CE colleague. Thank you to John for the help he gave me with his microscope and for always organising the ‘lunch-crew’. And to the tea-makers, Kate and Tom for doing such a magnificent job. A more general thank you goes to all the members of the Klug group that I’ve not mentioned here but appreciate the contribution that they’ve made to making a pleasant working environment.

Most of all, I would like to thank my family and friends who have cared for me and supported me throughout my doctoral studies. In particular, I give my heartfelt thanks to my parents and to Sarah, Chris, Martha and Anna - I could never have written my thesis without you.
# Contents

1 Introduction .................................................. 11
   1.1 Proteomics ................................................. 12
      1.1.1 Proteomic Technologies .......................... 13
   1.2 Capillary Electrophoresis ............................... 20
      1.2.1 Background ........................................ 20
      1.2.2 Principles of Capillary Electrophoresis .......... 22
         1.2.2.1 Overview ..................................... 22
         1.2.2.2 CE System ..................................... 23
         1.2.2.3 Mobility ....................................... 24
         1.2.2.4 Electroosmotic Flow ......................... 26
         1.2.2.5 Separation Parameters: Efficiency and Resolu-
                     tion ....................................... 30
      1.2.3 Instrumental and Operational Aspects ............. 34
         1.2.3.1 Bandbroadening ............................... 34
         1.2.3.2 Sample Injection ............................... 37
         1.2.3.3 Detection ..................................... 40
   1.3 The Ribosome .............................................. 43
   1.4 Thesis Outline ............................................. 48

2 Ribosome Affinity Purification ............................... 50
   2.1 Introduction .............................................. 51
   2.2 Methods: Ribosome Affinity Purification (RAP) ....... 54
CONTENTS

2.2.1 Yeast Strains ........................................... 55
2.2.2 Lysate Preparation ..................................... 56
2.2.3 Affinity Purification ................................... 57
2.2.4 Assessing the purification ............................... 58

2.3 Results of RAP .............................................. 63
2.4 Conclusions .................................................. 68
2.5 Appendix A .................................................. 69

3 CE Instrumentation ............................................ 73
3.1 Introduction .................................................. 74
3.2 Capillary Electrophoresis Instrument ........................ 74
3.2.1 Overview of CE instrument ............................... 74
3.2.2 Pressure System .......................................... 75
3.2.3 Instrument Control and Data Acquisition ............... 75
3.2.4 General Experimental Methods used in CE Separations 78
3.3 Validation of the Instrument ................................. 80
3.4 Conclusions .................................................. 86

4 CE of Ribosomal Proteins ..................................... 89
4.1 Introduction .................................................. 90
4.2 Materials and Methods ..................................... 92
4.2.1 Method Development for Ribosomal Protein Separations: Separation of a 6 Protein Mixture 92
4.2.2 Separation of Ribosomal Proteins ........................ 93
4.2.3 Increase in Ribosome Sample Loading ................... 94
4.3 Method Development for Ribosomal Protein Separations: Separation of a 6 Protein Mixture 95
4.4 Separation of Ribosomal Proteins ........................... 97
4.5 Increase of Sample Injection Quantities for Readout Methods 104
4.6 Conclusions .................................................. 110
4.7 Appendix A .................................................. 112
5 Track-Writing

5.1 Introduction ..................................................... 116
5.2 Experimental Apparatus ........................................... 118
  5.2.1 Overview of Interface ......................................... 118
  5.2.2 Capillaries ..................................................... 120
  5.2.3 Track-Reader .................................................. 123
5.3 Method Development .............................................. 123
  5.3.1 Calculation of Stage Velocity .................................. 125
  5.3.2 Improved Pressure Control ..................................... 131
  5.3.3 Effect of Background Electrolyte .............................. 132
  5.3.4 Capillary Tapering ............................................. 135
5.4 Track-writing of protein samples ................................. 137
  5.4.1 Separation of 3 protein mixture .............................. 137
  5.4.2 Track-writing of ribosomal proteins ......................... 139
5.5 Conclusions ....................................................... 141

6 CE-MALDI  ..................................................................... 144

6.1 Introduction .......................................................... 145
  6.1.1 Matrix Assisted Laser Desorption/Ionisation (MALDI)-
       MS ................................................................. 145
  6.1.2 CE-MALDI Interfaces ............................................ 147
6.2 CE-MALDI-TOF/MS ................................................... 149
  6.2.1 Experimental Procedure and Method Development for
       CE-MALDI .......................................................... 151
  6.2.2 Results and Discussion .......................................... 155
    6.2.2.1 Proof of principle for CE-MALDI interface .......... 155
    6.2.2.2 Development of a matrix deposition method .......... 155
    6.2.2.3 Detection of separated proteins in tracks using
           MALDI-TOF/MS ............................................. 159
  6.2.3 Conclusions ..................................................... 161
6.3 CE-MALDI Imaging .................................................... 163
CONTENTS

6.3.1 Introduction ................................................. 163
6.3.2 Materials and Methods ..................................... 165
6.3.3 Results and Discussion ..................................... 167
6.3.4 Conclusions .................................................. 170
6.4 CE-MALDI Applications for Ribosomal Proteins ............... 171
   6.4.1 Development of an On-Plate Tryptic Digest Protocol . 172
      6.4.1.1 Materials and Methods .......................... 172
      6.4.1.2 Results and Discussion .......................... 173
   6.4.2 CE-MALDI Imaging of Ribosome Tracks .................. 175
      6.4.2.1 Materials and Methods .......................... 175
      6.4.2.2 Results and Discussion .......................... 176
   6.4.3 Conclusions .............................................. 178
6.5 Summary .......................................................... 178

7 CE-2DIR .......................................................... 180
   7.1 Introduction .................................................. 181
   7.2 EVV 2DIR Spectroscopy ...................................... 182
      7.2.1 Protein Identification via EVV 2DIR Spectroscopy . 185
      7.2.2 The Optical Setup .................................. 187
      7.2.3 State of Development of EVV 2DIR .................. 187
   7.3 Experimental Procedure for CE-2DIR ....................... 188
   7.4 EVV 2DIR Analysis of Protein Tracks ...................... 190
   7.5 EVV 2DIR Analysis of Peptide Tracks ...................... 194
      7.5.1 Method Development ................................. 194
      7.5.2 Analysis of Tracks with EVV 2DIR .................. 200
      7.5.3 CE-2DIR: Proof of Principle ......................... 215
   7.6 Conclusion ................................................... 215

8 Conclusions ...................................................... 219
   8.1 Separation of Ribosomal Proteins ......................... 220
   8.2 CE-MALDI .................................................... 221
CONTENTS

8.3 CE-2DIR ...................................................... 221
8.4 Future Work ............................................... 222
Chapter 1

Introduction
CHAPTER 1. INTRODUCTION

1.1 Proteomics

Life requires thousands of different molecules. Proteins are one of the most complex molecules involved and they participate in almost every process within the cell. They have a complex three dimensional structure and can form complexes with other proteins or other molecules to make the machinery of the cell. They are enzymes, structural components and play roles in cell signalling - for example.

In the post-genomic era, with recognition that to truly understand the biology of a cell it is necessary to study the molecules that the genetic code is encoding, proteomics has become the next big challenge. The term ‘proteome’ was introduced in 1995 as ‘the total protein complement of a genome’ [2] and thus, proteomics has emerged as the study of the proteins expressed by a genome. The complexity of proteomics is enormous. Unlike the genome, the proteome of a cell is not constant in time. Instead, it is dependent on gene expression, post-translational modifications (PTMs) and interactions which are affected by the microenvironment in which the cell resides and which vary throughout the life of a cell or organism. This diversity is clear from the numerous cell types that exist within the human body, each performing widely different functions despite containing the same genome. Furthermore the dynamic range in relative protein abundances is vast; within the eukaryotic cell, this is typically seven or eight orders of magnitude.[3, 4]

These complexities make proteomics biologically interesting and also a powerful platform for medical advancements, due to the clear relationship between protein dynamics and disease. Proteomics has the potential to revolutionise biomarker discovery and identify new drugs targets. Furthermore, studying the molecular mechanisms of disease on the protein level, for example by investigating protein signalling pathways and quantifying protein levels, can answer fundamental questions in cancer research.[5]
CHAPTER 1. INTRODUCTION

1.1.1 Proteomic Technologies

Proteins have been studied for over one hundred years, but definition of the field of proteomics marked a new era requiring high throughput tools. A range of complementary analysis tools are demanded by the variety of biological problems. Thus, a variety of technologies exist for the separation, identification and characterisation of proteins. The complexity of the proteome poses a great analytical challenge, and separation methods such as electrophoresis and chromatography are a crucial component of sample preparation and identification. In particular, powerful separation methods are required firstly to detect proteins that would otherwise be masked by higher abundance proteins and secondly for the unambiguous identification of proteins of which there exist multiple forms. Thus, proteomic approaches commonly feature a separation technique followed by an identification technique.

Mass spectrometry (MS) is the method of choice for protein identification. Protein identification with MS become possible because of genomic sequence databases and because of the invention of the soft ionisation methods - in particular, matrix assisted laser desorption/ionisation (MALDI) and electrospray ionisation (ESI). The most commonly used forms of MS in proteomics are defined by the methods of ionisation as either MALDI-MS or ESI-MS.[6] These methods are described in Chapter 6. Proteins are usually identified following proteolytic digest, most commonly with trypsin. Thereafter, identification can be achieved using tandem MS (MS/MS) or, in small data sets, peptide mass fingerprinting (PMF).[7] These schemes are explained in further detail in Figure 1.1.1.

Traditionally, proteomics has been vastly reliant on two dimensional gel electrophoresis (2DE) with identification of proteins performed with MS (see Figure 1.1.2). Proteins are separated according to isoelectric point (pI) using isoelectric focussing in the first dimension and according to molecular weight using sodium dodecyl sulphate polyacrylamide gel electrophoresis
Figure 1.1.1 – Tryptic digest cleaves proteins on the carboxy terminal side of the arginine and lysine residues producing tryptic peptides which have a basic residue at the carboxy terminus and are in the preferred mass range for sequencing. [8] Peptides are then identified by a) peptide mass fingerprinting (PMF) or b) tandem mass spectrometry (MS/MS). To perform PMF, obtained masses of enzymatically or chemically cleaved peptides are compared to mass fingerprints with theoretical cleavage for proteins sequences stored in a database, and a cut off score for similarity set for selecting candidate proteins. MALDI is the most commonly used ionisation technique for PMF. [7] For identification using tandem MS (MS/MS), masses of tryptic peptides are measured in the first mass analyser. Peptides qualifying specified conditions such as mass range and signal to noise ratio (‘precursor ion’) are then selected for MS/MS, where they are fragmented by collision with uncharged gas atoms (‘product ions’). [9] Identification of the peptides takes place as a process of database matching between theoretically generated product ion spectra and experimentally generated spectra. Algorithms, such as Mascot [10] and SEQUEST [11] are used to search within sequence databases, such as those of UniProt, to enable identification of a protein from only a small proportion of its peptides. A requirement of sequence database searching is that the organism in question must have had its genome sequenced so that all possible peptides are known.[8]
(SDS-PAGE) in the second. It has become a popular and commercialised approach capable of resolving tens of thousands of proteins and identifying hundreds of these.[12] However, even with semi-automation of some aspects, the method is low throughput and laborious: a typical protocol will entail separation of proteins using 2DE (2 days), excising selected bands, trypsin digestion (hours-overnight), further purification and/or concentration of the sample (hours), mass spectrometry using either electrospray ionisation (ESI) or matrix assisted laser desorption/ionisation (MALDI). Crucially, it is not an ideal tool for many proteomic studies. Low abundance proteins are frequently lost, and some classes of proteins, such as the very basic, very acidic, the large, the small and membrane proteins are excluded from analysis.

Today, 2DE is still hugely popular and useful, but an understanding of its weaknesses along with the challenge of a global proteomic analysis has led to a drive towards the development of novel, non-gel based, high throughput technologies. Many focus on high performance liquid phase separation techniques as alternatives to gel electrophoresis, which are coupled to MS. These most commonly use forms of liquid chromatography such as high performance liquid chromatography (HPLC) [13], reversed-phase liquid chromatography (RPLC) [14], multidimensional separations which integrate RPLC with another technique [15] and, less commonly, capillary electrophoresis (CE) [16]. In a popular alternative to 2DE approach known as ‘shotgun’ proteomics, a protein sample is enzymatically digested into its constituent peptides, these are then separated by one dimensional or multidimensional liquid chromatography, before identification of the peptides using tandem MS and database searching to determine the protein content of the original sample. [17] This is depicted in Figure 1.1.3.

Shotgun proteomics is a ‘bottom up’ proteomic approach. Proteomic approaches are frequently discussed in two groups: bottom-up and top-down. Bottom-up proteomics describes analysis of the peptides generated by enzymatic cleavage of a protein; top-down proteomics describes analysis of the
Figure 1.1.2 – 2DE-MS proteomic analysis. A protein sample is separated according to isoelectric point and molecular weight using 2DE, spots of interest are excised, trypsin digested and identified with MS. Typically MALDI-MS is used with PMF.
Figure 1.1.3 – Shotgun proteomics LC-MS/MS workflow. A complex protein sample is trypsin digested, peptides are separated using one or more dimensions of LC and subjected to MS/MS analysis. Following data analysis and scoring, a ranked list of protein components of the original sample is compiled.
intact proteins. For MS based proteomics, this categorisation refers to the state of the proteins when they enter the mass spectrometer. [18] Bottom-up approaches are the most mature and popular for protein identification. Fully automated instrumentation and bioinformatics tools are widely available commercially. Using multidimensional LC-MS/MS, thousands of proteins from whole cell lysates can be identified. [15] However, there are limitations to bottom-up proteomics approaches. Firstly, mis-identification of proteins can occur as a result of mis-cleavages, which arise during digestion. A disadvantage of the shotgun proteomic method is that digestion of proteins to peptides occurs at the beginning of the procedure, resulting in a loss of information as to the origins of a peptide. This is in contrast to the 2DE approach where a group of peptides can be correlated with a particular protein band and the additional protein molecular weight and isoelectric point information is therefore accessible. Most importantly, as proteins are only identified from a portion of their sequence, information on protein isoforms resulting from, for example, post-translational modifications is frequently lost yet many play crucial regulatory roles in the cell.

There is growing interest in top-down proteomics approaches, which can give additional information to primary structure and enable, for example, the characterisation of PTMs. As yet, it is a much younger field, and MS instrumentation to study intact proteins requires a high resolution and high mass accuracy which is achieved using an instrument such as a Fourier Transform Ion Cyclotron Resonance (FT-ICR) or hybrid ion trap FT-ICR - instruments which are expensive to purchase and to operate.[18] In terms of separation, the field is less developed. Intact protein separations are in general more challenging and the gel plays a bigger role in top-down proteomics than liquid phase separation techniques such as HPLC. The efficiency and resolution of LC are often reduced as a result of unwanted interactions of proteins with the stationary phase or denaturation of proteins in the organic mobile phase.[19]

It is apparent that the development of high throughput separation tech-
niques has not matched the demand, which is demonstrated by the fact that gel electrophoresis is still essential to so much of proteomics. Improved high throughput methods for intact protein separations are required - even if later followed by a digest as in 2DE-MS, thus forming a combined top-down/bottom-up proteomic platform. Capillary electrophoresis (CE) is a liquid phase separation technique that was developed with the growth of high throughput and miniaturised separation techniques. However, although systems are available commercially, it has not reached its full potential in proteomics. It offers many of the benefits of HPLC but can be performed under aqueous conditions and without a stationary phase. As conditions can be chosen to be non-denaturing, CE can be applied to study non-covalent protein complexes, protein-protein interactions and drug-protein interactions. CE has many other advantages, such as short analysis times, small sample quantities and high efficiencies, which are described in detail in Section 1.2.

Furthermore, in this age of MS-based proteomics it is essential that development of alternative technologies continues: although it has transformed proteomics, MS is not perfect. It is clear that the levels of proteins are of crucial importance in proteomics, yet absolute quantification of proteins is difficult with MS approaches. MS does not easily lend itself to quantitative analyses as variations in physicochemical properties, such as size, charge and hydrophobicity, lead to wide differences in mass spectrometric response. To achieve quantification with mass spectrometry, typically stable isotope labelling is required. Label free methods are emerging, but at present these are less accurate. These methods are difficult, time consuming and are furthermore not suitable for clinical samples.[20] Quantitative techniques by 2DE using fluorophores, dyes or radioactivity do not provide identification themselves. There is therefore a strong case for the development of novel protein identification tools to complement MS. Spectroscopic techniques are promising for the ability to achieve quantification due to the usually simple relationship between signal intensity and the quantity of sample, and
at present do not feature much in the mainstream tools used for proteomic analysis.

This thesis aims to use capillary electrophoresis as a separation strategy to deliver samples to a substrate so they can be analysed by a range of detection methods. The main aim, which has guided the direction of this work, has been analysis with a novel spectroscopic technique, electron-vibration-vibration two-dimensional infrared spectroscopy (EVV 2DIR) but the format of the interface enables additional or alternative readout techniques to be used. Given the complexity of the proteome, flexibility to choose a readout technique and the ability to gather information from a variety of sources is of crucial importance.

1.2 Capillary Electrophoresis

1.2.1 Background

Electrophoresis can be defined as the movement of charged particles through a conducting medium under the influence of an electric field. Its use as a separation technique is facilitated by the differential migration of molecular species based on their mobilities, which in many circumstances is determined by charge-to-size ratio. Today, electrophoresis can be performed in several different formats, but all require a physical support, whether it is the apparatus for a gel or a thin capillary or a microchip, a carrier electrolyte, electrodes, a power supply and a method of detection.

Traditionally, electrophoresis has been performed in a slab gel format. The gel is typically made of polyacrylamide or agarose, with dimensions of approximately 5-25 cm × 5-25 cm × 1-2 cm and is immersed in an electrolyte. One purpose of the gel is to suppress convective flow, which results because the passage of the electric current generates heat, known as Joule heating. As heat is removed at the edges of the medium, temperature gradients develop, which lead to convection.
viscosity in warmer regions results in faster analyte migration. These processes contribute significantly to band broadening, which is the term given to describe the dispersion of zones. As a result, the performance of slab gel electrophoresis is limited as the voltages that can be applied without significant heating are low, leading to long analysis times.

Several other characteristics of slab gel electrophoresis, together with developments in miniaturisation technology in chromatography, led to the demand for a more automated technique with a higher degree of reproducibility. Performing slab gel electrophoresis is labour intensive and time consuming, and the technique is difficult to automate. Detection typically requires staining and visual observation of the bands, followed by scanning or photographing to obtain a permanent record. Quantification is difficult, but possible with image analysis. The sensitivity of slab gel electrophoresis is poor, and as a result large sample volumes in the region of microlitres are required and low abundance proteins cannot be detected. The most frequently used stain, Coomassie blue, has a detection limit of approximately 10 ng of protein per band [24], whilst the more sensitive silver staining can detect down to approximately 0.5 ng of protein per band. [24]

The development of capillary electrophoresis, where electrophoresis is performed in extremely narrow, long tubes, solved many of the problems of slab gels. Narrow capillaries have a low conductance, limiting the generation of Joule heating and, furthermore, effective heat dissipation can take place due to the high surface to volume ratio. As a result, high voltages (up to 30kV) can be applied, resulting in exceptionally high efficiencies and short analysis times. This concept is explained further in Section 1.2.2.5. Furthermore, since the walls of the capillary offer support for the electrolyte and no anticonvective media is necessary, no gel is required. Automation is achieved. Detection is online and on-capillary, with output in the form of an electropherogram, from which data analysis and quantification are simplified. Sample introduction (known as injection) is repeatable, and injection
volumes are of the order of nanolitres.[25]

Early research on free solution capillary electrophoresis was performed by Hjertén in 1967 [26], but only 3 mm bore capillaries were available, which were rotated to reduce the effects of convection. Free solution electrophoresis was later performed with 200 \( \mu \)m i.d. capillaries.[27] In 1981, Jorgenson and Lukacs [28] published work using 75 \( \mu \)m i.d. glass capillaries, which defined the technique of capillary electrophoresis today. Capillaries were 100 cm in length and 30 kV were applied. Crucially, they solved difficulties surrounding sample injection and detection, injecting the sample by applying the voltage for a short period of time (see Section 1.2.3.2) and using ‘on-column’ fluorescence detection to record an electropherogram.

1.2.2 Principles of Capillary Electrophoresis

1.2.2.1 Overview

Today, capillary electrophoresis is typically performed in fused silica capillaries, usually 25 – 100 \( \mu \)m in diameter [25]. The instrumentation in CE is relatively simple. Capillary electrophoresis is performed in a number of different formats, each with a different mechanism of separation and therefore a suitability to a certain sample type. The most common formats are capillary zone electrophoresis (separation based on size to charge ratio), capillary isoelectric focussing (separation based on isoelectric point), capillary gel electrophoresis (separation based on molecular weight), capillary isotachophoresis (separation based on size to charge ratio) and micellar electrokinetic chromatography (separation based on hydrophobicity). In this thesis, capillary zone electrophoresis is the format used.

Capillary Zone Electrophoresis (CZE), also known as free-solution capillary electrophoresis, is CE performed with a homogeneous BGE. This is in contrast to capillary isoelectric focussing and capillary isotachophoresis. There is no gel, unlike in capillary gel electrophoresis. In CZE, sample is
injected at the inlet, an electric field is applied and solutes are separated into
discrete bands according to their individual mobilities, which are defined by
their size to charge ratios. This is explained in further detail in Section
1.2.2.3. CZE can be used for the separation of both anions and cations.

CZE can be applied to a many samples types, with separations performed
ranging from small molecules to viruses. It is well suited to the analysis of
peptide and protein samples, as the charge can be carefully manipulated by
small adjustments of the pH of the buffer solution. Protein adsorption to
the capillary surface can lead to well documented problems with reduced re-
producibility, poor peak shape and lost peaks. However, considerable de-
velopments in capillary electrophoresis technology (discussed in Section 1.2.3)
enable these effects to be minimised. CZE was chosen for this work for
these reasons and because flexibility in choice of buffer with the ability to
use aqueous buffers makes it well suited to offline interfacing with a range
of identification methods. In particular, for EVV 2DIR analysis, the pres-
ence of surfactants or organic solvents would interfere with the identification
process.

1.2.2.2 CE System

A schematic representation of a generic CE system is shown in Figure 1.2.1.
The capillary is filled with an electrolyte and the ends placed in reservoirs
containing an electrolyte. In capillary zone electrophoresis, the electrolyte is
an aqueous buffer, known as the background electrolyte (BGE). For sample
introduction into the capillary, known as sample injection, the sample vial is
briefly exchanged with the buffer vial and injected either hydrodynamically
or with the application of a voltage, and then exchanged back for the pur-
poses of the separation. Within the buffer reservoirs, electrodes are placed.
These electrodes are connected to a high voltage power supply which typ-
ically applies voltages up to 30 kV. [29] Often a reversible power supply is
used. The term ‘normal polarity’ describes the situation where the cathode
is at the outlet and the anode is at the inlet. Near the outlet reservoir, a
detector provides on-capillary detection, and the information is then sent to
a data handling device, such as a computer.

For capillary electrophoresis, information from the detector takes the form
of an electropherogram (Figure 1.2.2). An electropherogram displays the de-
tector response as a function of time. Peaks appear according to the migra-
tion times of separated analytes. The migration time is the time taken for
an analyte to travel from the sample end of the capillary to the detection
window.

1.2.2.3 Mobility

To perform an electrophoretic separation, an electric field is applied across
a separation medium, which causes molecules to accelerate to either the
anode or cathode under the Lorentz force. The medium through which the
molecules are moving presents an opposing frictional force proportional to
the velocity at which the molecules are moving. As a result, the molecules
reach a terminal velocity, known as the electrophoretic velocity.

The Lorentz force, \( F \), is given by
Figure 1.2.2 – Drawing of an electropherogram, a plot of signal response as a function of time. The drawing represents a separation of three components, a, b and c with the migration times are $t_a$, $t_b$ and $t_c$ respectively, measured at the peak maximum.

\[ F = qE \]  
\[ f = 6\pi \eta rv \]

where $q$ is the net charge and $E$ is the electric field strength. The frictional force, $f$, is expressed by Stokes’ Law

where $\eta$ is the viscosity, $r$ is the ionic radius and $v$ is the velocity of the ion.

Therefore the electrophoretic velocity, $v_{ep}$, is defined as

\[ v_{ep} = \frac{qE}{6\pi \eta r} \]

The mobility of a molecule, $\mu_{ep}$ is the electrophoretic velocity per unit...
field strength. Hence

\[ \mu_{ep} = \frac{v_{ep}}{E} = \frac{q}{6\pi \eta r} \]  \hspace{1cm} (1.2.4)

The observed electrophoretic velocity and mobility of an analyte can be calculated experimentally from the migration time.\[22\] To calculate the mobility in terms of the migration time and applied voltage, two length parameters are considered: the total capillary length and the effective capillary length. The total capillary length is the length from the capillary inlet to outlet. Effective capillary length is the length from the capillary inlet to the detection window. The observed electrophoretic mobility, \( \mu \), can be calculated using Equation 1.2.5, where \( L_{eff} \) is the effective capillary length, \( t \) is the migration time, \( L_{tot} \) is the total capillary length and \( V \) is the applied voltage. The first term gives the observed velocity \( (v) \) and the second term gives the inverse of the field \((\frac{1}{E})\).

\[ \mu = \frac{L_{eff} L_{tot}}{t V} \]  \hspace{1cm} (1.2.5)

1.2.2.4 Electroosmotic Flow

Crucial to the theory of capillary zone electrophoresis is the phenomenon of electroosmotic flow (EOF). EOF is the bulk flow of the liquid through the capillary under the influence of the electric field. When a capillary is filled with an electrolyte, the capillary wall can acquire a charge due to the adsorption of ions onto the surface or the ionisation of certain species present in the wall. As a consequence, counter-ions cluster around the wall forming what is known as the electrical double layer.\[30, 31, 32\] In the case of bare fused silica capillaries, in a buffer of pH above approximately 2 \[33\] the deprotonation of silanol groups results in a net negative charge on the surface. This results in a distribution of cationic species from the bulk fluid at the capillary walls, forming the electrical double layer (Figure 1.2.3). The electrical double layer consists of a layer of tightly bound cations (known
as the Stern layer), followed by a layer of hydrated cations (known as the compact layer), which are mobile (Figure 1.2.3 (a)). The formation of the double layer gives an electrical potential which drops off exponentially with distance called the zeta potential ($\zeta$) (Figure 1.2.3 (b)).[34] At a greater distance from the capillary wall, the zeta potential of the capillary wall can no longer be sensed and the solution becomes electrically neutral in what is known as the diffuse region. When an electric field is then applied, the subsequent migration of the hydrated cations results in a bulk movement of the fluid towards the cathode. The bulk movement of the fluid which results from the formation of an electrical double layer is known as electroosmotic flow.[25]

As the charge density drops off with distance from the surface, so does the zeta potential. The zeta potential has values ranging from 0 to 100 mV. The thickness of the double layer is defined as the distance to a point in the bulk liquid at which the potential is 0.37 times the potential at the interface between the the Stern and diffuse layers.[34] The thickness of the double layer, $\delta$, is defined by the equation

$$\delta = \left[ \frac{\varepsilon RT}{2eF^2} \right]^{1/2} \quad (1.2.6)$$

The zeta potential is dependent on the charge density of the bulk solution and is defined by the equation

$$\zeta = \frac{4\pi\delta e}{\varepsilon} \quad (1.2.7)$$

where $e$ is the total excess charge in the buffer per unit area, $\delta$ is the thickness of the double layer and $\varepsilon$ is the dielectric constant of the buffer.

The magnitude of the electroosmotic flow velocity is proportional to the zeta potential. The velocity of electroosmotic flow, $v_{eo}$, is given by the equation
Figure 1.2.3 – (a) Representation of the electrical double layer at the capillary wall and (b) the distance dependence of the zeta potential moving away from the capillary wall across the double layer.
Figure 1.2.4 – Electroosmotic and hydrodynamic capillary flow profiles

\[ v_{eo} = \frac{\varepsilon \zeta E}{4\pi \eta} \]  

(1.2.8)

where \( \varepsilon \) is the dielectric constant of the buffer, \( \zeta \) is the zeta potential, \( E \) is the field strength and \( \eta \) is the viscosity of the buffer. The zeta potential is the electrical potential which exists at the double layer with respect to the bulk fluid, and in Equation 1.2.8, the value for the zeta potential is taken at the plane of shear. From Equations 1.2.6, 1.2.7 and 1.2.8 it can be seen that an increase in buffer concentration results in a decrease in EOF velocity for a constant field. Note that Equation 1.2.8 is only valid for capillaries of sufficiently large inner diameter that the double layers (which typically range from 1-10 nm [34]) on opposite walls do not overlap.

A feature of EOF is that it ensures a flat flow profile along the capillary (Figure 1.2.4). This is advantageous as the whole fluid is transported evenly and dispersion of separated zones is minimised. Parabolic flow profiles, as typical of hydrodynamic fluid motivation, suffer non-uniform fluid velocities across the capillary, causing band broadening and impaired efficiency and resolution.[31]

The migration velocity of an analyte is a sum of its electrophoretic velocity and the electroosmotic flow velocity. Consequently, the observed mobility
of the particle, $\mu$, is a sum of the absolute electrophoretic mobility of the particle, $\mu_{ep}$, and the mobility of electroosmotic flow, $\mu_{eo}$, as given by Equation 1.2.9

$$\mu = \mu_{ep} + \mu_{eo}$$  \hspace{1cm} (1.2.9)

To calculate the absolute electrophoretic mobility of a particle, knowledge of the electroosmotic velocity is required. This can be obtained through using a neutral marker that moves solely due to electroosmotic flow.[29]

Often the magnitude of the electroosmotic velocity of the fluid is greater than the magnitude of the electrophoretic velocity of the analyte, and therefore species will migrate towards the cathode irrespective of their charge. This enables the simultaneous separation of cations, anions and neutral ions using a single detector.[32] However, high electroosmotic flow reduces resolution as if it is too rapid it can result in elution of analytes before separation has occurred.

If reproducibility between runs is important then electroosmotic flow must remain constant, as variations in the electroosmotic flow alter the migration times of analytes. The electroosmotic flow velocity can be affected by a number of factors, including buffer pH and concentration, temperature, adsorption of protein to capillary walls and capillary wall modifications.[32, 29]

### 1.2.2.5 Separation Parameters: Efficiency and Resolution

By performing free zone electrophoresis in narrow bore capillaries, Jorgenson hypothesised that thermal effects become negligible and molecular diffusion becomes the predominant cause of band broadening [28, 35]. For this hypothetical situation, Jorgenson derived equations for the efficiency of the system. Efficiency, $N$, given in theoretical plates, is a term that comes from HPLC and relates to how narrow peaks are.

If diffusion is the sole cause of spreading of an initially infinitely sharp zone over a time, $t$, the spatial variance, $\sigma^2$, will be defined as
CHAPTER 1. INTRODUCTION

\[ \sigma^2 = 2Dt \quad (1.2.10) \]

Where D is the diffusion coefficient of the solute.

For a separation performed in a capillary of length \( L \), with detection at the far end, the velocity of migration, \( v \), of a species is given by

\[ v = \mu E = \mu V/L \quad (1.2.11) \]

where \( \mu \) is the apparent mobility, \( E \) is the field strength and \( V \) is the voltage.

The time, \( t \), for the solute to migrate the length \( L \) is given by

\[ t = \frac{L^2}{\mu V} \quad (1.2.12) \]

Substituting Equation 1.2.12 for \( t \) in Equation 1.2.10 gives

\[ \sigma^2 = \frac{2DL^2}{\mu V} \quad (1.2.13) \]

The total number of theoretical plates, \( N \), as defined by Giddings [36], is given by

\[ N = \frac{L^2}{\sigma^2} \quad (1.2.14) \]

Substituting Equation 1.2.13 into Equation 1.2.14 gives the result

\[ N = \frac{\mu V}{2D} \quad (1.2.15) \]

The most important aspect of Equation 1.2.15 is that the separation efficiency, \( N \), is directly proportional to the applied voltage. This suggests that to achieve the high efficiency, the highest possible voltage should be used. The efficiency is also proportional to the ratio of the mobility to the diffusion coefficient. These factors have a large dependence on the solute species, and
thus are less easily controlled to increase efficiency. As large molecules, such as proteins, have small diffusion coefficients, they are particularly well suited to analysis with CE and high plate numbers can be achieved. This is in contrast to HPLC, where efficiency is inversely proportional to molecule size.

It is also significant that Equation 1.2.15 is independent of capillary length. Equation 1.2.12 shows that the analysis time is proportional to the square of the capillary length, which suggests that to achieve the highest efficiency in the shortest possible time, high voltages should be applied across short capillaries. However, the resistance of a shorter capillary is reduced, thus producing more Joule heating. Furthermore, the surface area over which heat is dissipated is also reduced.

The efficiency can be determined from an electropherogram using Equation 1.2.16

\[ N = 5.54 \times \left( \frac{t}{w} \right)^2 \]  

(1.2.16)

Where \( t \) is the migration time and \( w \) is the peak full width at half maximum. The efficiencies achieved with CE are extremely high in comparison to high performance liquid chromatography (HPLC), which offers similar advantages to CE in terms of data acquisition and automation with applicability to many of the same sample types. In Jorgenson’s initial work, efficiencies exceeding 400 000 theoretical plates were achieved. Today, efficiencies of this level are regularly achieved and can even exceed \( 10^6 \) theoretical plates.

A high efficiency separation is meaningless if the species are not resolved. Jorgenson also derived equations which demonstrate how resolution should be optimised in CE.

The resolution between two solutes in electrophoresis, \( R_s \), was earlier defined by Giddings as
CHAPTER 1. INTRODUCTION

\[ R_s = \frac{N^{1/2}}{4} \frac{\Delta v}{v_{ave}} \] (1.2.17)

where \( \Delta v \) is the velocity difference between the two solutes and \( v_{ave} \) is their average velocity. The relative velocity difference between the two solutes, \( \frac{\Delta v}{v_{ave}} \), is equal to

\[ \frac{\Delta v}{v_{ave}} = \frac{\mu_1 - \mu_2}{\mu_{ave}} \] (1.2.18)

where \( \mu_1 \) and \( \mu_2 \) are the mobilities of the two solutes.

In the presence of an EOF, the relative velocity difference becomes

\[ \frac{\Delta v}{v_{ave}} = \frac{\mu_1 - \mu_2}{\mu_{ave} + \mu_{eo}} \] (1.2.19)

Substituting the relative velocity difference (Equation 1.2.19) and the number of theoretical plates (Equation 1.2.15) where \( \mu = \mu_{ep} + \mu_{eo} \) into Equation 1.2.17 gives

\[ R_s = \frac{1}{4} \left[ \frac{(\mu_{ave} + \mu_{eo})V}{2D} \right]^{1/2} \left[ \frac{\mu_1 - \mu_2}{\mu_{ave} + \mu_{eo}} \right] \] (1.2.20)

Rearranging Equation 1.2.20 yields Jorgenson’s equation for resolution [28]

\[ R_s = 0.177(\mu_1 - \mu_2) \left[ \frac{V}{D(\mu_{ave} + \mu_{eo})} \right]^{1/2} \] (1.2.21)

From Equation 1.2.21, it is apparent that increasing the voltage is not a very effective route to improving resolution. Whilst efficiency is directly proportional to voltage (Equation 1.2.15), resolution is only proportional to the square root of voltage. Limitations are placed on the voltage that can be applied by the power supply (usually operating up to 30kV) and by Joule heating. It is clear that a more effective route to improving resolution is by maximising the mobility differences of the solutes. This is best achieved by
changing the pH of the BGE.\[29\]

Equation 1.2.21 also predicts resolution to increase with a decreasing EOF. Infinite resolution will be obtained with $\mu_{\text{ave}}$ and $\mu_{\text{eo}}$ equal and opposite, which clearly will not result in species migrating any distance in the capillary. Similarly, large, reverse, EOF will result in long separation times. Therefore, a compromise must be made to give an EOF which gives short separation times whilst still maintaining resolution.

1.2.3 Instrumental and Operational Aspects

1.2.3.1 Bandbroadening

In Section 1.2.2.5, the efficiency of a separation is described as first derived by Jorgenson \[28\], which assumes that diffusion is the sole cause of bandbroadening. In practice, efficiencies achieved during separations do not meet the value calculated using Equation 1.2.15. This is because several processes in addition to diffusion contribute to bandbroadening. These processes include Joule heating, sample injection, electrodispersion, adsorption (and subsequent desorption) of sample to capillary walls, detection and laminar flow in the capillary generated by unlevelled buffer vials \[22, 29\], some of which are discussed in further detail. The total spatial variance is more realistically described by the sum of these contributing variances.\[29, 22\] Hence, the total spatial variance, $\sigma_T^2$ is given by

$$\sigma_T^2 = \sigma_D^2 + \sigma_{\text{temp}}^2 + \sigma_{\text{inj}}^2 + \sigma_{\text{electrodisp}}^2 + \sigma_{\text{ads}}^2 + ... \quad (1.2.22)$$

where subscripts describe diffusion, temperature gradients as a result of Joule heating, sample injection, electrodispersion and adsorption respectively.

Joule heating in CE has been studied extensively \[38, 39, 40, 41\], with the conclusion that it is negligible for typical separations in narrow bore capillaries ($\lesssim 75 \mu m$) (unless very high buffer concentrations are used).\[41,
However, in wider bore capillaries, Joule heating can be a cause of band broadening and an Ohm’s law plot should be performed to determine the maximum field strength that can be applied without Joule heating. This is carried out by filling the capillary with the BGE, gradually increasing the voltage and recording the current. A plot of current as a function of voltage will yield a straight line, but if Joule heating is taking place the resistance decreases, causing an increase in current and hence a deviation from linearity.

The length of the initial sample zone, as determined by sample injection, contributes to band broadening as it defines a spread of lengths to migrate to the detector. The ideal is to inject an infinitely narrow zone. In practice, detector sensitivities dictate the limits of injection volumes. Techniques can be employed to narrow the sample plug whilst injecting larger sample quantities. A discussion of sample injection is given in Section 1.2.3.2.

Large injection amounts can lead to another source of band broadening, known as electrodispersion. The presence of sample ions in a zone affects the local electric field strength in that zone, as the conductivity of this zone differs to that of the BGE. As the current must be constant over the entire capillary, where the conductivity is higher the local field strength is lower, and hence the analyte migrates slower, and vice versa. Hence a relationship exists between the velocity of the sample and the sample concentration. The concentration of an analyte in a zone typically initially follows a Gaussian distribution, and the velocity of analyte through a zone will vary accordingly causing deformation of peaks and band broadening, known as electrodispersion. Characteristic features of electrodispersion are tailing of fronting peaks, and migration times that shift with increasing injection volumes. As it is more pronounced with higher sample amounts, it is often termed sample overloading. The problem can be alleviated somewhat by increasing the concentration of the BGE. With a suitable BGE this effect is usually small.

Adsorption, and subsequent desorption, of analyte contribute to band-
CHAPTER 1. INTRODUCTION

broadening. Adsorption is particularly problematic for protein separations as they adsorb strongly to silica due to their numerous charges and hydrophobic moieties. The inner capillary walls of a bare fused silica capillary are negatively charged in a BGE of pH above the pI of fused silica, which is about 2.[33] In a buffer with a pH below the pI of constituent proteins, the proteins are positively charged and will therefore interact with the capillary walls. Furthermore, due to their complex structure, localised surface charges and hydrophobic interactions can cause adsorption even at a BGE pH equal to or above the protein pI.[45, 25] In CE, because the surface-to-volume ratio is very high, these interactions can be significant. Bandbroadening is a consequence of the desorption of adsorbed analyte from the capillary walls, and leads to peak tailing. Evaluation of the adsorption behaviour using adsorption/desorption kinetics of an analyte reveal that even modest retention of the analyte can have substantial impact on bandbroadening.[25, 22] In addition to bandbroadening, which results from protein adsorption and desorption, problems are associated with irreversible protein adsorption. Adsorption of proteins onto the capillary walls changes the charge on the wall, and thus affects the electroosmotic flow and hence the migration time, leading to irreproducibility. Peak area is reduced if adsorption occurs, which can lead to errors in quantitative analysis. Furthermore, damage to capillaries and blocking can occur.

There are several techniques that can be used to reduce protein adsorption. These include modifications of the capillary walls, use of extremes of pH for the BGE, high BGE concentrations and BGE additives. Capillary walls can be modified either with dynamic or permanent coatings. Dynamic coatings are either added to the BGE throughout the run or as pre-coatings that are flushed through prior to beginning the separation and are easily removed and regenerated. Capillaries can also be permanently coated, some of which are commercially available. Capillary coatings are typically polymers which screen the proteins from the negatively charged capillary walls. Examples
of commonly used coatings are polyacrylamide [46], commercially available from Beckman Coulter and polyvinyl alcohol [47], available from Agilent Technologies. These coatings give uncharged capillary walls and therefore eliminate EOF. Capillary functionalisations are also available that result in negatively charged walls and positively charged walls, and in these cases will give an EOF in the normal or reverse direction, respectively.

The advantage of using capillary coatings is that it leaves flexibility for the choice of BGE, and does not interfere with the separation to the same extent that extremes of BGE pHs, concentrations or additives can. Furthermore, as permanent coatings and dynamic precoatings do not enter the BGE, compatibility with a wider range of readout methods is possible, such as MS and EVV 2DIR.

1.2.3.2 Sample Injection

Sample injection is the process of introducing a small quantity of sample into the capillary. The ideal is to introduce a small, sharp plug of sample into the capillary to minimise zone spreading. Injection must also introduce sufficient sample quantities to surpass detection limits of the detection technique used (number for UV). Typically, volumes of a few nanolitres are injected [29, 25]. The requirement of a small sample quantity is of great advantage if limited sample is available, such as is often the case in the field of proteomics. There are two commonly used methods of sample injection: electrokinetic and hydrodynamic.

Electrokinetic injection is performed by placing the capillary and electrode in the sample vial and applying a voltage (usually 5kV – 15kV) for a given period of time (a few seconds). Analytes are carried into the capillary by electroosmotic flow and by electrophoretic mobility.

The quantity of a solute injected, $Q_{\text{inj}}$, is given by

$$Q_{\text{inj}} = \frac{V \pi c t r^2 (\mu_{ep} + \mu_{eo})}{L_{\text{tot}}}$$

(1.2.23)
where $V$ is the voltage, $c$ is the solute concentration, $t$ is the injection time, $r$ is the capillary inner radius, $\mu_{ep}$ and $\mu_{eo}$ are the mobilities of the solute and the EOF respectively and $L_{tot}$ is the capillary total length.

Electrokinetic injection has the advantage that the instrumentation is extremely simple, requiring only components that are necessary for the separation itself. It the method of injection first used in CZE by Jorgenson in 1981.[28, 35] However, the disadvantage is that the injection is biased, as ions with greater mobility enter the capillary more quickly than those with a lower mobility.[29, 35, 28]

Hydrodynamic injection can either be performed by siphoning or by pressure. For siphoning injection, the sample vial is raised above the buffer vial, and thus siphoning occurs. For pressure injection, either a pressure is applied to the sample vial or a vacuum is applied to the destination vial.

The volume of liquid injected hydrodynamically, $V_i$, can be calculated using the Poiseille Equation,

$$V_i = \frac{\Delta P r^4 \pi t}{8\eta L_{tot}}$$  \hspace{1cm} (1.2.24)

where $\Delta P$ is the pressure difference across the capillary, $r$ is the capillary inner radius, $t$ is the injection time, $\eta$ is the viscosity of the sample and $L_{tot}$ is the capillary total length.

For siphoning injections,

$$\Delta P = \rho g \Delta h$$  \hspace{1cm} (1.2.25)

where $\rho$ is the viscosity of the run buffer, $g$ is the gravitational constant and $\Delta h$ is the distance that the sample is raised above the outlet vial during injection.

In contrast to electrokinetic injection, hydrodynamic injection is unbiased. However, both forms of hydrodynamic injection are instrumentally more difficult than electrokinetic injection, with siphoning requiring a method
to reproducibly raise and lower the sample vial, and pressure injection requiring precisely regulated vacuum or pressure source to achieve reproducible injection. [29]

Due to the short path length of capillaries, the concentration detection limit in CE is low in comparison to HPLC. Therefore, analysis of dilute samples requires large injection volumes. However, a long sample plug impairs efficiency. It is widely accepted that to achieve high plate numbers, the plug length should not exceed 1% of the capillary length. [22, 29, 25] To enable analysis of dilute samples without a loss in efficiency, online sample preconcentration techniques, termed ‘sample stacking’, can be performed that reduce the length of the injected plug.

There are several techniques for online sample preconcentration, most of which exploit differences in properties of the sample buffer and BGE. These techniques include transient isotachophoresis (t-ITP) [48, 49], pH changes [50, 51], sweeping [52, 53] and Field Amplied Sample Stacking (FASS) [54]. In this work, the technique of FASS was performed.

FASS uses a continuous BGE with the sample buffer of lower conductivity. Given that the current density must be constant for the total capillary length, Ohm’s law (Equation 1.2.26) dictates that a decrease in conductivity in a zone results in an increase in field strength in that zone.[55] Hence a reduced sample buffer concentration results in a higher electric field strength in the sample zone than in the rest of the capillary. Analytes migrate quickly until they cross the boundary between the sample zone and the BGE, at which point the electric field abruptly drops and they slow down. Ions at the middle or rear of the injected plug move further through the high field in the sample zone and so catch up with ions from the front of the plug. Analyte ions are stacked in a narrow zone in front of the original boundary between the sample zone and the BGE.[31, 25] When the gradient of conductivity disappears, separation proceeds as usual.
\[ j = \sigma E \]

(1.2.26)

where \( E \) is the electric field strength, \( \sigma \) is the conductivity and \( j \) is the current density.

Typically, FASS describes this method of online concentration with hydrodynamic sample injection, and concentration factors of 10 are typically achieved.\cite{56} For electrokinetic injection, a plug of water or dilute BGE can be introduced prior to injection. As a result, a high field is produced at the inlet during injection. This method is known as Head Column-Field Amplified Sample Stacking (HC-FASS) and can result in up to 1000 times enrichment of the sample.\cite{56}

If the conductivity of the sample solution is much lower than the BGE (for example, if water is used), differences between the EOF in the sample and BGE zones arise which can generate laminar flow. Furthermore, as the electric field in the sample zone is much higher than the average across the capillary, the effect of Joule heating can cause temperature rises which lead to degassing or boiling of the sample. These effects both cause ‘destacking’ to take place. The optimal conditions for stacking have been found to be dilution of the sample in a 10-fold dilution of the BGE.\cite{57, 25, 29}

### 1.2.3.3 Detection

One of the great advantages of CE over slab-gel electrophoresis is the ability to perform on-capillary detection. The most common types of on-capillary detection are absorption and fluorescence (typically laser induced fluorescence, LIF \cite{58}), although a host of other techniques are also used including conductivity \cite{59}, chemiluminescence \cite{60} and refractive index \cite{61}.

High sensitivity detection techniques are required in CE because of the small injection volumes. An added difficulty for on-capillary detectors exists because of the short optical path length of the capillary. The limits of detection (LOD) can either be described in terms of concentration or mass, and
are referred to as a concentration LOD or mass LOD respectively.[25] The concentration LOD relates to the initial sample concentration, and therefore depends on the amount of sample injected. The mass LOD is a more accurate measurement of the detector capabilities, but is still affected by the quality of the separation. For a given analyte, the LOD will improve as efficiency increases. This is because, for a given mass of sample injected, a narrower, taller peak will contain a higher concentration than a wider, lower peak. The detection limit is typically estimated from the electropherogram at a 3:1 signal to noise ratio.[62, 63]

Ultraviolet/Visible absorption detection typically achieves sensitivities of $10^{-13} - 10^{-16}$ moles.[29] The best detection limits are achieved with LIF detection, where yoctomole sensitivities have been reached.[64, 65] Most proteins have intrinsic fluorescence due to tryptophan residues, but many peptides and other molecules do not and require derivitisation.[25]

Despite the advantages of both absorbance and fluorescence detection, such as the on-column nature and the outstanding sensitivities that can be achieved with LIF detection, neither technique can provide qualitative information about the nature of the analyte beyond migration times. Techniques that can be employed in absorbance or fluorescence detection to identify a compound, such as spiking, require a prior knowledge of the components of the initial sample and access to the individual compounds. Furthermore, these techniques are time consuming and are not suitable for samples of high complexity, such as many proteomic samples.

The demand for qualitative information from separations has led to the coupling of CE to mass spectrometry (CE-MS). CE-MS provides mass information and the identification not easily obtained by absorbance or fluorescence detection. It has become a popular technique, with particular importance in proteomics. Coupling of CE with MS is most often achieved online using Electrospray Ionisation (ESI), but advantages exist for offline interfaces with Matrix Assisted Laser Desorption Ionisation (MALDI).[19]
This is discussed in more detail in Chapter 6. Other detection methods which have been interfaced with CE and reveal information of the chemical nature of the analyte include online Raman spectroscopy [66], surface enhanced Raman spectroscopy (SERS) [67, 68] and NMR [69, 70].

In this thesis a number of spectral techniques are used in conjunction with an offline interface, which include mass spectrometry and spectroscopic techniques. Furthermore, on-line UV absorbance detection is used throughout capillary electrophoresis separations and track writing to monitor protein and peptide elution. The offline interface is used with fluorescence detection, MALDI-MS and MALDI-imaging, and EVV 2DIR which has great promise for proteomic applications. EVV 2DIR is a nonlinear spectroscopy which probes the coupling between vibrational resonances. A two dimensional infrared spectrum displays cross peaks where signal is enhanced for coupled vibrational modes. A protein spectrum displays cross peaks which are the vibrational signatures of certain amino acid side chains. Protein identification can be made from the intensities of these cross peaks with respect to an internal reference. This is explained in more detail in Chapter 7.

Fluorescence occurs when an excited state electron relaxes to a lower energy state, emitting a photon. To perform fluorescence spectroscopy, the species is first excited to one of the vibrational states of the excited electronic state by absorbing a photon. The excited molecule then relaxes to the lowest vibrational state of the excited electronic state before emitting a photon and returning to a lower electronic level. Fluorescence experiments are characterised by the intensity, the wavelength at which maximum emission occurs, \( \lambda_{\text{max}} \), the fluorescence lifetime (the time the species stays in the excited state before returning to the ground state) and the quantum yield (the fraction of molecules that relax by fluorescence). For proteomics, three aromatic amino acids, tryptophan, tyrosine and phenylalanine, exhibit useful fluorescence. Of the three, tryptophan has the highest quantum yield and has an emission maximum of approximately 348 nm.[71] The fluorescence of these residues
enables intrinsic fluorescence detection of proteins to be performed, where no derivitisation is required. Intrinsic fluorescence detection of proteins deposited from CE via tryptophan residues was used in this work as a method to image the deposited protein sample. This is discussed in Chapter 5.

1.3 The Ribosome

In all life forms, the journey from gene to protein passes through two main stages: transcription of DNA into the intermediate molecule, messenger RNA (mRNA) and translation, where the mRNA sequence is decoded and amino acids are assembled according to this sequence to produce the corresponding proteins. The ribosome is often referred to as the ‘protein factory’, as it is within the ribosome that translation is performed and proteins are constructed.[72]

Studying the ribosome is fascinating for biology as protein biosynthesis (the process of synthesising proteins) is a complex process that is central and essential to life. Studies of the ribosome are also medically relevant. Ribosomes are a prime target of antibiotics.[73] Furthermore, translational control is of fundamental importance to cancer development and progression.[74]

The ribosome is an enormous macromolecular machine (~2.6 MDa in prokaryotes [75] and ~4.3 MDa in eukaryotes [76]) consisting of a large and a small subunit, formed from ribosomal RNA (rRNA) and proteins (see Figure 1.3.1a). The small subunit is responsible for decoding the information in the mRNA, and the large subunit is the site at which peptide bond synthesis occurs.[77] Ribosomes are described in terms of the sedimentation coefficient, given in Svedberg units, S. Prokaryotic cytosolic ribosomes have a total sedimentation of 70S and consist of a large subunit with sedimentation coefficient 50S, a small subunit with sedimentation coefficient 30S. Eukaryotic cytosolic ribosomes are larger and more complex with sedimentation coefficients of 80S, 60S and 40S, respectively. Despite the structural differences, ribosomes
are functionally well conserved throughout all life forms.

The prokaryotic ribosome contains 3 strands of rRNA and >50 proteins; in *Escherichia coli*, the most widely studied, the 30S subunit consists of 16S rRNA (1542 nucleotides) and 21 unique proteins, and the 50S subunit consists of 23S rRNA (2904 nt), 5S rRNA (120nt) and 33 unique proteins.[78] Eukaryotic ribosomes contain 4 strands of rRNA and approximately 80 proteins: in *Saccharomyces cerevisiae*, the 40S subunit consists of 18S rRNA (1798 nt) and 46 proteins and the 60S subunit consists of 25S rRNA (3392 nt), 5.8S (158 nt) and 5S (121 nt) and 33 proteins.[79, 80, 14]

The understanding of the ribosome structure and function has been led largely by structural studies using x-ray crystallography and cryo-electron microscopy (cryo-EM). A dramatic increase in the understanding of prokaryotic ribosomes resulted from atomic structures of the small and large ribosome subunits from bacteria and archaea, published in 2000, separately by Ada Yonath et al [81], Thomas Steitz et al [75] and Venki Ramakrishnan et al [82], achievements which earned them the Nobel prize. Eukaryotic ribosome structure and function have been less well understood due to the lack of high resolution crystal structures [83], though high-resolution cryo-EM has given essential insight [80, 84]. However, this year, a high-resolution crystal structure of the eukaryotic ribosome 40S subunit from *Tetrahymena thermophila* [76], and in addition an intermediate resolution structure of the 80S yeast ribosome [85] have been published.

In all ribosomes, the rRNA chains, which account for approximately two thirds of the mass in prokaryotes and somewhat less in eukaryotes [75], fold to form the core which gives the ribosome its three dimensional structure. When the prokaryotic atomic structures were complete, it was astonishingly apparent that the peptidyl transferase centre was devoid of protein, with no visible peptide chain within 18Å.[87] For many [88], this confirmed a growing view that the ribosome was a ribozyme - an RNA enzyme. Most proteins could be seen to nestle between the rRNA folds, with many carrying long
(a) Cryo-EM structure of the 70S ribosome with EF-Tu bound. [86]

(b) The atomic structure of the 50S subunit of *H. marismortui*. [75] Proteins are shown in orange and labelled L1-L44.

Figure 1.3.1
chains going between the folds.[75] This can be seen in Figure 1.3.1b.

These proteins consist of a globular region, which sits on the surface of the ribosome, and a long segment, largely devoid of secondary structure, which creeps between the rRNA helices.[89] The long segments of the protein molecules are rich in the basic residues lysine, histidine and arginine, which form salt bridges with the rRNA, hence preventing the repulsion of nearby (negatively charged) rRNA molecules.[90] Thus, many ribosomal proteins are crucial in stabilising the structure of the ribosome. Eukaryotic ribosomal proteins range in size from approximately 10-50 kDa.[91]

Traditional ribosome knowledge classed the function of ribosomal proteins as a solely structural one. Today it is known that ribosomal proteins play a more complex role and as a result research on ribosomal proteins has intensified. Indeed, it is known that many ribosomal proteins have extraribosomal functions. In 1996, Wool described a range of ribosomal proteins, across species, exhibiting extraribosomal functions.[92] These extraribosomal functions include roles in transcription, translation, DNA repair, mRNA processing, apoptosis and developmental regulation. This original list of proteins stood at just over 30. In the last decade, the list has grown, and a large number of tentative cases also exist.[93] Not surprisingly, research is ongoing to confirm these as exhibiting bonafide extraribosomal activity and to find new candidates.

A contemporary issue in ribosomal proteomics is the question of stoichiometries, or, rather, substoichiometries. Conventional understanding states that ribosomal proteins exist at 1 to 1 stoichiometries in the ribosome, although they are not synthesised at stoichiometric quantities.[94] However, evidence for contradictory cases is amassing. Substoichiometries arise from scenarios relating to ribosomal proteins present at more than one copy (namely those of the stalk [95]), ribosome interacting proteins [96] or ribosomal proteins exhibiting extraribosomal functions which result in them being absent in the ribosome for periods of time [97]. A more recent hypoth-
esis, dubbed the ‘protein depot hypothesis’, suggests that ribosomes may act as hubs for ‘releasable regulatory proteins’ [98], with certain proteins of non-essential function in the ribosome performing their main role elsewhere in the cell. There are a number of candidates for release which have no known functional role in the ribosome and reside mainly on the surface, without the protrusions into the rRNA core to stabilise ribosome structure.[75, 98, 99] There are also regulatory proteins, which are not involved in the core function of the ribosome and may interact with complex signalling networks. An example is the Receptor for Activated C-Kinase, RACK1 (also known as Asc1p in S. cerevisiae), which has been confirmed as a core eukaryotic 40S ribosomal protein [14] but is also a scaffold protein, acting to recruit a range of kinases and membrane-bound receptors to the ribosome.[100] Proteins such as RACK1 potentially provide a physical link between signalling pathways and translation.[101] Additionally, there are the numerous non-core ribosomal components that are present in active ribosomes at various stages of translation, such as the initiation factors, elongation factors and termination factors.[78] Last, but not least, are enzymes responsible for the post-translational modification of some ribosomal proteins such as the well-known ribosomal s6 kinases, which regulate a diverse range of cellular processes and have been a subject of interest for nearly 30 years.[102]

Similarly to the field of proteomics in general, much of the data gained on ribosomal proteomics has been through low-throughput techniques of gel electrophoresis with mass spectrometry as well as structural studies. With the wide range of interactions of ribosomal proteins within and outside the ribosome becoming apparent, high throughput and robust tools to study intact ribosomal proteins are necessary. This thesis aims to address this by separating the intact ribosomal proteins using CE and applying this separation to the developed technology for readout from CE.
1.4 Thesis Outline

The aims of this project were two-fold: to use CE to separate the eukaryotic ribosomal proteins and to develop a novel proteomic technology based on the readout of capillary electrophoresis (CE) to electron-vibration-vibration (EVV) 2DIR spectroscopy. The basis of this was a CE instrument which was constructed by the author in collaboration with colleagues and used to perform robust, high efficiency separations of complex protein mixtures such as the ribosomal proteins. An interface enables deposition of the eluate from CE onto polished stainless steel plates for analysis with EVV 2DIR. Following several years of development of CE, the interface and EVV 2DIR spectroscopy, analysis of peptide samples from CE with EVV 2DIR was successfully achieved. The interface was adapted to enable readout with MALDI-MS. This was applied to the analysis of intact proteins from CE with MALDI-MS and a collaboration with Waters Corporation, Manchester enabled demonstration of the coupling of CE to MALDI-imaging for the first time.

**Chapter 2** describes the affinity purification method used for preparing ribosome samples.

**Chapter 3** describes the capillary electrophoresis system and characterisation experiments.

**Chapter 4** describes the separation of eukaryotic ribosomal proteins using CE. This is the only separation of the eukaryotic ribosomal proteins using CE. Over 26 peaks were resolved in less than 10 minutes. An outstanding RSD migration time of < 0.5% was achieved, demonstrating that the readout could provide a ribosomal ‘fingerprint’.

**Chapter 5** describes the sample deposition interface. The term ‘track-writing’ is used to describe the method of sample deposition, since
sample is deposited in the form of continuous tracks. This chapter describes how the interface works and the modifications made to develop it into a proteomic tool.

**Chapter 6** describes readout from CE with MALDI MS. The track-writing interface is adapted to enable matrix assisted laser desorption/ionisation (MALDI) - mass spectrometry (MS) to be used to analyse the samples in the tracks. Separations of proteins are successfully analysed using a standard MALDI-MS instrument. Through a collaboration with Waters Corporation, Manchester, this work was advanced by the offline coupling of CE to MALDI-imaging and applied to the ribosomal proteins.

**Chapter 7** describes the successful use of EVV 2DIR as a readout technique for CE for the first time. This is a strategy for protein fingerprinting which is currently under development in the research group of D.R. Klug at Imperial College London.[103, 1] Two-dimensional infrared spectroscopy (2DIR) probes vibrational coupling to achieve de-congested spectra compared with those of conventional optical spectroscopies, and can be understood as an optical analogue to two-dimensional NMR. There are several variants of 2DIR spectroscopy, and the form of 2DIR used is termed ‘electron-vibration-vibration’ (EVV) 2DIR. Spectra are successfully obtained with peptides and separated proteins in tracks are imaged, offering great promise as a proteomic tool.
Chapter 2

Ribosome Affinity Purification
CHAPTER 2. RIBOSOME AFFINITY PURIFICATION

2.1 Introduction

It was necessary to establish a robust method for preparing ribosome samples for CE experiments. A high throughput technique was desirable as the purposes of the technologies presented in this thesis is that they are, or have the potential to be, high throughput. Thus a laborious sample preparation technique would dwarf later downstream processes in terms of timescale. This chapter describes the preparation of ribosome samples for capillary electrophoresis using an affinity purification technique. This provides the ideal starting point for a high throughput platform for study ribosomal proteomics. A number of biochemical techniques are used to validate the affinity purification and these are discussed.

Most ribosome studies commence with a sucrose density gradient to prepare the ribosomes. Sucrose gradient centrifugation is frequently used in biochemistry to isolate large cellular components such as protein complexes, organelles or viruses. Sucrose gradients are made by layering lower concentrations of sucrose upon higher concentrations in a centrifuge tube. A sample, such as a cell lysate, is added to the top of the gradient and centrifuged at very high forces (typically in excess of 100 000×g) using an ultracentrifuge. Cellular components migrate through the gel until their density matches that of the surrounding sucrose. Fractions are then collected from the gradient and subjected to various analysis techniques to identify fractions of interest.[104] Ribosome purifications are one of the most frequent applications of sucrose gradients, and numerous different protocols can be found in the literature.[105, 106, 107, 108] In the case of ribosome purifications, following ultracentrifugation, to identify ribosomal fractions, the absorbance at 260 nm of each fraction is measured. Ribosome rich fractions exhibit a strong absorbance at 260 nm due to their high rRNA content. A ribosome purification performed using sucrose gradient centrifugation typically achieves separation between the small (40S) subunits, the large (60S) subu-
nits, the 80S monosomes and various length polysomes (depending on the sucrose density gradient and conditions).

Initial experiments towards the work presented in this thesis used sucrose gradients to purify ribosomes. A method was developed based on the ribosome purifications of Pestova et al [107] and Link et al [108]. As predicted, the procedure succeeded in purifying ribosomes with a level of separation between the major ribosomal components, as demonstrated in Figure 2.1.1. However, although the separation between ribosomal subunits, monosomes and polysomes achieved using a sucrose gradient is essential for some ribosome studies, such as structural studies, for many fields of ribosome work (including the investigations presented in this thesis) it is not essential, or necessarily desirable. There are a number of drawbacks of the sucrose gradient preparation procedure, some of which were highlighted whilst performing the gradients.

Firstly, sucrose gradients are extremely laborious: gradients require a day of preparation and, in this procedure, ultracentrifugation took 16 hours (overnight), following which fractions must be collected. Including gradient preparation, a sucrose gradient ribosome purification requires three days. Secondly, the procedure requires specialist equipment (an ultracentrifuge and ideally a fractionation device), which is expensive and not available in every lab. Thirdly, the initial experiments found that yields were too low (estimated concentration per ribosomal protein of 0.6 \( \mu \text{g/ml} \)), and would have required further concentration steps if the sample were to be used for capillary electrophoresis (CE). Large quantities of starting material are nonetheless typically required for sucrose gradients, with little scope for scaling down the procedure. Furthermore, dilution of the sample in the sucrose gradient yields final fractions that may not only suffer from low concentration, but also contain sucrose which can cause difficulties in further analyses if not removed. Additionally, ribosomal fractions can be contaminated by other high molecular weight complexes.
Figure 2.1.1 – Lysate from *Saccharomyces cerevisiae* was loaded onto 15-40% sucrose gradients and spun at 220 000 rpm in a Beckman Coulter SW28 rotor for 16 hours. a) The absorbance at 260 nm of each fraction from the sucrose gradient. The top of the gradient (lowest sucrose concentration) is fraction 22 and the bottom of the gradient (highest sucrose concentration) is fraction 1. The ribosomes can be seen in fractions 8-12 of the gradient. The absorbance measurements at 260 nm can be used to estimate the RNA concentration; for fraction 10 it equates to an RNA concentration of 80 ± 20 µg/ml. This equates to a concentration per ribosomal protein of approximately 0.6 µg/ml, which is too low for CE experiments and thus further concentration of the fractions would be required to use such samples for CE analyses. b) RNA gel electrophoresis showing lanes 8-12 in which ribosomal RNA bands are seen. RNA was purified, separated on a 1% agarose gel and visualised with SYBR®Safe DNA gel stain according to the procedure described in Section 2.2.4. As the 18S rRNA is from the 40S subunit and the 25S rRNA is from the 60S subunit, the gel showed that separation occurred between the subunits, with the small subunit settling in fractions further towards the top of the gradient than the large subunit.
Ribosome affinity purification methods can provide one-step methods achieving a highly purified sample with vastly improved yields, which require no specialist equipment and can easily be scaled up, scaled down and even miniaturised. However, bizarrely, their use is not commonplace. Ribosome affinity purifications are based on the use of either a tagged ribosomal protein or rRNA strand to capture the attached subunit or even the whole ribosome (often with associating proteins and mRNAs) onto a matrix. Ribosome affinity chromatography techniques have been developed most widely in *Escherichia coli* [110, 111, 112] and *Saccharomyces cerevisiae* [109, 113].

### 2.2 Methods: Ribosome Affinity Purification (RAP)

Ribosome affinity purification (RAP) was performed on the budding yeast, *Saccharomyces cerevisiae* using a FLAG affinity purification method. The recombinant strain expressed a C-terminally FLAG tagged version of a large subunit ribosomal protein. Purification was achieved using anti-FLAG affinity resin and competitive elution with a FLAG peptide. Lysate preparation and ribosome affinity purification were both performed according to an adaptation of the method by Inada *et al* [113].

As a major structural component of the ribosome is rRNA, degradation by ribonucleases would lead to dissociation of the ribosomal complex, preventing purification of intact 80S ribosomes. Therefore, ribosome preparations were all performed under RNase free conditions. Additionally, RNA purifications, RNA gels and native gels were all run under RNase free conditions. All water was either treated with 0.1% diethylpyrocarbonate (DEPC) and autoclaved or Ultra Filtered water (Chester Beatty Laboratories, Institute of Cancer Research, London) was used, which is RNase free. Glassware was DEPC treated and autoclaved. Disposable plasticware was purchased at a grade certified RNase free. Filter barrier pipette tips were used and sets
were designated for RNA work. Where possible, reagents were purchased at molecular biology grade that was certified RNase free. Apparatus, bench surfaces, pipettes and gloves were cleaned using RNaseZAP (Sigma-Aldrich, Dorset, UK).

Preparation of yeast strains was according to standard yeast practices. A detailed description of yeast handling techniques is beyond the scope of this thesis and can be read in detail elsewhere.[114]

2.2.1 Yeast Strains

Yeast strains YIT613 and YIT617 were a kind gift of Toshi Inada at Nagoya University, Japan. YIT613 (CB012 rpl25::LEU2) is the recombinant strain used for affinity purification. It contains the plasmid BIT700 [pRPL25-FH-URA3CEN], and so expresses a FlagHis$_6$ epitope tagged version of the large subunit ribosomal protein, Rpl25. YIT617 (CB012) is a wild type strain. The insertion of the gene, URA3, into the plasmid BIT700 makes YIT613 uracil autotrophic. As the wild type is a uracil auxotroph, this enables the use of uracil minimal media (media that is incomplete with respect to uracil) to prevent the growth of plasmid free cells.

On arrival from Japan, several tests were performed on the yeast strains and glycerol stock cultures were prepared for long term storage of the strains. The viability of both yeast strains was tested in terms of their growth on YPD (complete yeast growth media containing yeast extract, peptone and dextrose) agar plates. Both strains were found to be viable and glycerol stocks were made for each. To prepare the glycerol stocks, YPD liquid media was inoculated with single colonies of either YIT613 or YIT617 and grown to an optical density at 600 nm (OD$_{600}$) of approximately 1.5. 800 µl of sterile glycerol were then added to 800 µl of culture in Nunc cryotubes to give a final glycerol concentration of 20%. The resulting glycerol stock cultures were frozen for 30 minutes on dry ice and stored at -80 °C. To check that the strain YIT613 had not rejected the plasmid BIT700, both strains were plated
out and grown on uracil deficient media plates. As expected, YIT613 grew on uracil deficient media, whilst YIT617 did not, confirming the presence of BIT700 in YIT613 (Figure 2.2.1).

2.2.2 Lysate Preparation

For cell growth from glycerol stocks, YIT613 was plated onto a uracil deficient plate and grown at 30°C for 2 days. 50 ml of liquid YPD media was inoculated and grown overnight at 30°C. The following morning, 200 ml of YPD at 30°C were added to each culture, and the cells were grown to an OD$_{600}$ of 0.8.

Lysis was performed mechanically, using a bead beating method. The cell cultures were transferred to 500 ml plastic, conical bottomed flasks and the cells pelleted at 2500 rpm for 5 minutes in a Sorvall RC SC Plus centrifuge. The supernatant was poured away and the cells resuspended in 400 µl of lysis buffer (20 mM HEPES pH 7.4, 2 mM Mg(OAc)$_2$, 100 mM KOAc, 100 µg/ml cycloheximide, 0.5 mM dithiothreitol) per flask by repeated pipetting with a 5 ml pipette. Each cell culture was then transferred to 2×2 ml tubes and the cells pelleted by spinning at 7000 rpm for 5 minutes in an Eppendorf 5415 centrifuge. The supernatant was removed and the cells resuspended in an equal volume of lysis buffer + protease inhibitor (1×Roche EDTA-free complete protease inhibitor cocktail). Cell lysis was performed using
the FastPrep System (Qbiogene, Inc., Carlsbad, CA). The resuspended cells were transferred to 2 ml FastProtein Red Matrix\textsuperscript{TM} tubes and lysed by 6 repetitions of 20 second runs followed by 1 minute on ice. The tubes were then spun at 10000 rpm for 1 minute in an Eppendorf 5415 centrifuge to pellet the beads, and the lysate transferred to 2 ml tubes and cleared by spinning at 10000 rpm for 1 minute. The cleared lysate was transferred to 2 ml tubes, frozen in liquid nitrogen and stored at -80° C.

2.2.3 Affinity Purification

RAP was performed by incubating the lysate with an anti-FLAG resin (which is in the form of beads), following this, washing unbound material and eluting by competitive elution with a FLAG peptide. The procedure is as follows:

Anti-FLAG M2-agarose affinity resin (Sigma-Aldrich) was used for capture of FLAG-tagged ribosomes from YIT613 lysate. 200 \( \mu l \) of resin (100 \( \mu l \) bead volume) was washed twice in 4 ml TBS (50 mM Tris-HCl, 150 mM NaCl, pH 7.4) and then washed three times in 4 ml 2× binding buffer (100 mM Tris-HCl pH 7.5, 24 mM Mg(OAc)\textsubscript{2}, 50 U/ml RNAsin (Promega, Southampton, UK), 1 mM PMSF). Following this, 125 \( \mu l \) of lysate (approximately 4 A\textsubscript{260} units) was added to an equal volume of ice cold 2× binding buffer and 100 \( \mu l \) of resin (bead volume) in a 2 ml tube. Binding proceeded for 2\( \frac{1}{2} \) hours at 4°C with gentle rocking. The resin was washed five times with 0.2 ml of ice-cold IXA-100 buffer (50 mM Tris-HCl pH 7.5, 12 mM Mg(OAc)\textsubscript{2}, 100 mM KCl, 1 mM PMSF). Two consecutive elutions were performed, each time, by incubating the resin for 20 minutes at 4°C with gentle rocking in IXA-100 buffer containing FLAG peptide (Sigma-Aldrich) at a concentration of 100 \( \mu g/ml \). Washes and elutions were carried out by adding buffer to the resin, centrifuging at 3000 rpm in a Centra CL3R Thermo centrifuge at 4°C for 2 minutes and removing and retaining the supernatant. The absorbance at 260 nm of post resin lysate, washes and elutions was measured (Nanodrop Spectrometer, Nanodrop, Wilmington, DE, USA). Extracts were aliquotted
into 25 µl units, frozen in liquid nitrogen and stored at -80°C until required for gels or for capillary electrophoresis analysis.

After the procedure had been established with several repetitions, it was scaled up to produce a greater number of 25 µl aliquots for downstream developmental work and the volume of lysate loaded was increased to increase the ribosome concentration. For this, simply 1600 µl of lysate were loaded onto 800 µl of resin (bead volume) and the procedure carried out as described. All other volumes were used in equal ratios to those described. For experiments into deposition of ribosomal proteins (Chapters 5 and 6), an increased ribosome sample concentration was required to meet detection limits of the downstream analytical techniques under development. Therefore, 3 ml of lysate were loaded onto 800 µl of resin and left to bind overnight to ensure saturation of the resin.

2.2.4 Assessing the purification

To establish that the affinity eluate contained ribosomes and assess the quality of the sample, a number of biochemical techniques were used, which either characterised the ribosomal RNA or the proteins. An RNA gel is the clearest method to determine the presence of ribosomes due to the presence of two characteristic bands produced by the 25S and 18S strands. Measurements of the absorbance at 260 nm enable estimation of the concentration of nucleic acid in the sample. Studies on the proteins present were performed with denaturing sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Typically, these assays were performed on the post resin lysate, the washes and the eluate from each elution. The development of a capillary electrophoresis method as demonstrated in Chapter 4 would provide a way in the future to assess purifications, and in the later purifications performed, once the CE method was robust, the characteristic fingerprint seen was used in addition to gels to verify the integrity of the sample. For the purposes of the ribosome purifications, however, as the CE platform was in a develop-
mental stage, established biochemical techniques were used where the results could be compared to those published.

**RNA Gel Electrophoresis**

The presence of ribosomes in a sample can be clearly determined using an RNA gel due to the presence of two characteristic rRNA bands produced by the 25S and 18S strands at 3392 nucleotides (nt) and 1798 nucleotides (nt) respectively. Since the 25S rRNA is from the 60S subunit and the 18S rRNA is from the 40S subunit, the presence of both bands confirms the presence of both subunits. Before performing RNA gels, it was necessary to perform an RNA purification.

RNA was isolated from aliquots of sample from the RAP set aside for RNA studies using TRI reagent LS (Sigma-Aldrich), which is a solution containing guanidine thiocyanate and phenol. The 25 µl aliquots of sample were made up to 250 µl with RNase free water and mixed with 0.75 ml TRI-reagent LS in a 2 ml tube. To this, 0.2 ml of chloroform was added and the tubes shaken for 15 seconds. Tubes were then centrifuged at 12000 × g for 15 minutes at 4°C. The centrifugation separates the mixture into three phases: a dark pink organic phase at the bottom, which contains the proteins; a pink interphase, which contains DNA; and a colourless aqueous phase at the top, which contains the RNA. The colourless aqueous phase at the top was transferred to another tube and 0.5 ml isopropanol added to precipitate the RNA. Tubes were centrifuged at 12000 × g for 10 minutes at 4°C and the supernatant removed to leave an RNA pellet at the side and bottom of the tube. The RNA pellet was washed in 1 ml of 75% ethanol and centrifuged at 12000 × g for 5 minutes at 4°C. The ethanol supernatant was then removed and the RNA pellet air-dried for 10 minutes. The RNA pellet was solubilised by adding 20 µl of water, pipetting up and down 5 times with a P200 Gilson pipette, incubating at 60°C for 10 minutes and repeating the pipetting and incubation.
CHAPTER 2. RIBOSOME AFFINITY PURIFICATION

Native RNA gels were performed on 1% agarose gels in either 12 well trays or 30 well trays. Gels were made by heating 1% agarose in 1×Tris-acetate-ethylenediamine tetraacetic acid (TAE) to 200°C and, after the gel had cooled to 55°C, adding 0.1 µl/ml SYBR®Safe DNA gel stain (Invitrogen, Paisley, UK) before pouring and allowing the gel to set for at least 30 minutes at room temperature. A run buffer of 1×TAE was then poured onto the gel until the wells were filled and the gel covered. Samples were prepared by adding 5 µl of RNA sample loading buffer (Sigma-Aldrich) to 10 µl of the purified RNA solution and heating at 65°C for 7 minutes. A 0.2-10 kb (kilo base) RNA marker (Sigma-Aldrich) was prepared by adding 3 µl of DEPC treated water and 3 µl of RNA sample loading buffer (Sigma-Aldrich) to 2 µl of marker and heating at 65°C for 7 minutes. After samples had cooled to room temperature, 10 µl of each sample, or 8 µl of marker, was loaded per well. Gels were run at 80 V until the xylene cyanol dye (a component of the loading buffer) had run halfway down the gel. Following this, the gel was visualised using a UV light source and photographed.

SDS-PAGE

SDS-PAGE was performed using upright Criterion cell apparatus (Bio-Rad, Hertfordshire, UK) with pre-made Criterion XT Bis-Tris 4-12% gels (Bio-Rad). The run buffer used was Criterion XT MES, of which 60 ml was used per inlet tank and 400 ml for the main tank. Samples were prepared by adding 5 µl of XT sample loading buffer (Bio-Rad) and 1 µl of XT Reducing Agent (Bio-Rad) to 14 µl of the sample and heating at 95°C for 5 minutes. Precision Plus Protein All Blue Standard, 10-250 kDa molecular weight range (Bio-Rad) was used as a protein marker and 10 µl of marker was loaded. After samples had cooled to room temperature, 16 µl of each was loaded per well and the gel was run for 45 minutes at 200 V. Following this, the gel was stained with InstantBlue Stain (Expedeon, Cambridge, UK) for 2 hours with gentle shaking and then destained in ultrapure water overnight with gentle
shaking before photographing.

2D Blue Native / SDS - PAGE (BN/SDS-PAGE)

This technique is an unusual type of two dimensional gel, of which one of the applications is the study of protein complexes. It was used in this work partly as an alternative to native gels because it was found that ribosomes were too large to migrate through a native gel, instead migrating only a short distance past the loading well. In this method of two dimensional gel electrophoresis, blue native polyacrylamide gel electrophoresis (BN-PAGE) is run in the first dimension, where coomassie is added to the cathode buffer to add charge to complexes to aid in their migration through the gel and enable some visualisation of bands before staining. Following separation under non-denaturing conditions, the lane of interest is cut out, incubated in reducing agent and SDS sample loading buffer, and then loaded horizontally for SDS-PAGE. This enables components of bands in the native gel to be analysed. The gel added evidence that ribosomes were purified and that substantial purification was achieved. Due to time constraints and since information gained from other assays was sufficient, 2D BN/SDS-PAGE was not regularly performed for affinity purifications.

2D BN/SDS-PAGE was performed using a basic adaptable upright apparatus with pre-made gels for both dimensions. In the first dimension, NativePAGE Novex 3-12% Bis-Tris gels (Invitrogen) were used. Two gels were run so that, following the first dimension, one gel could be stained to visualise the bands whilst a lane from the other gel could be cut out and run in the second dimension. As intact ribosomes were required for the first dimension, to prevent the degradation of ribosomes by ribonucleases, native gels were run under RNase free conditions. Two types of NativePAGE Running Buffers (Invitrogen) were used to perform BN-PAGE: NativePAGE 20X Running Buffer and NativePAGE 20X Cathode Buffer Additive, which contains 0.4% Coomassie G-250. NativePAGE 20X Cathode Buffer Additive
was added to NativePAGE Running Buffer to give final proportions of 0.1X Cathode Buffer Additive in 1X Running Buffer to generate Light Blue Cathode Buffer (containing 0.002% Coomassie G-250). The Light Blue Cathode Buffer was used for the inlet tank and the running buffer was used for the main tank. Samples were simply prepared by adding 5 µl of NativePAGE Sample Buffer (4X) (which contains BisTris buffer, pH 7.2, NaCl, glycerol, and Ponceau S.) to 15 µl of the sample and loading. NativeMark Unstained Protein Standard, 20-1200 kDa molecular weight range (Invitrogen) was used as a marker and prepared by adding 5 µl of NativePAGE Sample Buffer (4X) (Invitrogen) and 10 µl of ultra-pure water to 5 µl of marker. Marker was loaded in the lanes at both ends of the gel. The gel was run at at 4°C at 150 V for 60 minutes, and then increased to 250 V for 60 minutes. Following this, one gel was stained with InstantBlue Stain (Expediton) for 2 hours with gentle shaking and then destained in ultrapure water overnight with gentle shaking before photographing.

To perform the second dimension, the eighth lane (corresponding to the first elution of the RAP) of the second gel was cut out using a scalpel blade and placed in a sterile 15 ml conical tube with 5 ml of reducing solution (4.75 ml 1X NuPAGE LDS Sample Buffer (Invitrogen) + 250 µl XT Reducing Agent (Bio-Rad)) and incubated at room temperature with shaking for approximately 1 hour. This was then inserted lengthways into the sample well of a NuPAGE Novex 4-12% Bis-Tris gel (2D well) (Invitrogen). 60µl of 1X NuPAGE LDS Sample Buffer (Invitrogen) was added to the gel strip in a continuous line using a gel loading tip. The marker was Precision Plus Protein All Blue Standard (Bio-Rad), prepared as described for the standard SDS-PAGE gels but using 1X NuPAGE LDS Sample Buffer. The run buffer was 1X NuPAGE MES SDS Running Buffer (Invitrogen). The gel was run for 35 minutes at 200 V. Following this, the gel was stained with InstantBlue Stain (Expeideon), destained in ultrapure water and photographed.
Mass Spectrometry

Eluate from an affinity purification was analysed by liquid chromatography matrix assisted laser desorption ionisation-time of flight (MALDI-TOF)/TOF-MS in the Mass Spectrometry Centre for Integrative Systems Biology at Imperial College London. Briefly, the fully prepared sample was desalted and trypsin digested. Fragments were then separated by reversed-phase HPLC and robotically spotted onto a metal MALDI target using an Ultimate 3000 nanoLC system (Dionex, Sunnyvale, CA, USA) coupled to a Probot automated spotting device (LC Packings, Sunnyvale, CA, USA). MALDI-MS/MS was performed using an Applied Biosystems 4800 TOF/TOF. All MS/MS data were analysed using the Mascot search engine (Matrix Science, London, UK). Mascot was set up to search the SwissProt *Saccharomyces cerevisiae* database (UniProt release 2010_09) for the digestion enzyme Trypsin. The database was searched with a fragment ion mass tolerance of 0.30 Da and a parent ion tolerance of 100 ppm. Oxidation of methionine residues and carboxymethylation of cysteine residues were specified as variable modifications. Scaffold (version 3.0, Proteome Software Inc., Portland, OR, USA) was used to validate MS/MS based protein identifications. Protein identifications stated in the supplementary information have greater than 50.0% probability and contain at least one identified peptide. Protein probabilities were assigned by the Protein Prophet algorithm.[115]

2.3 Results of RAP

Ribosome affinity purification (RAP) was performed on the recombinant *Saccharomyces cerevisiae* strain, YIT613, which expressed a C-terminally FlagHis\textsubscript{6} epitope tagged ribosomal protein Rpl25. Purification was achieved using anti-FLAG affinity resin and competitive elution with a FLAG peptide. This method yielded highly purified, intact ribosomes as were validated by a number of criteria.
In the first indication for the success of the purification, the absorbance measurements at 260 nm (recorded for the five washes and the two elutions in succession) decrease during the successive washes until elution, when the FLAG peptide is added (Figure 2.3.1). A large increase is seen in the values of the absorbance at 260 nm of the eluate, indicating the presence of RNA (which has a strong absorbance at 260 nm) and the successful elution of ribosomes, a trend which was observed in all ribosome affinity purifications performed.

RNA gels demonstrate that the high absorbance at 260 nm is a result of ribosomal RNA, as the characteristic 25S and 18S rRNA bands (from the 60S and 40S subunits respectively) are clear at 3392 nt and 1798 nt in eluate from purifications whereas bands corresponding to no other types of RNA (such as tRNA) are present (Figure 2.3.2a). The presence of an rRNA band from the 40S subunit in addition to an rRNA band from the 60S subunit confirms
Figure 2.3.2 – a) RNA gel of eluate from ribosome affinity purification. 25S and 18S rRNA bands can be seen. RNA associated with purified ribosomes were prepared resolved on a 1% agarose gel, and visualised by staining with SYBR Safe DNA gel stain. b) Denaturing protein gel of successive washes (lanes 2-6) and elutions (lanes 7-8). Proteins were resolved on 4-12% SDS-PAGE and visualised by coomassie staining. c) 2D BN/SDS-PAGE of eluate from RAP. In the first dimension, complexes from washes and elutions were resolved on two 3-12% native PAGEs. One gel was visualised with coomassie staining (shown, left). The lane from the unstained gel corresponding to the first elution was cut out, denatured and loaded horizontally onto a 4-12% SDS-PAGE. Resolved proteins were visualised with coomassie staining.
the purification of intact 80S ribosomes, in agreement with the work of Inada et al.[113]

The SDS-PAGE (Figure 2.3.2b) shows that the majority of proteins found in the eluate fall within the size range of ribosomal proteins of 10 - 50 kDa.[116, 117] These centre on 20-30 kDa, which is close to the number average, which for the proteins of both subunits is approximately 21 kDa.[116] As expected, distribution of protein bands over the full mass range can be seen in the first wash, which subsequently fade through successive washes. The subsequent appearance of bands in the size range of ribosomal proteins in the eluate therefore indicates substantial purification.

Further evidence that the RAP achieves a high degree of purification is seen from the two dimensional BN/SDS-PAGE (Figure 2.3.2c). The 2D BN/SDS-PAGE uses the principles of a first dimension blue native gel but follows this by performing a denaturing step on the lane of interest; this lane is then turned around and placed in the top of a denaturing polyacrylamide gel and therefore the components of each band in the native gel can be analysed. Two gels were run in the first dimension (BN-PAGE) and only one of them was subsequently used for the second dimension (SDS-PAGE), so that the second gel could be stained and compared to the second dimension of the first gel.

In the native first dimension, eluate runs as a large band with a molecular weight exceeding the 1.2 MDa band of the marker. On denaturation, this band yields a series of bands between 10 and 50 kDa in the second dimension, which is the size range of eukaryotic ribosomal proteins. Importantly, very few spots can be detected across the remainder of the two dimensional gel despite the wide mass range, which indicates that a high degree of purification was obtained.

The success of the ribosome purification was proven by identification of the proteins within the eluate with LC-MS/MS. Ribosomal proteins were identified from both the 40S and the 60S subunit, giving further confirma-
tion that the purification procedure led to capture and elution of intact 80S ribosomes. Interestingly, the data showed that regulatory components of the ribosome, such as Asc1p, and non core ribosomal components, such as elongation factors, were purified. Elongation factors are proteins that facilitate translational elongation and bind to 80S ribosomes at a number of sites [118, 119, 120], thus can be assumed to be copurifying proteins rather than contaminants. This demonstrates that the procedure is gentle in that it is capable of purifying ribosome associated proteins and ribosomal proteins exhibiting extraribosomal activities, which has implications for further downstream studies into these areas.

The results are shown in Section 4.7 Appendix A, Table 2.5.1. Out of 73 hits above a 53% protein identification probability, 61 were ribosomal proteins and 2 were elongation factors. With a cutoff of 100%, 43 out of 47 hits were ribosomal proteins or elongation factors. The ribosomal protein hits included Asc1p (‘Guanine nucleotide-binding protein subunit beta-like protein’) with a protein identification probability of 100%. A small number of non-ribosomal related proteins are present in the mass spectrometry data, the majority of which are highly abundant, glycolytic enzymes, for which it could be expected to be found a small number at some levels despite extensive washes.

The yield of the purifications, based on RNA concentrations in the first elution, was dramatically higher than that measured in ribosomal fractions in sucrose gradient purifications. Furthermore, effective scaling was simple to perform to increase both the volume of the sample purified and the concentration.

Initial purifications (n=3) gave an average RNA concentration of 419 ± 219 µg/ml in the first elution, calculated from the absorbance at 260nm. With an increase in scaling to produce a greater volume of sample and increase the concentration (1600 µl lysate loaded per 800 µl resin), an average RNA concentration of 1072 ± 297 µg/ml was measured (n=2). With an
increase in lysate and incubation time (3 ml lysate loaded per 800 µl resin, binding overnight) to give a higher concentration as required by downstream analytical techniques, an RNA concentration of 2277 µg/ml was measured (n=1). This is over 25 times greater than the maximum concentration achieved in the ribosomal fractions of sucrose gradients.

Protein concentrations can be calculated from these RNA values. Assuming that the rRNA accounts for approximately 60% of the mass of the ribosome and the yeast ribosome contains 79 proteins at stoichiometric levels [121], concentrations per protein of 4 µg/ml, 9 µg/ml and 19 µg/ml can be approximated for the stated RNA concentrations respectively. These are all above the concentrations required by the detection limits of CE.

The affinity purification procedure required just one day to produce samples of this concentration, or two half days including an overnight incubation to achieve the highest concentration sample. At all sample concentrations obtained with the RAP, no additional procedures were necessary to concentrate the sample for CE analysis, whereas to use the sucrose gradient fractions for CE sample concentration would have been essential.

2.4 Conclusions

Ribosome affinity purification was successfully used to purify ribosomes from *Saccharomyces cerevisiae*. Although the tagged protein is on the 60S subunit, the purification method resulted in purification of 80S ribosomes, as is demonstrated by the RNA gel (Figure 2.3.2a), which shows both the 25S and 18S rRNA strands from the 60S and 40S subunits respectively. Additionally, ribosome associated proteins were co-purified. In terms of concentration, the yields were over an order of magnitude greater than samples produced using sucrose gradient centrifugation. The process could be completed within 1 day and did not require further purification or concentration, saving several days of sample preparation. This method of ribosome preparation is the ideal
starting point for the development of a high throughput ribosome analysis platform.

### 2.5 Appendix A

**Table 2.5.1** – Mass spectrometry results from ribosome affinity purification. Mascot used MS/MS data to identify proteins from the SwissProt *Saccharomyces cerevisiae* database. Scaffold was used to validate MS/MS based protein identifications. Protein identifications shown have greater than 50.0% protein identification probability and contain at least one identified peptide.

<table>
<thead>
<tr>
<th>Identified Proteins</th>
<th>Molecular Weight (kDa)</th>
<th>Protein Identification Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>60S ribosomal protein L20</td>
<td>20</td>
<td>100%</td>
</tr>
<tr>
<td>Phosphoglycerate kinase</td>
<td>45</td>
<td>100%</td>
</tr>
<tr>
<td>Enolase 2</td>
<td>47</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L3</td>
<td>44</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S4</td>
<td>29</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L15-A</td>
<td>24</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L4-A</td>
<td>39</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S22-A</td>
<td>15</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L13-B</td>
<td>23</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L8-B</td>
<td>28</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L19</td>
<td>22</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S3</td>
<td>26</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L17-A</td>
<td>21</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S1-A</td>
<td>29</td>
<td>100%</td>
</tr>
<tr>
<td>Elongation factor 1-alpha</td>
<td>50</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S7-A</td>
<td>22</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L2</td>
<td>27</td>
<td>100%</td>
</tr>
<tr>
<td>Identified Proteins (\textit{S. Cerevisiae})</td>
<td>Molecular Weight (kDa)</td>
<td>Protein Identification Probability</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>40S ribosomal protein S18</td>
<td>17</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S9-A</td>
<td>22</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L10</td>
<td>25</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S11</td>
<td>18</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L16-B</td>
<td>22</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L31-A</td>
<td>13</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S8</td>
<td>22</td>
<td>100%</td>
</tr>
<tr>
<td>Glyceraldehyde-3-phosphate dehydrogenase 2</td>
<td>36</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L30</td>
<td>11</td>
<td>100%</td>
</tr>
<tr>
<td>Guanine nucleotide-binding protein subunit beta-like protein</td>
<td>35</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L24-A</td>
<td>18</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L14-A</td>
<td>15</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L33-A</td>
<td>12</td>
<td>100%</td>
</tr>
<tr>
<td>Elongation factor 2</td>
<td>93</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L35</td>
<td>14</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L27-A</td>
<td>16</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S5</td>
<td>25</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L7-A</td>
<td>28</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S6</td>
<td>27</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S16</td>
<td>16</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S0-A</td>
<td>28</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S14-A</td>
<td>15</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L38</td>
<td>9</td>
<td>100%</td>
</tr>
<tr>
<td>Fructose-bisphosphate aldolase</td>
<td>40</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L28</td>
<td>17</td>
<td>100%</td>
</tr>
</tbody>
</table>
### Identified Proteins (S. Cerevisiae)

<table>
<thead>
<tr>
<th>Identified Proteins</th>
<th>Molecular Weight (kDa)</th>
<th>Protein Identification Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>60S ribosomal protein L5</td>
<td>34</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L9-A</td>
<td>22</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S13</td>
<td>17</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L18</td>
<td>21</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S24</td>
<td>15</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S7-B</td>
<td>22</td>
<td>99%</td>
</tr>
<tr>
<td>60S ribosomal protein L36-A</td>
<td>11</td>
<td>99%</td>
</tr>
<tr>
<td>60S ribosomal protein L21-A</td>
<td>18</td>
<td>99%</td>
</tr>
<tr>
<td>Pyruvate kinase 1</td>
<td>55</td>
<td>98%</td>
</tr>
<tr>
<td>Pyruvate decarboxylase isozyme 1</td>
<td>61</td>
<td>98%</td>
</tr>
<tr>
<td>60S ribosomal protein L22-A</td>
<td>14</td>
<td>98%</td>
</tr>
<tr>
<td>40S ribosomal protein S20</td>
<td>14</td>
<td>53%</td>
</tr>
<tr>
<td>40S ribosomal protein S19-A</td>
<td>16</td>
<td>53%</td>
</tr>
<tr>
<td>40S ribosomal protein S21-B</td>
<td>10</td>
<td>53%</td>
</tr>
<tr>
<td>40S ribosomal protein S29-A</td>
<td>7</td>
<td>53%</td>
</tr>
<tr>
<td>60S ribosomal protein L12</td>
<td>18</td>
<td>53%</td>
</tr>
<tr>
<td>60S ribosomal protein L16-A</td>
<td>22</td>
<td>53%</td>
</tr>
<tr>
<td>60S acidic ribosomal protein P0</td>
<td>34</td>
<td>53%</td>
</tr>
<tr>
<td>40S ribosomal protein S17-A</td>
<td>16</td>
<td>53%</td>
</tr>
<tr>
<td>40S ribosomal protein S2</td>
<td>27</td>
<td>53%</td>
</tr>
<tr>
<td>60S ribosomal protein L33-B</td>
<td>12</td>
<td>53%</td>
</tr>
<tr>
<td>Glyceraldehyde-3-phosphate dehydrogenase 1</td>
<td>36</td>
<td>53%</td>
</tr>
<tr>
<td>60S ribosomal protein L23</td>
<td>14</td>
<td>53%</td>
</tr>
<tr>
<td>40S ribosomal protein S25-A</td>
<td>12</td>
<td>53%</td>
</tr>
<tr>
<td>Alcohol dehydrogenase 1</td>
<td>37</td>
<td>53%</td>
</tr>
<tr>
<td>Chaperone protein dnaK</td>
<td>67</td>
<td>53%</td>
</tr>
</tbody>
</table>
## Identified Proteins (S. Cerevisiae)

<table>
<thead>
<tr>
<th>Identified Proteins</th>
<th>Molecular Weight (kDa)</th>
<th>Protein Identification Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enolase 1</td>
<td>47</td>
<td>53%</td>
</tr>
<tr>
<td>60S ribosomal protein L25</td>
<td>16</td>
<td>53%</td>
</tr>
<tr>
<td>60S ribosomal protein L43</td>
<td>10</td>
<td>53%</td>
</tr>
<tr>
<td>40S ribosomal protein S23</td>
<td>16</td>
<td>53%</td>
</tr>
<tr>
<td>40S ribosomal protein S26-A</td>
<td>13</td>
<td>53%</td>
</tr>
</tbody>
</table>
Chapter 3

Capillary Electrophoresis
Instrumentation
3.1 Introduction

This chapter describes the home-built capillary electrophoresis instrument that forms the basis of the developed technology platform. Although excellent CE systems are available commercially, a custom made system offers the advantage of being easily adaptable to our needs; in particular those of a sample deposition interface. The capillary electrophoresis instrument was constructed by the author in collaboration with M. Ludwig and C.J. Barnett, with designs of the instrument drawn by B. Gauthe and LabVIEW software written by C.J. Barnett.

The instrument was characterised in terms of analytical resolution, efficiency, reproducibility and limits of detection using aniline derivatives and protein standards. The mode of separation chosen was capillary zone electrophoresis in order to avoid organic solvents, surfactants or gels to facilitate later interfacing with EVV 2DIR readout.

3.2 Capillary Electrophoresis Instrument

3.2.1 Overview of CE instrument

The capillary electrophoresis system is depicted in Figures 3.2.1 and 3.2.2 (without the sample deposition interface). The ends of a fused silica capillary are placed in glass vials (4 ml with screw top, Supelco, Bellefonte, PA, USA) filled with buffer. Platinum electrodes (0.5 mm diameter, Sigma-Aldrich, Dorset, UK) are also placed in these vials. The inlet buffer vial is manually exchanged for a sample vial for sample injection, which can be performed either hydrodynamically or electrokinetically. A reversible high voltage power supply (HW030P, Applied Kilovolts, Sussex, UK) provides voltages of up to 30 kV. A Perspex interlock safety box houses the power supply, electrodes and capillary set up. This also forms the structure of the CE system (see Figure 3.2.2) and includes a steel frame with an adjustable cross bar hous-
ing a plastic inlet vial holder with pressure injection and rinsing apparatus. An outlet vial holder is mounted in the wall of the safety box. A UV/Vis absorption detector (Model 500 variable UV/Vis absorption detector fitted with a deuterium lamp operating from 190-380 nm, Scientific Systems Inc, PA, US) is positioned approximately 11 cm from the capillary outlet. An on-column capillary cell (Model 9550-0155, Scientific Systems Inc) positions the capillary in the detector directly in the optical path such that the capillary is used as the flow cell. Thus the path length is simply the capillary inner diameter. The on-column capillary cell, fitted with the capillary, is shown in Figure 3.2.2c.

### 3.2.2 Pressure System

A pressure system was constructed for capillary rinsing and pressure injection of the sample. High pressure, between 1.5 and 2 Bar, is used for rinsing and filling the capillary. A precision pressure regulator (Model 10, Fairchild, NC, USA) with a sensitivity of below 0.3 mBar is used for hydrodynamic injection (requiring pressures in the region of 30-60 mBar) and sample deposition (requiring pressures between 1 and 20 mBar).

### 3.2.3 Instrument Control and Data Acquisition

Instrumental control and data acquisition are realised using the virtual instrumentation software, LabVIEW (National Instruments, TX, US). The software is linked to the system using data acquisition and instrumental control hardware (DAQ unit), also from National Instruments (USB-6211). The DAQ unit connects to a control unit, which forms the interface of the data acquisition and instrumental control with the main capillary electrophoresis instrument. Data are acquired every 0.2 seconds.
Figure 3.2.1 – Schematic diagram of the capillary electrophoresis system. The direction of arrows is used to indicate instrument control or data acquisition. The ends of a fused silica capillary are positioned in glass buffer vials. Platinum electrodes are placed in the vials and connected to a reversible high voltage power, supplying voltages up to 30 kV. Sample can be injected hydrodynamically or electrokinetically and vials at the inlet end are changed manually. Detection is performed using an online UV absorption detector positioned to fit the capillary detection window 11 cm from the capillary outlet.
CHAPTER 3. CE INSTRUMENTATION

Figure 3.2.2 – a) Photograph of the capillary electrophoresis system showing the main components as described in the text and in Figure 3.2.1; b) Design for vial assembly and c) Photograph of on-column capillary cell. The power supply, electrodes, capillary and vials are housed within a perspex safety cage with an interlock door which forms the structure of the system, enabling the positioning of the detector and the vial assembly.
CHAPTER 3. CE INSTRUMENTATION

3.2.4 General Experimental Methods used in CE Separations

Chemicals All chemicals used were of reagent grade or higher. 1M sodium hydroxide was purchased from Sigma-Aldrich (Dorset, UK). Ethanol was purchased from VWR (Leicestershire, UK). Buffer components (sodium phosphate, sodium citrate and ammonium acetate) were purchased from Sigma-Aldrich. Unless otherwise stated, all solutions were prepared using MilliQ water (Millipore, Milford, USA).

Capillaries Capillaries were 50-100 µm i.d. × 375 µm o.d. and were either bare fused silica with a polyimide outer coating (Composite Metal Services, Charlestown, UK), which does not have a pre-burnt detection window, or purchased with a neutral inner coating of linear polyacrylamide (LPA) (Beckman Coulter, Fullerton, CA, USA) or polyvinyl alcohol (PVA) (Agilent Technologies, Cheshire, UK), which both have pre-burnt detection windows. The LPA and PVA coated capillaries were used for protein separations as the neutral surface minimises protein adsorption. Furthermore, the neutral surface results in a negligible EOF.[122] LPA coated capillaries were used in the initial experiments described in this chapter and the change to PVA capillaries was made for later experiments (due to a much shorter delivery time from the supplier). Comparable performance was obtained, which is demonstrated by the migration time reproducibility of < 0.5% using PVA coated capillaries in Chapter 4. Using neutrally coated capillaries to prevent protein adsorption is ideal because this does not lead to unwanted components in the run buffer which would interfere with downstream readout methods (in particular, it was desirable to avoid buffer additives containing CH₂ or CH₃ groups as these would interfere with the protein fingerprinting procedure of EVV 2DIR). Additionally, the capillaries are stable over a wide pH range (pH 3-10).[122]

Capillaries were cut using a ceramic capillary cutter (Supelco, Bellefonte,
PA, USA). The inlet was examined under a microscope to ensure a straight cut to the capillary tip, as a poor cut can lead to problems with injection and thus cause zone broadening.\[123\] For bare fused silica capillaries, a detection window (approximately 5 mm long) was made at approximately 11 cm from the outlet end by burning away the polyimide coating with a hot filament (in initial experiments a cigarette lighter was used and in later experiments the Window Maker, MicroSolv, Eatontown, NJ, US was used) and cleaning with ethanol. Prior to use, new bare fused silica capillaries were cleaned by flushing (at 2 Bar) for 10 minutes with ethanol, 2 minutes with water, 10 minutes with 1M sodium hydroxide, 2 minutes with water followed by 2 minutes with the background electrolyte (BGE). For commercially coated capillaries, prior to use, capillaries were flushed for 5 minutes with water and 5 minutes with the BGE. A baseline reading was then taken with the high voltage to be used applied for 30 minutes. The high voltage baseline recording was found to be particularly important in the case of the bare fused silica capillaries as it also served as a ‘voltage conditioning’ step, to equilibrate the capillary and achieve a stable EOF.\[124\] Between runs, bare fused silica capillaries were rinsed by flushing for 2 minutes with water, 2 minutes with 1M sodium hydroxide, 2 minutes with water and 2 minutes with the BGE. Coated capillaries were rinsed between runs by flushing for 2 minutes with water and 2 minutes with the BGE.

**Background Electrolyte Preparation**  The BGEs used in this thesis are sodium phosphate, sodium citrate and ammonium acetate. A strict buffer preparation protocol was followed. Buffers were prepared using both acid and base components and the pH adjusted if necessary using a strong acid or base. Sodium phosphate buffers were prepared from sodium phosphate monobasic dihydrate (NaH$_2$PO$_4$.2H$_2$O) and sodium phosphate dibasic heptahydrate (NaHPO$_4$.7H$_2$O). Sodium citrate buffers were prepared from citric acid anhydrous (C$_6$H$_8$O$_7$) and monobasic sodium citrate anhydrous (C$_6$H$_7$NaO$_7$)
for buffers at pH 3.0 and pH 3.5 or monobasic sodium citrate anhydrous and dibasic sodium citrate sesquihydrate \((C_6H_6Na_2O_7,1.5H_2O)\) for buffers at pH 4.0. Ammonium acetate buffers were prepared from ammonium acetate \((CH_3CO_2NH_4)\) and glacial acetic acid. Buffers were prepared and used at room temperature. After preparation, buffers were degassed for 30 minutes by sonication. Preliminary experiments found that this was important to remove air bubbles which can lead to current instabilities in the CE run. This degassing step was performed for all buffers used in CE experiments in this thesis.

**Sample preparation** To perform experiments with protein standards, stock solutions at 2 mg/ml were prepared of each component in high purity water and stored at -20°C. These were then diluted for CE as appropriate and stored at 4°C. Samples for CE were prepared in a 10× dilution of the run buffer, unless otherwise stated.

**Sample Injection** Both electrokinetic and hydrodynamic injection can be performed using the system. Electrokinetic injection was performed for the experiments described in this chapter due to the instrumental simplicity. For both injection methods, following injection to avoid sample carryover the capillary inlet was immersed in a vial of MilliQ water for 1-2 s.

### 3.3 Validation of the Instrument

The performance of a CE instrument is typically assessed in terms of separation efficiency, reproducibility and detection limits.[124] High efficiency is one of the fundamental characteristics associated with capillary electrophoresis and is crucial in enabling complex mixtures containing many components potentially with small differences in mobility to be resolved. It is fundamentally important to achieve a high degree of reproducibility for validation of
a separation. To identify peaks by their migration times and compare the constituents of samples, high migration time reproducibility is essential. As discussed in Chapter 1, for a given applied voltage and capillary length, the migration time is dependent on the electroosmotic flow within the capillary and the electrophoretic mobility of the analyte. Variations in either, or both, of these will affect the reproducibility. As was discussed in Chapter 1, often low reproducibility is caused by variations in electroosmotic flow which results from factors such as changes in buffer pH, concentration, temperature and analyte adsorption. Many of the same factors affecting electroosmotic flow affect analyte mobilities.[124, 29]

Efficiencies and reproducibilities are highly dependent on the sample, CE separation conditions and the user, particularly for a home-built instrument with limited automation. Nevertheless, values for these parameters are frequently compared; with efficiencies of the order of $10^5 - 10^6$ typical [29, 22] and migration time relative standard deviations (RSDs) below 1-2% considered precise.[123]

For ribosome separations, where initial sample concentrations are low, it was necessary to have an understanding of the approximate detection limits (LOD) of the system. The LOD of a CE system can either be described in terms of the concentration or mass. However, as the concentration limit of detection depends upon the volume of sample injected, the mass limit of detection gives a better description of the system capabilities.[25] The LOD is estimated from the signal-to-noise ratio on the electropherogram and is generally taken to be the mass (or concentration) injected which achieves a 3:1 signal to noise ratio.[62, 63] LODs are determined by both the detector sensitivity and by the efficiency of the separation, because higher efficiency equates to taller, narrower peaks.[29] Thus the LOD is a property of the system and not solely of the detector. Additionally, the LOD in these experiments will also depend on the absorbance of analyte. Approximate detection limits for absorbance detection in CE are $10^{-13} - 10^{-16}$ moles.[29]
Initial validation of the instrument was performed using three aniline derivatives: aniline, N-methylaniline and N,N-dimethylaniline. These are small organic molecules and are well suited to the initial testing of a CE system as they circumvent some of the difficulties associated with protein separations such as adsorption and sample degradation. Aniline, N-methylaniline and N,N-dimethylaniline have molecular weights of 93.13, 107.15 and 121.18 respectively and pKas of 4.60 [125], 4.85 [125] and 5.15 [126] respectively. Following this, validation of the instrument was performed for a sample of three standard proteins, lysozyme from hen egg white (Sigma-Aldrich), trypsinogen from bovine pancreas (Sigma-Aldrich) and α-chymotrypsinogen A from bovine pancreas (Sigma-Aldrich). The isoelectric points and molecular masses, as stated by the supplier, are 11.4 and 14.3 kDa [127] for lysozyme, 9.3 and 23.7 kDa [128] for trypsinogen and 9.0 and 25.6 kDa [129] for α-chymotrypsinogen A.

Validation with these sets of molecules were used to assess if separations achieving high efficiencies, resolution and reproducibility could be performed on the developed instrument. Efficiencies were calculated using the peak full width at half maximum (FWHM). The aniline derivatives, aniline, N-methylaniline and N,N-dimethylaniline, were baseline resolved by capillary zone electrophoresis (Figure 3.3.1) with efficiencies obtained between $204 - 520 \times 10^3$ plates/m for the three components. Migration time reproducibility was achieved with a precisions of <0.54% for a series of 6 runs. These results are displayed in Table 3.3.1. Separations of model proteins, lysozyme, trypsinogen and α-chymotrypsinogen A were also baseline resolved under the conditions explored (Figure 3.3.2) achieving high efficiencies of $229 - 818 \times 10^3$ plates/m for the three components. Migration time precision was also extremely high, achieving precisions of <0.35% over 6 runs. These results are displayed in Table 3.3.2. The increase in efficiency with migration time observed for the three proteins (Figure 3.3.2) is explained by the electrokinetic injection, which results in the injection of smaller sample
Figure 3.3.1 – Separation of aniline, N-methylaniline and N,N-dimethylaniline, showing 3 consecutive repeats. The peaks were identified as: 1. aniline, 2. N-methylaniline, 3. N, N-dimethylaniline. The sample contained concentrations of 11 µM aniline, 9 µM N-methylaniline and 8 µM N,N-dimethylaniline. The sample was injected at 5 kV for 6 s in a 50 µm i.d. bare fused silica capillary with a total length of 60 cm and separated at 30 kV using a pH 3 sodium phosphate buffer. t = 0 corresponds to the start of the high voltage.
Figure 3.3.2 – Separation of lysozyme, trypsinogen and α-chymotrypsinogen, showing 3 consecutive repeats. The peaks were identified as: 1. lysozyme, 2. trypsinogen, 3. α—chymotrypsinogen. Proteins were used at a concentration of 40 µg/ml. The sample was injected at 5 kV for 9 s in a 50 µm i.d. LPA capillary with a total length of 55.5 cm and separated at 20 kV using a pH 3 sodium phosphate buffer.
### Table 3.3.1 – Summary of performance parameters for CE separations of aniline derivatives. *Run-to-run reproducibility, n=6

<table>
<thead>
<tr>
<th></th>
<th>Aniline</th>
<th>N-methylaniline</th>
<th>N,N-dimethylaniline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency</td>
<td>520 ± 29</td>
<td>204 ± 9</td>
<td>348 ± 19</td>
</tr>
<tr>
<td>Migration time precision (RSD)*</td>
<td>0.54%</td>
<td>0.44%</td>
<td>0.48%</td>
</tr>
</tbody>
</table>

### Table 3.3.2 – Summary of performance parameters for CE separations of lysozyme, trypsinogen and α–chymotrypsinogen. *Run-to-run, n=6

<table>
<thead>
<tr>
<th></th>
<th>Lysozyme</th>
<th>Trypsinogen</th>
<th>α–chymotrypsinogen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency</td>
<td>229 ± 24</td>
<td>544 ± 20</td>
<td>818 ± 10</td>
</tr>
<tr>
<td>Migration time precision (RSD)*</td>
<td>0.25%</td>
<td>0.28%</td>
<td>0.35%</td>
</tr>
</tbody>
</table>

quantities for analytes of lower mobility. Correspondingly, decreasing peak heights can be observed in the electropherogram (Figure 3.3.2) due to the mobility bias introduced through electrokinetic injection.

The values for performance characteristics are well matched with published values obtained for similar capillary zone electrophoresis separations on commercial systems for both aniline [130, 131] and protein separations [132, 133] in terms of efficiency and migration time reproducibility. Migration time reproducibility was lower when measured on a day-to-day basis (<3.5%), which is a common trait of CE systems, in particular, those that operate without temperature control and can be explained by fluctuations in the ambient temperature which result in a change in the viscosity of the run buffer.[134, 135, 136, 22] To ensure high reproducibility using this system, all future experiments consisting of a set of runs were completed within one day.
CHAPTER 3. CE INSTRUMENTATION

The extremely high intra-day migration time reproducibilities were achieved through careful operation and control of run conditions. In particular, capillary preparation and inter-run rinsing were found to be crucial in minimising error arising due to fluctuations in EOF.[137] The migration time reproducibilities calculated based on a series of 6 runs were improved from RSDs for initial experiments of over 6% to less than 0.5%, reflecting a well established protocol which generated a stable EOF and migration times for the aniline experiments (Figure 3.3.1), and for the protein experiments, using a neutral LPA capillary which eliminated EOF and prevented protein adsorption (Figure 3.3.2). Furthermore, buffer vials were replenished with 0.5 ml of fresh buffer before each run to improve reproducibility by preventing buffer depletion and to avoid siphoning due to uneven buffer levels.

The detection limit of the system was estimated by injecting a series of lysozyme samples of descending concentration from 1 \(\mu\)g/ml to 0.25 \(\mu\)g/ml (Figure 3.3.3). The data shows non-linearity between concentration and peak height, which indicate potential low precision in electrokinetic sample injection. Nevertheless, the purpose of these experiments were to provide an estimate of the concentration LOD, and this can be established from this data. The concentration LOD was calculated to be \(0.26 \pm 0.02 \mu\)g/ml, which equates to a mass LOD of 460 amol and \(2 \times 10^8\) molecules of lysozyme.

3.4 Conclusions

In this chapter the construction and operation of the capillary electrophoresis system that forms the core of the work presented in this thesis are described. The instrument operates as a standard CE instrument but with the additional flexibility of one custom made. Validation experiments performed with aniline derivatives and a selection of model proteins yielded high efficiencies with a maximum of \(818 \times 10^3\) plates/m for \(\alpha\)-chymotrypsinogen, and excellent migration time precision (run-to-run) with RSDs of <0.35% for the protein
Figure 3.3.3 – Electropherograms for lysozyme injected at descending concentrations. A 100 µm i.d. LPA capillary was used as the highest LOD is achieved with the longest pathlength. The total length was 44 cm with an effective length of 33 cm. Sample was injected at 5 kV for 9 s and separated at 20 kV using a pH 3 sodium phosphate buffer.
samples. These experiments demonstrate that the performance for capillary
zone electrophoresis is comparable in terms of efficiency and migration time
reproducibility with a commercial instrument (through comparison to liter-
ature values). This enables it to be operated successfully both as a standard
CE instrument, as is demonstrated with the separation of the highly complex
ribosomal protein sample in Chapter 4, or with the track-writing interface
and the coupling of identification techniques, as is described in Chapters 5,
6 and 7.
Chapter 4

Separation of Ribosomal Proteins using Capillary Electrophoresis
CHAPTER 4. CE OF RIBOSOMAL PROTEINS

4.1 Introduction

In Chapters 2 and 3, the development of a capillary electrophoresis system and ribosome affinity purification (RAP) are described. Here these two techniques are brought together with the goal of separating out the ribosomal proteins. This forms a novel high throughput and robust separation technique for the ribosomal proteins and is the first reported separation of eukaryotic ribosomal proteins using capillary electrophoresis. Affinity ribosome purification couples extremely well to the separation method of capillary electrophoresis as both have small sample requirements and are high throughput methods. Using affinity purification as the first dimension of separation, as opposed to a sucrose gradient, offers increased flexibility, facilitating prospects such as the analysis of copurifying proteins. The potential to scale down purification will be essential in future studies where, for example, growth conditions may be changed in several samples and the results compared.$^{1}$

The proteomics workflow of ribosome affinity purification-capillary electrophoresis (RAP-CE) can be considered as an alternative top-down proteomic approach to the more commonly used workflows, which use sucrose gradient purified ribosomes and various forms of gel electrophoresis with mass spectrometry.$^{[138, 139, 140]}$ Although some great successes have been had using these methods, for many ribosomal studies the demand is for automated methods with a higher degree of reproducibility. Accordingly, liquid chromatography methods with mass spectrometry have become increasingly popular. However, only a limited number of separations on intact ribosomal proteins have been reported, most likely because mass spectrometry instrumentation (fourier transform ion cyclotron resonance (FT-ICR), hybrid ion trap FT-ICR or hybrid ion trap-orbitrap [18]) for top-down proteomic analyses are not yet common-place and are expensive to purchase and maintain. Lee et al.$^{[121]}$ successfully used capillary reversed-phase liquid chromatog-

$^{1}$This work is in preparation for the journal Journal of Chromatography A.
raphy (RPLC)/FT-ICR to identify 42 out of the 43 core yeast 60S subunit proteins. Additionally, a two-dimensional HPLC system, comprising ion-exchange followed by RPLC, was used by Cohen et al [141] to separate yeast ribosomal proteins, with >80% of the proteins in the complex identified using electrospray ionisation (ESI)-MS. Capillary electrophoresis has been infrequently used to study ribosomal proteins. Moini et al [142] separated out intact proteins from a bacterial ribosome with an ESI-MS interface enabling 55 out of the 56 ribosomal proteins to be resolved in two runs. The same research group also analysed a tryptic digest of ribosomal proteins using a novel technique to improve sequence coverage, (CE-MS/MS)$^n$, whereby multiple injections and analyses were performed in one run.[16]

Perhaps the paucity of studies of ribosomal proteins using capillary electrophoresis is because of predicted difficulties in both separation and sample preparation. The separation of ribosomal proteins using capillary electrophoresis is challenging because of the relatively large number of proteins that exist within a small range of pI values, and the fact that separation in free solution is based on the size to charge ratio. However, capillary electrophoresis is an excellent separation technique for resolving complex protein mixtures.[142, 143] With careful optimisation of sample preparation and separation conditions, the advantages of small sample consumption, short separation times and exceptional efficiency can be exploited for the yeast ribosome.

To achieve the CE separation of ribosomal proteins, the separation of a more complex mixture of model proteins within a similar mass and pI range to the ribosomal proteins was first optimised. A sample preparation method was developed to reproducibly dissociate the ribosome and remove the rRNA. Furthermore, in this chapter, preparations are made to couple the ribosome separation to deposition onto a substrate by investigating methods to increase the loading capacity, which is complicated for a dilute, multi-component system.
4.2 Materials and Methods

4.2.1 Method Development for Ribosomal Protein Separations: Separation of a 6 Protein Mixture

Capillaries

Capillaries used were 50 μm i.d. × 375 μm o.d polyvinyl alcohol (PVA) coated fused silica with polyimide outer coating (Agilent Technologies) trimmed to 44 cm total length and 33 cm effective length.

Sample and buffer preparation

Stock solutions of each protein at 2 mg/ml were prepared in MilliQ water and stored at −20°C. From these, mixtures of the proteins at 100 μg/ml in a 10× dilution of the run buffer were made and stored at 4°C. Peaks were identified by running samples with each component added in turn. 20 mM sodium citrate at pH 3.0, pH 3.5 and pH 4.0 was made by mixing appropriate proportions of either citric acid and monobasic sodium citrate (pH 3.0 and pH 3.5) or monobasic sodium citrate and dibasic sodium citrate (pH 4.0) and making up to the desired volume with MilliQ water.

Sample Injection and CE separation

Sample was injected hydrodynamically at 60 mBar for 3 s, injecting approximately 7 nl of sample (calculated using Equation 1.2.24). Hydrodynamic injection was chosen to avoid the mobility bias of electrokinetic injection. Separation was performed with an applied voltage of 20 kV.
4.2.2 Separation of Ribosomal Proteins

Sample preparation

A sample preparation method (Figure 4.4.3) was developed through performing runs of CE on sample after three successive stages of preparation: incubation of the ribosome sample with RNase, precipitation of the RNA with acetic acid and finally dialysis with a 10 times dilution of the BGE. The final preparation procedure, thereafter used throughout this thesis, was as follows: 25 µl of affinity purification eluate was incubated for 15 minutes at 37°C with 2.4 µg/ml RNase A (Sigma-Aldrich, molecular biology grade) and then placed immediately on ice. Following this, RNA precipitation was performed as a modification of work by Hardy et al [144], by adding 40 µl of ice cold glacial acetic acid and centrifuging at 12000×g for 6 minutes at 4°C. The RNA pellet was usually visible on the side of the tube. The liquid phase was then transferred to a 500 µl tube and diluted with 60 µl ice cold MilliQ water. This was then dialysed with 2 mM sodium citrate, pH 4 for 3 ½ hours with 4 changes of buffer at 4°C (Slide-A-Lyser Mini Dialysis Units with Microtubes MWCO = 7 kDa, Thermo Scientific Pierce Protein Research Products, Rockford, IL, USA) to give the final sample.

Sample Injection and CE separation

Unless otherwise stated, all materials and methods are as described in Section 4.3. The initial run buffer used was 20 mM sodium citrate, pH 4.0. Sample was injected at 60 mBar for 15 s. Separation was performed at 20 kV.

MALDI-MS/MS of ribosomal protein sample  Following capillary electrophoresis, the sample was analysed by liquid chromatography matrix assisted laser desorption ionisation-time of flight (MALDI-TOF)/TOF-MS in the Mass Spectrometry Centre for Integrative Systems Biology at Imperial College London. Briefly, the fully prepared sample was desalted and trypsin
digested. Fragments were then separated by reversed-phase HPLC and robotically spotted onto a metal MALDI target using an Ultimate 3000 nanoLC system (Dionex, Sunnyvale, CA, USA) coupled to a Probot automated spotting device (LC Packings, Sunnyvale, CA, USA). MALDI-MS/MS was performed using an Applied Biosystems 4800 TOF/TOF. All MS/MS data were analysed using the Mascot search engine (Matrix Science, London, UK). Mascot was set up to search the SwissProt \textit{Saccharomyces Cerevisiae} database (UnitProt release 2010_09) for the digestion enzyme Trypsin. The database was searched with a fragment ion mass tolerance of 0.30 Da and a parent ion tolerance of 100 ppm. Oxidation of methionine residues and carboxymethylation of cysteine residues were specified as variable modifications. Scaffold (version 3.0, Proteome Software Inc., Portland, OR, USA) was used to validate MS/MS based protein identifications. Protein identifications stated have greater than 20.0\% probability and contain at least one identified peptide. Protein probabilities were assigned by the Protein Prophet algorithm.[115]

### 4.2.3 Increase in Ribosome Sample Loading

Three methods are discussed: increasing the diameter of the capillary to a 100 µm capillary, increasing the sample volume injected in a 50 µm capillary and head-column field amplified sample stacking (HC-FASS). All experiments were conducted using the full sample preparation procedure. The affinity purified ribosome sample used in these experiments had a starting RNA concentration of 862 µg/ml.

For experiments with an increased capillary i.d., ribosome separations were performed in a 100 µm capillary with a 71 cm total length and 60 cm effective length. Sample was injected at 60 mBar for 6 s and separated at 12 kV. Four repeats were performed. To determine the appropriate voltage, an Ohm's law plot was made by increasing the voltage in increments of 1 kV from 5 kV to 15 kV and recording the current at 45 s after application of the voltage.
For experiments to increase the injection volume with FASS, separations were performed in 50 µm capillaries with a 71 cm total length and 60 cm effective length. Experiments were performed by increasing the injection time incrementally until peaks were no longer resolvable for 20 mM, 30 mM and 40 mM sodium citrate pH 4.0 BGEs. Injection was performed at 60 mBar and increased in increments of 24 s from 24 s to 120 s for 20 mM sodium citrate and increments of 48 s from 72 s to 168 s for 30 mM sodium citrate and 168 s to 216 s for 40 mM sodium citrate. Separation was performed at 25 kV. Three repeats were performed under each set of conditions.

For electrokinetic injection with HC-FASS, separations were performed in a 50 µm capillary with a 71 cm total length and 60 cm effective length. A plug of 2 mM sodium citrate, pH 4 was hydrodynamically injected at 60 mBar for 15 s, following which sample was injected at 4.3 kV, 2 s (for an injection of 10^8 molecules); 10 kV, 8.6 s (10^9 molecules); 10 kV, 86 s (10^{10} molecules) or 15 kV, 287 s (5×10^{10} molecules). Separations were performed with an applied voltage of 20 kV. Three repeats were performed under each set of conditions.

4.3 Method Development for Ribosomal Protein Separations: Separation of a 6 Protein Mixture

To develop a method for the separation of ribosomal proteins, and demonstrate the capabilities of the system for the separation of a more complex protein mixture, six proteins were selected that possess isoelectric points and masses similar to those of the ribosomal proteins. These proteins were basic and had molecular weights in the range of 10-50 kDa. The proteins chosen were as follows: Lysozyme (pI 11.4, MW 14.3 kDa), Trypsinogen (pI 9.3, MW 23.7 kDa), α-chymotrypsinogen A (pI 9.0, MW 25.6 kDa), Cytochrome
Figure 4.3.1 – Separation of lysozyme, trypsinogen and α-chymotrypsinogen, cytochrome C, RNase A and myoglobin. The peaks were identified as: 1. lysozyme, 2. trypsinogen, 3. α-chymotrypsinogen, 4. cytochrome C, 5. RNase A and 6. myoglobin. Proteins were used at a concentration of 100 µg/ml. Three consecutive repeats are shown at increasing BGE pHs of 20 mM sodium citrate, where sample was injected at 5 kV for 9 s and separated at 20 kV in a 50 µm i.d. PVA capillary with a total length of 44 cm.
C (pI 9.6, MW 11.7 kDa), RNase A (pI 8.7, MW 13.7 kDa) and Myoglobin (pI 6.8/7.2, MW 17.0 kDa).

A run buffer of 20 mM sodium citrate, pH 3.0 was initially used and all 6 proteins were resolved but not to baseline. To optimise the resolution of the separation, the pH was increased in increments of 0.5 units until all 6 proteins were baseline resolved at pH 4.0, as is demonstrated in Figure 4.3.1. This strong dependency of resolution on buffer pH is demonstrated by Equation 4.3.1 (as described in Chapter 1 or 2), where $R$ is the resolution between two species, $\Delta \mu_{ep}$ is the difference in mobility between the two species, $\mu_{ep}$ is their average mobility, $\mu_{eo}$ is the electroosmotic flow mobility and $N$ is the number of theoretical plates.[25] Adjusting the buffer pH to closer to the sample pI values will increase $\Delta \mu_{ep}$, thus providing the demonstrated resolution.

$$R = \frac{1}{4} \frac{\Delta \mu_{ep} N^{\frac{1}{2}}}{\mu_{ep} + \mu_{eo}}$$ (4.3.1)

The optimised conditions were then used as a starting point for CE conditions for ribosomal separations. These experiments demonstrated the capacity of the CE system for separating a more complex sample and that separations can be manipulated through optimisation of separation parameters. These are factors that are essential for separation of a real biological sample such as the ribosomal proteins.

### 4.4 Separation of Ribosomal Proteins

The separation of ribosomal proteins using CE is challenging because of the number of proteins, the relatively small range in size, pI and other similarities which will arise as a result of all proteins originating in the same cellular molecular environment (almost all are nucleic acid binding proteins, for example). Using the CE set-up developed and with optimisation of buffer conditions, ribosome purification and sample preparation conditions, over 26
Figure 4.4.1 – Separation of ribosomal sample after digestion with RNase and treatment with acetic acid and dialysed. Sample was injected at 60 mBar for 15 s in a 50 µm i.d. PVA capillary, total length 44 cm and separated at 20 kV.
peaks were resolved (Figure 4.4.1). To the best of my knowledge, this is the first time that proteins from the eukaryotic ribosome have been separated using CE.

Based on the series of 7 consecutive runs, high reproducibility of the separation was observed, with RSD migration times < 0.5%. With this degree of reproducibility it is evident that the method is not only high throughput, but also gives quantifiable data that can be compared across experiments. The profile was found to be reproducible across batches of affinity purified ribosomes (Figure 4.4.2), providing a ‘fingerprint’ for ribosomal proteins prepared under these conditions.

Considering the entire workflow, this can be viewed as affinity chromatography/capillary electrophoresis. The two stages in this workflow, affinity chromatography and capillary electrophoresis, share the advantages that they are fast, flexible and have small sample requirements in comparison to more commonly used techniques. Furthermore, the use of ribosome affinity purification rather than sucrose gradients, combined with the outstanding CE RSDs achieved, paves the way for future studies with modified cell growth conditions (such as dietary restriction, which has been observed to correspond with a reduction in 60S subunits [145]) and probing of the attached ribosome on the affinity matrix before elution to see alterations in the CE profile. Therefore, this method is a remarkably quick route from cell to separation, providing a novel strategy for high throughput studies of the ribosomal proteome.

A crucial stage in this protocol was the preparation of ribosome sample prior to CE to ensure that ribosomes are reproducibly dissociated into ribosomal proteins. The intention of the first stage of the ribosome sample preparation method is to use RNase to cleave the rRNA which maintains the structure of the ribosome. Experiments without this step were performed and no peaks were present during a 60 minute run (Figure 4.4.3a). Following incubation of the sample with RNase, over 20 peaks are baseline resolved
Figure 4.4.2 – Separation of ribosomal sample after digestion with RNase and treatment with acetic acid and dialysed, performed 25 days later under identical conditions to Figure 4.4.1 but using a different affinity purified ribosome sample. Sample was injected at 60 mBar for 15 s in a 50 µm i.d. PVA capillary, total length 44 cm and separated at 20 kV.
but separations were irreproducible (Figure 4.4.3b). The acetic acid RNA precipitation step is essential to then remove rRNA from the sample and improve the reproducibility. Separations performed following this step had an improved reproducibility but poor resolution and efficiency (Figure 4.4.3c). Dialysis into a 10 times dilution of the BGE enabled the high efficiencies, resolution and reproducibility of the final result to be achieved (Figure 4.4.1). To establish the preparation method for the ribosome sample, a series of runs were performed after each of the steps in the preparation (Figure 4.4.3).

Mass spectrometry analysis of the ribosomal protein sample following capillary electrophoresis confirms that ribosomal proteins were injected into the capillary (see Section 4.7 Appendix A, Table 4.7.1). Simply, following CE runs the remaining sample as prepared and used for CE experiments was analysed by LC-MS/MS. Out of 54 hits above 20%, 39 were ribosomal proteins. With a stringent cutoff of 100%, 19 out of 22 were ribosomal proteins. Hence it is reasonable to assume that the separation is indeed a separation of almost entirely ribosomal proteins. A small number of non-ribosomal proteins were identified in the mass spectrometry data and could account for peaks seen, many of which were highly abundant glycolytic enzymes. However, the molecular weights and pI values (as predicted with ExPasy [146]) of these proteins are such that, with the exception of phosphoglycerate mutase 1 (MW 27.5 kDa, pI 8.86), it is likely that these proteins would migrate much later than the bulk of the ribosomal proteins. Corresponding with MS results of the raw sample prior to CE (Chapter 2), once again, regulatory components of the ribosome are also present, notably guanine nucleotide-binding protein subunit beta-like protein (also known as Asc1p) with a protein identification probability of 100%. It is important to note that MS did not identify the presence of all of the yeast ribosomal proteins. The missing proteins may be due to co-precipitation with RNA, although two acidic proteins were identified. It is more likely that the missing ribosomal proteins were simply not identified due to the relatively low concentration of the final sample used for
Figure 4.4.3 – CE separation of the ribosomal sample at various stages of the preparation procedure: a) Without sample preparation b) After digestion with RNase c) After digestion with RNase and treatment with acetic acid. Sample was injected at 60 mBar for 15 s in a 50 µm PVA capillary, total length 44 cm and separated at 20 kV.
MS, which was approximately 2 µg/ml per protein. However, the purpose of incorporating mass spectrometry analysis at this stage in the methodology was not to assess the affinity purification method, but rather to verify that ribosomal proteins were being injected into the capillary. The same MALDI-MS/MS analysis performed on eluate after affinity purification (but prior to preparation for CE) identifies 61 ribosomal proteins out of 73 hits above 53% probability, and, 41 ribosomal proteins out of 47 hits above a cutoff of 100%, including Asc1p (Chapter 2). This may be due to the higher concentration of the sample.

To rule out the possibility of peaks being due to the FLAG peptide, controls were performed with the FLAG peptide (Figure 4.4.4). CE runs under identical conditions to those use for ribosomal separations were performed for the FLAG peptide on its own at 40 µg/ml, which is in excess of 5 times the maximum concentration of FLAG peptide that could be contained within the sample in the unlikely event that no FLAG peptide had bound (which would have prevented ribosomes from eluting during the affinity purification). No peaks were observed for the FLAG peptide.

It should be noted that some single peaks in the electropherograms shown in Figure 4.4.1, may contain more than one protein. The use of laser induced fluorescence detection, which can achieve unrivalled detection limits, would enable injection of smaller sample plugs and is likely to result in the resolution of a greater number of peaks. Mass spectrometry identification with reconstruction of the electropherogram would also enable a greater number of peaks to be resolved.[147, 148] In later work (see Chapter 5), for track writing of the separation, the separation was performed with a volatile BGE (50 mM ammonium acetate, pH 4.0) which demonstrates the potential for CE-ESI-MS of eukaryotic ribosomal proteins.
Figure 4.4.4 – Control with FLAG peptide. FLAG peptide was diluted to 40 µg/ml in 2 mM sodium citrate and injected at 60 mBar for 15 s in a 50 µm PVA capillary, total length 44 cm and separated at 20 kV. The first window shows a baseline check, followed by 3 consecutive runs. The baseline check is displayed to demonstrate that the slight drop in absorbance at ~10 minutes occurs in the baseline check.

4.5 Increase of Sample Injection Quantities for Readout Methods

In Figure 4.4.1, I demonstrate the separation of eukaryotic ribosomal proteins using CE for the first time. One of the aims of this work has been EVV 2DIR readout of ribosomal proteins separated by CE. The sensitivity of the EVV-2DIR system in the Klug laboratory is between $10^{11}$ and $10^{12}$ protein molecules.[1] In the electropherograms displayed in Figure 4.4.1, sample has been injected at 60 mBar for 15 seconds, corresponding to the injection of 35 nl of sample. The number of molecules of an individual ribosomal protein injected can be estimated from the initial RNA concentration of the affinity purified ribosome sample. The RNA concentration in the affinity purified sample was measured at 1282 ng/µl (Nanodrop Spectrometer, Nanodrop, Wilmington, DE, USA). Based on the premise that rRNA accounts for 60%
of the mass of the ribosome [79], the concentration of protein can therefore be approximated to be 855 ng/µl. As there are 79 yeast proteins, this equates to a concentration of 11 ng/µl per protein in the initial sample. In the ribosome preparation for CE, a fivefold dilution was performed, giving the concentration of sample injected as 2 ng/µl per protein. The average mass of a ribosomal protein can be estimated as 20.5 kDa, based on a yeast ribosomal mass of 4 MDa, a protein content of 40% [79] and 79 proteins [80]. This gives an estimate of $2 \times 10^9$ molecules of each ribosomal protein injected into the capillary.

There is a disparity of between 2 and 3 orders of magnitude between the amount of sample injected and the sensitivity of EVV 2DIR spectroscopy. At the time of performing the work, it was thought that a sensitivity of $10^{10}$ protein molecules would be achievable within the timescale of this PhD. Unfortunately, due to experimental difficulties, no increase in EVV 2DIR sensitivity was accomplished; however, it is hoped that in the near future these targets will be met and analysis of complex mixtures from CE with EVV 2DIR spectroscopy may be possible. Aiming to tackle this disparity from the perspective of the CE, which would be accompanied by an increase in the sensitivity of the 2DIR spectrometry, methods to increase the loading of ribosomal proteins were explored.

Sample concentration techniques are the immediate solution, and this can be performed by a number of methods. However, concentration of the sample post-preparation for CE using a vacuum concentrator (Eppendorf Concentrator Plus, Eppendorf, Hamburg, Germany) caused sample degradation, which was recognised as irreproducible electropherograms with no resemblance to the fingerprint electropherogram. A concentration increase of approximately two fold was achieved by scaling at the affinity purification stage, which did not affect CE profiles.

Increasing the inner diameter of the capillary enables a larger volume of sample to be injected for a given length of sample plug. This method
was investigated by using a 100 µm capillary. Injection conditions were chosen to inject a plug of equal length to that used in the separations of Figure 4.4.1, leading to a 4 times increase in the volume of sample injected. Using this method, $6 \times 10^9$ molecules were injected and separations showed a reproducible profile (see Figure 4.5.1) showing the characteristic features of the ‘fingerprint’ electropherogram shown in Figure 4.4.1. However, to avoid excessive Joule heating, an Ohm’s law plot was performed which determined a maximum separation voltage of 12 kV and this resulted in a slow separation time of $\sim 42$ minutes.

Online sample concentration techniques enable the use of narrow capillary bores with the injection of large sample volumes, as were discussed in Chapter 1. A number of other online preconcentration methods, although frequently used to achieve a high degree of concentration, were not considered appropriate because of constraints on sample matrix and BGE which would either require radical changes (transient ITP or dynamic pH junction) or the use of reagents not compatible with 2DIR (such as in the case of sweeping).
The simplest method of online concentration, field amplified sample stacking (FASS), was performed in all ribosome separations by performing hydrodynamically injection of the ribosomal protein sample dialysed into a 10 times dilution of the BGE.

FASS resulted in interesting effects when increasing the injection time of ribosome separations, as it was found that during increased injection times migration times decreased and resolution was lost. As sample is stacked further in front of the boundary between the injection zone and the BGE, as the injected plug became increasingly long, sample began the separation further along the capillary and therefore had a shorter distance to migrate for longer injection times. However, by increasing the concentration of the BGE, the mobility of the proteins were reduced and resolution improved, enabling large volumes of sample to be hydrodynamically injected and separated up to a plug length of 16 cm.

The pattern of decreasing migration times at longer injection times is shown in Figure 4.4.1. Hydrodynamic injection was performed at 60 mBar for 24 s (corresponding to the injection of $1.5 \times 10^9$ molecules), and the injection time then increased in 24 s intervals up to 120 s (corresponding to the injection of $8.9 \times 10^9$ molecules) with runs performed in triplicate for each set of injection conditions. Migration times were stable for each set of 3 runs (1 run from each set shown in Figure 4.5.2), but as injection times were increased, migration times were shifted towards decreasing values.

Increasing the concentration of the BGE increased migration times and hence improved the resolution (Figure 4.5.3), thus enabling a greater volume of sample to be injected. The proteins are surrounded by a double layer of counterions. Migration of the counterions are in the opposite direction to the direction of the migration of the protein, and hence an increased buffer concentration increases the drag on the protein and reduces its mobility, thus improving resolution.[25] The same sample was injected in 30 mM and 40 mM sodium citrate buffers, pH 4.0 with conditions as described for the
Figure 4.5.2 – Increasing the sample plug hydrodynamically injected. Sample was injected at 60 mBar for 24 s (injecting $1.5 \times 10^9$ molecules), 48 s ($3.0 \times 10^9$ molecules), 72 s ($4.5 \times 10^9$ molecules), 96 s ($6.0 \times 10^9$ molecules) and 120 s ($7.6 \times 10^9$ molecules). A BGE of 20 mM sodium citrate, pH 4.0 was used. Separations were performed in a PVA capillary with a total length of 71 cm and a 50 μm i.d at 25 kV. Each condition was performed in triplicate and stable migration times were observed (1 run shown for each).
Figure 4.5.3 – Increasing the buffer concentration enables the peaks to be resolved when a greater volume of sample is hydrodynamically injected. In the electropherograms shown, at the top, sample was injected at 60 mBar for 120 s (injecting $1.5 \times 10^9$ molecules) and a 20 mM sodium citrate, pH 4.0 BGE was used, in the middle, sample was injected at 60 mBar for 168 s (injecting $1.1 \times 10^{10}$ molecules) and a 30 mM sodium citrate, pH 4.0 BGE was used and in the third, sample was injected at 60 mBar for 216 s (injecting $1.4 \times 10^{10}$ molecules) and a 40 mM sodium citrate, pH 4.0 BGE was used. Separations were performed in a PVA capillary with a total length of 71 cm and a 50 µm i.d at 25 kV.
20 mM sodium citrate separations and separations performed in triplicate. For 30 mM sodium citrate, injection was performed at 72 s, 120 s, and 168 s and using 40 mM sodium citrate, injections were performed for 168 s and 216 s. At an injection of 216 s, \(1.4 \times 10^{10}\) molecules were injected in a 16 cm plug, yet the peaks in the ribosomal electropherogram are resolved and the electropherogram is reproducible. A 50 mM sodium citrate was also tested with injection times of 216 s and 264 s, but the ribosomal CE profile was not maintained. It is possible that by a concentration of 50 mM, Joule heating became significant.

Head column-field Amplified Sample Stacking (HC-FASS) was also performed, by performing hydrodynamic injection of a plug of 2 mM sodium citrate prior to electrokinetic sample injection. This method enabled injections of up to \(10^{10}\) molecules, but better resolution was obtained with FASS. Furthermore, for large injections the sample bias introduced by electrokinetic injection is substantial and would severely complicate any studies into protein stoichiometries.

The challenge is to maintain the separation of the ribosomal proteins whilst loading the levels required for downstream readout methods. Studies found that increasing the injection time whilst performing FASS with a higher buffer concentration (40 mM sodium citrate) was the most effective way to increase the amount of sample loaded and maintain the ribosomal protein separation. Using this method, \(10^{10}\) molecules of each ribosomal protein can be loaded and separated.

4.6 Conclusions

In this Chapter the first use of capillary electrophoresis for a separation of yeast ribosomal proteins was demonstrated. Over 26 peaks were resolved with outstanding reproducibility and efficiency. The challenge facing the separation of ribosomal proteins was the sheer number of proteins many of
which fall within a small range of molecular weights and are highly basic. In addition, the ribosomal proteins are bound within a complex containing large amounts of RNA. A crucial stage in this protocol was the development of the preparation method for the ribosomal sample prior to CE, which ensured that ribosomes were reproducibly dissociated into ribosomal proteins. This used ribonuclease A to cleave the rRNA strands followed by acetic acid to precipitate the RNA and dialysis to optimise the separation.

The ribosomes used were purified using affinity chromatography, which is both simple and efficient and thus it is an ideal match with the advantages of CE, creating a novel top-down workflow for preparation and separation of ribosomal proteins. The readout provides a ribosomal protein fingerprint, from which a wide variety of applications are possible. The high level of reproducibility inherent in this method in terms of migration time and peak height suggests application in studies of proteins exhibiting extraribosomal functions and thus absent from the ribosome, substoichiometries, posttranslational modifications resulting in shifting migration times and fluctuations resulting from changes in growth conditions. To study these, the use of laser induced fluorescence detection would be required to enable the injection and detection of reduced sample quantities. The affinity purification method results in the co-purification of proteins interacting with the ribosome. As the entire method is of high throughput and sample requirements are low, this would provide an ideal platform to scan for ribosome interacting proteins. Output to an identification technique would then create a novel platform for fast analyses of the ribosomal proteome. The most obvious next step will be to interface with an identification technique, such as mass spectrometry. For an electrospray-ionisation interface, a volatile BGE would be required and accordingly we have demonstrated application of an ammonium acetate buffer. The yeast ribosomal protein separation could also be applied to CE-MALDI interfaces.

Developments were made to increase the loading of the ribosomal sample
in preparation for readout by EVV 2DIR spectroscopy following deposition onto a metallic surface. Increasing the inner diameter of the capillary, increasing the concentration of the buffer and performing head-column field amplified sample stacking were successful techniques, enabling $10^{10}$ molecules to be injected whilst maintaining the resolution, which was not possible by increasing the injected volume alone.

### 4.7 Appendix A

**Table 4.7.1** – Mass spectrometry results for the CE sample used in Figure 4.4.1. Mascot used MS/MS data to identify proteins from the SwissProt *Saccharomyces cerevisiae* database. Protein identifications were validated using Scaffold. Protein identifications shown have greater than 20.0% protein identification probability and contain at least one identified peptide.

<table>
<thead>
<tr>
<th>Identified Proteins (S. Cerevisiae)</th>
<th>Molecular Weight (kDa)</th>
<th>Protein Identification Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>60S ribosomal protein L8-B</td>
<td>28 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S7-A</td>
<td>22 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S3</td>
<td>27 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L4-A</td>
<td>39 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S22-A</td>
<td>15 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L20</td>
<td>20 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S16</td>
<td>16 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L17-A</td>
<td>21 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S20</td>
<td>14 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L30</td>
<td>11 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L14-A</td>
<td>15 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>60S ribosomal protein L6-B</td>
<td>20 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>Phosphoglycerate mutase 1</td>
<td>28 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S0-A</td>
<td>28 kDa</td>
<td>100%</td>
</tr>
</tbody>
</table>
### Identified Proteins

*(*S. Cerevisiae*)

<table>
<thead>
<tr>
<th>Identified Proteins</th>
<th>Molecular Weight (kDa)</th>
<th>Protein Identification Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enolase 2</td>
<td>47 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>Guanine nucleotide-binding protein subunit beta-like protein</td>
<td>35 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>Phosphoglycerate kinase</td>
<td>45 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S19-A</td>
<td>16 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S5</td>
<td>25 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S21-B</td>
<td>10 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S29-A</td>
<td>7 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S10-A</td>
<td>13 kDa</td>
<td>100%</td>
</tr>
<tr>
<td>40S ribosomal protein S11</td>
<td>18 kDa</td>
<td>93%</td>
</tr>
<tr>
<td>60S ribosomal protein L10</td>
<td>25 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>60S ribosomal protein L16-B</td>
<td>22 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>Glyceraldehyde-3-phosphate dehydrogenase 2</td>
<td>36 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>60S ribosomal protein L33-A</td>
<td>12 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>60S ribosomal protein L7-A</td>
<td>28 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>60S ribosomal protein L36-A</td>
<td>11 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>60S ribosomal protein L12</td>
<td>18 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>60S ribosomal protein L16-A</td>
<td>22 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>60S acidic ribosomal protein P0</td>
<td>34 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>40S ribosomal protein S17-A</td>
<td>16 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>40S ribosomal protein S2</td>
<td>27 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>40S ribosomal protein S15</td>
<td>16 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>60S ribosomal protein L8-A</td>
<td>28 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>60S acidic ribosomal protein P1-beta</td>
<td>11 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>40S ribosomal protein S21-A</td>
<td>10 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>40S ribosomal protein S27-A</td>
<td>9 kDa</td>
<td>72%</td>
</tr>
</tbody>
</table>
## Identified Proteins

*(S. cerevisiae)*

<table>
<thead>
<tr>
<th>Identified Proteins</th>
<th>Molecular Weight (kDa)</th>
<th>Protein Identification Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>40S ribosomal protein S28-A</td>
<td>8 kDa</td>
<td>72%</td>
</tr>
<tr>
<td>40S ribosomal protein S1-A</td>
<td>29 kDa</td>
<td>71%</td>
</tr>
<tr>
<td>Suppressor protein STM1</td>
<td>30 kDa</td>
<td>64%</td>
</tr>
<tr>
<td>60S acidic ribosomal protein P2-alpha</td>
<td>11 kDa</td>
<td>57%</td>
</tr>
<tr>
<td>60S ribosomal protein L35</td>
<td>14 kDa</td>
<td>52%</td>
</tr>
<tr>
<td>60S ribosomal protein L33-B</td>
<td>12 kDa</td>
<td>47%</td>
</tr>
<tr>
<td>60S ribosomal protein L1</td>
<td>24 kDa</td>
<td>41%</td>
</tr>
<tr>
<td>60S acidic ribosomal protein P2-beta</td>
<td>11 kDa</td>
<td>36%</td>
</tr>
<tr>
<td>Uncharacterized protein YHR032W-A</td>
<td>7 kDa</td>
<td>33%</td>
</tr>
<tr>
<td>40S ribosomal protein S4</td>
<td>29 kDa</td>
<td>29%</td>
</tr>
<tr>
<td>Glyceraldehyde-3-phosphate dehydrogenase 1</td>
<td>36 kDa</td>
<td>29%</td>
</tr>
<tr>
<td>COP9 signalosome complex subunit 10</td>
<td>76 kDa</td>
<td>29%</td>
</tr>
<tr>
<td>40S ribosomal protein S14-A</td>
<td>15 kDa</td>
<td>27%</td>
</tr>
<tr>
<td>60S ribosomal protein L23</td>
<td>14 kDa</td>
<td>23%</td>
</tr>
<tr>
<td>40S ribosomal protein S25-A</td>
<td>12 kDa</td>
<td>22%</td>
</tr>
</tbody>
</table>
Chapter 5

Capillary Electrophoresis
Track-Writing Interface
Here I describe a platform of technologies incorporating capillary electrophoresis with continuous deposition of proteins onto metallic substrates and analysis by fluorescence microscopy. Since sample deposition and identification are decoupled, there is also the opportunity for a multitude of other analysis techniques including matrix assisted laser desorption/ionisation mass spectrometry and EVV 2DIR spectroscopy, which are demonstrated in the proceeding chapters. The term ‘track-writing’ is used to describe the method of sample deposition, since sample is deposited in the form of continuous tracks.

5.1 Introduction

Capillary electrophoresis can offer unrivalled efficiencies and outstanding resolving power [29]; however, its use is not as commonplace as other separation techniques such as HPLC and gel electrophoresis. Although they offer different advantages and disadvantages, one of the reasons that capillary electrophoresis is not more widely used is the difficulty associated with interfacing it to a downstream readout technique that can enable further sample analysis. This is due to the need to preserve an electric current through the capillary for the separation to take place.

Initial design and construction of the interface was made by C. Loeffeld and C.J. Barnett, with LabVIEW programs written by F. Fournier and C.J. Barnett. Later work to the system was performed by the author, which included several constructional changes, including a modification in the use of pressure to avoid the use of siphoning as this would create band broadening throughout the separation process, the derivation of stage speed equations, and writing of associated LabVIEW programs, and operational changes such as buffers, stage speeds and voltages.

The interface described in this thesis was designed primarily with the intention of application for EVV 2DIR spectroscopy, which required the sample
to be dried onto a reflective or transmissive substrate (ie. metal or glass). In the case of the interface from CE, a reflective stainless steel substrate was chosen, and an explanation of how EVV 2DIR spectroscopy is performed in reflection mode is given in Chapter 7. As the interface enables readout offline, it is applicable to a range of other techniques, some of which are demonstrated in this thesis, such as intrinsic fluorescence and matrix-assisted laser desorption/ionisation-mass spectrometry (MALDI-MS). The use of complementary readout techniques increases the amount of information that can be gained from a sample. Furthermore, decoupling sample deposition from identification enables the use of identification techniques, such as EVV 2DIR spectroscopy, that require bulky instrumentation and could not easily be incorporated into an online set-up.

There are a number of interface styles, which can be divided into those in which the current is interrupted for deposition and those in which the current is maintained. A wide range of readout methods are used with these interfaces; most often a sample deposition interface is designed with a specific readout method in mind. Commercial CE machines, such as the Beckman P/ACE MDQ, typically include a fraction collection function by which the outlet of the capillary is moved from the buffer vial into a collection vial containing a small amount of water or dilute buffer (2-10 µl). With this method, each analyte is substantially diluted, fractionation is arbitrary, and thus multiple analytes can elute in one fraction and one analyte can be spread across multiple fractions. Further to these disadvantages, this type of sample collection simply would not have been suitable for use in this work as the readout methods desired required the sample to be dried onto the substrate. A number of sample deposition interfaces have been developed for the offline coupling of CE and MALDI-MS, including direct deposition onto a conductive substrate [149, 150], electrospray onto a target [151] and spotting [152, 153], as are described in further detail in Chapter 6. Surface enhanced Raman scattering (SERS) has also been used as a form of detection
with CE, enabling structural information to be gained for analytes deposited onto a SERS active substrate.\[68, 67\]

5.2 Experimental Apparatus

5.2.1 Overview of Interface

The track-writing interface, depicted in Figure 5.2.1, operates on the principle of continuous deposition from the capillary, either hydrodynamically or electrokinetically, onto a stainless steel plate which forms the cathode of the CE high voltage power supply. The stainless steel substrate is mounted on x,z translational stages adjusted so that the capillary outlet is just above the surface (typically 10-50 µm) (Figure 5.2.2a) and on application of a voltage (500 - 2000 V) a stable liquid cushion is formed between the capillary outlet and metal surface (Figure 5.2.3). As the analytes elute, the plate is moved in the x direction, forming a track of liquid on the surface, which rapidly dries (Figure 5.2.2b).

Two motorised translational stages (Newport, Irvine, CA, USA) are assembled at 90 degrees to give movement in two dimensions, with the x-stage also forming the base of the interface equipment. The x-stage translates along the axis of the track, while the vertical z-stage provides translation in height, for adjustment of distance between the stainless steel substrate and capillary.

A metal bracket with a plastic sample mount attached is screwed onto the z-stage in which the stainless steel substrate is fixed. The stainless steel substrate is connected to the power supply simply by a crocodile clip attached to a power cable. Translational stages are controlled using LabVIEW software, written by F. Fournier and C.J. Barnett.

The z-stage may also be operated by a hand controller which is used for adjustment of the stage to the distance below the capillary for a stable liquid cushion to form (10-50 µm). The adjustment of the stage and the track-
Figure 5.2.1 – Schematic Diagram of Sample Deposition Interface. A stainless steel substrate, at the outlet of the capillary, is connected to ground and mounted on robotic stages. A liquid droplet of capillary effluent completes the circuit and as the stage moves with the stainless steel substrate the capillary eluate dries in a track along the x-direction.
CHAPTER 5. TRACK-WRITING

writing process can be monitored using a video microscope (Dino Lite Digital Microscope, IDCP, Naardan, Netherlands). Alternatively, simply monitoring the current was found to be an accurate way to first adjust the stage to the correct height. The formation of a liquid cushion would cause a rapid increase in the current. The current was also monitored during track-writing as a stable current indicated a stable liquid cushion.

Before track-writing, a complete separation is first performed to determine approximate migration times. This is carried out without the stainless steel plate as the cathode, and instead using the standard CE outlet buffer reservoir and platinum wire cathode as in the standard technique described in Chapter 3. To perform the track-writing process, a separation is run at high voltage (20-30 kV) with the standard CE outlet buffer reservoir in place until the first peak is near the UV detector, at which point the high voltage is turned off, the outlet buffer reservoir removed and the stainless steel substrate placed as the new cathode of the system. A pressure, typically of 1-2 mBar (unless otherwise stated) is applied throughout the track-writing process to help maintain stability of the liquid cushion using the low pressure part of the pressure system controlled by the precision pressure regulator (Fairchild) (see Chapter 3). A voltage of 500-2000 V is then applied and the metal substrate is raised in the z-direction using the hand controller until a liquid cushion has formed between the capillary outlet and surface (as viewed through the video microscope) and a stable current is maintained. At this point the x-stage motion is initiated and a track of eluting material is deposited onto the plate. The separation continues to take place and a track-writing electropherogram is obtained from the online UV detector.

5.2.2 Capillaries

To create narrow tracks, capillaries were heat tapered at the outlet. The capillary was initially cut to approximately 15 cm longer than the desired effective length. A region of the capillary was held in the flame of a butane
Figure 5.2.2 – The sample deposition interface at the capillary outlet. a) Shows the capillary outlet, the video microscope and construction of motorised stages and b) shows the capillary outlet meeting the stainless steel substrate.
burner whilst it was pulled along its length on either side of the heated region so that it thinned until it pulled apart at this point. The capillary was then trimmed using a ceramic capillary cutter to the appropriate length at the inlet end and, once in the CE system, trimmed at the heat tapered capillary tip by removing approximately 1-5 mm (depending on the pulling process) so that an opening formed allowing liquid to flow through when applying a pressure of 2 Bar. For all track-writing experiments, unless otherwise stated, the capillaries used were either 50 or 75 µm i.d. × 375 µm o.d. UV transparent fused silica (Composite Metal Services, Charlestown, UK).

Capillaries were prepared as described in Chapter 3; however, for protein separations, the dynamic capillary pre-coating UltraTrol LN (Target Discovery Inc., Palo Alto, CA, USA) was used as commercially coated capillaries could not be used because they were pre-cut to a length too short to enable tapering, and heat tapering the end would remove the coating. The coating, UltraTrol LN, gives highly reduced cathodic EOF and minimises protein adsorption. Prior to use, new capillaries were prepared by flushing at 2 Bar for 10 minutes with ethanol, 2 minutes with water, 10 minutes with 1M sodium hydroxide, 2 minutes with water, 2 minutes with the dynamic pre-coating (UltraTrol LN) 1 minute with water and 2 minutes with the background
electrolyte (BGE). Between runs the capillary was flushed at 2 Bar for 2 minutes with water, 2 minutes with 1M sodium hydroxide, 2 minutes with water, 2 minutes with UltraTrol LN, 1 minute with water and 2 minutes with the BGE.

5.2.3 Track-Reader

To image the tracks using intrinsic tryptophan fluorescence (excitation maximum 295 nm, emission maximum 348 nm [71]), a custom built ‘track-reader’ was constructed by C.J. Barnett and A. Salehi-Reyhani. This comprised an LED with a spectral output centered at 285 ± 5 nm (LED285W, Thorlabs, Ely, Cambridgeshire, UK), an excitation filter with a centre wavelength of 280 nm (FF01-280/20, Semrock, Rochester, NY, USA), homemade focusing tubes (components from Thorlabs with UV-optimised antireflection coated lenses), dichroic mirror (FF310-Di01, Semrock), 25 mm diameter plano convex lens with 50 mm focal length (Thorlabs) used as the objective lens and an emission filter with a centre wavelength of 357 nm (FF01-357/44, Semrock). The fluorescence was imaged onto a CCD (Hamamatsu C4742-80-12AG) with image processing by LabVIEW (National Instruments, Austin, Texas, USA). The microscope stage was motorized using a ProScan II motorized stage (Prior Scientific, Cambridge, UK), controlled by LabVIEW. A schematic diagram of the track-reader is shown in Figure 5.2.4. Spectral profiles from the supplier (Semrock) of the excitation and emission filters and the dichroic mirror are shown in Figure 5.2.5.

5.3 Method Development

By controlling the velocity of the x stage, the surface area over which the sample is deposited in the direction along the axis of movement of the stage can be controlled. This enabled control over the concentration of the sample deposited onto the substrate. For example, for the purposes of deposition of
Figure 5.2.4 – Schematic diagram of the track-reader in the set-up for tryptophan fluorescence.

Figure 5.2.5 – Spectral profiles of excitation and emission filters and dichroic mirror (all Semrock) used in the track-reader set-up for tryptophan fluorescence. Tryptophan has excitation and emission maxima of $\sim 295$ nm and 348 nm respectively.
a sample for 2DIR analysis where the beam spot is approximately 100 µm in diameter, it is desirable to set the stage speed such that a single peak is deposited over a track length of 100 µm. The calculation of the stage speed is described in Section 5.3.1.

In general, it is desirable not to have the analyte spread over too great an area as it may then end up below the detection limit of the particular readout technique in question. The optimum length of the eluted spot is dependent on the readout method. In the case of EVV 2DIR spectroscopy, a beam size of 100 µm is typically used, and thus a peak is deposited over a focal area of 100×100 µm². The stage velocity is calculated to deposit the peak over 100 µm in length using Equation 5.3.11. To achieve a width of 100 µm for EVV 2DIR experiments requires experimental determination of optimal track-writing conditions through choice of the track-writing voltage and pressure, BGE, and careful tapering of capillary tips. For method development, an inverted microscope (Nikon TE 2000-U) with 10× objective was used to study and measure deposited tracks and capillary tips.

5.3.1 Calculation of Stage Velocity

The length of a spot of analyte ($l_{\text{spot}}$) deposited in a track will be a product of the selected stage velocity ($v_{\text{stage}}$) and the time it takes for a peak to elute from the capillary ($t_{\text{el}}$), hence:

$$v_{\text{stage}} = \frac{l_{\text{spot}}}{t_{\text{el}}} \quad (5.3.1)$$

The time it takes for a peak to elute ($t_{\text{el}}$) can be defined as the length of the zone at the capillary outlet ($l_{\text{outlet}}$) divided by the speed at which the analyte is migrating within the capillary under the track-writing voltage ($v_{\text{analyte}}$).

$$t_{\text{el}} = \frac{l_{\text{outlet}}}{v_{\text{analyte}}} \quad (5.3.2)$$
Hence the stage speed can be rewritten as

\[ v_{\text{stage}} = l_{\text{spot}} \frac{v_{\text{analyte}}}{l_{\text{outlet}}} \]  \hspace{1cm} (5.3.3)

The analyte velocity has two components: the electrokinetic and the hydrodynamic components, as a low pressure (typically 1-2 mBar) is applied during the track-writing process.

The electrokinetic component is calculated using the mobility of the analyte.

In Chapter 1, it was shown that for a standard CE separation,

\[ \mu = \frac{v_{\text{CE}} E_{\text{CE}}}{V_{\text{CE}}} = \frac{L_{\text{tot}}}{t_{\text{mig}}} \frac{L_{\text{eff}}}{V_{\text{CE}}} \]  \hspace{1cm} (5.3.4)

where \( v_{\text{CE}} \) is the velocity of the analyte, \( E_{\text{CE}} \) is the electric field, \( L_{\text{eff}} \) is the effective capillary length, \( L_{\text{tot}} \) is the total capillary length, \( t_{\text{mig}} \) is the migration time and \( V_{\text{CE}} \) is the separation voltage.

With a change in voltage to the track-writing voltage, \( V_{\text{track}} \), the mobility is unchanged but the analyte moves with velocity \( v_{\text{ek}} \).

Hence

\[ \mu = v_{\text{ek}} \frac{L_{\text{tot}}}{V_{\text{track}}} \]  \hspace{1cm} (5.3.5)

Inserting Equation 5.3.4 in place of mobility in Equation 5.3.5 and rearranging gives

\[ v_{\text{ek}} = \frac{L_{\text{eff}} V_{\text{track}}}{t_{\text{mig}} V_{\text{CE}}} \]  \hspace{1cm} (5.3.6)

The hydrodynamic velocity, \( v_{\text{hyd}} \), is defined as

\[ v_{\text{hyd}} = \frac{\Delta p r^2}{8 L_{\text{tot}} \eta} \]  \hspace{1cm} (5.3.7)

where \( \Delta p \) is the pressure applied during track-writing, \( r \) is the capillary inner radius and \( \eta \) is the run buffer viscosity (a value of \( 1 \times 10^{-3} \) Pa.s is used as the viscosity of water at 20°C).[29]
CHAPTER 5. TRACK-WRITING

The total analyte velocity, $v_{\text{analyte}}$, is the sum of the electrokinetic and hydrodynamic contributions:

$$v_{\text{analyte}} = v_{\text{ek}} + v_{\text{hyd}}$$ (5.3.8)

The analyte velocity during the track-writing process can therefore be defined as follows:

$$v_{\text{analyte}} = \frac{L_{\text{eff}} V_{\text{track}}}{t_{\text{mig}}} \frac{V_{\text{CE}}}{V_{\text{CE}}} + \frac{\Delta pr^2}{8L_{\text{tot}}\eta}$$ (5.3.9)

where $V_{\text{track}}$ is the voltage applied during track-writing, $V_{\text{CE}}$ is the voltage applied during the separation stage and $t_{\text{mig}}$ is the migration time of the analyte at the separation voltage.

The length of the zone at the capillary outlet can approximated as the length of the zone at the detector ($l_{\text{det}}$), which is straightforward to approximate from the electropherogram:

$$l_{\text{det}} = \frac{L_{\text{eff}}}{t_{\text{mig}}} \Delta t$$ (5.3.10)

where $\Delta t$ is the base peak width at the detector. Assuming a constant velocity for the entire plug equivalent to that at the peak maximum introduces a small error, which can be neglected.

Hence, the stage velocity can be approximated as

$$v_{\text{stage}} = l_{\text{spot}} \frac{L_{\text{eff}} V_{\text{track}}}{t_{\text{mig}}} \frac{V_{\text{CE}}}{V_{\text{CE}}} + \frac{\Delta pr^2}{8L_{\text{tot}}\eta}$$ (5.3.11)

In Equation 5.3.11, the length of the zone at the detector was used as an approximation for the length of the zone at the capillary outlet, which assumes that no bandbroadening occurs during track-writing. In reality, bandbroadening will occur during the track-writing process as a result of the processes discussed in Chapter 1. Equation 5.3.11 thus places an upper limit
on the stage velocity.

To enable quick calculations, a program for calculation of the required stage speed for a given spot length, was written using LabVIEW.

For example, for the separation of ribosomal proteins shown in Figure 5.3.1 the calculation of stage speed would be made as follows:

![Graph](image)

**Figure 5.3.1** — Separation of ribosomal proteins. Sample was injected hydrodynamically, introducing a 1.1 cm plug of sample and separated at 20 kV using a 20 mM sodium citrate BGE, pH 4.0 in a 50 µm i.d. PVA coated capillary, with a total length of 44 cm.

For the peak shown, the base peak width is 0.13 minutes and the migration time is 6.92 minutes. The capillary had a total length of 44 cm and an effective length of 33 cm, and an i.d. of 50 µm. The separation was performed at 20 kV. The calculations are made with the representative values of a track-writing voltage ($V_{\text{track}}$) of 500 V and a pressure of 2 mBar. Using Equation 5.3.11, it can be calculated that to deposit a peak over a length of 100 µm, a stage velocity of 54 µm/min is required. Table 5.3.1 shows the length over which a peak would be deposited for this separation as a result of several stage velocities under these conditions.

To deposit larger quantities, as greater peak broadening often occurred
Table 5.3.1 – Spot lengths resulting from various stage speeds calculated using Equation 5.3.11, with a track-writing voltage of 500 V, a track-writing pressure of 2 mBar and parameters from the separation of ribosomal proteins displayed in Figure 5.3.1.

<table>
<thead>
<tr>
<th>Stage Speed (µm / min)</th>
<th>Spot length (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>93</td>
</tr>
<tr>
<td>75</td>
<td>140</td>
</tr>
<tr>
<td>100</td>
<td>187</td>
</tr>
<tr>
<td>125</td>
<td>233</td>
</tr>
</tbody>
</table>

due to electrodispersion [22], slower stage velocities or higher voltages were required.

Additionally, equations were derived to estimate the length of a track and time required to deposit an entire separation, and a corresponding LabVIEW program was written.

The length of a track, $l_{\text{track}}$, is the time it takes for track-writing, $t_{\text{track}}$, multiplied by the stage velocity. Hence

$$l_{\text{track}} = v_{\text{stage}} t_{\text{track}} \quad (5.3.12)$$

To calculate the time at which track-writing needs to proceed for a separation containing a number of components, to begin with, it can be estimated that at the time when the separation voltage ($V_{\text{CE}}$) is turned off when the first peak reaches the detector, the last peak in the separation has migrated some distance $l'_{\text{mig}}$ (see Figure 5.3.2). The migration time of this peak is $t_{\text{mig}}'$, and the time at which the separation voltage is stopped is $t_{\text{CE}}$.

$$l'_{\text{mig}} = \frac{L_{\text{eff}}}{t_{\text{mig}}'} t_{\text{CE}} \quad (5.3.13)$$

The distance left for the analyte in the last peak to migrate is $L_{\text{tot}}$ minus this distance ($l'_{\text{mig}}$). The time it takes for the track to be written is the
CHAPTER 5. TRACK-WRITING

Figure 5.3.2 – Calculation of track-writing length. When the first peak reaches the detector and the high voltage is switched off, the last peak has migrated a distance $t'_\text{mig}$. The length of capillary left for this peak to migrate is $L_\text{tot} - t'_\text{mig}$.

remaining distance that the analyte in the last peak has to migrate divided by its velocity under the conditions during track-writing. The velocity for track-writing is composed, once again, of the hydrodynamic part and electrokinetic part. Hence

$$t_\text{track} = \frac{(L_\text{tot} - \frac{L_\text{eff}}{t'_\text{mig}} \cdot t_\text{CE})}{v_\text{analyte}}$$

where $v_\text{analyte}$ refers to the last peak, of which the migration time is $t'_\text{mig}$. And thus the time to deposit a separation is given by

$$t_\text{track} = \frac{(L_\text{tot} - \frac{L_\text{eff}}{t'_\text{mig}} \cdot t_\text{CE})}{\Delta \mu \frac{d^2}{8 \eta L_\text{tot}}} + \frac{L_\text{eff} \cdot V_\text{track}}{t'_\text{mig} \cdot V_\text{CE}}$$

The length of a track can therefore be calculated as
\[ l_{\text{track}} = v_{\text{stage}} \left( \frac{L_{\text{tot}} - \frac{L_{\text{eff}}}{t_{\text{mid}}}}{t_{\text{CE}}} \right) \frac{\Delta \rho r^2}{8L_{\text{tot}} \eta} + \frac{L_{\text{eff}}}{t_{\text{mid}}} \frac{V_{\text{track}}}{V_{\text{CE}}} \]  

(5.3.16)

5.3.2 Improved Pressure Control

In the initial construction of the track-writing interface, the inlet vial was elevated 5 cm from the outlet vial or metal substrate to ensure stability of the liquid droplet during track-writing but, as a consequence, siphoning occurred throughout the separation process which was equivalent to the application of a pressure of 4.9 mBar (assuming a value of \( \rho = 0.9972 \text{ g cm}^{-3} \) for water at 20°C [29]). Siphoning is a major cause of bandbroadening in capillary zone electrophoresis [123], and it is frequently suggested by investigators to level volumes of liquid in buffer reservoirs to avoid it. [155, 156] For a 50 µm i.d., 50 cm capillary, this would create an unnecessary hydrodynamic flow of 10 nl/min during the course of the separation. Therefore, the system was reconstructed so that the separation prior to track-writing was performed with the buffer vials level (as they are in the set-up for standard CE operation, see Chapter 3) and a low pressure (typically 1-2 mBar) was applied only at the start of the track-writing phase, following the main stage of the separation. As well as giving the advantage of having no siphoning effect during the high voltage stage, additional control was given as the precision pressure regulator has a sensitivity of 0.3 mBar.

A series of experiments were performed to ascertain an optimum applied pressure during track-writing. Pressures were systematically decreased from 25 mBar to atmospheric pressure and track-writing performed at each pressure with a voltage of 500 V applied. The appearance of the resulting tracks were studied using an inverted microscope (Nikon TE 2000-U) with 10× objective. It was found that at pressures above 10 mBar a large amount of liquid rapidly came out of the capillary at the start of the track (shown in Figure 5.3.3b) producing a large bulge (862 µm in width at the widest point),
CHAPTER 5. TRACK-WRITING

(a) Start of track using 25 mBar

(b) Capillary tip using 20 mBar

(c) Track at atmospheric pressure

Figure 5.3.3 – Tracks deposited at a) 25 mBar, b) 20 mBar and c) atmospheric pressure at 500 V using a 50 µm i.d. UV transparent capillary with a total length of 50 cm and a 40 mM sodium citrate, pH 4.0 BGE. a) and c) were viewed with the inverted microscope using the 10× objective and b) was viewed with the video microscope (10× magnification).

as shown in Figure 5.3.3a. At lower pressures a small bulge (< 200 µm in width) was still visible near the start of the track, but this was largely due to a small delay between the substrate meeting the capillary effluent and the x-stage motion starting. It was found that at atmospheric pressure, by performing electrodeposition alone (Figure 5.3.3c), a current was not reproducibly maintained. It was concluded that an optimum pressure of 1-2 mBar ensured a constant current and hence flow without an undesirable increase in track width.

5.3.3 Effect of Background Electrolyte

In terms of chromatographic parameters, in theory, reducing the stage velocity should increase efficiency as it would lead to narrower spots in the track. However, in practice, broadening of the track and increase in inhomogeneity was observed at the slower stage velocities required to deposit spot lengths of ≲ 200 µm when using a 40 mM sodium citrate buffer, pH 4.0.

From the images shown in Figure 5.3.4, taken with an inverted microscope,
and from monitoring the track-writing process with the video microscope it was apparent that bubbles are forming in the track, either due to electrolysis occurring on the surface in the deposited material or excessive heating. At slower stage velocities, a greater amount of liquid is deposited per unit area without drying, enhancing this effect. In the images shown, larger bubbles clearly correspond to the greater track width and standard deviations at slower stage speeds, with a maximum bubble diameter of approximately 100 µm at a stage velocity of 50 µm/min. Standard deviations are taken from 5 width measurements along the length of a track (taken at 200 µm intervals) and so indicate the level of uniformity of the track.

Due to the problems associated with 40 mM sodium citrate, it was hypothesised that a volatile BGE, such as ammonium acetate, would dry faster when deposited, before the processes forming the bubbles occurred. Tracks of lysozyme, trypsinogen and α-chymotrypsinogen were initially deposited using a BGE of 100 mM ammonium acetate, pH 4.0, but, although the bubbling in the track was prevented, it was observed that the buffer droplet dried too quickly leaving an uneven track with sections missing. A 50 mM ammonium acetate, pH 4.0, BGE was therefore trialled and found to dry slower yet deposit narrow, regular width tracks without the bubbles associated with sodium citrate tracks at the stage velocity required to deposit a 100 µm spot of 100 µg/ml lysozyme (48 µm/min) (Figure 5.3.5). A peculiar drying feature resembling a repeated, overlapping, coffee ring effect was observed to occur in some portions of tracks, potentially due to the speed of drying. However, this effect did not appear to be associated with any problems during fluorescence imaging of tracks of lysozyme, trypsinogen and α-chymotrypsinogen samples (Section 5.4.1). Using this method, track widths of ≤61±5 µm (calculated from 5 measurements along the length of the track) were recorded.

Clearly the fluid processes at the interface are complex. The BGE must dry quickly enough to avoid the production of bubbles on the surface, but it
CHAPTER 5. TRACK-WRITING

(a) Stage velocity of 50 $\mu$m/min

(b) Stage velocity of 75 $\mu$m/min

(c) Stage velocity of 100 $\mu$m/min

(d) Stage velocity of 125 $\mu$m/min

(e) Plot of the track width as a function of stage velocity. Errors are standard deviations in the track width calculated from 5 measurements along the length of the track.

Figure 5.3.4 – Reducing the stage speed results in an increase in track width. A 40 mM sodium citrate buffer at pH 4.0 was used in a UV transparent capillary with total length of 50 cm. Deposition was performed at a voltage of 500 V with a pressure of 2 mBar.
Figure 5.3.5 – Track deposited using 50 mM ammonium acetate, pH 4.0 as BGE. A UV transparent tapered outlet capillary with 50 µm i.d., 50 cm total length was used. Deposition was performed at a voltage of 500 V with a pressure of 2 mBar and a track-writing speed of 48 µm/min.

must not dry too quickly as the liquid cushion must remain stable and ensure a flow of liquid that forms regular tracks before drying.

5.3.4 Capillary Tapering

Tapering the capillary tip reduces the deposited track width, and can enable tracks with widths of 60 µm to be deposited (see Figure 5.3.5). The initial aim was to deposit the sample within a track width that correlated with the beam size in EVV 2DIR spectroscopy (typically 100 µm). However, it was observed that tapering of capillaries was a cause of capillary blocking for the high injection quantities, such as were required for EVV 2DIR spectroscopy (≥10^{11} molecules of a protein in a band) and track-writing of ribosomal proteins. When a band of protein containing ≥10^{11} molecules entered the capillary tip, the narrowing caused blocking and the current to drop to zero. This was confirmed by viewing these tips under the inverted microscope, as shown in Figure 5.3.6a, where blockages at the tips can be seen. Furthermore, the problem is enhanced if heating occurs in the capillary tip. Since the tapered tip is narrower than the rest of the capillary, it has a higher resistance and therefore more heat is generated in this region. From Ohm’s law, the
resistance, \( R \), for a cylindrical capillary can be calculated as

\[
R = \frac{V}{I} = \frac{L}{\sigma \pi r^2} \tag{5.3.17}
\]

where \( I \) is the current, \( V \) is the voltage, \( L \) is its length, \( \sigma \) is the conductivity of the solution filling it and \( r \) is the tip inner radius.

and Joule heating, \( J \), as

\[
J(t) = I^2 R t \tag{5.3.18}
\]

where \( t \) is time. Therefore, a localised reduction in diameter causes an increase in resistance, resulting in heating in the tip.

\[\text{Figure 5.3.6} \] – Heat tapered capillary tips from UV transparent capillary, photographed with a Nikon TE 2000-U inverted microscope with a trans-illumination light source using a 10× objective. a) Blocked capillary tips b) Capillary tip tapered to a less narrow width to enable injection of a larger volume without blocking. Capillaries were UV transparent with o.d. and i.d. prior to tapering of 375 \( \mu \text{m} \) and 75 \( \mu \text{m} \) respectively.

Nevertheless, without tapering, the outer diameter of the capillary is 375 \( \mu \text{m} \), and capillary outer diameter was observed to correlate with track width. Therefore, tapering the capillary, which can lead to a reduction in track width and therefore an enhancement in concentration of the analyte deposited in the spot, is advantageous, so long as the capillary is not overly tapered so
as to block (see Figure 5.3.6). It was found that using a capillary that was
tapered to an o.d. of approximately 200 ± 30 µm (Figure 5.3.6b) was an
appropriate compromise for analyses with the larger injection requirements
for ribosome separations and EVV 2DIR of peptide separations (Chapter 7),
enabling the injection of 9×10^{13} peptide molecules in the latter case.

5.4 Track-writing of protein samples

The adjustments described enabled the track-writing interface to successfully
be used for deposition of proteins from CE separations. This was performed
first with a set of standard proteins and then with the ribosomal proteins.
Protein intrinsic fluorescence in tracks was imaged using the track reader.

5.4.1 Separation of 3 protein mixture

Separations and depositions of lysozyme, trypsinogen and α−chymotrypsinogen
were performed using a 50 mM ammonium acetate BGE, at various pH val-
ues. As the resolution of trypsinogen and α−chymotrypsinogen was poor at
pH 4, the pH was increased in increments of pH 0.5 up to pH 5.5 at which
point the proteins were resolved in the electropherogram under a 20 kV sep-
aration without track-writing. These were deposited, with 3 repeats, and
the resulting tracks imaged with intrinsic fluorescence. In all cases, a clear
correlation was visible between the position of spots in the track and the
position of peaks in the electropherogram (Figure 5.4.1).

This demonstrates the first readout technique: intrinsic fluorescence. This
method of readout gives information about the deposition on the track in
terms of position and intensity. An application for this readout method with
the developed track-writing platform, which is currently under development
in the Single Cell Proteomics Group at Imperial College is the study of the
co-localisation of drugs and binding partners using dual wavelength detec-
tion. For many studies, qualitative information about proteins leading to
Figure 5.4.1 – Track-writing of separation of lysozyme, trypsinogen and \( \alpha \)-chymotrypsinogen (100 \( \mu \)g/ml) using a 50 mM ammonium acetate, pH 5.5 BGE, showing the electropherogram and corresponding intrinsic fluorescence image. Sample was injected at 5 kV for 9 s on a 50 \( \mu \)m i.d., 55 cm total length UV transparent tapered outlet capillary pre-coated with UltraTrol LN. Separation was performed at 20 kV. Track-writing was performed at 1 kV with a pressure of 2 mBar and a stage speed of 60 \( \mu \)m/min.
identification is required and this is demonstrated by the readout methods of MALDI-MS and EVV 2DIR provided in Chapters 6 and 7.

5.4.2 Track-writing of ribosomal proteins

The next stage was to use the track-writing interface to deposit ribosomal proteins. To develop a working protocol for the track-writing of ribosomal proteins, as previous experiments had found that 40 mM sodium citrate was not a suitable BGE for track-writing and 50 mM ammonium acetate gave consistent narrow tracks, the first stage was to optimise ribosomal protein separations in a 50 mM ammonium acetate BGE. Best resolution and efficiency were obtained using a 50 mM ammonium acetate pH 4.0 BGE and thus this was used for subsequent experiments (Figure 5.4.2).

Deposition of the ribosomal proteins was performed under the conditions described for the ribosomal separation in the ammonium acetate BGE but using a tapered capillary with UltraTrol LN dynamic precoating. The main challenge was to deposit sufficient quantities for downstream identification. Methods to increase the loading capacity for the ribosomal sample are discussed in Chapter 4, relating to the loading of dilute, complex samples. In practice the highest injection amounts demonstrated in Chapter 4 could not be implemented as they resulted in blocking of the capillary due to the tapered end. Trimming of the capillary ends to o.d.s of $\sim 200 \mu m$ enabled injections of 100 s with successful track-writing of ribosomal proteins, as shown in Figure 5.4.3. This demonstrates the application of the track-writing interface to a complex biological sample, where fluorescence is used as the readout method. Track-writing provides a chemical record of the separation in the capillary and additional analysis techniques such as MALDI-MS or EVV 2DIR spectroscopy can later be used to obtain further information about the sample.

The electropherogram and corresponding fluorescence image of the deposited ribosomal proteins are shown in Figure 5.4.3, which is representative
Figure 5.4.2 – Separation of ribosomal proteins using a 50mM ammonium acetate, pH 4.0 BGE (without track-writing). Three consecutive repeats are shown, where sample was injected at 60 mBar for 40 s and separated at 25 kV in an untapered PVA coated capillary with a 50 μm i.d. and total length of 71 cm.
of the depositions of ribosomal proteins. A large, sharp edged peak appears after all the other peaks, which can be correlated with a bright spot in the track. This was apparent in all track-writing electropherograms and resembles a narrow peak seen at the end of electropherograms of ribosomal protein separations. This peak has the appearance of a system peak, but the fluorescence image of the track provides evidence that it is a component of the sample. It may be aggregation of several proteins, intact ribosomes or co-elution of several proteins.

From both the electropherogram and fluorescence image of the deposited ribosomal proteins it is apparent that resolution has been lost in comparison with standard CE separations. As resolution loss is observed in the track-writing electropherogram, it can be deduced that the loss of resolution is a result of processes in the capillary rather than during the deposition. Some loss of resolution may have occurred after the separation at high voltage, while the system is set up for the low voltage track-writing stage (typically < 3 minutes). Effects occurring at the capillary tip such as siphoning, droplet formation and evaporation and physical movement of the capillary could have caused disturbances to the separation. For a complex separation such as the ribosomal proteins, track-writing at a low voltage takes several hours and bandbroadening from processes such as diffusion and hydrodynamic flow will have a greater impact on the separation than for tracks that can be written in a shorter period of time. In the future, advances of the interface will be made to reduce bandbroadening by running the track-writing process at a high voltage.

5.5 Conclusions

In this chapter a sample deposition interface from capillary electrophoresis onto metal substrates is advanced to enable deposition of separations of proteins. Traditionally, sample is injected and separated with CE but dis-
Figure 5.4.3 – Track-writing of separation of ribosomal proteins showing the electropherogram and corresponding intrinsic fluorescence image. Sample was injected at 60 mBar for 100 s on a 50 µm i.d., 55 cm total length UV transparent tapered outlet capillary pre-coated with UltraTrol LN. A 50 mM ammonium acetate, pH 4.0 BGE was used and separation was performed at 21 kV. Track-writing was performed at 1 kV with a pressure of 3 mBar and a stage speed of 50 µm/min.

carded after analysis; this interface enables a host of complementary readout methods to be used to add to the information gained from the CE alone.

This type of interface has great potential as it does not fractionate the sample, being instead continuous with the ability to preserve the integrity of the separation. Advances were made in several aspects, in particular, the use of a volatile BGE and the optimisation of capillary tapering. The locations of the separated and deposited proteins were imaged through intrinsic tryptophan fluorescence using a custom built track-reader. These images replicate the separations, and the track is thus shown to provide a chemical record
of the separation in the capillary. Ribosome separations were deposited and fluorescence images corresponded with the electropherograms. The samples are storable and once deposited, dried ribosomal proteins can be probed using a host of techniques. The offline nature of the deposition interface offers the advantage of flexibility, and this is demonstrated in subsequent chapters by the applications of MALDI-MS and EVV 2DIR spectroscopy, there truly showing the development of a CE platform technology for proteomics.
Chapter 6

Capillary Electrophoresis - Matrix Assisted Laser Desorption/Ionisation
6.1 Introduction

Chapter 5 saw the advancement of a sample deposition ‘track-writing’ interface to flat metal substrate from capillary electrophoresis. In this chapter, the track-writing interface is adapted to enable matrix assisted laser desorption/ionisation (MALDI) - mass spectrometry (MS) to be used to analyse the samples in the tracks. A method of matrix application post sample deposition is developed, which enables the analysis of separated species in tracks and allows multiple readout methods to be employed. This is demonstrated for tracks of lysozyme and α-chymotrypsinogen, where fluorescence imaging is used to visualise proteins in tracks prior to MALDI-MS. This work is then advanced by the offline coupling of CE to MALDI-imaging for the first time and applied to the ribosomal proteins.

6.1.1 Matrix Assisted Laser Desorption/Ionisation (MALDI) - MS

MALDI is a method of producing ions from samples on surfaces for mass spectrometry. The focusing of a laser on the sample induces desorption of both ions and neutral molecules. The analyte is mixed with a matrix, usually a small organic molecule dissolved in a solvent, and matrix and analyte co-crystallise as the solvent evaporates. The matrix has a strong absorption at the laser wavelength. The matrix serves two main purposes: firstly, to absorb the energy from the laser, and, secondly, to isolate analyte molecules from one another.[157] Irradiation of the matrix-analyte mixture by the laser results in excitation of matrix molecules causing localised disintegration and evaporation of the solid matrix-analyte deposit. The desorbed material contains clusters consisting of analyte molecules surrounded by matrix and salt ions. Energy is transferred from matrix molecules to the analyte by proton transfer, resulting in ions of both analyte and matrix.[158]

MALDI is a ‘soft ionisation’ technique because little or no fragmentation
takes place. It is thought that this is because the matrix protects the analyte from taking ‘direct hits’ from the laser, with analyte ions created indirectly by transfer of kinetic energy from matrix ions. A nitrogen laser at 337 nm is typically used, which is not absorbed by the aromatic amino acids. This helps further to avoid fragmentation.[159]

This soft laser desorption technique takes place in a vacuum in the mass spectrometer and creates a cloud of ions near the surface of the metal. MALDI is typically coupled to a Time Of Flight (TOF) mass analyser, and an electrical potential is applied to the plate to accelerate the ions into the mass analyser.[8]

The matrix plays a crucial role in the analyte desorption and ionisation process, but performance is dependent on the chemical composition of the analyte leading to the development of a range of different matrices for different types of molecules.[160, 157] For proteins, the matrix 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid) is frequently used, and for peptides α-cyano-4-hydroxycinnamic (CHCA) is the matrix of choice.[161, 160] Other biological matrices include 2,5-dihydroxybenzoic acid (DHB), which is commonly used for peptides, glycopeptides, glycoproteins and small proteins [157] and 2-(4-hydroxyphenylazo)benzoic acid (HABA) [162].

The process of matrix application is important in determining the quality of the signal in MALDI [163], and problems often arise from heterogeneous matrix-analyte crystal formation leading to ‘sweet spots’ and areas where ionisation does not occur [162]. The most common method of matrix application is the dried droplet technique, first introduced by Karas and Hillenkamp in 1988 [164], whereby a droplet of matrix dissolved in an aqueous solution (and usually a high proportion of a volatile solvent) is mixed with a droplet of sample and deposited. The solvent evaporates and the matrix and the sample co-crystalise. Various methods of matrix deposition by precoating of targets followed by sample deposition are frequently used for coupling of MALDI to LC or CE, including matrix precoated membranes [165, 166], fast
solvent evaporation method [167] and spin coating [168]. Pneumatic sprays are used either to precoat targets [169] or to prepare samples for MALDI imaging [170], as is described in Section 6.3.

6.1.2 CE-MALDI Interfaces

Most commonly, CE-MS is performed with online coupling of an electrospray ionisation (ESI) interface. In an ESI interface, liquid elutes from the capillary at atmospheric pressure whilst a high voltage (usually 1-4kV [171]) is applied between the capillary exit and the inlet of the mass spectrometer. This causes a spray of charged microdroplets, spreading outwards due to repulsion in a so-called Taylor cone.[172] On entering the vacuum area of the mass spectrometer, the solvent evaporates causing shrinkage of the droplets, and droplets eventually disintegrate to form gas-phase ions.[19] Using ESI, ions with multiple charges are generated, allowing the detection of very large molecular species within the $m/z$ range of most mass analysers [172]; furthermore, it can be interfaced with a host mass analysers.

A CE-ESI interface makes perfect sense at first consideration as a voltage is already applied at the capillary outlet where liquid is exiting the capillary. However, there are in practice a number of disadvantages of this approach. Firstly, despite the advantages associated with the generation of multiply charged ions, it makes the interpretation of incompletely separated peaks in the electropherogram difficult.[171] Secondly is an issue of timing, particularly if sequencing (using CE-MS/MS) is required, as in this case, for a typical ESI-MS experiment, a full scan across the mass range will be performed followed by selection of up to five precursor ions for MS/MS analysis.[173, 174] If peaks are narrow and close together, the time for each analysis is limited, and quadrupole mass analysers, which are often used, can be slow at scanning the entire mass range.[171] The limitation of timing can be circumvented somewhat through modifications to the separation or through similar methods to peak parking techniques used in liquid LC-MS interfaces.[175]
CHAPTER 6. CE-MALDI

However, this results in reduced throughput. Finally, an ESI interface places strict restrictions on the BGE which can be used for CE - a volatile BGE is generally a requirement [19] and a number of the salts commonly used in CE buffers must be avoided [176].

An offline CE-MS interface is therefore an attractive alternative and, with the rival popular soft ionisation technique to ESI being MALDI, a logical next step is an interface based on sample deposition onto a solid surface for MALDI analysis. Typically, deposited samples are stable for extended times so analyses can be performed weeks after deposition, and MS can be performed followed by later MS/MS analysis of selected analytes. Furthermore, MALDI is tolerant of a much wider range of the common constituents of CE buffers than ESI.[176] A number of CE-MALDI interfaces have been previously described. The primary difficulties arise from maintaining electrical contact between capillary outlet and deposition surface (as discussed in Chapter 5) and application of matrix.[173]

CE-MALDI interfaces have been designed using a variety of techniques, including those based on the principles of fraction collection, spotting, continuous deposition interfaces and electrospray-MALDI interfaces. Many interfaces deposit the sample in a discontinuous way, using devices that operate by fraction collection [177, 178], spotting [152, 179] or a related technique known as droplet electrocoupling (DEC) [180, 181]. Femtomole [153, 182], or even sub-femtomole [152] detection limits (depending highly on the analyte) can be achieved with a number of these techniques. Nevertheless, a flaw for these techniques (unless careful voltage programming takes place [182]) is the arbitrary nature of fractionation which bears little relation to separation. Thus the sampling rate is the limiting factor for efficiency in deposited CE eluent. Furthermore, as a result of sample carryover, it impacts the resolution.

For a separation technique that prides itself on outstanding separation efficiencies and for future applications to complex proteomic problems with
large dynamic range, an interface that preserves the separation achieved in
the capillary is in many ways more desirable. Several continuous deposition
interfaces have been developed. A number of these are online interfaces,
which use rotating ball [183] or moving belt-like systems at the inlet [184,
185, 186]. Offline interfaces have been developed that operate under a similar
principle to track-writing, where deposition is onto a moveable target that
is part of the electrical circuit.[149, 187, 171] The work presented in this
chapter offers advantages in flexibility and simplicity to these techniques.
Other interfaces require custom made or adapted MALDI targets, and online
interfaces also require adaptation of the mass spectrometer as the sample
must enter the mass spectrometer under vacuum. Application of the matrix
takes place either before deposition of the sample via a sheath flow; therefore
in no cases are parallel analysis techniques employed.[174]

In the CE-MALDI interface presented henceforth, the track-writing equip-
ment is used to deposit tracks onto standard MALDI target plates, requiring
no modifications. Matrix is applied by spray coating after sample deposi-
tion, and there is the therefore the option for tracks to be studied with a
range of other techniques prior to MALDI analysis such as fluorescence mi-
croscopy, surface enhanced Raman scattering (SERS) or EVV-2DIR, as is
demonstrated in Chapter 7. The workflow is demonstrated in Figure 6.1.1.
Three major developments are presented. Firstly, MALDI-MS as a read-
out technique is applied using a standard MALDI-TOF/MS instrument to
proteins in tracks (Section 6.2). Secondly, MALDI-imaging is used to study
tracks (Section 6.3). Thirdly, this work is then applied to the ribosomal
proteins (Section 6.4).

6.2 CE-MALDI-TOF/MS

In this section it is shown how proteins in tracks can be analysed using a
standard MALDI-TOF/MS instrument. Proof of principle results are fol-
Figure 6.1.1 – Application of track-writing to CE-MALDI. Separations are written onto standard MALDI plates, sprayed with matrix and analysed using MALDI-MS or MALDI imaging. Prior to spraying with matrix, tracks can be analysed with alternative readout methods as required.
lowed by further developmental work, such as the development of a matrix application method. The successful development of a spray coating method enables separated proteins in tracks to be identified with MALDI-MS and correlated with the fluorescence images.

6.2.1 Experimental Procedure and Method Development for CE-MALDI

The track-writing equipment described in Chapter 5 was used to deposit tracks of standard proteins onto Waters MALDI imaging target plates. Combinations of the proteins lysozyme, trypsinogen and α-chymotrypsinogen A were used as tracks had been successfully written of these in previous experiments. Deposition onto the MALDI plates required no modification of the equipment, and CE and track-writing conditions used were similar to those described in Chapter 5.

Separations and Track-writing

Protein stock solutions at 2 mg/ml were prepared from lysozyme from hen egg white (Sigma-Aldrich), trypsinogen from bovine pancreas (Sigma-Aldrich) and α-chymotrypsinogen A from bovine pancreas (Sigma-Aldrich) in water and stored at -20°C. The masses of the proteins as stated by the supplier are displayed in Table 6.2.1. Mixtures of the proteins at final concentrations of either 400 µg/ml or 100 µg/ml (as stated in the Results) each in a 10× dilution of the CE run buffer were made and stored at 4°C.

Capillaries were 50 µm i.d. UV transparent fused silica, 60-70 cm in length, with a heat tapered outlet. Capillaries were prepared as described in Chapter 5, including a 2 minute rinse with the dynamic pre-coating reagent UltraTrol LN (Target Discoveries). A 50 mM ammonium acetate run buffer pH 4.0 run buffer was used for separations and depositions. Sample was injected electrokinetically at 5 kV for 20 s or 9 s (as stated). Separation was
<table>
<thead>
<tr>
<th>Protein</th>
<th>Mass / kDa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysozyme</td>
<td>14.3</td>
</tr>
<tr>
<td>Trypsinogen</td>
<td>23.7</td>
</tr>
<tr>
<td>α-chymotrypsinogen A</td>
<td>25.6</td>
</tr>
</tbody>
</table>

Table 6.2.1 – Masses of proteins stated by the supplier

performed at 20 kV.

Track-writing was performed with stage velocities of either 50 \( \mu m/min \) or 60 \( \mu m/min \) (as stated). A voltage of 1 kV was applied with a pressure of 1 mBar. Images of intrinsic tryptophan fluorescence in protein tracks were obtained using the track-reader as described in Chapter 5. The MALDI imaging plate (Figure 6.2.1b) is engraved with an array of numbered crosses. To record the position of a spot, the plate was viewed in brightfield and the position of the crosses recorded.

Matrix Application

The matrix sinapinic acid (Sigma-Aldrich) was used, which is a frequently used matrix for protein experiments. It was made up at 10 mg/ml in 50/50 acetonitrile/ 0.1% trifluoroacetic acid (TFA) in MilliQ water.

For initial proof of principle experiments, two tracks of lysozyme, trypsinogen and α-chymotrypsinogen were spotted with 1 \( \mu l \) spots of matrix solution along the length of the track. Controls of lysozyme at 1mg/ml and also with the background electrolytes of 40 mM sodium citrate and 50 mM ammonium acetate were performed by mixing 1 \( \mu l \) of the solution with 1 \( \mu l \) of matrix solution and spotting onto a standard Waters MALDI target plate.

For spray coating, a paint sprayer (Model No. AB93, Sealey Power Tools, Colchester, UK) was used. A spray coating method was developed which was similar to that used with manual sprayers in MALDI imaging for tissue sections [188], where cycles of matrix coating were performed, each cycle
depositing a small amount of matrix and barely wetting the surface. The spray coating method used was developed empirically on stainless steel slides, as it was observed that these conditions did not overly wet the surface as to cause diffusion in the track but gave an even, dense film of matrix. For the matrix spraying method used for all MALDI applications (whether to pre-spray the plate or for post-deposition spraying of the track) the sprayer was held approximately 30 cm from the target plate and passed over the plate twice before allowing the solvent to evaporate. This process was then repeated three further times.

To develop the spray coating method as part of a CE-MALDI interface, comparisons were made between tracks deposited on pre-sprayed MALDI target plates, tracks deposited onto uncoated plates and then sprayed and tracks deposited in a method similar to the sandwich technique, where the matrix was applied both before and after deposition. MALDI-MS was performed on continuous lysozyme tracks trialling these three matrix deposition conditions. Lysozyme was prepared at 500 µg/ml in 5 mM ammonium acetate. The capillary was filled with lysozyme by applying a pressure of 1 Bar for 2 minutes. A 12 mm track was then deposited on either an uncoated or coated MALDI plate by applying a pressure of 30 mBar and with a stage speed of 1 mm/min. A total of two tracks were deposited on an uncoated Waters MALDI imaging target plate and two tracks on a Waters MALDI imaging target plate pre-sprayed with matrix. For the target plate pre-sprayed with matrix, once the track had dried, parafilm was used to cover half of both tracks (without touching them) and the remaining half was sprayed again with matrix.

In all later experiments, tracks were deposited onto uncoated MALDI imaging target plates and spray coated post deposition using the method described above.
MALDI-TOF/MS Data Collection and Analysis

A Waters MicroMass MALDI micro MX time-of-flight mass spectrometer was used in positive ion/linear mode with the software MassLynx. The mass range was set at 10000-29000 $m/z$ for initial experiments in which matrix was spotted onto tracks and 9000-30000 $m/z$ for all further experiments. The instrument was calibrated using bovine serum albumin (BSA).

The instrument was operated for manual data collection. The instrument is set up to read standard Waters MALDI target plates, where sample is deposited in discrete spots in specified positions (Figure 6.2.1a). The laser can be fired in a rectangular area outside these spots but cannot be run continuously across the length of the plate. To acquire data from tracks deposited continuously onto Waters MALDI imaging plates (Figure 6.2.1b), first the built-in camera was used to locate the start of the track. Subsequently, for each ‘spot’ position, the laser was fired along the length of the section of track enclosed within that region and spectra were recorded at places along the tracks where peaks in the signal occurred.

An estimate of the sensitivity of the MALDI instrument for proteins was performed using the protein standards deposited with matrix using the dried droplet method. The sensitivity was estimated to be in the high femtomole range.
6.2.2 Results and Discussion

Tracks were deposited onto Waters MALDI imaging target plates and analysed using a Waters MicroMass MALDI-MS. Through the development of a spray coating matrix application method, the analysis of separated proteins in tracks is achieved with MALDI-MS.

6.2.2.1 Proof of principle for CE-MALDI interface

The potential to use MALDI as a readout technique for CE through the track-writing equipment was demonstrated in initial experiments in which matrix was applied in 1 µl spots along the length of tracks. This is shown in Figure 6.2.2, which displays the mass spectrum for the track of a separation of lysozyme, trypsinogen and α-chymotrypsinogen A at 400 µg/ml, injected at 5 kV for 20 s equating to 281 fmol lysozyme, 109 fmol trypsinogen and 98 fmol α-chymotrypsinogen A. All three proteins were identified, thus demonstrating the potential to use MALDI as a readout technique with the track-writing equipment.

However, despite being baseline resolved in the track-writing electropherogram, the separation was not maintained as all three proteins were identified at one point in the track. This is because spots of matrix merged before drying. Thus this work showed that proteins from the capillary electrophoresis system could be detected using MALDI-MS; however, it demonstrated the need to develop a method of applying matrix with limited effect on the separation.

6.2.2.2 Development of a matrix deposition method

Initial experiments in which the matrix was spotted onto tracks (Figure 6.2.2) gave proof of principle evidence that MALDI-MS could be used as an identification technique with the track-writing interface to CE. However, the loss of chromatographic resolution demonstrated the need to develop a method
Figure 6.2.2 – Mass spectrum obtained from initial track of lysozyme, trypsinogen and α-chymotrypsinogen A spotted with matrix. As would be expected from non-MS grade samples, all three proteins produce at least one peak away from the [M+H]+ and [M+2H]2+ m/z values which correspond to the stated molecular weights for these proteins (Table 6.2.1) as a result of trace impurities which are also observable as small extra peaks or shoulders close to the main peak in electropherogram. These were observed in mass spectra recorded for the CE sample and individual proteins deposited using the dried droplet method for comparison.
of coating that did not destroy the separation in the track. A spray coating method was chosen as small droplets of matrix can be applied evenly across the surface. Experiments were performed to optimise the method of spray matrix deposition, which were trialled on continuous tracks of lysozyme.

An effective protocol for spray coating was developed, which gave an optimised degree of wetting of the surface, was described in Section 6.2.1. Furthermore, three techniques of matrix spray application were tested: firstly, track-writing onto MALDI plates pre-sprayed with matrix, secondly, track-writing onto bare MALDI plates with matrix sprayed on after deposition and thirdly, track-writing onto MALDI plates pre-sprayed with matrix with matrix also sprayed on after track-writing (termed the ‘sandwich’ method).

Two important conclusions were made from the resulting mass spectra and from observations made during the track-writing process, which are demonstrated in Figure 6.2.3. Firstly, a crucial finding was that desorption-ionisation of protein occurred when matrix was sprayed onto dried tracks. As previously discussed (Section 6.1.1), the formation of matrix-analyte crystals is a critical factor in the MALDI process. The possibility for signal to be seen when matrix is sprayed onto dried tracks must be due to it achieving adequate wetting of the sample for sufficient co-crystallisation to occur. Interestingly, an increase in signal was observed with both post-deposition spray coating methods (post-deposition spraying alone and sandwich method, 5 times greater) than with pre-spraying alone. Secondly, the matrix pre-coating caused spreading of the track. This was observed through monitoring the deposition process with the digital microscope camera and studying the tracks after track-writing with the brightfield microscope described in Chapter 5. Track widths were 300-500 µm when deposited onto coated plates as opposed to approximately 100 µm for uncoated plates, and additionally liquid was observed to spread along the direction of track-writing, in front of the capillary (see Figure 6.2.3d). The layer of matrix prevented a tight liquid cushion from forming at the plate surface; instead, as liquid eluted from the
Figure 6.2.3 – Methods of matrix application. Mass spectra of lysozyme deposited in continuous tracks onto MALDI plates with a) post-deposition, b) pre-deposition and c) sandwich spray coating matrix application methods. d) Track-writing onto matrix coated and uncoated stainless steel plates, monitored using the digital microscope camera.
capillary, it rapidly wetted the matrix and spread out across the surface. This could cause mixing of components eluting from the capillary, and spreading of the analyte across a wider surface which may partially account for the lower signal intensity observed.

As a result of these experiments, post-deposition spraying was selected as the method of matrix application used in this work. In addition to the enhanced signal seen for the pre-sprayed tracks and smaller track width, post-deposition spraying without precoating with matrix is important to the concept of the technology platform presented in this thesis. Depositing matrix before writing tracks would constrain the readout method to MALDI-MS. With deposition of the matrix post track-writing possible (and, indeed, preferable), as has been shown, multiple readout methods can be performed.

6.2.2.3 Detection of separated proteins in tracks using MALDI-TOF/MS

Following the development of a matrix spraying protocol, proteins could be successfully detected at separate locations in tracks in the order predicted from the electropherogram, as is shown in Figure 6.2.4. Capillary electrophoresis separations of lysozyme and α-chymotrypsinogen A at concentrations of 400 µg/ml, injected for 5 kV for 20 s were deposited onto MALDI plates and sprayed with matrix following deposition. This equated to approximately 332 fmol of lysozyme and 118 fmol of α-chymotrypsinogen on target. No analyte ions were detected at other points in the track. As expected, the α-chymotrypsinogen A signal is at a lower intensity than lysozyme due to the higher molecular mass [189, 190] and the smaller number of moles deposited.

In experiments using the fluorescence track-reader, correlation could be shown between UV absorbance electropherograms, fluorescence images of tracks and MALDI mass data of proteins at these points, demonstrating the application of the developed matrix method to enable multiple readout methods to be employed for a separation of proteins. Correlation between
Figure 6.2.4 – Detection of separated proteins using MALDI-MS, showing the electropherogram of the track-writing process for lysozyme and α-chymotrypsinogen A and the corresponding MALDI-MS spectra. Proteins at 400 µg/ml were injected for 5 kV for 20 s, separated at 20 kV in a 50 µm i.d., 60 cm total length UV transparent tapered-outlet capillary pre-coated with UltraTrol LN and written onto MALDI target plates. Track-writing was performed using a voltage of 1 kV, a pressure of 1 mBar and a stage speed of 60 µm/min. Following deposition the track was sprayed with sinapinic acid matrix made up at a 10 mg/ml solution in 50/50 acetonitrile/0.1% TFA. Mass spectra were acquired using a Waters MicroMass MALDI-MS operated in positive ion/linear mode.
fluorescence images and MALDI-MS could be made using the engraved numbered crosses on the Waters MALDI imaging plates. The position of fluorescent spots in relation to the crosses was recorded using a torch to increase light levels. The crosses could be seen with the camera on the MALDI-MS and at the points on the track at which desorption-ionisation occurred the co-ordinates in terms of the crosses were noted.

The three forms of readout: UV electropherogram, fluorescence image and mass spectra are demonstrated in Figure 6.2.5. To test the sensitivity of the system, reduced sample concentrations of 100 µg/ml lysozyme and α-chymotrypsinogen A were used with injection at 5 kV for 9 s. This corresponded to 31 fmol of lysozyme and 11 fmol of α-chymotrypsinogen A. This is an improvement over sensitivity measurements performed using the dried droplet method of matrix application, in which lysozyme and α-chymotrypsinogen could not be detected at concentrations of 1 µg/ml which corresponded to 70 fmol of lysozyme and 39 fmol of α-chymotrypsinogen on target. The differences in sensitivity could arise from the matrix application method or from track-writing, which results in concentration of the sample over a smaller surface area.

6.2.3 Conclusions

This section describes the successful development of the track-writing equipment as an offline continuous deposition CE-MALDI interface which was applied to a group of proteins. The CE-MALDI interface was used to obtain qualitative information about proteins at separate locations in the tracks. By combining the track-writing interface with an identification technique, it becomes a real proteomic tool.

The CE-MALDI interface is simple and versatile, which enables its use with complimentary readout methods to MALDI. There are no size restrictions for the substrate and no customisation of MALDI plates is required; the only restriction is that it must be conducting, and, accordingly, a MALDI
Figure 6.2.5 – Multiple readout methods for lysozyme and α-chymotrypsinogen A: UV electropherogram, intrinsic fluorescence and MALDI mass spectra. Proteins at 100 µg/ml were injected for 5 kV for 9 s, separated at 20 kV in a 50 µm i.d., 70 cm total length UV transparent tapered-outlet capillary pre-coated with UltraTrol LN and written onto MALDI target plates. Track-writing was performed using a voltage of 1 kV, a pressure of 1 mBar and a stage speed of 60 µm/min. Following deposition the track was sprayed with sinapinic acid matrix made up at a 10 mg/ml solution in 50/50 acetonitrile/0.1% TFA. Mass spectra were acquired using a Waters MicroMass MALDI-MS operated in positive ion/linear mode.
plate for almost any MALDI mass spectrometer can be used. This contrasts with other CE-MALDI interfaces, which require specially designed plates, for example, with grooves [150, 149, 187] and/or a matrix coated membrane [149, 187]. Matrix application is performed after track-writing. As a result, multiple readout methods can be used to study a single sample, rather than confining the readout from the start of the separation to MALDI-MS. The practicality of this was demonstrated by performing fluorescence imaging of the track prior to applying matrix and performing MALDI-MS.

6.3 CE-MALDI Imaging

Previous experiments showed that tracks could be analysed with MALDI-TOF/MS, an essential part being the finding that analyte ions could be generated when matrix was sprayed onto dried tracks. The next stage was to use MALDI-mass spectrometry imaging (MALDI-MSI) to determine a spatial mass spectrometric image of the distribution of proteins or peptides in tracks using the Waters MALDI SYNAPT G2 HDMS.

To do this a collaboration was established with Waters Corporation, Manchester together with Dr Judit Nagy at Imperial College. In the following section all separation optimisation and deposition experiments were performed by myself. Matrix application and MALDI-MSI were performed by Dr Ian Edwards at Waters Corporation, Manchester.

6.3.1 Introduction

MALDI-mass spectrometry imaging (MALDI-MSI) is a technique which has mainly come into emergence in the last 15 years, with its greatest use for in situ analysis of peptides and proteins in tissue sections.[191] Matrix is uniformly deposited over the surface of the tissue section, usually either by spraying with a thin layer chromatography (TLC) sprayer or airbrush or using a spotting device.[188] The sample is rastered in front of the laser
and mass spectra are recorded according to pixels on the surface. Image reconstruction creates a map of masses of desorbed ions over the surface. Hundreds of different proteins or peptides can be mapped in a single measurement [161]. The lateral resolution ranges from 30-200 µm [170, 161], and is limited by the diameter of the laser (150 µm in these experiments), the size of the matrix crystals deposited (≲ 100 µm for spray coating [192]) and the step size of the imaging process (set at 50 µm in these experiments).

Despite excitement in the field of MALDI-MSI, it has seldom been explored as an identification technique for liquid phase separation techniques as is described here. The principle of CE-MALDI-MSI was demonstrated in 1999 by Zhang et al [187], where sample was deposited in a continuous track onto a matrix precoated membrane and a standard MALDI instrument was used with software developed to enable scanning of the plate and construction of images. However, this work was not furthered, and since the developments in MALDI-imaging instrumentation and software to create the technique of MALDI-MSI as it is known today, there have been no examples of readout from CE with MALDI-MSI until this year, during the writing of this thesis.[193] To the best of my knowledge, the CE-MALDI-MSI presented in this thesis was a novel advance at the time.

The Waters MALDI SYNAPT G2 HDMS combines MALDI-MSI with state of the art MS technologies, achieving resolutions of > 40 000 FWHM and exact mass (1 ppm RMS). The instrument obtains fragmentation information using a novel alternative to MS/MS, termed MS$^E$. Parallel alternating scans are performed at high and low collision energy to generate both precursor and product ions in a single run. This eliminates the need to re-run a sample following MS for MS/MS and results in an increase in the data acquired as information is not discarded if it falls below an MS/MS threshold. An additional advantage of the MALDI SYNAPT G2 HDMS for the future of complex protein mixtures in this work is ion mobility separations, ('Tri-Wave' technology) which enables separation of species based on size, charge,
6.3.2 Materials and Methods

To show proof of principle the coupling of capillary electrophoresis and MALDI-MSI for the first time, tracks of separations of angiotensin II, leu-enkephalin and met-enkephalin were deposited onto Waters MALDI imaging target plates and analysed using the Waters MALDI SYNAPT G2 HDMS. These peptides were chosen because the instrument has a mass cut off of approximately 10 kDa as it can only be operated in reflectron mode, and cannot be operated in the linear mode which is required for protein mass determination.

Separations and Track-writing

Angiotensin II, leu-enkephalin and met-enkephalin were purchased from Sigma-Aldrich and were of analytical reagent grade. 2 mg/ml stock solutions of each peptide were prepared in water and then used at 100 µg/ml in a 10× dilution of the BGE.

Initial separations were performed with 50 mM ammonium acetate, pH 5.5 and using UltraTrol LN capillary precoating. Sample was injected at 5 kV for 9 s and separated at 20 kV. However, it was found that sharper peak shapes were obtained without UltraTrol LN and consequently the BGE pH was lowered to reduce the EOF where separations were performed without UltraTrol LN. The optimised conditions used a BGE of 50 mM ammonium acetate, pH 4.0 and uncoated 50 µm i.d. UV transparent, tapered-outlet capillaries, as shown in Figure 6.3.1. Capillaries were prepared and rinsed between runs including a sodium hydroxide rinse, as described in Chapter 3.

Two tracks of angiotensin II, leu-enkephalin and met-enkephalin were deposited on Waters MALDI imaging target plates. Separation and deposition were performed using a background electrolyte of 50 mM ammonium acetate,
Figure 6.3.1 – CZE separation of angiotensin II, [Leu]-enkephalin and [Met]-Enkephalin. The sample, containing each peptide at 100 µg/ml, was injected at 5 kV for 9 s on a 50 µm i.d. bare UV transparent, tapered-outlet capillary with a total length of 55 cm. Separation was performed at 20 kV using a 50 mM ammonium acetate, pH 4.0 BGE.

pH 4. A fused silica capillary without inner wall coating was used with a 50 µm i.d. and 55 cm total length. Sample was injected at 5 kV for 15 s, injecting approximately 0.8 pmoles angiotensin II, 1.0 pmoles leu-enkephalin and 0.9 picomoles met-enkephalin. Separation was performed at 20 kV. Track-writing was performed at 1 kV with 2 mBar applied pressure and a stage speed of 50 µm/min.

Matrix Application

The matrix CHCA was used at a concentration of 5 mg/ml in 30/70 acetonitrile/0.1% TFA. Spray coating of tracks was performed using an Eclipse HP-CS Artist’s Airbrush (Iwata Studio Series, Yokohama, Japan). The airbrush was held 15 cm from the target plate and 4 coats of matrix were made before the solvent had dried. Following drying, a further two cycles were made.

Prior to performing MALDI-MSI of tracks, sensitivity tests were performed with angiotensin I and leu-enkephalin at Waters Corporation, Manch-
ester. Angiotensin I and leu-enkephalin were spotted onto Waters MALDI imaging target plates, sprayed with matrix and mass spectra recorded. Angiotensin I was detected at signal to noise ratios of $\sim 2:1$ for 500 fmol on target, 2:1 for 1 pmol and $>5:1$ for 10 pmol. However, leu-enkephalin could only be detected at very low signal to noise ratios ($\lesssim 2:1$).

**MALDI-MSI of Tracks**

Matrix was applied to tracks using the spray coating method described. Tracks were analysed using the MALDI SYNAPT G2 HDMS (Waters Corporation, Manchester, UK). For each track, an area covering the track area (approximately $14\, \text{mm} \times 1\, \text{mm}$ for each track) was selected to be imaged using MALDI Imaging Pattern Creator (Waters Corporation, Manchester, UK), as shown in Figure 6.3.2. The MALDI SYNAPT G2 was run in resolution mode for MS-TOF analyses with a mass range set at 100-2000 $m/z$. The spatial resolution was set at 50 $\mu\text{m}$.

Following data acquisition, data was evaluated using MassLynx. Data were converted into Analyse file format using MALDI imaging converter (Waters Corporation, Manchester, UK) and the MALDI imaging software, BioMAP (Novartis, Basel, Switzerland) was used to reconstruct spatial MS images of the distribution of angiotensin II and a matrix adduct (693.97 $m/z$) across the tracks. Due to low signal to noise ratio, leu-enkephalin and met-enkephalin were not analysed.

**6.3.3 Results and Discussion**

Ion intensity images of the distribution of angiotensin II in tracks showed a localisation at the start of the track. This is demonstrated in Figure 6.3.3a. The matrix adduct at 693.97 $m/z$ was spread evenly along the length and breadth of the track, in contrast to angiotensin II, highlighting the localisation of angiotensin II in the track. As MALDI-MSI acquires positional
Figure 6.3.2 – MALDI imaging pattern creator is used to select areas to be automatically imaged. An area covering the first track has been selected and is highlighted in red.

Information and identity of the analytes simultaneously, MALDI-MSI can be used as a single readout technique without the need for fluorescence imaging prior to identification (in this experiment, it was also not possible to use the track-reader in the described setup because the peptides did not contain tryptophan). Furthermore, the information obtained is more informative and requires less processing as both the location and the mass of a species are conclusively correlated since they were recorded in the same experiment.

Imaging tools such as MALDI-MSI are extremely useful for developmental work for this reason. Analysis of the BioMAP images shows that the area over which angiotensin II is deposited is greater than predicted by theory. Through analysis of the spreading and comparison to theory and electropherograms it can be deduced it occurs largely due to the application of the matrix. Track-writing equations (Chapter 5), give a theoretical spot length of approximately 133 \( \mu \)m for the CE and track-writing conditions from this experiment. The deposition electropherogram (Figure 6.3.3b) can be used to estimate that angiotensin II would have been deposited over a length of 135 \( \mu \)m (track 2) through a calculation of the times at which the front and tail of the angiotensin II peak will arrive at the outlet (106.3 and 109.0 minutes.
Figure 6.3.3 – a) Spatial MS image showing the distribution of angiotensin II (1046.5438 m/z) and matrix adduct (693.97 m/z) in the tracks. Images were generated using BioMAP MALDI imaging software (Novartis). b) Electropherogram for deposition of angiotensin II, leu-enkephalin and met-Enkephalin. The sample, containing each peptide at 100 µg/ml, was injected at 5 kV for 15 s on a 50 µm i.d. bare UV transparent, tapered-outlet capillary with a total length of 55 cm. Separation was performed at 20 kV using a 50 mM ammonium acetate, pH 4.0 BGE. Deposition was performed with a voltage of 1 kV, an applied pressure of 2 mBar and a stage speed of 50 µm/min.
CHAPTER 6. CE-MALDI

respectively) and the stage speed of 50 µm/min.

Analysis of the BioMAP images shows that a much greater amount of spreading has occurred on the surface than calculated from the electropherogram and from theory. Angiotensin II is deposited over a length of 2.65 mm in track 2 (and 1.77 mm in track 1). Although additional bandbroadening will occur during track-writing, importantly, the width of both tracks is 400 µm but initial track widths prior to matrix deposition were ~100 µm. The increase in track width in addition to track length signifies that matrix application is a major cause of spreading rather than the track-writing process.

Advancement in matrix application is necessary. This may be achieved simply by further optimisation of the spray coating method. Indeed, during a later visit, it was observed that using the artist’s airbrush and the described matrix application method in this section, a greater degree of wetting of the sample was observed than in the author’s developed spray coating method described in Section 6.2, which had one fewer coating per cycle. Furthermore, the sensitivity tests gave poorer sensitivities than the experiments discussed in Section 6.2 despite using a superior instrument and lower mass peptides. An alternative would be to use a membrane, as used by Zhang et al [149, 187]. This would not be compatible with the principle of multiple readout methods, but MALDI-MSI provides both imaging and identification reducing the need for fluorescence microscopy.

6.3.4 Conclusions

To the best of my knowledge, this was the first time that MALDI-mass spectrometry imaging had been used to identify samples from capillary electrophoresis. MALDI imaging is an ideal readout technique for an interface of this type because it supplies both positional information, equivalent to that gained by intrinsic fluorescence, and mass information, and hence identification. Once further advancements in matrix application are achieved to preserve the separation efficiency, it will be particularly suited to samples of
high complexity because of the automated nature of data collection from the area of track and generation of the corresponding image.

6.4 CE-MALDI Applications for Ribosomal Proteins

The aims of this work include the development of a novel CE based proteomics platform and to apply this to the analysis of the ribosomal proteins, as separated by capillary electrophoresis in Chapter 4.

Typical protocols within proteomics include preparation of a sample (in this case, affinity purification of ribosomes), separation by some form of chromatographic or electrophoretic procedure and mass spectra generated using ESI or MALDI. I have demonstrated this with offline coupling of CE-MALDI. However, for a full proteomics workflow to be demonstrated by the developed technology platform, other essential elements of a typical protocol must be possible in the track if required; namely chemical processing of the proteins. In a proteomics setting, one of the first steps in characterisation is often tryptic digest of the proteins. In this approach, tryptic digest is performed after separation.

To achieve MALDI imaging of ribosome tracks using the Waters MALDI SYNAPT G2 HDMS (as most intact ribosomal proteins exceed the mass limit of 10 kDa) and to enable sequence determination with tandem MS, tryptic digest of the ribosomal proteins was necessary. In this work, on plate tryptic digest of tracks was performed by spotting droplets of trypsin that did not merge over the area of the track. An initial on plate tryptic digest protocol was developed using an Applied Biosystems 4800 MALDI TOF/TOF Analyzer. This was then applied to ribosome tracks, which were analysed using the Waters MALDI SYNAPT G2 HDMS. MALDI-MSI was performed in collaboration with Waters Corporation, Manchester.
6.4.1 Development of an On-Plate Tryptic Digest Protocol

MALDI-imaging of ribosomal tracks required digestion of proteins as the instrumentation has a mass cutoff at approximately 10 kDa and thus it was necessary to develop a method to digest ribosomal proteins in tracks. This demanded reconciliation of the requirements of small volumes of liquid to limit resolution loss with a trypsin incubation period. The minimum is approximately 15 minutes [195], whilst overnight incubation is commonly used. The success of a tryptic digest was assessed by tandem MS analysis.

Trials were first performed on ribosome tracks and dried spots of ribosome sample. It was found to be simple to dispense an array of trypsin spots of the submicrolitre volume. Very small spots of trypsin (0.05 µl) dried too rapidly, but 0.5 µl spots stayed moist for long enough to give mass spectra showing tryptic peptides. Factors to optimise were primarily droplet volume, incubation environment, incubation time and trypsin concentration. Improvements in the incubation environment, such as increased humidity, enabled increases in signal to noise ratio to be made and sequencing of tryptic peptides. Trials were performed with trypsin concentrations of 2.5 ng/µl, 5 ng/µl, 7.5 ng/µl and 10 ng/µl, using previously optimised droplet volume and environment conditions. An optimised concentration of 10 ng/µl was chosen. The optimised method is described in this section.

6.4.1.1 Materials and Methods

Ribosome sample was prepared as for CE (as described in Chapter 4), spotted down in 0.5 µl spots onto an Applied Biosciences LC-MALDI plate and allowed to dry in air for 2 hours. Trypsin was sequence grade (Trypsin Gold Mass Spectrometry Grade, Promega, UK) and was made up at a concentration of 10 ng/µl in 50 mM ammonium bicarbonate, pH 8.4. The trypsin solutions were applied in 0.5 µl spots on top of the dried ribosomal protein
spots and the plate was placed in a plastic cassette with damp tissue, sealed with parafilm and incubated at 37 °C overnight. For application of matrix, following tryptic digest, 10 mg/ml CHCA in 50/50 acetonitrile/0.1% TFA was applied to the dried spots of trypsin in 0.5 µl droplets.

MS and MS/MS were performed using an Applied Biosystems 4800 MALDI TOF/TOF Analyzer in collaboration with Paul Hitchen at the Mass Spectrometry Centre for Integrative Systems Biology at Imperial College London. The mass range was set at 800-4000 Da. For MS data collection, 1000 laser shots were fired per spot. Precursor ions were automatically selected for MS/MS which had a signal to noise ratio of ≥50. MS/MS spectra were acquired with 1250 laser shots.

All MS/MS data was analysed using the Mascot search engine (Matrix Science, London, UK). Mascot was set up to search the SwissProt Fungi database for the digestion enzyme Trypsin. The database was searched with a fragment ion mass tolerance of 0.30 Da and a parent ion tolerance of 100 ppm. Oxidation of methionine and iodoacetic acid derivative of cysteine were specified as variable modifications.

6.4.1.2 Results and Discussion

An on-plate tryptic digest method was developed in which trypsin was applied in 0.5 µl spots and incubated at 37 °C in a sealed container of moist tissue overnight. By performing the digestion in a humid environment, trypsin spots did not dry out, tryptic digest was thus successful and ribosomal proteins could be identified using MS/MS (see Figure 6.4.1).

This demonstrates the potential to perform on plate trypsin digest with 0.5 µl droplets. Dispensing 0.5 µl spots along tracks is relatively simple and thus this method can be applied to enable digestion of ribosomal proteins in tracks for MALDI-MSI or sequencing with tandem MS. A further advantage of performing on-plate tryptic digest is for experiments where only tiny quantities of material are available and avoiding sample loss due to adsorp-
### Table 6.4.1

Mascot hits generated for the on-plate tryptic digest of ribosomal protein spots. Proteins shown have ion scores > 20 and contain at least 1 identified peptide. Individual ions scores > 28 indicate identity or extensive homology.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Protein Name</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60S ribosomal protein L13-A, <em>S. cerevisiae</em></td>
<td>95</td>
</tr>
<tr>
<td>2</td>
<td>40S ribosomal protein S29-A, <em>S. cerevisiae</em></td>
<td>66</td>
</tr>
<tr>
<td>3</td>
<td>40S ribosomal protein S8, <em>S. cerevisiae</em></td>
<td>47</td>
</tr>
<tr>
<td>4</td>
<td>40S ribosomal protein S21-A, <em>S. cerevisiae</em></td>
<td>40</td>
</tr>
<tr>
<td>5</td>
<td>40S ribosomal protein S27-B, <em>S. cerevisiae</em></td>
<td>34</td>
</tr>
<tr>
<td>6</td>
<td>60S ribosomal protein L4-A, <em>S. cerevisiae</em></td>
<td>32</td>
</tr>
<tr>
<td>7</td>
<td>40S ribosomal protein S3, <em>S. cerevisiae</em></td>
<td>30</td>
</tr>
<tr>
<td>8</td>
<td>60S ribosomal protein L27-A, <em>S. cerevisiae</em></td>
<td>29</td>
</tr>
<tr>
<td>9</td>
<td>40S ribosomal protein S24, <em>S. cerevisiae</em></td>
<td>25</td>
</tr>
<tr>
<td>10</td>
<td>60S ribosomal protein L2, <em>S. cerevisiae</em></td>
<td>24</td>
</tr>
<tr>
<td>11</td>
<td>40S ribosomal protein S11, <em>S. cerevisiae</em></td>
<td>22</td>
</tr>
<tr>
<td>12</td>
<td>Suppressor protein STM1, <em>S. cerevisiae</em></td>
<td>21</td>
</tr>
</tbody>
</table>
tion of proteins and peptides during an in-tube digestion method would be particularly desirable.

6.4.2 CE-MALDI Imaging of Ribosome Tracks

MALDI imaging was used to analyse ribosomal proteins separated by capillary electrophoresis and deposited by track-writing onto MALDI target plates. Ribosomal proteins localised in the track were identified by imaging ribosomal tryptic peptides generated by the on-plate tryptic digest of tracks using the developed method described in Section 6.4.1.

6.4.2.1 Materials and Methods

Tracks were deposited of separations of ribosomal proteins onto Waters MALDI imaging plates. Ribosome samples were prepared as previously described (Chapter 4, Section 4.2.2). CE separation and deposition were performed using a background electrolyte of 40 mM ammonium acetate, pH 4. UV transparent fused silica capillaries were used of dimensions 75 µm i.d. with a total length of 70 cm. Capillaries were prepared using the rinsing protocol described in Chapter 5, and included use of the dynamic precoating agent, UltraTrol LN. Sample was injected at 60 mBar for 120 s, injecting 897 nl. Separation was performed at 20 kV. Track-writing was performed at 1 kV with 1 mBar applied and a stage speed of 50 µm/min.

Following track-writing, tracks were digested using the protocol described in Section 6.4.1. Briefly, trypsin was used at a concentration of 10 ng/µl in 50 mM ammonium bicarbonate, pH 8.4 and was applied in 0.5 µl along tracks using a Hamilton syringe, as close together as was possible without merging. The plate was incubated in a sealed petri dish with moist tissue at 37°C overnight.

The matrix CHCA was used at a concentration of 5 mg/ml in 30/70 acetonitrile/0.1% TFA. It was applied spot-wise in spots of 0.2-0.3 µl covering
CHAPTER 6. CE-MALDI

the entire area of the track.

Tracks were analysed using the MALDI SYNAPT G2 HDMS (Waters Corporation, Manchester, UK) as described in Section 6.3. The MALDI SYNAPT G2 was run in resolution mode for MS-TOF analyses with a mass range set at 100-3000 m/z. The spatial resolution was set at 200 µm.

Following data acquisition, data was evaluated using MassLynx. Data were converted into Analyse file format using MALDI imaging converter (Waters Corporation, Manchester, UK) and the MALDI imaging software, BioMAP (Novartis) was used to reconstruct spatial MS images of the distribution of a number of ribosomal proteins and a porcine trypsin autocatalysis peak (1045.56 m/z) across the tracks. Ribosomal protein assignments were made by comparing experimentally determined tryptic peptide masses in the track to a list of theoretical masses of ribosomal tryptic peptides previously acquired in the LC-MS/MS analysis of a ribosome affinity purification.

6.4.2.2 Results and Discussion

Using MALDI imaging, ribosomal proteins were identified following capillary electrophoresis and track-writing. Ribosomal proteins were localised in certain regions of the track whereas the porcine trypsin autocatalysis peak was identified across the length of the track. This is demonstrated in Figure 6.4.1, which shows a photograph of the track following application of the matrix in addition to the MS images.

The ribosomal protein assignments were made using a single peptide per protein and are therefore tentative. In peptide mass fingerprinting, typically 3 or 4 peptides are required to identify a protein from a database of 50 000 proteins.[196] In this case, qualitative information in the form of knowledge of the protein complex was available and thus limiting the possible peptides to those from a ribosomal purification made identification from one peptide possible. To identify ribosomal proteins with greater certainty would require PMF or MSE. Tandem MS was performed on one track, but the signal
Figure 6.4.1 – a) Photographic image of Ribosome track for MALDI imaging after trypsin digest and spotting with matrix and b) Spatial MS image generated using BioMAP MALDI imaging software (Novartis) showing the distribution of ribosomal proteins and porcine trypsin autocatalysis peak (1045.56 m/z) in the tracks. Assignments were made by comparing experimental mass measurements to theoretical ribosomal tryptic peptide masses obtained from the S. Cerevisiae database using the Mascot search engine. The theoretical m/z values were as follows: 40S ribosomal protein S3, m/z = 1903.8435, 40S ribosomal protein S19-A, m/z = 1927.9428 and 40S ribosomal protein S11, m/z = 1855.9825.
intensity of tryptic peptides was too low to acquire any spectra.

Although further optimisation of sample preparation are required to achieve high resolution MALDI imaging and MS\textsuperscript{E} identification of ribosomal proteins, this work demonstrates the proof of principle of a MALDI readout for the ribosomal separations demonstrated in Chapter 4. This is a completely novel platform for studying ribosomal proteins incorporating affinity purification, capillary electrophoresis, trypsin digestion and MALDI-MSI.

6.4.3 Conclusions

In this section I develop a method to perform a tryptic digest in 0.5 µl volumes, enabling MS/MS analysis of ribosomal proteins and apply this to perform MALDI-MSI of ribosome tracks. Chemical processing, such as proteolytic digest, is a crucial part of a many proteomic studies and thus it is useful to add this to the post-track-writing analysis tool kit.

In this methodology, final analysis is of peptides but initial stages of analysis is of proteins. Although it would have been possible to separate and deposit a tryptic digest of ribosomal proteins, there are certain advantages of this 'middle-up' style of proteomic approach. With on-plate trypsin digest, MALDI-MS/MS does not have to be defined as the identification technique at the outset of the experiment; instead, other techniques, top-down techniques such as EVV 2DIR spectroscopy for example, can be used and trypsin digest with MALDI-MS/MS can be used as the final stage. Accordingly, the analysis of both intact protein and tryptic peptides from one sample would be possible.

6.5 Summary

In this chapter, the adaptation of the track-writing interface to CE-MALDI is described. MALDI was used successfully as a means to provide qualitative information (ie. their mass) about separated proteins in tracks. In
comparison with other continuous deposition interfaces, this platform is extremely versatile, with no customisation required for MALDI plates and matrix applied post-track-writing, which enables many applications. The use of complimentary readout methods to MALDI was demonstrated by performing fluorescence imaging prior to MALDI-MS. The interface was adapted for the novel application of coupling with MALDI imaging (CE-MALDI-MSI). Finally, an on-plate trypsin digest protocol was successfully devised for ribosomal proteins in tiny volumes to enable identification of unknown proteins in tracks and this was used with MALDI-imaging.

With some improvements in sample preparation, the CE-MALDI-MSI interface presented can be considered a high throughput and higher mass accuracy alternative to a 1D gel. In a 1D gel, following the separation of bands, staining of the gel is performed and masses are obtained through comparison to a marker, to store the gel the gel can be photographed or dried, and bands must be excised for more exact mass determination with MS or identification with tandem MS. By comparison, in this approach, following separation and deposition, bands are imaged and exact masses (SYNAPT HDMS can achieve up to 1 ppm [194]) or identifications using tandem MS achieved in the same experiment. Furthermore, all of the advantages of the Waters SYNAPT G2 HDMS can be exploited for the sample, which include ion mobility separation, hence additional structural information can be gained. The technique is highly automated, from the generation of the laser raster area using the image pattern creator to automated system control and simplified system setup such as calibration of plate using IntelliStart Technology (Waters Corporation, Manchester, UK), reducing the opportunity for human error and increasing data throughput. The instrument is only able to study molecules up to 10 kDa, but as MALDI-MSI becomes more widely used, instruments able to operate in linear mode may have an imaging function.
Chapter 7

Capillary Electrophoresis - Electron-Vibration-Vibration Two Dimensional Infrared Spectroscopy
7.1 Introduction

Today, there exists a toolkit of techniques for proteomics: separation techniques such as electrophoresis and chromatography, identification techniques such as mass spectrometry and techniques for studying structures, such as x-ray crystallography and nuclear magnetic resonance (NMR). Interestingly, optical approaches, such as infrared spectroscopy and Raman spectroscopy are not widespread in proteomics, in part due to spectral congestion.

In many respects, mass spectrometry (MS) can be considered the workhorse of proteomics today and, through coupling with liquid phase chromatography and the use of soft ionisation techniques, it has reached a high throughput, high sensitivity and wide applicability.[6] Bottom-up approaches dominate, although the scientific community is now well aware of the associated problems which can result in incorrect protein identifications, information loss and, from the separation aspect, associated difficulties with the added complexity of the separation. Top-down approaches show promise but are less widely used as they require expensive instrumentation.[18]

The quantification of protein levels is a crucial part of proteomics yet is difficult and time consuming to achieve with MS due to the differences in mass spectrometric response which result from variations in physicochemical properties between proteins, as was discussed in Chapter 1. Although techniques exist to achieve quantification using 2D gel electrophoresis, these do not implicitly provide protein identification and are low throughput.[20] There is a lack of high throughput and accurate quantitative proteomic approaches. Thus a strong case exists for the development of novel protein identification tools to complement MS. Spectroscopic techniques are promising for their ability to achieve quantification as (for most spectroscopies) there is a simple relationship between signal intensity and the quantity of sample. Electron-vibration-vibration two dimensional infrared (EVV 2DIR) spectroscopy is a non-linear spectroscopy that has been shown to be able to achieve absolute quantification of proteins.[1]
In this chapter, the successful use of EVV 2DIR as a readout technique for CE is demonstrated for the first time. This technology is termed ‘CE-2DIR’. EVV 2DIR is well suited to coupling with CE. EVV 2DIR can take samples in a number of forms, including those deposited in either solid or liquid form onto a polished metal slide. Hence it is ideal for CE-2DIR in which the sample is deposited onto a metal slide. EVV 2DIR is a top-down technique with no mass restrictions. It is label free and requires no extensive sample preparation that could compromise the resolution or efficiency of the separation within the track. It is non-destructive, so following analysis a track could in principle be analysed with, for example, MALDI-imaging. Crucially for the field of proteomics, EVV 2DIR has the ability to achieve absolute quantification.[1]

The work presented in this chapter was completed in conjunction with E.M. Gardner from the D.R. Klug research group at Imperial College, who performed the EVV 2DIR analyses presented.

### 7.2 EVV 2DIR Spectroscopy

To perform EVV 2DIR, three picosecond pulsed lasers are focused on the sample: two in the infrared and one visible beam \(^1\) (see Figure 7.2.1). The two infrared beams are individually tuneable in frequency and scan a selected region of the vibrational spectrum. The coherences excited in the sample by the infrared beams are probed by the visible beam, and signal is retrieved through coherent Raman scattering in the form of visible photons (see Figure 7.2.2). The output signal is given as \(\omega_\delta = \omega_\beta - \omega_\alpha + \omega_\gamma\), where \(\omega_\alpha\) and \(\omega_\beta\) are the infrared pulses, \(\omega_\gamma\) is the visible pulse and \(\omega_\delta\) is the output signal. When the two IR beams are in resonance with coupled vibrational modes, the signal is enhanced. A two-dimensional map of scanned wavenumbers reveals cross

---

\(^1\)the visible beam is an 800 nm beam which is referred to as the ‘visible’ beam to distinguish it from the infrared beams
Figure 7.2.1 – EVV 2DIR set-up in reflection mode. Three picosecond pulsed lasers are focused onto the sample, deposited onto a reflective substrate. Two beams are tuneable infrared beams ($\omega_\alpha$ and $\omega_\beta$), which scan a region of the vibrational spectrum. Coupled vibrations are probed by the 800 nm beam ($\omega_\gamma$). Output signal is in the visible ($\omega_\delta$).

Figure 7.2.2 – EVV 2DIR wave mixing energy level diagram. Rather than showing absorption and emission, arrows denote the creation and destruction of coherences. Level a is a vibrational fundamental, c is a combination band and e is a virtual electronic state. $\omega_\alpha$ and $\omega_\beta$ are the tuneable infrared pulses, $\omega_\gamma$ is the visible pulse and $\omega_\delta$ is the output signal. The nonlinear output signal is given as $\omega_\delta = \omega_\beta - \omega_\alpha + \omega_\gamma$. For further information, see work of Fournier et al. [103]
Figure 7.2.3 – EVV 2DIR pulse sequence. Independent timings of beams enhance spectral decongestion and enable the EVV 2DIR pathway displayed in Figure 7.2.2 to be selected over other pathways. Adapted from [103].

peaks for the vibrational couplings.

The timing of the two input IR beams and the visible beam are independent; delays specified in the results are denoted $T_{12}$ for the time between the first and second infrared pulses, $\alpha$ and $\beta$, and $T_{23}$ for the time between the second infrared pulse, $\beta$, and the visible beam, $\gamma$ (see Figure 7.2.3). The pulse ordering is required to select the desired EVV 2DIR pathway shown in Figure 7.2.2 over other coherence pathways that exist. Furthermore, delays between the pulses are crucial to suppress a strong electronic non resonant background that is generated when the beams overlap. As the lifetime of the electronic non resonant background is much shorter (femtoseconds) than that of the resonant signal produced from the vibrational couplings (picoseconds), picosecond delays between pulses can be used to enhance spectral decongestion. For the case of CE-2DIR experiments, non resonant background signal is produced from the BGE and metal substrate. Recording images and spectra with a delay of the order of picoseconds between the beams enables EVV 2DIR signal to be distinguished from the non resonant background signal. A detailed description of the principles and operating procedures of EVV 2DIR can be read in detail in the theses of E. M. Gardner [197] and P.M. Donaldson [198].
7.2.1 Protein Identification via EVV 2DIR Spectroscopy

Protein identification is performed as an optical form of amino acid composition analysis. A protein (or peptide) spectrum displays cross peaks due to the various amino acid side chains and CH$_3$ and CH$_2$ groups. Proteins are identified by the relative intensities of amino acid cross peaks with respect to an internal reference. To establish relative intensities of the amino acids, an internal reference peak is required, for which the vibrational signatures of CH$_3$ and CH$_2$ are used (CH$_x$). Ratioing the intensity of amino acid cross peaks to the intensity of the reference peak yields amino acid/CH$_x$ values. In EVV 2DIR spectroscopy, the signal intensity is proportional to the square of the number of molecules probed by the beams and, consequently, from the intensity ratios, amino acid/CH$_x$ number ratios can be calculated. Proteins are identified by comparing measured amino acid/CH$_x$ ratios to known amino acid/CH$_x$ number ratios for proteins in a database. To date, the vibrational signatures of the three aromatic amino acid side chains, tryptophan, phenylalanine and tyrosine, have been identified and validated. The phenylalanine (F), tyrosine (Y) and CH$_2$ cross peaks are depicted in the diagrammatic representation of, and observed spectrum of the peptide, YGGFF, shown in Figure 7.2.4.

An appealing feature of EVV 2DIR is absolute quantification of protein copy number, achieved simply from spectroscopic measurements of the proteins. As in most spectroscopic methods, the signal intensity has a simple relationship to the quantity of material (in this case, intensity is proportional to the square of the number of molecules probed by the beams). Therefore, the absolute quantity of a chemical group irradiated can be determined by comparing the integrated signal with a linear calibration for known quantities. It has been found that the CH$_3$ and CH$_2$ cross peaks are structurally insensitive (the signal intensity remains the same for all proteins studied) and so a previously obtained calibration can be used to for the quantification of other proteins.[1] This is highly valuable to proteomics, where copy numbers
Figure 7.2.4 –  a) Diagrammatic representation and b) EVV 2DIR spectrum of the peptide YGGFF. The tyrosine (try), phenylalanine (phe) and CH$_2$ cross peaks are clearly identifiable. (adapted from [197] with permission)
are known to be crucial but obtaining this data with MS is extremely non trivial. Absolute quantification with EVV 2DIR is appealing because it is achieved in a label free manner, requiring no additional sample preparation steps.

7.2.2 The Optical Setup

The EVV 2DIR optical setup consists of a commercial Ti:Sapphire laser system (Spectra Physics) with a Spectra Physics Spitfire amplifier, which provides the visible beam (800 nm and 1 µJ energy) and the input pump for two optical parametric amplifiers (OPAs). The spitfire amplifier outputs a 800 nm beam with a 1.5 ps pulse duration and an average energy per pulse of 0.7 mJ energy. This is used to pump two 800C OPA (Spectra Physics), which provide the tuneable infrared beams. One OPAs is set to a centre frequency of approximately 1500 cm\(^{-1}\) (1.5 µJ output) and the other is set to a centre frequency of approximately 3000 cm\(^{-1}\) (3.5 µJ output). A photomultiplier tube retrieves the output signal from the sample (ωδ) in the form of photons in the visible at approximately 700 nm.

7.2.3 State of Development of EVV 2DIR

At present, EVV 2DIR is in its infancy, requiring a full laser lab of instrumentation and expert operation. The group have achieved both peptide and protein fingerprinting by 2DIR spectroscopy. It was first shown that EVV 2DIR could be used to detect and quantify peptides [103], and later work was extended to the identification and quantification of proteins [1]. It has now also been demonstrated that it is possible to identify post-translational modifications, such as phosphorylation, using EVV 2DIR. [199] Work is also ongoing in the research group of D.R. Klug to investigate the use of EVV 2DIR as a structural analysis tool for probing intermolecular interactions. [200]
At the current state of development, in comparison with mass spectrometry (which regularly exceeds femtomole sensitivities), large sample quantities are required. The detection limits for proteins are between $10^{11}$ and $10^{12}$ molecules (~picomoles) [1], and for peptides, approximately $10^{13}$ molecules are required [personal communications with E. M. Gardner]. Detection limits are based on calculations made on the intensity of the CH$_3$ and CH$_2$ groups from a sample deposited over a 100 $\mu$m $\times$ 100 $\mu$m area (approximately the diameter of the laser beam), hence a greater number of molecules are required for peptides.

At present, vibrational signatures have been validated for three amino acids in addition to the CH$_x$ signatures. The use of EVV 2DIR in proteomics requires the position of further amino acid cross peaks to be found to enable unique identification of unknown proteins from large data sets, as three relative amino acid ratios are insufficient to uniquely identify an unknown protein. Therefore, work is ongoing to find the EVV 2DIR cross peaks for additional amino acids. It has been shown that 6-9 amino acid/CH$_x$ ratios are needed to unambiguously identify 90% of the proteins from a dataset of 33,000 human proteins.[197]

### 7.3 Experimental Procedure for CE-2DIR

The track writing equipment described in Chapter 5 was used to deposit either protein or peptide samples, as stated, onto polished stainless steel substrates (with the exception of Figure 7.4.1 in which a silver mirror was used). In terms of CE and track writing, the biggest challenge was to load sufficient sample to meet the detection requirements of EVV 2DIR whilst still resolving components of the sample and preventing blockage of the capillary.

CE and track writing conditions were similar to those described previously. Capillaries were either 50 $\mu$m or 75 $\mu$m i.d. UV transparent fused silica with a heat tapered end. Long lengths of 65-80 cm were used due
to the large loading requirements. Ammonium acetate BGEs were used at several pHs, as discussed in the coming text and figures. Samples injected were simple peptide or protein samples dissolved in a 10 times dilution of the BGE. Track writing was performed as previously described, but slightly higher pressures were sometimes required (up to 5 mBar) due to the large loading quantities. No preparation of tracks was required before EVV 2DIR analysis.

For EVV 2DIR analysis, imaging of the track was first performed to study its topography and to identify regions of high signal. Spectra were then taken at the co-ordinates of areas of high intensity. The position of a cross peak on a spectrum is defined in terms of its two IR frequencies as $\omega_\alpha/\omega_\beta$ cm$^{-1}$. To conduct EVV 2DIR imaging, two IR frequencies corresponding to a cross peak of interest, typically the CH$_2$ ($\sim$1440/2860 cm$^{-1}$) or CH$_3$ ($\sim$1495/2920 cm$^{-1}$) cross peak, are first selected. The sample is rastered in front of the input beams using two automated translational stages (LabVIEW controlled). A plot of signal intensity as a function of the scanned area reveals the regions of high signal. Both imaging and recording a spectrum can be performed with either short (typically pulses overlapping, termed ‘$T_0$’) or long delays (typically 0.5-1.5 ps) to enable the differentiation of EVV 2DIR signal from the non resonant background. Images were first recorded at $T_0$, and spectra were then recorded at longer delays from the regions of high intensity. Depending on the intensity of the signal and the available time (imaging takes several hours), images were sometimes recorded at longer delays prior to taking spectra.

The angle at which the beams cross is calculated in order to meet phase matching conditions. The beam pointing can vary slightly over the frequency ranges scanned. Typically, the beam pointing is optimised at a set of frequencies central to the spectral region being measured; however slight adjustments can be used to enhance a particular cross peak. For cross peak ratios from different spectra to be comparable and used for fingerprinting, all data must
have been recorded with the beam pointing optimised on the same frequencies. In the future, beam pointing will be automatically adjusted as the frequencies are scanned so that they are optimised over the whole spectral region.

Experiments can be performed in either transmission or reflection mode. The spectrometer is used in transmission for samples deposited onto glass slides. For this work, as tracks were deposited onto polished metal substrates, the spectrometer was used in reflection mode.

7.4 EVV 2DIR Analysis of Protein Tracks

As the successful development of protein fingerprinting with EVV 2DIR was well underway, work initially focused on the deposition of protein tracks for analysis with EVV 2DIR. For this aim, tracks containing model proteins were deposited for analysis with EVV 2DIR. A continuing challenge of the work remained the detection limit of EVV 2DIR, requiring a minimum of between $10^{11}$ and $10^{12}$ molecules of each protein to be loaded into the capillary.

Tracks were deposited of combinations of the proteins lysozyme, trypsinogen and $\alpha$-chymotrypsinogen, containing between $10^{11}$ and $10^{12}$ molecules of each protein. When EVV 2DIR analysis was performed, tracks either gave no signal or gave non resonant signal from which no fingerprint could be obtained. A key breakthrough was had with a track of lysozyme and $\alpha$-chymotrypsinogen, deposited onto a silver mirror (Figure 7.4.1). This demonstrated the possibility of imaging separated proteins with EVV 2DIR.

Images of the track were taken with the IR frequencies set on the protein CH$_3$ cross peak (1495 / 2920 cm$^{-1}$). Signal was obtained from the length of the track in the initial image taken at short delays ($T_0$, pulses overlapping). To establish whether signal from the track was background or resonant, the track was imaged at longer delays ($T_{12} = 0.5$ ps, $T_{23} = 0.5$ ps). At short delays, signal was obtained from the length of the track, but at the longer delays,
Figure 7.4.1 – Electropherogram and corresponding EVV 2DIR images for a separation and deposition of lysozyme and α-chymotrypsinogen (500 µg/ml). EVV 2DIR imaging was performed at both short delays ($T_0$) and long delays ($T_{12} = 0.5$ ps and $T_{23} = 0.5$ ps). Performing imaging at a longer delay isolates resonant signal due to the EVV 2DIR pathway, and it can be seen that at longer delays signal is localised in two regions corresponding to the electropherogram. CE and track writing were performed in a 75 µm i.d., 60 cm total length UV transparent tapered outlet capillary pre-coated with UltraTrol LN using a 50 mM ammonium acetate pH 4 BGE. Sample was injected at 5 kV for 20 s, introducing approximately $4 \times 10^{11}$ and $2 \times 10^{11}$ molecules of lysozyme and α-chymotrypsinogen respectively. Separation was performed at 20 kV. Track-writing was performed with a voltage of 500 V and a pressure of 2 mBar. The track was written onto a silver mirror.
Figure 7.4.2 – EVV 2DIR image and spectra recorded for the lysozyme and α-chymotrypsinogen track shown in Figure 7.4.1. The spectra both show non resonant background. The second spectrum was stopped part way through recording when it became clear that only background was generated and therefore only a screen shot is available. The spectra were recorded with delays of \( T_{12} = 0.5 \) ps and \( T_{23} = 0.5 \) ps for the first spectrum and \( T_{12} = 1 \) ps and \( T_{23} = 1 \) ps for the second spectrum.
signal was localized at two regions close to the beginning of the track. This demonstrates that the signal from these regions was resonant, potentially as a result of an EVV 2DIR pathway, and not due to the non resonant background. These regions correspond with the electropherogram from the track writing process in both the relative intensities of the two proteins using the two detection methods and in their positions on the track. Calculations from the migration times in the electropherogram predict lysozyme to be deposited at 4.5 mm along the track and α-chymotrypsinogen at 7.8 mm which is well replicated in the EVV 2DIR image in which the the two regions of high signal intensity are observed at 4.8 mm and 6.9 mm.

Spectra (Figure 7.4.2) were taken from the two pixels which showed the most intense signal at long delays. However, the spectra from both regions only showed non resonant background. This could either be due to the quantity of sample per pixel falling short of the detection limit or to the morphology of the track. Although $7 \times 10^{11}$ molecules were injected, it can be seen from the image taken at long delays that, for the lysozyme peak, these are in fact spread over approximately fifteen $100 \mu m \times 100 \mu m$ pixels and, given that the detection limit is estimated at $10^{11} - 10^{12}$ molecules per pixel, the number of molecules would therefore be likely to be below the detection limit. Hence it seems likely that lack of material is the prime reason for the lack of resonant signal in the spectra.

Nevertheless, this work demonstrates the full procedure for CE-2DIR analysis of a multi-component protein mixture, comprising separation, deposition and EVV 2DIR analysis, requiring no sample pretreatment. With a sensitivity increase in EVV-2DIR or increased loading, the full workflow would be in place. This is therefore the groundwork for CE-2DIR as a proteomic tool.
7.5 EVV 2DIR Analysis of Peptide Tracks

The strongest signals with EVV 2DIR have always been achieved with peptide samples rather than protein samples. It has been assumed that this is because peptides have a less defined third order structure than proteins which enables them to pack more closely.\[1\] Added to this is the ability to choose peptides containing large proportions of the three amino acids with characterised vibrational signatures. Therefore, due to ongoing difficulties in obtaining spectra from protein tracks, attention was focused on obtaining spectra from peptide tracks to achieve proof of principle of a CE-2DIR interface. Peptides were selected which contained a large number of tryptophan (W), phenylalanine (F) or tyrosine (Y) groups and a control, containing none of these residues. These peptides were designed by E. M. Gardner and had been synthesised earlier for previous EVV 2DIR experiments using an internal peptide synthesis service at Imperial College London.\[201\] The peptides chosen for use in CE-2DIR experiments were as follows: LRRAALG, LRRWWLG, FLRRFSLG, YLRRYSLG and YYPFF. These were selected as a track containing these peptides would theoretically contain 5 regions of resonant signal showing: just the \( \text{CH}_2 \) cross peak, W and \( \text{CH}_2 \), F and \( \text{CH}_2 \), Y and \( \text{CH}_2 \) and both Y and F and \( \text{CH}_2 \) respectively. A diagrammatic representation of the predicted EVV 2DIR spectra for a track of the first four of these peptides is shown in Figure 7.5.1. The peptides used in the CE-2DIR experiments and their corresponding molecular weights and isoelectric points are presented in Table 7.5.1.

7.5.1 Method Development

To achieve EVV 2DIR analysis of a CE deposited track, as demonstrated in Section 7.5.3, substantial experimentation was necessary with the deposition and analysis of tracks containing various permutations of the 5 peptides. This was to match the loading of CE to the current sensitivity of EVV
Figure 7.5.1 – Diagram of how a track of the four peptides LRRAALG, LRRWWLG, FLRRFSLG and YLRRYSLG would exhibit the reference peak and amino acid cross peaks in individual locations. LRRAALG contains none of the amino acids for which spectral signatures have been found and thus only the CH$_2$ cross peak would be expected. In addition (not shown), since certain residues contain CH$_3$ groups (notably alanine, A, and leucine, L), the CH$_3$ could be expected for all 4 peptides and would be likely to be seen most strongly for LRRAALG. Spectra adapted from [197].
2DIR. Separation and track writing with high loading requirements are more difficult for more complex samples, because the limit of detection must be exceeded for all the components in the sample. For a given detection limit, a greater total quantity of material must be loaded for a more complex sample, increasing the likelihood of problems such as capillary blockage.

It must be noted that the challenges experienced with the sensitivity requirements were different to those discussed in Chapter 4, where techniques to increase the quantity of the ribosomal protein sample separated were explored. During the ribosomal protein separations the aim was to increase the quantity of sample whilst retaining the resolution and efficiency, whilst for this proof of principle CE-2DIR work, at the current sensitivity levels of EVV 2DIR it emerged that the challenge was much more fundamental: the capillary must not be blocked. In the later experiments discussed in this chapter (tracks deposited of two peptides, Figure 7.5.8), it was found that trimming the capillary tip to a wider end (≈200 µm) could be performed to reduce blocking and increase the amount of sample deposited.

**Separation Optimisation**

Optimisation of the separation was not trivial as it was imperative that resolution was maximised due to the requirement to load large sample quantities. Initial separations were performed for a sample containing the 5 peptides

<table>
<thead>
<tr>
<th>Peptide</th>
<th>MW</th>
<th>pI</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRRAALG</td>
<td>755.92</td>
<td>12.00</td>
</tr>
<tr>
<td>LRRWWLG</td>
<td>986.19</td>
<td>12.00</td>
</tr>
<tr>
<td>FLRRFSLG</td>
<td>995.19</td>
<td>12.00</td>
</tr>
<tr>
<td>YLRRYSLG</td>
<td>1027.19</td>
<td>9.99</td>
</tr>
<tr>
<td>YYPFF</td>
<td>735.84</td>
<td>5.52</td>
</tr>
</tbody>
</table>

**Table 7.5.1** – Peptides selected for CE-2DIR and their respective molecular weights (MW, Da) and isoelectric points (pI).
but despite extensive modifications of the CE conditions, difficulties were had with obtaining a sharp peak from the peptide YYPFF. As a result, the separation was optimised for the 4 peptides, LRRAALG, LRRWWLG, FLRRFSLG and YLRRYSLG and a 50 mM ammonium acetate BGE at pH 9.0 was selected. As such a high pH was used, it was necessary to use the dynamic pre-coating reagent, UltraTrol LN, to reduce the EOF. These conditions were initially used for track writing. In later track writing experiments, a 50 mM ammonium acetate BGE at pH 4.0 was used with UltraTrol LN.

**Description of CE-2DIR Experiments**

Results from EVV 2DIR analysis of the peptide tracks directed the choice of subsequent peptide tracks to write and track writing conditions. Initial tracks were deposited of 4 peptides, to demonstrate the full range of EVV 2DIR spectral features used for identification (see Figure 7.5.1), but to increase sample loading to meet detection limits of EVV 2DIR it was found to be necessary to use simpler mixtures of peptides. Therefore complementary sets of 3 and 2 peptide tracks were deposited, which together would still show the full set of EVV 2DIR amino acid cross peaks. Finally continuous single peptide tracks were deposited. Furthermore, a change in buffer to 50 mM ammonium acetate BGE at pH 4.0 was made as spectra recorded from tracks in the pH 9.0 50 mM ammonium acetate BGE exhibited sum frequency generation (SFG) due to their increased crystalline nature.

CE-2DIR results are presented from the following sets of peptides:

- LRRAALG, LRRWWLG, FLRRFSLG and YLRRYSLG
- LRRAALG, FLRRFLSG and YLRRYSLG
- LRRAALG and YYPFF
- YYPFF
The preparation of these tracks is discussed below. The CE-2DIR results for these tracks are presented in the next section.

**LRRAALG, LRRWWLG, FLRRFSLG and YLRRYSLG**  Initial work concentrated on separation and deposition of tracks comprising of the 4 peptides LRRAALG, LRRWWLG, FLRRFSLG and YLRRYSLG without the 5th peptide, YYPFF, due to the discussed difficulties with YYPFF during preliminary separations. Using a 50 µm i.d., 72 cm total length capillary with sample injection for 15 s, two tracks of separation of the 4 peptides were deposited, depositing approximately $8 \times 10^{12}$ molecules of each peptide. Results from EVV 2DIR analysis of the tracks indicated that peptide levels fell short of the detection limit. Further tracks were written with increased sample loading, injecting up to $2 \times 10^{13}$ molecules of each protein but in all cases, blockage of the capillary occurred during the track-writing process, causing catastrophic current failure.

**LRRAALG, FLRRFSLG and YLRRYSLG**  Although the limits of loading with 4 peptides appeared to have been reached, a set of complementary tracks, between them depositing peptides with the different aromatic residues (such as tracks of LRRAALG, FLRRFSLG & YLRRYSLG and LRRAALG & LRRWWLG) would enable the principle of CE-2DIR to be demonstrated yet enabling greater loading. Three tracks of LRRAALG, FLRRFSLG & YLRRYSLG were deposited using the conditions described previously, using 75 µm capillaries and injecting in the region of $2 \times 10^{13}$ molecules of each peptide. Spectra obtained from the tracks did not exhibit features attributable to the EVV 2DIR pathway, but instead evidence of sum frequency generation (SFG), which is characteristic of crystalline structures, was observed. These results are discussed in Section 7.5.2. Tracks written using the pH 9 ammonium acetate (50 mM) BGE were observed to have a more crystalline appearance than those written with the pH 4 ammonium acetate (50 mM) BGE. Thereafter, tracks were written using a pH 4 amo-
nium acetate (50 mM) BGE to obtain an improved morphology. UltraTrol LN was used to maintain resolution.

**LRRAALG and YYPFF** Out of the initial 5 peptides chosen, YYPFF contains the highest ratio of the aromatic residues to CH$_2$ and was therefore most likely to give cross peaks. As a result, despite early difficulties with YYPFF separations, three tracks were written of separations of YYPFF and LRRAALG. Signal was obtained from all three tracks; however, in the first two the intensity of the signal was too low to obtain a spectrum. In these tracks, approximately $3 \times 10^{13}$ molecules of each peptide were injected and it may be that spreading on the metal plate resulted in the number of molecules per pixel falling below the detection limit. In the third track, approximately $9 \times 10^{13}$ molecules of each peptide were injected. To perform a separation with an injection quantity of this magnitude the tapered end of capillary was trimmed to give an o.d. of approximately 200µm and by doing this, blocking did not occur.

**YYPFF** To deposit tracks of higher peptide quantities and demonstrate proof of principle CE-2DIR, continuous tracks of single peptides were deposited using track-writing conditions and studied with EVV 2DIR. The capillary was filled with a high concentration solution of the peptide (2 mg/ml) and a track written with the peptide solution as the inlet vial (during which a voltage of 1 kV was applied). Tracks were deposited of the peptides YYPFF, LRRWWLG and LRRAALG. It is difficult to provide an exact value of the number of molecules deposited with this method, as electrokinetic injection of the peptide will have taken place during the track writing process. Approximate values can be obtained by calculating the number of molecules of peptide deposited onto the entire track ($5 \times 10^{15}$) and dividing this by the area of the track to obtain the number of molecules per pixel ($\sim 1 \times 10^{13}$).
7.5.2 Analysis of Tracks with EVV 2DIR

The tracks described in Section 7.5.1 were analysed by EVV 2DIR spectroscopy. EVV 2DIR was first used as an imaging tool to study the distribution of signal intensity over the length of the track. If signal intensity was great enough, spectra were recorded. The aims were therefore twofold: to record images of the distribution of signal in tracks of separated proteins and to record spectra of peptides deposited.

The results of the EVV 2DIR analysis, in terms of the spectroscopic features observed and signal intensity, is dependent on the quantity of sample deposited per pixel and the morphology of the track. In the results presented it is demonstrated how increases in sample deposition quantities and choices of buffer to give appropriate track morphology enabled improvements in the quality of the spectra, with the result that the full spectra of amino acid and CH$_2$ cross peaks were obtained for a peptide track. In the experiments leading up to this result, spectra contained several features in addition to the desired cross peaks generated by the EVV 2DIR pathway. These are summarised in Figure 7.5.2.

In some figures (in particular, Figures 7.5.4 and 7.5.7), two electropherograms are shown because to maximise efficiency, the high voltage separation phase was run as long as possible. Track writing was started after the first peak had passed the detector. Therefore, to show all peaks, two electropherograms are required: from separation at high voltage and track writing at low voltage.

**LRRAALG, LRRWWLG, FLRRFSLG and YLRRYSLG** Tracks of the 4 peptides, LRRAALG, LRRWWLG, FLRRFSLG and YLRRYSLG containing $\sim 8 \times 10^{12}$ molecules of each peptide were analysed with EVV 2DIR. Signal levels were too low to image at long delays or obtain useful spectra indicating that the amount of peptide deposited fell short of the detection limit.
Figure 7.5.2 – Examples of features due to EVV 2DIR, non resonant background, Raman process and sum frequency generation (SFG) observed in the spectra of the peptide tracks recorded using EVV 2DIR spectroscopy.
The results from EVV 2DIR analysis of a 4 peptide track are shown is Figure 7.5.3. The electropherogram from the track writing process, the EVV 2DIR images (measured with 2 different step sizes) and a photograph of the track are shown. Signal was obtained from the track with no delays between the pulses \(T_0\) but it was not intense enough to measure at long delays, either due to a lack of material or due to the morphology of the track causing, for example, scattering. The track can be seen as a shadow of negative signal, which is due to the reduced reflection from the buffer track in comparison to that from the polished metal slide. Comparison of the EVV 2DIR image to the photograph demonstrates that the EVV 2DIR image can give a representation of the topology of the track and, furthermore, that the background of the BGE is low. Due to the low signal intensity, spectra were not recorded at locations along the track. Increasing sample loading of the 4 peptides sample was not achieved as it resulted in capillary blockage.

**LRRAALG, FLRRFLSG and YLRRYSLG** Reducing the complexity of the sample injected enabled a greater number of molecules of each component peptide to be loaded and deposited. EVV 2DIR analysis from tracks of three peptides, LRRAALG, FLRRFLSG and YLRRYSLG \((\sim2 \times 10^{13} \text{ molecules of each peptide})\), showed increased signal levels, with EVV 2DIR images showing localised regions of intense signal. This is demonstrated in the results shown in Figure 7.5.5. The EVV 2DIR images show high signal localised in three regions along the length of the track. This corresponds well with the three peaks shown in the electropherogram. Calculations from the electropherograms predict elution along the track of LRRAALG at 4.4 mm, FLRRFLSG at 5.9 mm and YLRRYSLG at 7.7 mm. This is extremely well replicated in the EVV 2DIR image, which are deposited at approximately 4.2 mm, 5.4 mm and 8.9 mm respectively past the start of the track. Note that in the EVV 2DIR image, the track starts at 3 mm.

Spectra were taken at the three regions showing intense signal. The type
Figure 7.5.3 – Electropherogram, corresponding EVV 2DIR analysis and photograph for the 4 peptide sample: LRRAALG, LRRWWLG, FLRRFLSG, YLRRYSLG (500 mg / ml). CE and track writing were performed in a 50 µm i.d., 72 cm total length UV transparent tapered outlet capillary pre-coated with UltraTrol LN using a 50 mM ammonium acetate pH 9 BGE. Sample was injected at 60 mBar for 15 s, introducing approximately $8 \times 10^{12}$ molecules of each peptide. Separation was performed at 20 kV. Track-writing was performed with a voltage of 1 kV and a pressure of 1.2 mBar. EVV 2DIR imaging was performed with 2 stage increments: 0.1 mm and 0.05 mm, the latter to gather a more defined image of the track. The region of shadow corresponds with the photograph of the track and is explained by scattering from the track.
Figure 7.5.4 – Electropherogram, corresponding EVV 2DIR analysis and photograph for the 3 peptide sample: LRRAALG, FLRRFLSG and YLRRYSLG (400 mg / ml). EVV 2DIR imaging shows three regions of high intensity signal which correspond in position to the electropherogram. The track starts at 3 mm on the EVV 2DIR scale bar. CE and track-writing were performed in a 75 µm i.d., 77 cm total length UV transparent tapered outlet capillary pre-coated with UltraTrol LN using a 50 mM ammonium acetate pH 9 BGE. Sample was injected at 60 mBar for 10 s, introducing approximately $2 \times 10^{13}$ molecules of each peptide. Separation was performed at 20 kV. Track-writing was performed with a voltage of 1 kV and a pressure of 1 mBar.
Figure 7.5.5 – Image and corresponding spectra for EVV 2DIR analysis of the track of LRRAALG, FLRRFLSG & YLRRYSLG shown in Figure 7.5.4. Spectra were recorded at the three regions of high signal on the track. The vertical lines at $\omega_\alpha \approx 1400 \text{ cm}^{-1}$ are a result of sum frequency generation (SFG). The diagonal features are the appearance of a Raman part of the signal created by the visible beam. Spectra were recorded with delays of $T_{23} = 1 \text{ ps}$ and $T_{23} = 1 \text{ ps}$ and signal averaging of 1 s.
of signal was not due to an EVV 2DIR pathway. The spectra exhibited characteristic features of sum-frequency generation (SFG), a non-linear process observed on the spectra to be only dependent on one IR input beam, $\omega_\alpha$. SFG was apparent in EVV 2DIR spectra recorded for a number of other tracks deposited under these conditions. SFG is characteristic of ordered or crystalline structure and therefore signifies that the morphology of the track, possibly as a result of the BGE, might be a problem. Indeed, it was observed that the tracks deposited in pH 9 ammonium acetate (50 mM) were more crystalline in nature and wider than those previously deposited in lower pH ammonium acetate (50 mM). This is demonstrated by photographs taken of the tracks at pH 9 and pH 4 ammonium acetate (50 mM), which show a substantially different morphology (Figure 7.5.6).

SFG was not observed in any of the later tracks deposited in pH 4 ammonium acetate (50 mM) and analysed with EVV 2DIR that gave detectable signal (such as the track shown in Figure 7.5.6 d) and therefore it is likely that the improved morphology of the tracks written with the pH 4 BGE resulted in a reduction of SFG.

**LRRAALG and YYPFF** With the increased simplicity of the two peptide tracks, a further increase in the number of molecules deposited per peptide species was possible, and a corresponding improvement in the EVV 2DIR spectra was achieved. The EVV 2DIR spectroscopy of the YYPFF and LRRAALG track containing $\sim 1 \times 10^{14}$ molecules of each peptide were the most successful so far, with spectra demonstrating features characteristic of an EVV 2DIR pathway (Figure 7.5.8). This was an exciting result that demonstrated the feasibility of CE-2DIR on surpassing the detection.

Imaging showed an intense region of signal and three less intense regions of signal. Although only two peptides were used, difficulties were had in obtaining a peak from YYPFF, as can be seen from the electropherogram. It therefore follows that the track could be deposited as a band of LRRAALG
Figure 7.5.6 – Photographs of peptide tracks. All 4 tracks gave detectable EVV 2DIR signal when imaged. SFG was observed in spectra from both tracks shown at pH 9, but not in those at pH 4. The tracks deposited using 50 mM ammonium acetate, pH 9 clearly have a more crystalline nature than those deposited at 50 mM ammonium acetate, pH 4. Tracks a)-c) are depositions of LRRAALG, FLRRFLSG & YLRRYS; track d) is a deposition of YYPFF and LRRAALG. The number of molecules deposited per peptide are as follows: a) $1 \times 10^{13}$ molecules, b) $2 \times 10^{13}$ molecules, c) $2 \times 10^{13}$ molecules, d) $3 \times 10^{13}$ molecules.
with YYPFF spread at other regions in the track (see Figure 7.5.8).

At the most intense region shown in the EVV 2DIR image, spectra taken showed signal attributable to an EVV 2DIR pathway. Cross peaks close to the CH$_2$ and CH$_3$ characteristic frequencies can be observed. This indicates that the peptide LRRAALG was present at this point on the track, as both alanine and leucine contain CH$_3$ groups whereas the residues in the peptide YYPFF do not. Spectra were taken at this location at a number of delays (see Figure 7.5.8). At longer delays, the non resonant background, which was observed at delays of 1 ps was not observed, but the CH$_3$ cross peak was. As would be expected for a spectrum of LRRAALG, no cross peaks due to the aromatic residues were observed. The other regions of signal in the track, which can be assumed to be due to YYPFF due to the tentative assignment of the first region to LRRAALG, only gave non resonant background. The reason for the success of the EVV 2DIR of LRRAALG is likely to be due to the amount per pixel exceeding the detection limit, in addition to a buffer choice that gave an appropriate track morphology. It can be seen that this region is spread over approximately 10 pixels, and therefore the number of molecules per pixel will be close to the limit of detection of $10^{13}$ molecules. This demonstrates the first CE-2DIR analysis of a multi-component peptide mixture, comprising separation, deposition and EVV 2DIR analysis, requiring no sample pretreatment and achieving (albeit tentative) identification.

Through the developmental work of CE-2DIR, it is apparent that a general problem is simply a shortage of sample on the metal substrate. Furthermore, even if the quantity of sample injected exceeds the detection limit, spreading on the surface beyond the 100 $\mu$m $\times$ 100 $\mu$m pixel size frequently results in the quantity on the surface per pixel falling below the detection limit. Crucial to the success with the track of LRRAALG and YYPFF was the initial injection of $10^{14}$ molecules per peptide so that despite spreading of LRRAALG over 10 pixels, the concentration per pixel still exceeded the detection limit of $10^{13}$ molecules per peptide. Tracks deposited of the same
Figure 7.5.7 – Electropherogram, corresponding EVV 2DIR analysis and photograph for the 2 peptide sample: LRRAALG and YYPFF (500 mg / ml). CE and track writing were performed in a 75 μm i.d., 70 cm total length UV transparent tapered outlet capillary pre-coated with UltraTrol LN using a 50 mM ammonium acetate pH 4 BGE. Sample was injected at 60 mBar for 40s, introducing approximately $1 \times 10^{14}$ molecules of each peptide. Separation was performed at 20 kV. Track-writing was performed with a voltage of 1 kV and a pressure of 1.5 mBar. EVV 2DIR imaging shows an region of high intensity signal at $x \sim 2.5$mm and three further regions of lower intensity signal.
Figure 7.5.8 – EVV 2DIR image and associated spectra of a track of the peptides LRRAALG and YYPFF. The CH$_3$ and CH$_2$ peaks can be observed in spectra from the region of high intensity towards the beginning of the track. Spectra were recorded at increasing delays to distinguish cross peaks from non resonant background. The spectrum recorded at approximately 5.5 mm along the track shows only non resonant background.
peptides, under identical conditions (and in addition, a number of other depositions mentioned in the text in this section) but with $3 \times 10^{13}$ molecules per peptide injected showed an insufficient intensity of signal during imaging for spectra to subsequently be taken (see Figure 7.5.9). It is challenging to increase the amount of sample loaded into the capillary as capillary blocking is likely to occur using tapered capillaries; however, using non tapered capillaries increases the width of the track. As discussed in Chapter 5, and used in Figure 7.5.8 to load $\sim 1 \times 10^{14}$ molecules of each peptide, a capillary i.d. of approximately 200 $\mu$m appears to be the optimum. Further improvements in the deposition interface to reduce the longitudinal spreading and an increase of sensitivity in EVV 2DIR are necessary.

In addition to the quantity of material deposited, the track morphology appears to be implicated in the success or failure of EVV 2DIR analysis. In particular, the failure of some to give signal from the EVV 2DIR pathway in cases where signal intensities were high can be attributed not to a lack of material but to the crystalline nature of the track due to the observance of SFG type signatures in the spectrum. These were only observed in tracks deposited using pH 9 ammonium acetate (50 mM) BGEs, and not at pH 4
Figure 7.5.10 – EVV 2DIR image of a continuous deposition of the peptide YYPFF. The track was written at a voltage of 1 kV. EVV 2DIR imaging was performed with delays of $T_{12}=1\ \text{ps}$ and $T_{23}=1\ \text{ps}$. The delays isolate resonant pathways from the non resonant background, and considerable signal can be observed along the length of the track. The area of signal in the EVV 2DIR image follows the shape and dimensions of the track shown in the photograph.

**YYPFF**  As coordinating sensitivity of EVV 2DIR with loading restrictions of CE was a current constraint, to demonstrate the principles of CE-2DIR, continuous tracks of single peptides were deposited using track writing conditions and studied with EVV 2DIR. These tracks contained approximately $5 \times 10^{15}$ molecules in total. Dividing this by the area of the track gives a total number of molecules per pixel of approximately $1 \times 10^{13}$.

EVV 2DIR analysis of the LRRWWLG and LRRAALG tracks gave no signal. EVV 2DIR analysis of the YYPFF track gave intense signal at long delays (1 ps) along the length of the track, as is shown in Figure 7.5.10. Spectra were taken at the three locations of the most intense signal. The spectra revealed EVV 2DIR signal showing the tyrosine, phenylalanine and $\text{CH}_2$ cross peaks, and additionally a new cross peak at approximately $1595 / 3190\ \text{cm}^{-1}$ which is due to contributions from both aromatic side chains. EVV 2DIR spectra recorded at one point in the track are displayed in Figure 7.5.11.

The success of this result indicates that the reason for not obtaining
Figure 7.5.11 – EVV 2DIR spectra recorded for the YYPFF track shown in Figure 7.5.10. The tyrosine, phenylalanine and CH$_2$ cross peaks can all be observed. A new cross peak, which is due to contributions of the aromatic groups is also visible in the second spectrum at approximately 1595 / 3190 cm$^{-1}$. The spectra were recorded with delays of T$_{12}$ = 1.5 ps and T$_{23}$ = 1 ps. The intensity of cross peaks varied as spectra were recorded with the beam pointing optimised on two slightly different sets of frequencies, both in the centre of the spectral region studied.
<table>
<thead>
<tr>
<th>Peptide Track</th>
<th>Number of Molecules</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>LRRAALG, LRRWVLG, FLRRFSLG and YLRRYSLG</td>
<td>$8 \times 10^{12}$</td>
<td>Signal too weak</td>
</tr>
<tr>
<td>LRRAALG, FLRRFLSG and YLRRYSLG</td>
<td>$2 \times 10^{13}$</td>
<td>Image correlates with electropherogram but spectra show SFG</td>
</tr>
<tr>
<td>YYPFF and LRRAALG</td>
<td>$1 \times 10^{14}$</td>
<td>EVV 2DIR - CH$_2$ and CH$_3$ cross-peaks</td>
</tr>
<tr>
<td>YYPFF and LRRAALG</td>
<td>$3 \times 10^{13}$</td>
<td>Signal too weak</td>
</tr>
<tr>
<td>YYPFF</td>
<td>$1 \times 10^{13}$ per pixel</td>
<td>EVV 2DIR - CH$_2$, Y and F cross-peaks</td>
</tr>
</tbody>
</table>

*Table 7.5.2 – Summary of CE-2DIR results of peptide tracks*

Adequate spectra from previous tracks was not due to the morphology but a lack of material. In this instance, the quantity of YYPFF deposited per pixel ($\sim 10^{13}$ molecules were deposited per pixel on average) surpasses the detection limit where as on the other occasions it did not. This shows that it is possible to use EVV 2DIR as a readout method for CE if sufficient material is present. To date, this is the only example of EVV 2DIR to be applied to samples from CE which shows spectroscopic features from both amino acid side chains and CH$_x$ groups.
7.5.3 CE-2DIR: Proof of Principle

The successful spectra for the YYPFF track demonstrate the potential of EVV 2DIR as a readout method for CE. The advantages of EVV 2DIR, as discussed earlier, are immense. There exist two main challenges in this work; these relate to the sensitivity of EVV 2DIR and track morphology. At present, it is difficult to perform CE and track-writing with the sample quantities required for EVV 2DIR analysis, particularly with a multi-component system. The disparity in scale between CE and readout means that some of the most appealing features of CE cannot be utilised. However, the sensitivity of EVV 2DIR is improving, and it is aimed to be at the femtomole level within several years. With additional improvements to the focusing of the sample on the plate, the advantages of CE-2DIR will then be realisable.

Track morphology is an issue that can be tackled from the perspective of track-writing, although it is exacerbated by the requirement to deposit large quantities of material. The persistent appearance of SFG in tracks that showed strong signal implied that the crystalline nature of tracks could affect the signal. A number of options exist to tackle this, including experimentation into optimal buffer choice. A difficulty is that to separate large sample quantities, fairly high buffer concentrations are required to avoid electrodispersion but this may not be ideal for EVV 2DIR. Another option is to rehydrate tracks to a gel like phase using a controlled humidity cell.

7.6 Conclusion

This work has laid the foundations for CE-2DIR, which is a novel proteomic tool with the ability to achieve absolute quantification. EVV 2DIR spectroscopic features from tyrosine, phenylalanine and CH$_2$ were identified in a track from the peptide YYPFF. In a separation of the peptides LRRAALG and YYPFF, tentative assignment of LRRAALG was achieved from CH$_2$ and CH$_3$ spectroscopic features. Furthermore, EVV 2DIR was successfully used
to record images of tracks from separated proteins and peptides that corresponded with the electropherograms. This work demonstrates the feasibility of EVV 2DIR as an identification method for CE. EVV 2DIR is an ideal qualitative and quantitative analysis technique for coupling to CE, particularly for this style of interface, as no sample preparation steps are required that cause a loss of resolution. The non-destructive nature of the technique leaves open the possibility for analysis of the sample with complementary techniques after EVV 2DIR analysis. These are the first and to date only examples of EVV 2DIR as an analysis technique for a sample deposited with capillary electrophoresis.

At present, EVV 2DIR is at an early stage of its development. To fully employ the advantages of CE in terms of throughput and scale, EVV 2DIR will also need to be advanced in terms of throughput and scale. In terms of throughput, it is envisaged that when the cross peaks from a greater number of amino acids have been characterised, a protein will be analysed by measuring the amino acid/CH$_2$ ratios at those corresponding frequencies rather than scanning the whole spectral space, enabling a reduction of the identification time for a protein to between 4 and 45 s.\cite{103} The use of CE-2DIR as a proteomic tool will be reliant on an increase in the sensitivity of EVV 2DIR to enable the analysis of multi-component peptide or protein samples. I estimate a three order of magnitude improvement in sensitivity to be required, in order to analyse samples containing 100 components whilst maintaining the resolution of the separation. EVV 2DIR aims to improve the current picomole sensitivity to femtomole within the next 2 years. The application of CE-2DIR to tackle real proteomic problems will also require the spectral signatures of further amino acids to be characterised. Theoretical spectra have been calculated that give the positions of other amino acid cross peaks. These can be used to look experimentally for the cross peaks of amino acids which have long enough lifetimes to be usable.

Once further advances are made, an excellent application of CE-2DIR
CHAPTER 7. CE-2DIR

will be as a platform to study eukaryotic ribosomal proteins. For studies into ribosomal protein substoichiometries and extraribosomal activities, clearly a quantitative analysis technique would be highly advantageous. In comparison to the work with MALDI imaging, no trypsin digestion of the tracks would be required, no matrix application is necessary, and no additional sample preparation steps must be performed. Therefore loss of resolution following deposition is reduced. The simplicity of this strategy is demonstrated in Figure 7.6.1.

The approach of CE-2DIR can be seen as analogous to LC-MS. In a time of MS-based proteomics, there is value in the development of alternative technologies, especially those with such promise to tackle the fundamental questions in proteomics that are difficult to answer with existing technologies. Perhaps at present, both track-writing and EVV 2DIR need to be further advanced to meet in terms of scale and throughput. However, this work is a clear demonstration of a new strategy with outstanding potential. The development of new approaches is essential to take steps forwards in research to places where previously existing techniques cannot reach.
Figure 7.6.1 – Application of the CE-2DIR strategy to the study of ribosomal proteins. Ribosomes are affinity purified, which is in itself a high throughput method of ribosome purification. As the readout method is quantitative, it will be an ideal system to study questions in ribosomal proteomics relating to protein substoichiometries and extraribosomal activities.
Chapter 8

Conclusions
The aims of this thesis were to develop new approaches to capillary electrophoresis based proteomics: firstly, by separating out the intact eukaryotic ribosomal proteins, and secondly, by the development of a novel proteomic platform which was then applied where possible to the ribosomal proteins. The basis of this platform was separation with CE, track-writing and readout with either MALDI-MS or EVV 2DIR. The three main achievements of this thesis are summarised in this chapter. Specifically, three novel developments are the CE separation of the eukaryotic ribosomal proteins, the coupling of MALDI-imaging to CE (CE-MALDI-MSI) and the coupling of CE to EVV 2DIR (CE-2DIR).

8.1 Separation of Ribosomal Proteins

The core of the proteomic platform is a home-built CE instrument which can function both as a standard CE instrument and as a method of sample separation and deposition. In Chapter 4 I present a robust top down separation platform for yeast ribosomal proteins using affinity chromatography and capillary electrophoresis. FLAG tagged ribosomes were affinity purified, and rRNA acid precipitation was performed on the ribosomes followed by capillary electrophoresis to separate the ribosomal proteins. Over 26 peaks were detected with excellent reproducibility (<0.5% RSD migration time) in less than 10 minutes. This is the first reported separation of eukaryotic ribosomal proteins using capillary electrophoresis. The outstanding reproducibility enables the readout to provide a ribosomal fingerprint, suggesting applications to studies of proteins exhibiting extraribosomal functions, substoichiometries, posttranslational modifications resulting in shifting migration times and fluctuations resulting from changes in growth conditions. The two stages in this workflow, affinity chromatography and capillary electrophoresis, share the advantages that they are fast, flexible and have small sample requirements in comparison to more commonly used techniques. This method is a remark-
ably quick route from cell to separation, providing a novel strategy for high throughput studies of the ribosomal proteome.

8.2 CE-MALDI

In Chapter 6, I present the development of a CE-MALDI platform. The track-writing interface provides a chemical record of the contents in the capillary. MALDI provides qualitative information (ie. mass) about separated proteins in tracks, which is not available from a UV electropherogram alone. This was successfully achieved for separations of proteins. This platform is extremely versatile, with no customisation required for MALDI plates and matrix applied post-track-writing, which enables samples to be used with a host of mass spectrometers and multiple readout techniques to be used prior to MALDI, which was demonstrated by performing fluorescence imaging prior to MALDI-MS.

The interface was adapted for use with MALDI imaging (CE-MALDI-MSI). At the time, these techniques had never been coupled before and this advance was advantageous as it provided both imaging and identification in one measurement. CE-MALDI-MSI was applied to the analysis of ribosomal proteins. Ribosomal proteins localised in the track were identified by imaging ribosomal tryptic peptides generated by the on-plate tryptic digest of tracks. This is a novel platform for studying ribosomal proteins which incorporates all stages of a proteomic protocol from cell to readout, consisting of: affinity purification, capillary electrophoresis, trypsin digestion and MALDI-MSI.

8.3 CE-2DIR

The technique of CE-2DIR is demonstrated for the first time in Chapter 7. Proof of principle CE-2DIR was demonstrated by three achievements: firstly, the correlation between EVV 2DIR images and electropherograms of protein
and peptide tracks; secondly, the spectrum tentatively assigned to the peptide LRRAALG localised in a separation of two peptides and, thirdly, the outstanding spectra obtained for a track of the peptide YYPFF. Spectra for the peptide YYPFF showed EVV 2DIR cross-peaks from tyrosine, phenylalanine and CH$_2$. This is a clear demonstration of the applicability of EVV 2DIR as a readout method for CE, and shows the potential for CE-2DIR to be used as a proteomic tool. EVV 2DIR spectroscopy has great value in proteomics, not least because of its ability to achieve absolute quantification in addition to identification. CE-2DIR is crucial to its use as a high-throughput proteomic tool by providing the interface with a high performance separation technique to deliver purified proteins from a biological sample onto a stainless steel target, and therefore this work is an important step towards utilising the benefits of EVV 2DIR in proteomics.

8.4 Future Work

Through the work presented in this thesis, a novel proteomic platform has been developed. All the elements of this proteomic workflow have been demonstrated with a range of readout methods and using the ribosomal proteins as a real biological sample. To utilise this platform as a proteomic tool that can rival existing technologies requires further advances in several of these areas to enable smaller sample quantities to be studied. Improvements in sensitivity can be achieved at either the separation, deposition or readout stage. Development in these areas will enable application to questions in ribosomal proteomics.

Capillary zone electrophoresis was an excellent choice for a separation technique to couple to EVV 2DIR as sample is delivered in an aqueous buffer, free from detergents, organic solvents or gel constituents that would interfere with the identification process. Furthermore, it is well suited to the separation of intact proteins. To enable loading of higher sample quantities and
to enable more complex, real biological samples to be resolved, it may be necessary to add an additional dimension of separation. Multidimensional separations are essential in proteomics to enable high complexity samples to be separated. A form of LC-CE (typically performed with RPLC as the first dimension) [202] or CE-CE [203] would be the most likely option.

Improvements to the track-writing interface would provide concentration on the plate and ensure that the resolution of the separation was maintained in the track, which would be essential for high complexity samples. A significant advance to the system could be made by running the entire process at a high voltage, with the metal substrate in place from the beginning. This would theoretically improve the efficiency, maintain the resolution from the separation and vastly increase the throughput. The track would either need to be written in a serpentine geometry or a small well of BGE could be milled in the plate to enable the capillary to remain stationary until, for example, the first peak had reached the detector.

The application of CE-2DIR to real proteomic questions will require further developments in EVV 2DIR. Firstly, improvements in sensitivity from the current picomole sensitivity to femtomole sensitivity are required, and with developments that are underway in the group, this should be achievable within the next 2 years. Secondly, spectral signatures of further amino acids must be characterised. It has been shown that 6-9 amino acid/CH$_x$ ratios are needed to unambiguously identify 90% of the proteins from a dataset of 33,000 human proteins.[197] Theoretical spectra have been calculated that give the positions of these cross peaks. These are then being used to search experimentally for cross-peaks which have sufficiently long lifetimes to be usable.

One of the clear benefits of this platform is its flexibility. Once a track is written, it provides a chemical record of the separation and components can be analysed in time with a variety of techniques. CE-MALDI will continue, with some advancements required on the matrix application front to reduce
spreading. These may be simply achieved through further optimisation of spray coating or may require some customisation of plates, such as grooves. Forthcoming work in the research group aims to detect proteins in tracks using antibodies, performing a chip based Western Blot style assay.

The achievement of the CE separation of ribosomal proteins paves the way for a host of studies of ribosomal proteomics. This is particularly the case as ribosome purification was performed by affinity chromatography. Therefore, attachment of ribosomes to the affinity gel can be a means by which ribosomes can be ‘probed’ with particular compounds and eluting ribosomal proteins studied. For example, it would be interesting to probe ribosomes attached to the affinity gel with ribonuclease to study how the enzyme disrupts the ribosome. To fully use the affinity purification/capillary electrophoresis ribosomal protein method, identification is required and it would be ideal if it can be achieved using the developed proteomic tool. The ability to achieve absolute quantification would make CE-2DIR an ideal system to study many of the interesting questions in ribosomal proteomics which relate to protein substoichiometries and extraribosomal proteins.
Bibliography


[127] Sigma-Aldrich. Lysozyme grade 1 from chicken egg white - product information, 2008.


[131] John H.T. Luong, Pierre Bouvrette, Yali Liu, De-Quan Yang, and Edward Sacher. Electrophoretic separation of aniline derivatives using


[151] George S. McLeod, Jan Axelsson, Ron Self, and Peter J. Derrick. Comparison of novel sampling methods for the analysis of capillary
electrophoresis fractions by matrix-assisted laser desorption/ionization


[175] M.T. Davis and T.D. Lee. Variable flow liquid chromatography-tandem mass spectrometry and the comprehensive analysis of complex protein...


[201] Department of Chemistry, Imperial College London. *Tate and Leatherbarrow Research Group*.
