Ultra Scale-Down of Elution Chromatography

A thesis submitted to University College London for the degree of Doctor of Engineering

by

Paul James Beckett

Department of Biochemical Engineering
University College London
Torrington Place
London WC1E 7JE

Pall Europe Ltd.
Europa House
Portsmouth
Hampshire PO1 3PD

Acknowledgements

This thesis was brought into existence by the efforts of a great many people over and above myself and I would like to take the opportunity to thank all those who contributed moral and physical support to the project.

I would especially like to thank my parents, Jan and Brian, whose love, support and belief was unyielding even when things were very difficult.

These special thanks also extend to Thea and Emma, two ladies very close to my heart that drove this EngD forward with their care and love, even when I was being unreasonable.

I would also like to thank the staff at Pall Life Sciences for their continued support during the project. Many thanks go to Peter Levison in particular for the direction, advice and opportunities he has offered throughout the EngD. My eternal gratitude also goes to Simon Edwards-Parton, whose technical insight was invaluable throughout the project.

And last, but not least, I would very much like to thank Nigel for his unending patience, super-human efficiency at returning work and putting up with me for 6 years. A better supervisor I could not have had.

Thesis Abstract

It is highly advantageous to be able to develop bioprocesses early on in the product design lifecycle, where strategic options and alternatives can be considered cheaply and effectively. Resources are however often in very short supply at this point in the process, typically with very limited quantities of feedstock available and with minimal access to capital equipment. As such there is significant advantage in being able to develop chromatographic processes at very small scales that will only require small quantities of feedstock and can utilise common laboratory equipment.

If the height of the packed bed is not maintained during scale up then it becomes difficult to predict chromatographic behaviour and efficiency. This seriously limits how small a chromatography system can be and still be representative of how a process scale chromatographic separation will behave. Ultra scale-down (USD) methodologies for chromatography take a different approach. A small scale device, which may or may not have geometric similarity to the process scale equipment, generates data which when combined with a specific methodology and mathematical model allows the prediction of the large scale equivalent.

The work in this thesis sought to develop a USD methodology and model to accurately predict large scale chromatographic behaviour using very small scale devices.

The development of a model separation was required to act as a source of realistic experimental data for the development of a USD methodology. This model separation required a suitable feedstock, a suitable USD device and a suitable sorbent on which to perform the separation.

It was found that the most suitable feedstock of those tested was a FAb fragment containing periplasmic lysate produced in *E. coli,* due to the ease of pre-chromatographic processing and industrial relevance. Several designs of very small column were investigated and it was found that PRC pre-packed columns (Pall Life Sciences) had superior separation characteristics and were therefore selected as the USD device of choice. This was combined to produce a viable separation process using MEP HyperCel presented in the PRC pre-packed column format with FAb fragment containing periplasmic lysate as a feedstock. A linear pH gradient produced a clearly resolved two peak system with excellent FAb fragment purity that was deemed very suitable as a reference separation for the USD methodology development.

The premise of the USD methodology was the deconvolution of each relevant peak within an elution profile into its four curve coefficients, namely height, width at half height, skew and peak location. These four curve coefficients for each peak in the small scale chromatogram could then be individually modelled using a transformation function into the large scale equivalents and then recreate a large scale chromatogram prediction from these values. This was also used to predict how the chromatograms will change with respect to altering the packed bed height and linear velocity of the loading and elution steps. The methodology that was developed was shown to be effective and was typically accurate to under 5% difference normalised root mean square when the predicted large scale chromatograms and real large scale data was compared.

The methodology was further validated by testing with a range of different chromatographic systems and processes. These included changing the feedstock by reducing the FAb fragment titre by 50%, changing the chromatographic ligand to PPA and also a multi-variate change that altered the ligand, the sorbent backbone and the feedstock all at once. The transformation functions were found not to be generic and required system specific alterations. However, the methodology itself was shown to be very effective across a wide range of chromatographic conditions and systems and as such would be a good basis for ultra-scale down development in elution chromatography.

Contents

Abbreviations

BSA	Bovine Serum Albumin
CHO	Chinese Hamster Ovary
CM	Carboxymethyl
CQA	Critical Quality Attribute
CV	Column Volumes
DBC	Dynamic Binding Capacity
DEAE	Diethylaminoethyl
DoE	Design of Experiments
EDTA	Ethylenediaminetetraacetic Acid
ELISA	Enzyme-Linked Immunosorbent Assay
EMG	Exponentially Modified Gaussian
FAb	Fragment Antigen Binding
Fc	Fragment Crystallisable
GUI	Graphical User Interface
HCIC	Hydrophobic Charge Induction Chromatography
HEA	Hexyl-amine
HETP	Height Equivalent to Theoretical Plate
HIC	Hydrophobic Interaction Chromatography
HPLC	High Performance Liquid Chromatography
HTS	High-Throughput Screening
IgG	Immunoglobulin G
IMAC	Immobilised Metal-Ion Affinity Chromatography
ICH	International Conference on Harmonisation

IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	Kilo-Dalton
LLOQ	Lower Limit of Quantification
LB	Lysogeny Broth
MEP	4-Mercapto-Ethyl-Pyridine
mS	mili-Siemens
OPD	*o*-Phenylenediamine
PPA	Phenyl Propyl Amine
QbD	Quality by Design
RPC	Reverse Phase Chromatography
RPM	Revolutions Per Minute
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SEC	Size Exclusion Chromatography
SELDI	Surface Enhanced Laser Desorption/Ionisation
TMP	Trans-membrane Pressure
USD	Ultra Scale-Down
UV	Ultraviolet

1 Introduction

1.1 Historical Perspective

Chromatography (lit. “colour writing”) is a powerful chemical method for the separation of complex mixtures. Although work had been done with paper chromatography in the 1800s the first real column chromatography was carried out by Mikhail Semanovich Tsewett in 1903 (Tsewett et al, 1906) to separate plant pigments and it was he that named the technique due to the separation of colours. Over the next century various advances and new techniques were discovered and put to use in a range of industries and applications.

Chromatography, in a variety of forms, has proved to be an exceptionally powerful technique for the separation of a range of chemical species. In particular it has been developed extensively for bio-separation, particularly in the bio-pharmaceutical industry where purity is paramount.

1.2 Chromatography Concepts

Chromatography is an umbrella term for any technique that relies on the affinity of a sample, or its differential partition, for a stationary phase and a mobile phase. The mobile phase is by definition moving and can either be a liquid or a gas while the stationary phase, sometimes called the sorbent, is usually a porous solid or a liquid supported by a solid. Columns or plates are used to support the stationary phase itself and to provide a conduit for the mobile phase. The system varies depending on the type of chromatography employed, but generally the sample is applied or injected along with the mobile phase onto the stationary phase and moves along it. The constituents of a sample will have a range of affinities for both the stationary phase and the mobile phase and this influences the relative mobility of the components, causing them to separate as they progress down the column or the plate. Components in the sample which either have a high affinity for the mobile phase or a very low affinity for the stationary phase will move very quickly along the column and will exit the column first, while those with low affinity for the mobile phase or high affinity for the solid phase will exit later, leading to separation.

1.3 Industrial Relevance

Initially chromatography systems were used almost exclusively for analytical purposes but the preparatory applications have developed greatly over the decades since its inception. Engineering advances have allowed liquid and expanded-bed chromatography capital equipment to be scaled to hundreds of litres and still provide acceptable throughput, making chromatography an effective and cost-efficient method of purification.

This is particularly important within the biological pharmaceutical industries where often chromatography is the only way to purify the target molecule to regulatory standards; typical chemical separation techniques can involve harsh solvents, high temperatures and high pressures which will render labile biological products, particularly proteins with complex tertiary structures, inactive. As such chromatographic unit operations are present in almost all biotechnological purification processes and are actually a requirement of some regulatory authorities. (Sofer et al, 1989).

Analytical applications for chromatography should not be overlooked however. Biological pharmaceutical development and batch release for commercial supply requires a battery of analytical tests be conducted and chromatographic techniques have proven to be very powerful tools. High Performance Liquid Chromatography (HPLC) in particular has proven to be a reproducible and sensitive quantitative analytical method for analysis of product and impurity profiles, commonly utilising ion-exchange, reverse-phase and gel filtration methods for this purpose.

1.4 Types of Chromatography

The principles of chromatography can be applied in many forms and utilising many chemical interactions to achieve the desired separations.

1.4.1 Surface chromatography

1.4.1.1 Paper and Thin Layer chromatography

Paper chromatography was the forerunner of modern packed bed chromatography and is very simple in principle. The paper fibres act as the stationary phase and the flow of the mobile phase is accomplished by the diffusion of the sample and carrier solvent through the paper. This technique is actually still used in some industries for the test purity of compounds due to the process being inexpensive and fast (Sherma & Fried, 2003).

In most analytical applications however thin-layer chromatography (TLC) is preferred. A glass, metal or plastic plate is covered in a thin layer of a solid absorbent and the samples are spotted onto the plate. Diffusion of the mobile phase up the plate carries and separates the samples which then can be visualised with UV light or a chemical visualising agent, such as molybdic acid for lipids. This technique is still fast and inexpensive but offers improved resolution, though the sample is still essentially non-recoverable after this process, limiting its use to analytical applications (Sherma & Fried, 2003).

1.4.2 Column chromatography

The final type of chromatography is column chromatography, an extensive field with many industrial and analytical applications and as such has many different functions. Column chromatography splits broadly into three types, gas chromatography, liquid chromatography and liquid-liquid (or counter-current) chromatography.

1.4.2.1 Gas chromatography

Gas chromatography is used where the sample is in gaseous form and there are three types in existence. Gas adsorption utilises a packed bed of solid absorbent within the column which the sample is pumped through, commonly used to separate gaseous mixtures (Grob & Barry, 2004). In gas-liquid chromatography the solid part of the packed matrix is inert but it coated with an absorbent liquid, separating the gaseous analytes by order of volatility as affected by affinity for the stationary phase. Both of the above applications uses columns of 2-4mm in diameter which means the column needs to be 1.6-10m long under normal circumstances. Capillary chromatography uses very fine columns (approx 0.2mm) and coats the stationary phase on the inside surface of the capillary.

Detection is often done with thermal conductivity detectors or electron capture detectors, but other detection systems exist for particular analytes and applications.

1.4.2.2 Counter-current chromatography

Counter-current chromatography, sometimes referred to as partition chromatography, exploits the partition coefficient between two immiscible liquids (Foucault et al, 2001). There is no solid component in this form of chromatography so there is no irreversible adsorption (except for the column itself which is designed to be non-absorbing) so high product yields can theoretically be obtained. Modern commercial instruments work on a coil planet centrifuge method to mix the phases and then partition them. This technology is very much developing into the industry at the moment (Marston & Hostettmann, 2006).

1.4.2.3 Liquid chromatography

Liquid chromatography is where the sample, analyte or feedstock is in the liquid phase and the stationary phase is a solid sorbent or a gel supported by a solid sorbent. The solid phases are often made of silica gel, cellulose or agarose materials. The bead can act as the separating material by its physical nature (as in gel filtration) or be coated with ligands for binding the materials of interest. There are many different types and applications of this sort of chromatography and is used extensively in industry as a preparative and analytical tool, particularly within the biotechnology industry.

1.4.2.3.1 High performance liquid chromatography

High performance liquid chromatography (HPLC) uses a very narrow column and high pressures to reduce running time and diffusion within the column, increasing resolution and performance (Lillehoj & Malik, 1989). This system is very commonly utilised within the analytical chemistry/biochemistry fields as it gives excellent resolution with small samples volumes. Use as a preparative technique has been limited so far due to the relatively small capacity and the use of volatile and often highly flammable solvents (in the case of the very common reverse-phase HPLC) making it dangerous, difficult and generally uneconomical to use on an industrial scale.

1.4.2.3.2 Expanded bed chromatography

Expanded bed chromatography uses a sorbent with a variety of bead sizes that are far denser than normal and that can be derivatised with a range of potential ligands. Expanded bed columns are not packed but fluidised, and have an mobile adaptor that can be raised or lowered during operation (Anspach et al, 1999). The up flow, used for primary capture, allows the beads to become buoyant and rise up the column thus expanding the bed and increasing the bed voidage. This means very viscous feed materials, even an unclarified culture directly from a fermentation, can be run through the column and provide selective capture of proteins with negligible pressure drops. When the loaded column is run in down flow the beads settle and it acts more like a packed bed, suitable for concentration and elution (Chase et al, 1994; Fenneteau et al, 2003).

1.4.2.3.3 Packed bed chromatography

Packed bed techniques are the most common form of chromatography used within the biopharmaceutical industry. In this form of liquid chromatography the stationary phase, usually in the form of beads between 20-100μm, is packed under pressure (either with liquid flow or mechanically) into the column to form a consistent and close matrix. This method has the capability to scale up to industrially relevant volumes whilst still retaining, packing, ligand selection and operating conditions notwithstanding, good separation characteristics.

1.5 Packed bed chromatography – Bioprocess applications and ligands

The large variety of potential ligands that can be bound to a large variety of solid phases makes packed bed chromatography very versatile and this is why it is found so frequently within industry for both analytical and preparative sections of processes (Jungbauer et al, 2005). A diagrammatical representation of a typical rig for packed bed chromatography is shown in Figure 1.1.

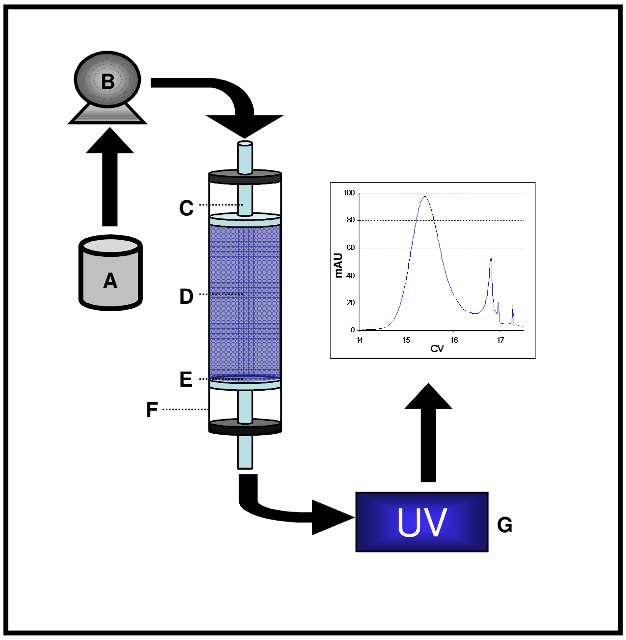


Figure 1.1: A diagrammatical representation of a packed bed chromatography rig. A: Buffer or feedstock, B: Pump, C: Column adaptor, D: Packed bed chromatography matrix, E: Frit, F: Column, G: UV detector.

1.5.1 Stationary phase

The stationary phase, or solid phase or matrix, within a chromatography system impacts heavily on the capabilities of a system. This is obviously the case with gel filtration where the solid matrix, rather than a ligand, is key to the separation potential. However, even within ligand based packed bed chromatography the stationary phase is important as the structure of the matrix, together with the geometries of the column and frit design, impact heavily upon the internal hydrodynamics and therefore separation efficiency of the chromatography system.

Ideally a stationary phase should be hydrophilic to allow wetting, chemically inert, be able to form macroporous structures and have good compressible strength so as to resist the pressures of packing and operation. Polysaccharides, such as cellulose and agarose, have proven to be good materials for chromatographic bead manufacture, but the compromise is a lack of compressible strength in their native form. Advances in silica gel and ceramic designs have begun to address the compressibility problems (Kawahara et al, 1989; Cortes et al, 1987), as have evolutionary advances to the traditional polysaccharide media (O'Neil & Sellick, 2004).

1.5.2 Stationary phase packing

The packing of the matrix into the chromatography column, and from this the bed porosity and matrix structure, is critical to the separation effectiveness of packed bed chromatography. Three forms of packing exist, gravity, flow and mechanical, and they are often used in combination on a single column.

Gravity packing is where the matrix slurry is simply left to settle over a period of time within the column. This method is not particularly effective as the bed height will reduce once a fluid flow is applied, but it can be appropriate for very small scale applications or with ceramic sorbents, which are heavy enough to settle reliably. However, gravity packing is usually taken as a starting point for other packing methods (Guiochon et al, 1997).

Flow packing is where the bed is subjected to a steadily increasing fluid flow, typically using the equilibration and loading buffer of the chromatography process itself to prevent unforeseen pressure or compression issues during later operation and to ensure comparability. Laboratory scale column packing almost solely utilises this method and is it commonly applied to pilot scale columns as well, although the high flow rates required can cause hydrodynamic flow issues with the large scale distributors (Yuan et al, 1999). Flow packing can be problematic with certain stationary phase materials however – cellulose sorbents are naturally elastic and as such require 1000 cm/h linear velocities to pack effectively, a linear velocity many manufacturing chromatography skids are incapable of producing. Ceramic hydrogel sorbents, such as the Ceramic HyperD range from Pall Life Sciences, have also been observed to swell under low ionic strength conditions, so the low ionic strength equilibration buffer for an ion-exchange ligand will create substantial pressure drops during packing. Typically packing flow rates are over 25% more than the fastest expected operation flow rate and within 80-90% of the critical velocity for bed compression. Feeds are often far more viscous than loading buffers, leading to higher pressure drops, therefore a safety factor has to be implemented to prevent bed compression during normal operation.

Mechanical packing is a technique where the bed is physically compressed using a mobile adaptor. This is very common for pilot scale chromatography rigs and larger and may be the only method of producing acceptable packing in very large chromatography columns. The mechanical pack compresses the bed slowly under a slow flow rate to ensure even packing throughout the entire bed.

Gases are a significant problem during the packing process, as bubbles will deform the matrix structure and cause channels within the packed bed that will increase apparent dispersion and peak band broadening. For this reason all buffers are degassed before entry into a packed bed system; this is however rarely appropriate for feed streams as the air liquid interfaces would cause an unacceptable level of protein shear.

The compressibility of the bead is a critical aspect of the packing regime, as is the stationary phase bead size. Bed compression is defined by Equation 1.1,

$$\lambda = \frac{L_0 - L}{L_0} \qquad (1.1)$$

where L_0 is the gravity settled bed height, L is the height of the bed and λ is the level of bed compression.

Bed porosity with a packed bed is then defined from the level of bed compression, as shown in Equation 1.2,

$$\varepsilon = \frac{\varepsilon_0 - \lambda}{1 - \lambda} \qquad (1.2)$$

where ε_0 is gravity settled bed porosity, λ is the level of bed compression and ε is the compressed bed porosity.

Packing quality is typically ascertained by an acetone pulse analysis and HETP calculation, as detailed in section 1.6.1.

1.5.3 Elution

The elution of proteins from a chromatography column is commonly achieved by modification of the mobile phase. This can work by modifying an aspect of the bound protein, an aspect of the stationary phase or by the introduction of a competing species which displaces the proteins of interest. There are two methods of elution commonly used.

Step elution is performed by rapidly changing the conditions to two or three set levels, eluting the proteins off in blocks. This method allows efficient, low volume capture at the expense of resolution.

Gradient elution is performed by altering the elution buffer concentration as a steadily increasing proportion of the mobile phase. This provides much higher resolution and separation of the proteins in the feed, but results in a higher process volume and less capacity.

A combination of the above elution methods is often used within a single bioprocess – step elution is applicable for primary capture where yield is the over-riding priority, with gradient elution being utilised for the purity driven polishing steps.

1.5.4 Negative Chromatography

Sometimes referred to as "contaminant capture" chromatography, this technique is utilised when it is more effective to bind the impurities in a feed stream rather than bind the target molecule itself, with the target molecule appearing in the flow through (Levison et al, 2003). This is commonly employed to remove particularly hazardous contaminants late in the purification process, for example endotoxins, or to remove undesirable nucleic material.

1.5.5 Ion exchange chromatography

Ion exchange chromatography is the most commonly used within industry due to its high capacity and good selectivity, as well as domestically in water filters and wastewater treatment. It is also relatively inexpensive, at least compared to affinity sorbents such as Protein A (Lillehoj & Malik, 1989).

Nearly every protein has charged groups on its surface, due to the acidic and basic amino acids within its structure. Highly polar groups also contribute a small amount of charge, and cumulatively this produces a protein's net charge. Each protein has a certain pH, its isoelectric point (pI), where it has no net charge, but these conditions are often avoided in protein preparation due to the increased likelihood of aggregation when electrostatic repulsion is removed. The polarity and magnitude of a charge found on a particular protein can be manipulated by modifying the pH and ionic strength conditions, although other factors such as the state of the protein's tertiary or quaternary structure and the relative location of the highly charged amino acids throughout the protein structure can also contribute. The differences in these net charges can be exploited for separation purposes by using charged ligands.

There are two types of ion-exchange ligand, cationic (binds positive charges) and anionic (binds negative charges) and these are categorised further into strong and weak charges, which is not a reference to the strength of binding. Strong ion-exchange resins work across a large range of pH and ionic strength values and often have a capacity independent of the buffer pH. On the other hand weak ion-exchange resins are dependent on pH conditions for the capacity they demonstrate and this range tends to be narrow. Examples of commonly used ligands are summarised in Table 1.1.

Proteins can be eluted from an ion-exchange matrix either by changing the pH, thus modifying the net charge, or by increasing the ionic strength of the buffer which introduces a competing ion to displace the protein from the matrix.

Ligand	Classification
Carboxymethyl (CM)	Weak cation
Sulphonic acid (S)	Strong cation
Diethylaminoethyl (DEAE)	Weak anion
Quarternary ammonium (Q)	Strong anion

Table 1.1 – Examples of strong and weak ion-exchange ligands commonly used in chromatographic applications

Ion-exchange systems have been proved sometimes to be unstable and can lead to different elution volumes found with small changes in ionic strength and pH. It has also been found that the pI of a protein does not necessarily dictate how it will interact with an ion-exchanger of a particular type and the resulting chromatographic behaviour can be unpredictable (Yamamoto et al, 1999).

Ion exchange is very common in industrial processes due to the low cost of the media and high binding capacities. However, low ionic strength is required in the feed (typically under 3mS) for adequate binding, which is usually far lower than the liquor from a bioreactor – as such the feed will need to diafiltered or more commonly diluted to achieve this. At larger scales this involves extremely large holding tanks and can make ion exchange processes infeasible.

1.5.6 Affinity Chromatography

Affinity chromatography has the highest specificity of all the modes of operation currently used and the ligand is highly selective. Adequate purification from a single affinity chromatography step is not unheard of, even in a primary capture capacity (Jungbauer et al, 2005).

Affinity chromatography has several forms. Immobilised metal-ion affinity chromatography (IMAC) is a very common form of affinity chromatography, where the ligand on the solid phase consists of a 2^+ ion, such as copper or nickel. Proteins to be captured are modified using recombinant techniques to include a his-tag, a loop of at least six histidine residues on the N or C terminus of the protein, which binds selectively to the 2^+ ion on the solid phase. However, this has caused some problems within the biopharmaceutical industry as the regulatory authorities are concerned about the potential immunogenicity of his-tags for human therapeutics and would like to see it removed at least by the start of human trials, adding to the cost of the process. This is however offset by the ligand itself being very cheap by chromatography standards (Hefti et al, 2003; Low et al, 2007).

Affinity ligands can also be more sophisticated, for example protein A ligands which selectively bind immunoglobulins, or even using recombinant antibodies as a ligand; this has been further evolved by the use of cheaper synthetic equivalents of such protein ligands (Roque et al, 2005). Affinity chromatography processes are designed to be highly selective, but are extremely expensive and tend to have lower capacities than other ligands. The use of a protein affinity ligand also restricts the cleaning regimen, as most proteins become unstable under the alkaline conditions a CIP process will require. The cost of purchasing enough media for a large production scale process can be extremely prohibitive but despite this affinity ligands, particularly Protein A and G, remain very popular for high cost separations involving antibodies.

1.5.7 Hydrophobic interaction chromatography

Hydrophobic interaction chromatography (HIC) works in an opposite fashion to ion-exchange chromatography, using hydrophobic ligands such as butyl or octyl (Roettger & Ladisch, 1989). Most proteins have hydrophobic patches on their surface due to long chains of hydrophobic amino acids, such as tryptophan and phenylalanine, though these are far less common than the charged groups used in ion-exchange chromatography. In this sense a protein could have a similar size and charge to another, making it difficult to separate with ion-exchange or gel filtration, but have enough hydrophobic differences to make resolution possible with HIC.

A high lyophilic salt concentration is used to get proteins to bind to the matrix and elution is performed by a reduction in the ionic strength of the mobile phase. High salt concentrations can be difficult to use in practice as many proteins will either precipitate or denature under such conditions, and binding to the ligands itself can also encourage denaturation (Ingraham et al, 1985). Logistical issues of disposal of large amounts of environmentally unfriendly salt can also be a factor, especially at larger scales.

1.5.8 Hydrophobic charge induction chromatography

Hydrophobic charge induction chromatography (HCIC) is a newer ligand system was designed to overcome some of the shortcomings of the HIC system whilst retaining selectivity based on hydrophobicity. The most common ligand for this mode of chromatography is 4-Mercapto-Ethyl-Pyridine (MEP), which is selected because the pyridine ring is hydrophobic at physiological pH, allowing hydrophobic patches to bind (Burton & Harding, 1998). Elution is performed by lowering the pH, at which point the pyridine ring becomes positive and repels the protein, leading to desorption. The advantage is hydrophobic binding is accomplished under low salt concentrations and at physiological pH, which stops protein aggregation and denaturation. This ligand is often used for IgG capture and purification, as the MEP ligand has a pseudo-affinity for this class of molecules (Schwartz et al, 2001).

As low levels of salt are required and elution is obtained by pH manipulation, HCIC chromatography is economical to scale up to process scales.

1.5.9 Reverse phase chromatography

Reverse phase chromatography (RPC) is similar to HIC in ligand selection, except organic solvents are used to elute the proteins. The use of organic solvents has limited its use for preparative processes within industry, due to contamination issues with therapeutics and safety issues with solvents that are both volatile and highly flammable (Lindsay & Barnes, 1992). However, many HPLC systems use this method and it gives excellent separation, making RPC a good choice for analytical applications.

1.5.10 Gel filtration

Gel filtration, or size exclusion chromatography (SEC), is a very commonly used mode of chromatography within industry for both smaller scale preparative and analytical applications. Unlike the other matrices found in packed bed chromatography gel filtration matrices have no ligands, as it is the structure and composition of the bead itself that produces the separation characteristics.

The solid phase is similar to the sorbent backbone used in most other packed bed chromatography materials, but the pore size has strict tolerances. Proteins larger than the pore size, taking stearic hindrances into account, will flow almost unimpeded through the column, and come out in the void volume (although in practice some stationary phase/protein interaction will have an effect on elution). Smaller proteins have restricted mobility as they penetrate the pores and therefore elute later, producing a separation based on the relative size of the protein. Conformation is however important, as globular proteins will have a different mobility to relatively unfolded proteins (Mori & Barth, 1999). The quality of the column packing is absolutely critical to size-exclusion chromatography performance, as the separation effectiveness is defined by the bead structure within the column.

This form of chromatography tends to produce excellent resolution compared to other elution chromatography methods, particularly in the case of aggregates or conformation changes where there may not be distinct charge or hydrophobicity differences. However, as elution cannot be controlled by the elution buffer concentration all protein components will elute within a single column volume, limiting resolution of highly complex mixtures and significantly restricting available capacity. This method is therefore practically restricted to small scale preparatory applications (such as final polishing steps) and, particularly when combined with HPLC, to analytical applications.

Size-exclusion chromatography also has uses for buffer exchange and desalting applications, as the separation is buffer independent (Chen et al, 1992).

1.5.11 Chromatographic Sequence Selection

A downstream process for a biological product will typically require at least two chromatography steps as part of the purification regime. The sequence of these steps, and indeed the chromatography methods that are selected, are of critical importance to the final yield, purity and ultimately the economics of the entire process.

Assuming a two step chromatography process, the steps will fall into two categories. The first step will be primary capture. In primary capture yield is the most important requirement, as a complex feed liquor will lead to a great deal of non-specific binding to a chromatography matrix and as such high purities values are unlikely. This chromatography step will also be key in increasing product concentration in the feed and reducing the liquor volume.

The second chromatography step will have purity as the most vital feature, as this "polishing" step will be critical in ascertaining final product purity.

The actual chromatographic mechanism used at each stage is highly dependent on the feed and target molecule properties. However, high capacity methods, such as ion-exchange, are typically deployed in a primary capture step as the binding capacity is so high, which reduces the quantity of matrix and scale of chromatographic capital equipment required. It is also important to take the eluent conditions into account – the eluent of an ion-exchange step that is eluted with salt will have a very high ionic strength and as such will be unsuitable for further ion exchange without diafiltration. However, this eluent would be highly suitable for HIC, as the high salt conditions promote binding to such a ligand.

1.6 Chromatography Operation and Performance

Ultimately chromatography processes are judged on yield and separation/purity efficiency for the target molecule. However, chromatography is a complex process so a number of system parameters have to be investigated to fully evaluate performance, using a variety of models and techniques. These system parameters should fully describe chromatography performance under specific sets of conditions and are important for comparison of sorbents and conditions in relation to a particular process. This is typically measured by analysis of the elution profile.

1.6.1 Theoretical plates

Theoretical plates are used to define the separation efficiency for a single species within a chromatography column. The concept of the theoretical plate was taken from distillation columns where the column was divided into a number of sections or plates. This is analogous to a packed bed (which has no physical plates, so are theoretical) where the equilibrium is the partition of the solute between the sorbent surface and mobile solvent. The more plates within the column the higher the efficiency and the greater level of separation that will be achieved. This is usually tested in practice by injecting a pulse of a non-retained tracer molecule (such as acetone or blue dextran) into the column and analysing the peak characteristics. The number of theoretical plates is defined by Equation 1.3,

$$N = 5.545\left(\frac{t_R}{W_h}\right)^2 \qquad (1.3)$$

where N is the number of plates, t_R is the peak distance from the injection point and W_h is the peak width at half peak height.

The number of theoretical plates within a column is defined by the length of the column, which leads to the less system specific height equivalent to one theoretical plate (HETP). HETP is calculated by dividing the length of the column by the number of plates and is expressed as a height of a single plate within the column, which is limited to being greater than the bead diameter. It is generally accepted that for acceptable packing and column efficiency the HETP should be less than four times the diameter of the sorbent bead.

The use of theoretical plates to predict chromatographic performance is only really useful as a reference however and the equation is empirical in nature. A number of assumptions also have to be made whilst using the HETP method to judge the standards of packing and separation capability. It is assumed that the diffusion coefficient for the marker pulse used will be identical to the components of interest during normal chromatography operation and this is, in practice, unlikely. It is also assumed that there is no time factor for approaching equilibrium with a theoretical plate, which is also unlikely for large proteins and molecules.

Unusual chromatographic conditions, such as electrically driven flows for HPLC (Grimes et al, 2002) or rectangular "columns" (Golay et al, 1981) significantly impairs the accuracy of Equation 1.3 as a method of determining plate height and as such new expressions have been derived for these cases. The equation for calculation of theoretical plates above is useful however as the values for equation 1.3 are easily acquired from a chromatogram, unlike the more accurate equations that are based on first principles and require parameters that are difficult to analyse and obtain.

1.6.2 Van Deemter Equation

Rate theory accounts for dispersion processes within a column that lead to peak broadening on the chromatogram. Rate theory in this case is taken from gas chromatography and has been applied to liquid chromatography. A number of equations have been derived for the calculation of dispersion and plate number, such as the Huber equation (Tallarek et al, 1996) and Knox equation (Berthod et al, 1989), but the 50 year old Van Deemter derived equation has been shown, in the absence of experimental artefacts, to be still the most accurate of the dispersion equations for the majority of processes (van Deemter et al, 1956; Usher et al, 2008). Three factors were identified to be critical to column efficiency; diffusion along the column, mass transfer resistance (such as gel layers on the beads and diffusion within the beads) and disturbances in the flow (wall effects and "path effects"). If these three factors are known then plate height can be calculated. The simplified Van Deemter equation is expressed in Equation 1.4,

$$H = A + \frac{B}{u} + Cu \qquad (1.4)$$

where H is plate height, A is eddy diffusion, B is longitudal diffusion, u is the linear velocity and C is the value for the mass transfer kinetics.

The Van Deemter equation is particularly useful for calculating optimal linear flow velocities and is used to this end during process development.

1.6.3 The General Rate Model

The general rate model utilises a series of partial differential mass balance equations to describe chromatography processes and is considered one of the most complex models to be used (Felinger & Guiochon, 2004). The model can be broken down into four sections; the behaviour of the analyte within the bulk mobile phase in the interstitial space, the mass transfer onto the external surface of the stationary phase, the diffusion into the pores of the bead and finally the absorption onto the surface/ligand within the bead itself.

The mass flux of the analyte within the bulk mobile phase is described in Equation 1.5,

$$\frac{\partial C}{\partial t}+u_h\frac{\partial C}{\partial z}+\frac{3}{r_p}FN_0=D_L\frac{\partial^2 C}{\partial z^2} \qquad (1.5)$$

where C is the concentration of the analyte in the mobile phase, u_h is the interstitial velocity of the mobile phase, r_p is the average radius of the stationary phase particles and D_L is the axial dispersion coefficient. F in Equation 1.5 is defined in Equation 1.6,

$$F=\frac{(1-\varepsilon)}{\varepsilon} \qquad (1.6)$$

where ε is the external porosity. N_0 in Equation 1.5 defines the mass transfer of the analyte to the surface of the sorbent bead and is defined in Equation 1.7,

$$N_0=k_f\left[C-C_p\left(r_p\right)\right]=D_p\left.\frac{\partial C_p}{\partial r}\right|_{r=r_p} \qquad (1.7)$$

where Cp is the concentration of the analyte within the pores, k_f is the external mass transfer coefficient, rp is the particle radius, r is the radial distance within the particle and Dp is the pore diffusivity coefficient. In this case the particles are assumed to be spherical and modifications must be made for other stationary phase structures, for example utilising a cylindrical geometry instead of spherical to model the pores within monolith columns.

The diffusion of the analyte within the pores of the bead itself is defined by the mass balance in Equation 1.8,

$$D_p\left(\frac{\partial^2 C_p}{\partial r^2}+\frac{2}{r}\frac{\partial C_p}{\partial r}\right)-N_p=\varepsilon_p\frac{\partial C_p}{\partial t} \qquad (1.8)$$

where ε_p is the internal porosity of the stationary phase and N_p is the mass flux of the analyte from the interior of the mesopore to the surface of the bead. N_p is further defined in Equation 1.9,

$$N_p=\left(1-\varepsilon_p\right)\frac{\partial q}{\partial t}=\left(1-\varepsilon_p\right)k_a\left(C_p-\frac{q}{K_a}\right) \qquad (1.9)$$

where q is concentration of the analyte that has been adsorbed on the stationary phase, k_a the adsorption rate constant and K_a the equilibrium constant for adsorption.

The general rate model accurately describes the mass transfer and adsorption properties which underpin the effectiveness and performance of packed bed chromatography. The disadvantages of this model lie mainly within its complexity and reliance on coefficient values that can be very difficult to obtain experimentally, especially in multi-component binding systems where the potential for multiple species to bind to the stationary phase need to be considered.

1.6.4 Yield determination

The relative protein yield of a process can be analysed in a multitude of ways, at least off-line from the chromatography equipment itself. The assays that are available are highly dependent on the material to be analysed and each assay will have a sensitivity and lower limit of quantification (LLOQ), which can be highly restrictive for small scale applications.

Typical off-line assays for determination of protein content include the Lowry, Bradford and BCA assays, but all have the limitation of requiring a standard curve derived from the assay of a highly purified protein, in contrast to the potentially complex mixture that is a chromatography eluent. UV spectrometry at 280nm can also provide an approximate gauge of how much protein is within a fraction, but without detailed knowledge of the extinction coefficients of the polypeptides within the sample this method will be inexact.

Specific protein concentrations can also be obtained by a great many methods, but the methods employed are highly dependent on the target protein of interest. Enzyme-linked immunosorbent assays (ELISA) can give rapid and extremely sensitive results but require an antibody that binds to the protein of interest and a suitable antibody-enzyme conjugate as a detection system; this also applies to Western blots, where proteins separated by electrophoresis are transferred to a membrane and treated with antibodies to identify bands (Baines et al, 2001). If the target protein has a biological activity then this can also be utilised in an assay and is particularly common for the analysis of enzymes; this is also true if the target protein has an associated metal ion, in which case spectroscopic techniques could be appropriate. Affinity HPLC has also proven to be a very effective and reproducible method for protein quantification, particularly for immunoglobulins where commercial affinity ligands in small scale presentations (Protein A, G and L) are readily available.

Although time-consuming and with necessary capital outlays, mass spectrometry methods are also highly effective at determining protein concentrations. This is particularly true for surface-enhanced laser desorption/ionisation (SELDI) methods which were designed for analysis of protein mixtures (Issaq et al, 2002).

On-line analytical methods are required to monitor the chromatographic process and generally consist of a UV spectrometry flow cell located at the terminus of the column. Typically these UV cells operate at a wavelength of 280nm, since aromatic amino acids absorb at this wavelength, although the 200nm wavelength will also detect the peptide backbone of the protein. In the case of primary capture and for complex mixtures 260nm is also analysed, as this will detect nucleic acid contamination.

1.6.5 Purity determination

The ease of determining purity is again highly dependent on the protein of interest and its intrinsic attributes. There are however several conventional methods that are applicable to all classes of proteins.

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is an very popular technique for analysis of complex protein mixtures, separating the proteins with respect to molecular weight (under reducing conditions) or with respect to molecular weight and conformation (under non-reducing conditions). Visualisation of the proteins is usually accomplished by staining, either with Coomassie blue or by silver staining. Closely related species in molecular weight can be difficult to visualise but with the right densitometry equipment it is possible to acquire a semi-quantitative concentration and purity for a selected protein (Goetz et al, 2004).

A range of other analytical techniques are also commonly employed to ascertain purity, such as SEC HPLC, affinity HPLC, isoelectric focussing and mass spectrometry.

1.7 Scale up and Down of Chromatography Processes

1.7.1 Scaling up from laboratory to manufacturing scale

There are a number of occasions during manufacturing development where a chromatography process will need to be scaled up. This is generally due to transfer from the research and development processes to a pilot scale equivalent or to a new facility, or in cases where the entire production process is scaled up to provide more material for later phases of clinical trials or for additional commercial supply.

However, scale up of chromatography unit operations present significant challenges. Much of chromatographic performance is dictated by particular hydrodynamic phenomena and the influence of each phenomena can change, sometimes quite drastically, at different scales. Unfortunately current understanding of these phenomena is still partial and the complexity and large number of variables makes predicting chromatographic behaviour at different scales difficult to model by conventional methods. This provided the motivation for this EngD study.

Conventionally scaling is done empirically and rarely in a single step, with ten-fold increases in interior column volume at each stage being typical from laboratory to manufacturing facility. This is expensive and requires large quantities of feedstock, which delays downstream process development until much later in the product development cycle.

1.7.2 Non-chromatographic scale up factors

In considering scale-up it is not just the column and quantity of matrix that needs to be scaled up but also the associated pipe work, pumps and skids. This can create a very large discrepancy in the extra-column volume of pipe work and internal instrumentation that the process liquor will have to travel in order to reach and leave the column at different scales. Moreover, the proportion of internal column volume compared to the extra-column volume in a small scale chromatography rig can be orders of magnitude different to those seen at manufacturing scales; the extra-column volumes can even be equal or larger than the internal column volume due to engineering limitations with the chromatography apparatus. As diffusion and dispersion effects are observed within this extra-column volume this can lead to non-comparable chromatographic behaviour.

The materials of construction are also likely to be dissimilar between laboratory and manufacturing scale chromatography apparatus. Laboratory scale chromatography columns and associated equipment are conventionally made of glass and various plastics. Manufacturing scale equipment beyond a certain scale is produced in electro-polished stainless steel, including the pipe work. Although these materials are selected for being inert, any interaction with solvents, feed liquor or the stationary phase will potentially have an effect on chromatographic effectiveness.

1.7.2.1 Extra-column band broadening

In large columns extra-column peak broadening (or band broadening) is negligible, as the internal column volume far exceeds the extra-column volumes. This is not the case for small columns however (Kaltenbrunner et al, 1997), as the small volumes are liable to contain dead spaces and require proportionally larger instrumentation volumes to be effective, particularly with micro-chromatographic systems (Beisler et al, 2003).

The total band broadening is a combination of dispersion in the pipe work, dead volumes, instrumentation volumes and an electronic response (for detectors), as shown in Equation 1.10,

$$\sigma_{ex}^2 = \sigma_{pw}^2 + \sigma_{inst}^2 + \tau_{dead}^2 + \tau_{er}^2 \qquad (1.10)$$

where pw is dispersion from the pipe work, inst is the dispersion from the instrumentation, dead is the dispersion from the dead volume and er is the dispersion from the electric response

Although this equation is used to determine extra-column derived dispersion, dead volumes also need to be taken into account within the column itself, in the event of poor packing or poor column/frit design. In practice it is found that modern apparatus only has a very small instrumentation dead volume, electric response rates are too fast to be an issue and properly designed equipment for fluid flow will have minimal dead space, therefore only leaving pipe work and hold up volume as a significant contributor to extra column dispersion.

1.7.3 Chromatographic factors

Scale up of the chromatographic system itself has a number of variables that need to be considered. As with all scale up procedures, or indeed any experimental process, as many of the potential system parameters as possible need to be maintained as constant but it is recognised that some of the parameters will need to be modified to suit the larger scale design requirements.

Parameters such as buffers and the chemical conditions of the separation and feed, together with the quantity of sample loaded per mass of matrix, should be maintained and there should be few reasons why these should change at larger scales. This also applies to such parameters as relative gradient volumes and linear velocity/residence time, as these are critical for chromatographic performance and will make comparability of the two scales very challenging if not maintained. The volumetric flow rate through the column will obviously increase substantially as the process is scaled up, but the residence time within the column should remain equivalent.

The main parameter that will need to be changed to scale up a chromatographic process is the size of the column itself. It has long been recognised that scaling up of chromatographic processes is facilitated if the bed height is kept constant and extra intra-column volume is generated by an increase in column diameter (Sofer and Hagel, 1997). These changes also have important economic considerations which are only applicable for commercial manufacture (Joseph et al, 2006).

There are a number of major complicating factors associated with increasing the bed diameter. In a conventional laboratory scale column the column walls provide structural support to the matrix inside, allowing higher linear velocities and the use of potentially more viscous feeds before critical pressure drop is observed, so a high linear velocity parameter may not be able to be preserved during scale up. It also has been shown that the packing structure of the matrix is different within a certain range of the wall, the so-called wall effect (Shalliker et al, 2000) and this can influence dispersion within the column.

The distributor design is also far more complex in a large scale column with a large diameter. In laboratory scale columns the frit and distributor are not particularly complex, as its relatively easy to ensure an even flow across the whole axial cross-section of the column. However, in large columns the fluid flow has to be distributed across a much large surface area and at high flow rates this can cause dead volumes and jetting effects. The different designs for large scale distributors can potentially interact with process materials and affect sample injection (Broyles et al, 1999) so would therefore need to be monitored during scale up procedures.

1.8 Thesis Aims

The aim of the work contained within this thesis was to produce an ultra scale-down methodology that allows accurate prediction of large scale chromatographic behaviour using data derived from ultra scale-down devices. There were several factors and constraints that directed this work:

- All feedstocks utilised had to be industrially relevant and realistic, rather than the mixture of various purified proteins typically used in such studies.
- As a number of highly effective high-throughput chromatography systems already exist, any methodology produced should favour the accuracy of the large scale chromatogram prediction over the speed of obtaining such data.
- Any model separations used to develop the ultra-scale down methodology should industrially applicable and should be directed to a valuable and realistic target molecule.
- The methodology to be developed should have an experimental premise and be based on apparent data, with minimal reliance on existing complex theoretical models. In this way the methodology should "bridge" between purely experimental and experience derived process development and the mostly theoretical approaches.

- The methodology and model to be developed needs to be validated and tested for robustness by exposure to different separations utilising different ligands and sorbent backbones.

1.9 Thesis Synopsis

This thesis details the investigations that took place between October 2005 to September 2009 within the Biochemical Engineering Department of UCL and in collaboration with Pall Life Sciences.

The results of this investigation are contained within 5 chapters as summarised in Figure1.2.

1.9.1 Chapter 2 – Feedstock Selection

This chapter details the investigation into various biological feedstocks considered as suitable candidates for use with a model chromatographic system

1.9.2 Chapter 3 – Ultra Scale-Down Device Selection

This chapter is concerned with the investigation into three different small scale (1mL internal volume) pre-packed chromatographic column designs, to identify the best candidate for use in the formulation of a ultra scale-down methodology.

1.9.3 Chapter 4 – Separation of FAb fragments by Mixed Mode Chromatography

The feedstock (FAb fragment containing periplasmic lysate from *E. coli*) and the ultra scale-down device (Pall Pre-Packed PRC columns) were selected in Chapter 2 and 3. This chapter investigated the separation of a FAb fragments from the periplasmic lysate and the most suitable chromatographic conditions to take forward.

1.9.4 Chapter 5 – Formulation of an Ultra Scale-Down Methodology for Chromatography Scale Up

This is the critical keystone chapter of the thesis. The data obtained from the studies in Chapters 2-4 were incorporated into an experimental system that was utilised to develop a methodology and mathematical model, to allow the accurate prediction of large scale chromatographic behaviour using small scale data.

1.9.5 Chapter 6 - Validation of Methodology and Model for the Ultra Scale-Down of Elution Chromatography

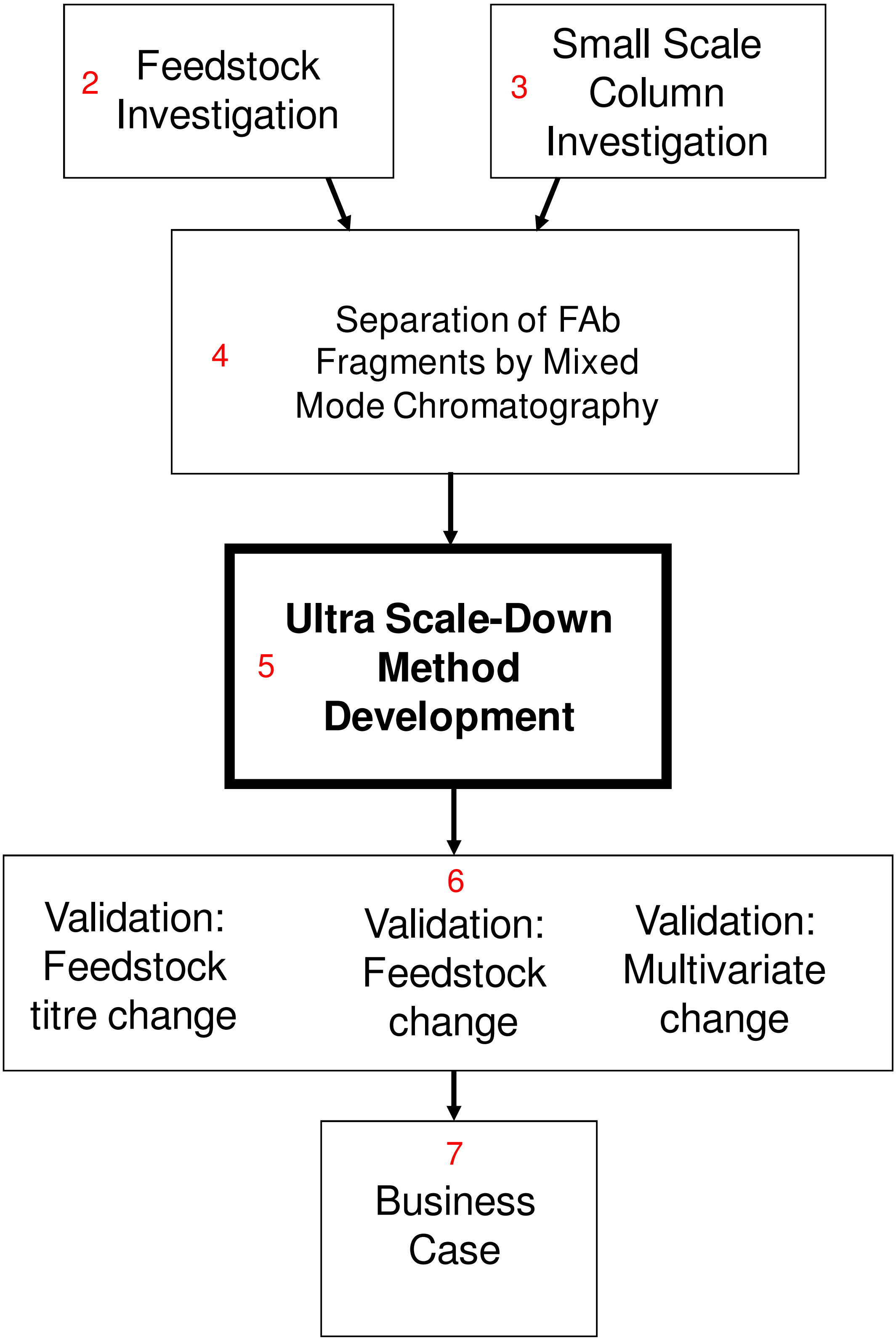
This chapter consists of three separate validation studies. These studies were applied the methodology and model developed in Chapter 5 to other experimental chromatographic systems.

1.9.6 Business Case

This chapter addresses commercialisation potential and proposes routes for bringing the concepts investigated in this EngD to market.

Figure 1.2

Ultra Scale-Down of Elution Chromatography



2 Feedstock Selection

2.1 Abstract

A suitable feedstock was required as a feed for packed bed elution chromatography. This feed had to be realistic, industrially relevant, inexpensive and contain a relevant target molecule to give direction to the chromatographic separation. Four forms of bovine whey (differing in how the whey was extracted from the milk) and a recombinant FAb fragment containing *E. coli* periplasmic lysate were investigated as to their suitability for use as an industrially representative model feed stream. The relevant target molecule was bovine IgG for whey and FAb fragments for the periplasmic lysate.

Whey produced by acidification, treatment with rennet, spray-drying and tangential microfiltration ("ideal whey"), together with the periplasmic lysate feed, were separated by reducing SDS-PAGE. Target molecule concentration was analysed by ELISA (for IgG) and by Protein G HPLC (for FAb fragments).

It was found that the spray drying process caused extensive denaturation of the protein components of the whey and IgG concentrations were 60% less than the 0.5mg/mL IgG concentration found in whey obtained by the use of rennet and acidification. A concentration of 0.55mg/mL FAb fragments was found within the periplasmic lysate feed stream.

Ease of processing the potential feed streams to produce a liquor suitable for packed bed chromatography was also investigated. Ideal whey proved to be very difficult to process in significant quantities dues to microfiltration membrane fouling, while whey produced by acidification and rennet required a time-consuming and variable manual skimming process. The periplasmic lysate feed however proved to be straight-forward and quick to process via centrifugation and 0.22µm depth microfiltration. It was concluded that FAb fragment containing periplasmic *E. coli* lysate, despite its cost, was the preferred feedstock for use in further experiments.

2.2 Introduction

2.2.1 Feedstock Selection

A feedstock suitable for scale-down experiments was required. This feedstock required a number of attributes:

- Realistic – it is common to use largely synthetic or highly processed feeds for chromatography modelling work (Pedersen et al, 2003). This can lead to unforeseen issues in application of the research to complex industrially relevant feeds with a diverse range of contaminants. A feedstock that derives from a potential industrial process was therefore sought containing a wide range of protein contaminants; however, the components of this feedstock will have to be easily quantifiable to facilitate analysis, so feeds highly contaminated with host cell proteins and nucleic material, such as bacterial homogenates, would not be suitable.

- Inexpensive and available – even experiments at the sub 10mL scale can require substantial quantities of feedstock and expensive feedstocks will limit experimental scope. An experimental feed derived from food sources, such as cow's milk or hen's eggs, would be suitable on grounds of cost, as would an already developed fermentation process that has been optimised for high titres of expressed protein.
- Contain realistic target products – the feedstock must contain a high value and specific product to give a sense of purpose to the separation. If any further investigation is to involve the highly IgG selective MEP HyperCel then an IgG or IgG fragment would be an excellent and relevant target protein.

Two feedstocks were examined for suitability, namely whey in various forms and a periplasmic lysate from *E. coli* containing recombinant FAb fragments, produced in-house using a developed and optimised method.

2.2.2 Whey as a Model Protein System

Dairy whey is of use within industry as it contains several proteins of note, the major ones being defined in Table 2.1.

There are two kinds of whey commonly produced within the dairy industry; sweet whey and acid whey. These are defined by the extraction process used to separate the whey from the casein proteins, commonly identified as curds.

Protein	**Concentration (g/L)**
β-lactoglobulins	3-4
α-lactalbumin	1.2-1.5
Serum albumin	0.3-0.6
Immunoglobulins (including IgG)	0.6-0.9

Table 2.1: The protein composition of bovine whey.

The majority of the whey in use in the dairy industry is sweet whey which is produced by the use of rennet, an enzyme located in the fourth bovine stomach. Acid extracted whey is produced by precipitating the curds at pH 4.6 and due to the selectivity of the precipitation reaction it has a different protein composition to sweet whey. Whey also contains a great deal of lactose, some minerals, vitamins and lipids. Whey is treated as a waste product in the cheese making industry, therefore substantial quantities, in spray dried form, can be obtained for under £50 per kilogram.

Although whey is usually a waste product of the more profitable cheese manufacturing it can still be used for animal feedstocks or for the production of lactose (Pedersen et al, 2003). The proteins contained in whey also have a number of commercial uses and have typically been recovered by precipitation. The research that has gone into using transgenic milk to produce therapeutic proteins has also increased interest in the bio-processing of whey (Pollock et al, 1999) from bovine, caprine or ovine sources. More recently a combination of membrane filtration and ion exchange chromatography is used as a purification route to the same end in an effort to increase efficiency. Yields have been shown to be poor using conventional ion-exchange chromatography but the membrane ion-exchange chromatography systems have been shown to improve this (Bhattacharjee et al, 2006) and in particular cationic ion-exchange resins have proved useful in separating out the valuable immunoglobulin G (Gerberding et al, 1998).

Even with the difficulties with processing whey on an industrial scale it has merit for use as a model system (Pederson et al, 2003), particularly for newer technologies such as simulated moving beds (Andersson et al, 2006). It is a complex mixture made up of different proteins, lipids and polysaccharides and therefore provides a more realistic feedstock than artificially produced protein mixtures. The proteins include very valuable commercial proteins such as IgG which make it relevant to potential industrial processes. Whey is also very economical to use as its a waste product; even whey produced in-house comes from inexpensive bovine milk.

2.2.3 FAb Fragment Containing *E. coli* Periplasmic Lysate

E. coli based expression systems have traditionally been favoured for its extensively characterised genetic make-up and its ability to grow extremely rapidly with minimal specialist fermentation nutrient needs (Makrides et al 1996; Baneyx et al 1999). Small scale expression of plasmid encoded recombinant proteins within *E. coli* are typically regulated by use of *lac* promoters. Significant limitations have been shown with this promoter at larger scales however, including the relatively high cost of non-hydrolysable induction agents such as IPTG. The T7 bacteriophage promoter system (via pET vectors) addresses this problem by being a strong promoter hence massively over-expressing the target protein within the cell. However, strong promoter systems can produce a high relative proportion of inclusion bodies and have ramifications for correct target protein folding (Schein et al, 1989; Hannig et al, 1998). A reduction in fermentation temperature has been utilised to counteract inclusion body formation (Sorensen et al, 2005) but this has other metabolic impacts upon the cell and the fermentation progression.

Conventional recombinant protein expression in *E. coli*, as well as Gram-negative prokaryotes generally, produces an intracellular product that is then liberated via a homogenisation or lysis step. This produces a complex, viscous liquor with a multitude of host cell protein contaminants, nuclear material and other cell debris that can inhibit the effectiveness of proceeding downstream purification (Middelberg et al, 1991; Geciova et al, 2002). This has been addressed by the use of periplasmic targeting and lysis systems, the concept being that the product is transported to the periplasm and only the periplasm is lysed to liberate the product into the extra-cellular space. This leaves most of the host cell protein and nuclear contaminants still within the cell which are removed effectively with the primary capture clarification step.

The engineering of the molecular biology so as to ensure that a functional and correctly folded protein is produced via the periplasm is not a trivial task. An 18-30 amino acid leader sequence directs proteins to the periplasm and to fold correctly they need to arrive in a loosely folded form, as assisted by chaperone proteins (Langer et al, 1992). Although the periplasm is an excellent environment for the finalising of the folding of proteins of eukaryotic origin (Missiakas et al, 1997), this can lead to high rates of mis-folding and inclusion body formation in non-native protein products. Localised proteases are also in abundance in the periplasm leading to protein degradation (Gottesman et al, 1996) and the process of transporting an over-expressed protein in the periplasm can lead to slower growth rates and cell damage if the cells walls are not reinforced with additional magnesium and phosphate during fermentation. Research into increasing the permeability of the outer membrane so that the product can be transported into the bulk liquor, much like mammalian and eukaryotic yeast systems, has had some success using a variety of approaches (van der Wal et al, 1998; Wan et al, 1998).

2.2.4 Aims of This Chapter

This chapter investigates whey and a periplasmic *E. coli* lysate for suitability as a economical, realistic and reproducible feedstock for packed bed chromatography experiments at small scale.

2.3 Materials and Methods

Unless otherwise stated all chemicals were sourced from Sigma-Aldrich (Dorset, UK).

2.3.1 Whey

Sweet whey was obtained from Colston Bassett Dairy (Colston Bassett, Nottinghamshire, UK) and was stored at -20°C. Spray-dried whey solution was produced using spray-dried whey (Sigma) at 10% W/W in distilled water. This solution/suspension was centrifuged at 3000 RPM within a Beckman JE-26 centrifuge to remove solid particulates and stored at 4°C. Acid-extracted whey was produced using standard supermarket skimmed milk acidified using 7M sulphuric acid to pH 4.6. This was then heated in a 65°C water bath for 5 minutes to accelerate precipitation and then centrifuged for 5 minutes at 3000 RPM in a Beckman GS 6R centrifuge. The liquid fraction was then removed using a Pasteur pipette into a clean Falcon tube and stored at 4°C.

2.3.2 FAb lysate feedstock production - Fermentation

The host cell was *E. coli* W3110 containing the A33 FAb fragment plasmid (UCB Celltech, Slough, UK). Primary seed cultures stored at -80°C were thawed and grown up in 2L shake flasks containing LB broth at 37°C until OD_{600nm} >1. Flask contents are aseptically transferred to secondary culture flasks containing defined media utilising glycerol as the carbon source. The cultures were then incubated at 30°C until the OD_{600nm} was between 2-4.

The contents of the secondary culture flasks were then transferred aseptically into a 20L LH fermenter containing 12L sterile $SM6G_c$ defined media with 112g glycerol per litre. Initial fermenter conditions are pH 6.95 with 400RPM agitation linked to a DOT value of 30%. Air was sparged until 26 hours from inoculation at which point the air is replaced with 40% oxygen (BOC, UK).

When the OD_{600nm} reached 40 the temperature was reduced to 25°C and 160mL of 1M $MgSO_4$ was added via an injection port. At 30 hours from inoculation and again at 34 hours from inoculation 120mL of 1M sodium phosphate was added via an injection port. The 80% glycerol feed was switched on at 24mL/h during the DOT spike and the recombinant protein production was induced by the addition of 30mL 15g/L IPTG. Harvesting took place at 80 hours after inoculation.

Primary clarification of the liquor was performed in a Carr P6 Powerfuge (Carr Centritech Franklin, MA, USA) at 5000g. The cell paste was stored at -20°C.

Final concentration of FAb fragments within the lysate was 0.55mg/mL.

2.3.3 FAb lysate feedstock production – Periplasmic lysis

Liberation of the FAb fragments by periplasmic lysis of the centrifuged biomass was performed by stirred incubation for 16 hours at 60°C in 1L of lysis buffer (50mM Tris-HCl, 20mM EDTA, pH 7.4) per Kg of biomass. The resulting lysate was centrifuged in a Beckman JE-26 fixed rotor centrifuge at 10,000 RPM for one hour to remove cells/cell debris and the supernatant was stored in a conventional freezer at -20°C. Before application to packed bed chromatography the lysate was filtered using a Nalgene 0.22µm depth filter with the filtrate being retained. Without pH adjustment the feed is pH7.4; this was adjusted to pH 6 for CM Ceramic HyperD experiments using concentrated HCl and to pH 8.8 with 4M NaOH for PPA and MEP HyperCel.

The single fermentation batch from which the feedstock for the study in this chapter was derived will be referred to as "Batch 2".

2.3.4 ELISA Assay for Detection of Bovine IgG

The ELISA assay was performed in NUNC 96 well microtitre plates by coating with anti-bovine IgG antibody (B9884, Sigma), blocking with bovine antibody free BSA and then binding the samples and controls for 1 hour. A secondary anti-bovine IgG antibody conjugated with peroxidase (A5295, Sigma), which reacts with the OPD and produces a quantifiable colour change, was incubated for 1 hour. Washes were performed with PBS and 0.2% PBS-Tween. The detection reagent was OPD and was read at 490nm in a Tecan Safire II plate reader.

2.3.5 Protein G HPLC

Protein G HPLC for analysis of FAb fragment concentration was performed on an Agilent 1200 HPLC using Protein G packed Sepharose Fast-Flow Hi-Trap columns (GE Healthcare) at a flow rate of 2mL/min. Elution was via pH 7.4 to 2.2 gradient over 10 minutes. The controls consisted of pure FAb fragment purified using mixed-mode chromatography.

2.3.6 Tangential Flow Filtration

Tangential flow filtration was performed on skimmed milk to produce ideal whey. A Millipore Labscale system was utilised according to manufacturer's instructions, with a 0.1µm Pelicon microfilter run at a constant TMP of 1.2 bar. The permeate was collected and stored.

2.3.7 General Analysis

Absorbance at 280nm was measured in a conventional laboratory spectrophotometer in UV compatible cuvettes. SDS-PAGE was performed using 4-20% tris-glycine gels (Invitrogen EC6025BOX) under reducing conditions at 140 volts for 1 hour then stained with Coomassie blue utilising Invitrogen's SeeBlue system.

2.4 Results and Discussion

Different forms of whey and a FAb fragment containing E. coli periplasmic lysate were analysed for their protein composition and to the level of pre-processing required to produce a feed liquor suitable for packed bed chromatography.

2.4.1 Protein Composition – Whey

The proteins in processed whey sourced by acidification, spray dried and rennet (sweet) methods were separated using reducing SDS-PAGE. This is shown in Figures 2.1 and 2.2.

This gel was prepared and run under reducing conditions so the IgG appears as two bands; the light chains at 25KDa and the heavy chains at 50KDa. The lanes with the spray-dried whey samples show poor resolution and smearing indicating extensive denaturation and breakdown of the proteins found within the sample. The spray-drying process involves the vaporisation of the slurry under a stream of hot air and this would involve a great deal of shear at the air/liquid interfaces and high temperatures. Under these conditions protein denaturation and breakdown would be expected.

A comparison of acid-extracted whey and sweet whey is shown in Figure 2.2.

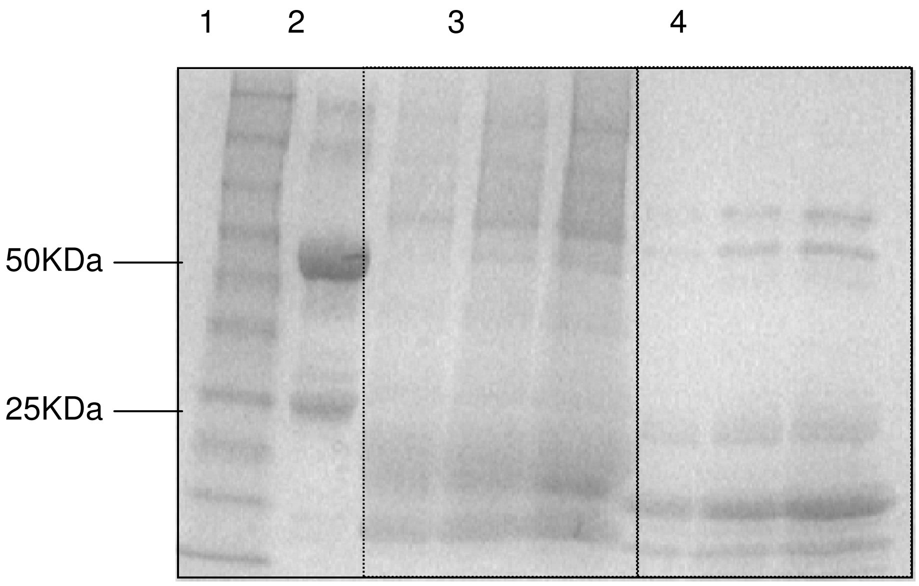


Figure 2.1: A reducing SDS-PAGE gel showing samples from both acid-extracted whey and spray-dried whey feedstocks. The marker is pre-stained Benchmark (Invitrogen) and the general staining is with Coomassie blue. Key: 1: Marker 2: Bovine IgG control 3: Spray-dried whey 4: Acid extracted whey.

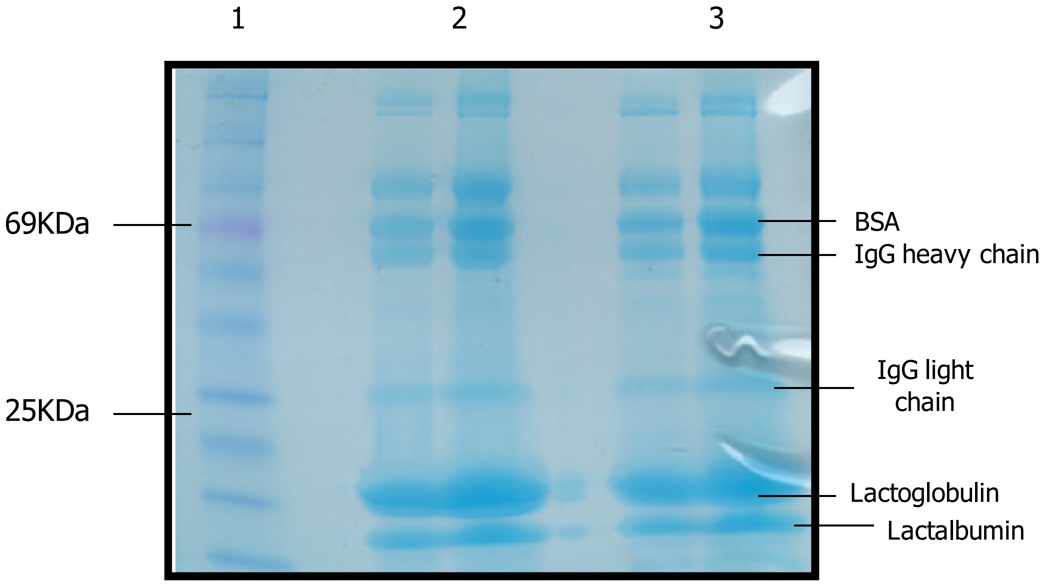


Figure 2.2: A reducing SDS-PAGE gel showing samples from both acid-extracted whey and sweet whey feedstocks. The marker is pre-stained Benchmark (Invitrogen) and the general staining is with Coomassie blue. Key: 1: Marker 2: Acid-extracted whey 3: Sweet whey.

Figure 2.2 shows clearly there is almost no qualitative difference between the protein composition of sweet whey and acid-extracted whey; protein bands are clear and degradation seems limited for both samples. Although extra proteins are theoretically added to sweet whey due to the addition of rennet they cannot be visualised at such a low concentration and should have minimal impact on any chromatographic separation. Acid-extracted whey and sweet whey both show a good number of discrete protein contaminants and good concentrations of the target molecule and therefore, on this basis, would make a good model system for chromatographic separation.

Table 2.2 shows the bovine immunoglobulin concentration as determined by ELISA. Literature values are approximately 0.6mg/mL IgG in conventional bovine whey (Pedersen et al, 2003) and the acid-extracted and sweet whey values broadly agree with this. Spray-dried whey shows only 30% of the literature value, presumably due to high levels of protein denaturation created by the spray drying process.

Whey	**IgG Concentration (g/L)**
Spray-dried	0.2
Acid-extracted	0.5
Sweet	0.5

Table 2.2: IgG concentration in the whey samples tested, as determined by ELISA.

2.4.2 Protein Composition - FAb Fragment Containing Lysate

The protein composition of the periplasmic lysate is shown in Figure 2.3.

The actual identity of the majority of the proteins within the lysate is difficult to elucidate without extensive analysis but the major contaminant at approximately 50kDa is believed to be periplasmic alkaline phosphatase

In this batch of periplasmic lysate the FAb concentration was resolved to be 0.55mg/mL by Protein G HPLC (section 2.3.7). This titre can vary to a limited degree between batches but this is controllable by manipulation of the volumes of buffer used in the periplasmic lysis (section 2.3.3).

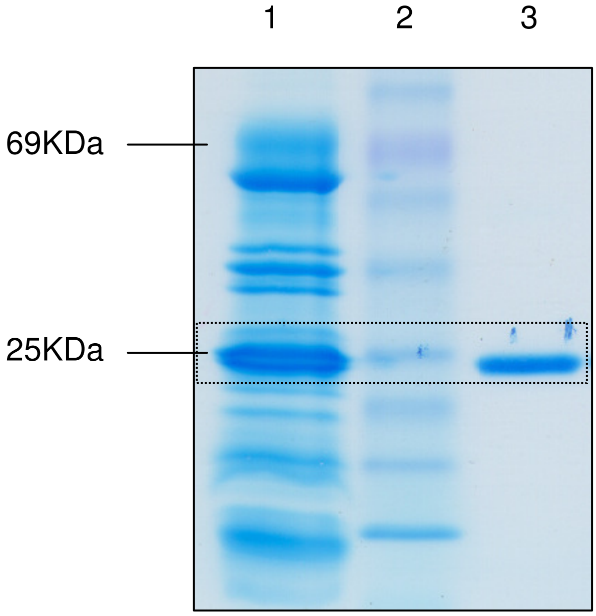


Figure 2.3: A reducing SDS-PAGE gel showing samples from an E. coli *periplasmic lysate containing recombinant FAb fragments. The marker is pre-stained Benchmark (Invitrogen) and the general staining is with Coomassie blue. Key: 1: Periplasmic lysate 2: Marker 3: FAb control. FAb band is boxed.*

2.4.3 Pre-Processing of Feedstocks

There were considerable pre-processing issues with both the acid-extracted and sweet whey. Spray dried whey was quick and easy to clarify, but the high lipid content of the full fat sweet whey created considerable difficulties for the removal of supernatant from centrifuged samples, as the fats coagulate and float to the surface. The same problem was encountered by the curdling acidification process with skimmed milk leading to a necessary and none too effective skimming process to remove. This skimming process was manual, time-consuming and, because of the manual intervention required, creates unacceptable levels of batch to batch variability.

FAb containing periplasmic lysate on the other hand did not require pre-processing once the fermentation and upstream clarification was complete.

2.4.4 Ideal Whey

"Ideal whey" is essentially demineralised whey produced by tangential flow micro-filtration. In this work a 50cm^2 Pelicon TFF cassette (Millipore, Watford, UK) was used to treat skimmed milk in order to limit lipid content. Observed fluxes were extremely low, in the region of a 0.8mL/min/m^2 at the start of operation. This was further reduced as the membrane fouled, rendering the filtration process unacceptably slow. As a result this option was not pursued any further.

2.5 Conclusions

Four potential feedstocks were investigated for use as a feed for the model chromatographic system; spray-dried whey, acid-extracted whey, sweet whey and a *E. coli* derived periplasmic lysate containing FAb fragments.

All three whey products were relatively inexpensive to obtain. However, the time and effort required to process the whey feeds into something suitable for packed bed chromatography was substantial. In the case of ideal whey it proved impossible to process substantial quantities of feed at all, let alone enough for laboratory scale chromatographic separations, so this feed would not be suitable for any thorough study that would require large quantities of feedstock.

Periplasmic FAb containing lysate was much more expensive to produce, as it required a 96 hour fermentation process, however the extent of pre-processing needed to produce a suitable feed for packed bed chromatography, after the pilot scale centrifugation, was very short.

All the feedstocks investigated could be considered realistic. Although research into transgenic milk continues the biopharmaceutical industry is far more focussed on recombinant feeds from prokaryotic or eukaryotic hosts, so the periplasmic lysate would have to be considered the more relevant of the feeds. All the feedstocks contained a realistic target molecule, either a FAb fragment or a complete immunoglobulin molecule.

The additional cost of producing the periplasmic lysate is more than compensated by the ease of processing the feed ready for packed bed chromatography. To this end the FAb fragment periplasmic lysate was adopted as the basis for future investigations involving the selection of an ultra scale-down experimental device and the formulation of an ultra scale-down methodology.

3 Small Scale Chromatography Device Selection

3.1 Abstract

It is not feasible during early process development to run investigative chromatographic separations at large scale. A large number of individual experiments may be required to fully analyse and optimise all the operational chromatographic parameters; even at laboratory scales this would require a great deal of highly expensive feedstock that may not be available in sufficient quantities. A number of different small-scale chromatographic techniques and platforms have been developed in order to study chromatographic behaviour very rapidly and only requiring minimal quantities of feedstock, known collectively as high-throughput screening methods (HTS). These include microwell based systems, microfluidic systems and very small columns amongst others. The very small volumetric needs of these systems can however introduce significant engineering constraints and produce unpredictable chromatographic behaviour that does not correlate with the expected chromatographic outputs at laboratory or process scale.

Three separate very small column designs containing pre-packed ion-exchange and HCIC media, namely Pall Life Sciences PRC, a Pall Inc. AcroSep prototype and GE Healthcare HiTrap, were investigated in this chapter with

regard to a range of chromatographic parameters. The PRC format showed superior separation characteristics and higher total plate numbers than the HiTrap or AcroSep designs. It was concluded that the PRC small column format would be the most appropriate for further investigations.

3.2 Introduction

3.2.1 Small Scale Chromatography

Scale in chromatography, and bio-process unit operations generally, is subjective. Although the terms laboratory, pilot and process scales are freely used they can apply to a range of unit operation sizes and depend more on application than scale; typically however column volumes of less than 250mL are considered laboratory scale operations, with systems larger than this being referred to as pilot or process scales depending on whether that sized unit operation will be used for production or as a stepping stone for scale up procedures.

For industrial purposes only large laboratory scale to preparatory scale is used for reasons of economics. In chromatography the volumes given above do not just apply to the volume of the chromatography device or the volume of the feed liquor, but is a mean value of volumes required for the effective elution of the product stream and the completion of a full chromatographic cycle, including loading, washing, elution and regeneration. For example, although a 1mL column would only potentially require 5mL of feed the process would require 10-15 CV for washing and perhaps 5 CV for elution, requiring the handling of much larger volumes of liquid than would be initially apparent.

It is however neither feasible nor economic to run every chromatography investigation or process at large scale. In early phase process development the process feed is usually produced in shake flasks or in very small bioreactors and often using a non-optimised process, so the feed will only be available in low volumes and/or titres. There are also a range of chromatographic process options available for the purification of any target molecule and a thorough investigation requires that many individual experiments that would take considerable time and resources at laboratory scale of operation. A key function of scale-down devices is to produce high-throughput screening (HTS) methods that allows for the very rapid or simultaneous analysis of dozens of chromatography sorbents and operating conditions; a degree of automation could be envisaged for these screening methods that would add significantly to the rate of investigation (Rege et al, 2006).

There have been a number of approaches to performing chromatography at small scales, including microfluidic approaches, microwells, batch methods, specialist pipette tips and very small columns (Chhatre & Titchener-Hooker, 2009).

3.2.2 Microfluidic Devices

Microfluidic devices are extremely small, typically working with volumes of under 500µL and as low as a few µL. Microfluidic devices were inspired by microchips and microdot chips systems that have been used for affinity ligand scouting (Regnier et al, 1999; M. Schena et al, 1998) and for analysis of DNA hybridisation libraries (Heller et al, 1997) in the form of microarrays.

This led to the micro-machining of channels on quartz or plastic chips which could act as a capillary for electrophoresis and liquid chromatography. Using current engineering methods it is possible to pack micro-capillary channels with under 1µL matrix volume for HPLC processes, either driven by conventional pump/vacuum systems (Vanhoutte et al, 1997) or by electro-osmotic flow (Bartle & Myers 2001). Most of the issues with microfluidic liquid chromatography systems are not conceptual but in the engineering of such small devices; they are difficult to pack in a reproducible fashion and have a tendency to leak at operating pressures and during decoupling from the pumping system. Due to the huge differences in the hydrodynamics of such systems compared to process scales it is currently very difficult to scale up process chromatographic parameters from microfluidic devices alone. However, significant data such as dynamic and maximum protein binding capacities have been obtained with 1.5µL microfluidic systems which compared well with conventional sized chromatography columns (Shapiro et al, 2009).

3.2.3 Non-Column Chromatography

Non-column chromatographic methods have no flow parameters but instead rely on agitation to mix the solid and solvent phases. There are very few industrial process uses for non-column chromatography as it offers few advantages over a typical packed bed system. Elution volumes are typically very large and restricted to isocratic procedures. Engineering constraints exist (such as shear damage to the matrix) due to the need to agitate large numbers of beads. Historically non-column incubation methods were used extensively for screening chromatographic conditions (Scopes et al, 1987) and even now non-column methods are still useful for small scale experimental investigations. Contaminant removal chromatography is viable as it requires no elution of valuable product and non-column methods have been used for feedstock pre-processing before packed bed chromatography (Levison et al, 1989) due to it being inexpensive and fast. Although batch methods can give rudimentary binding data, actual process data is difficult to obtain using these methods.

3.2.4 Microwell Chromatography

The typical 200µL volume 96-well plate format is highly standardised among laboratory equipment/analytical systems as well as those responsible for robotic automation so is an attractive base for high-throughput chromatographic systems. These systems can be used in two ways: as a batch/non-column system (albeit with a much smaller volume than conventional non-column chromatographic systems) where the wells are self-contained, or as a rudimentary flow-through system utilising filter plates.

The use of 96 well plates for solid phase applications have been in use for over 15 years (Venn et al, 2005). Typically these systems are used for the analysis of multiple products on a single purification or affinity technique, for example the screening of protein and DNA libraries as well as affinity libraries (Lesley et al, 2001; Lowe et al, 2001). A significant limitation of the 96 well plate format is that although many experiments can be run simultaneously the actual mechanism is still non-column chromatography and as such it is more difficult to obtain flow related process data, although this method has still been used to great effect with affinity and ion-exchanger resins to determine binding capacities (Bergander et al, 2008). The lack of a flow characteristic can be addressed by the use of filter plates, 96 well plates that contain a filter instead of a base which acts as a frit. The liquid phase is drawn through the base of the plate either by vacuum or by centrifugation.

This has been adapted to a number of purification applications such as ion-exchange and HIC (Shih et al, 2002; Rege et al, 2006) and has been linked to SELDI for rapid analysis of multiple samples (Wierling et al, 2007). The 96 well plates have been stacked to give acceptable mimics of chromatographic behaviour (Coffman et al, 2008) and there has been success with automating the process using Tecan robotic systems (Bensch et al, 2005).

Although far more process data can be obtained using stacked microwell high-throughput techniques, including elution profiles and potentially purity and yield estimations, the capacity to predict a chromatogram with any degree of accuracy is still poor. In practice an approximation of the chromatogram is formed by taking data from a series of single well experiments at discrete data points and as such resolution of the whole peak is limited. The key advantage of the 96 microwell format is for effective high-throughput screening, which is highly applicable to very early phase experiments such as rapidly screening of sorbents and conditions.

3.2.5 Pipette Tips

A novel response to the challenge of high-throughput small scale but predicative chromatography is the use of pipette tips packed with chromatographic resin, which can be as small as 10µL in internal volume, but can be as large as 500µL. Early work used a similar device to a pipette tip but loaded into bespoke equipment for gas chromatography purposes (van Hout et al, 1999). More recently resin packed tips combined with a robotic system, utilising microwell plates for buffers, have been demonstrated to be capable of processing a multi-stage chromatographic purification (Creasey et al, 2001) and allowed reasonably accurate predictions of final purity and yield of a given chromatographic separation (Wenger et al, 2007). These chromatography tips are commercially available from a range of suppliers, including the key supplier PhyNexus (Chhatre & Titchener-Hooker, 2009).

3.2.6 Very Small Columns

A detailed review of the scaling down of preparative column chromatography processes are presented in Chapter 1 Section 1.7. In summary small columns that do not maintain bed height of the process scale rig give poor prediction of chromatographic performance due to a number of factors, including extra-column volumes influencing peak broadening, wall effects and other hydrodynamic influences (Kaltenbrunner et al, 1997).

3.2.7 Aims of This Chapter

Although the microwell, microfluidic and pipette tip based small scale chromatography methods have much to recommend them, for these particular experiments a precise process chromatography mimic is required, rather than the pseudo-chromatographic outputs typically produced by the scale-down systems detailed above.

Very small columns are capable of producing comparable data to a large scale chromatographic equivalent with minimal amounts of feedstock, without the serious engineering issues that are apparent at the microfluidic scale. As such small scale columns were selected as the ultra scale-down method of choice. A number of pre-packed small scale (1mL) chromatography columns are commercially available and three different column designs were investigated in this chapter, with regard to critical chromatographic parameters such as pressure drop, HETP and dynamic binding capacity.

3.3 Materials and Methods

Unless otherwise stated all reagents and chemicals were sourced from Sigma-Aldrich (Dorset, UK).

3.2.1 Small Scale Columns

Three different small scale columns were tested: Pall Pre-Packed PRC , prototype Pall AcroSep and the GE Healthcare Hi-Trap. All columns had an internal volume of 1mL and were pre-packed with the sorbent to manufacturer tolerances. Each was supplied sealed in 20% v/v ethanol. The aspect ratio of the columns tested are shown in Table 3.1.

Column	Aspect Ratio
PRC	1:10
AcroSep	1:1.57
Hi-Trap	1:3.5

Table 3.1: The aspect ratio values for each of the column designs tested

3.2.2 Equipment

All experiments were performed using an ÄKTA Basic chromatography systems utilising an A280nm UV cell for detection and Unicorn control software for the operation of the separations (GE Healthcare, Chalfont St Giles, U.K.).

3.2.3 Height Equivalent to a Theoretical Plate (HETP) Analysis

Height equivalent to a theoretical plate was measured by introducing a 50µL 5% v/v acetone/50mM NaCl pulse via a loop directly onto the column. A flow rate of 0.1mL/min was used for all columns and the resulting 280nm absorbance peak was recorded. Integration of said peak and quantification of the HETP was performed using Unicorn Evaluation software.

3.2.4 Dynamic Binding Capacity

The dynamic binding capacity was measured by analysing breakthrough protein concentrations, utilising a 280nm UV flow cell, for each column type at the terminus of the column. The flow rate was kept constant for all column types at 1mL/min.

The different sorbents/ligands were tested with commercially available analytical grade proteins, solubilised in the appropriate loading buffer at 5mg/mL as described in Table 3.2.

Ligand type	**Binding Protein**	**Loading Buffer Conditions**
CM	Lysozyme	50mM sodium acetate pH 4.8
Q	BSA	50mM Tris-HCl pH 9

Table 3.2: The loading buffer conditions for the cationic and anionic ion-exchange experiments

Columns were equilibrated and prepared with 20CV of loading buffer before the protein feed was loaded. Cleaning was performed with 5CV 1M NaOH. Columns were stored in 20% v/v ethanol.

3.2.5 Test Separation

A test separation was performed using a complex feedstock, a clarified periplasmic lysate from *E. coli* containing recombinant FAb fragments (Section 2.3.2). The periplasmic lysate contained 0.55mg/mL FAb fragments. Before application to packed bed chromatography the lysate was filtered using a

Nalgene 0.22µm depth filter to remove particulates, with the filtrate being retained.

Due to the relatively high pI of the FAb fragments (pI 9) binding to a Q ligand would only be achievable at a loading buffer pH value significantly above this. Such a high pH buffer causes precipitation effects within the periplasmic lysate and causes unacceptable fouling of the chromatography columns, therefore the test separation was only performed upon the columns packed with CM sorbents.

3.3.1.1 Chromatographic Separation Procedure

The loading, washing and elution steps were performed at 150cm/h. Each pre-packed column was equilibrated with an equilibration buffer consisting of 3CV 20mM pH 4.8 sodium phosphate. The pH of the FAb fragment feed was altered with 0.1M HCl to pH 4.8 and then 1 CV was loaded onto the column. The column was then washed with 5CV of 20mM sodium acetate equilibration buffer. Elution was performed via a linear pH gradient over 10CV, utilising an elution buffer consisting of 50mM diethylamine/1M NaCl at pH 12. Peak fractions were taken for further analysis. Cleaning was performed with 5CV NaOH and the columns were stored in 20% ethanol.

Elution chromatograms for each column design were recorded and fractions were analysed using SDS-PAGE. SDS-PAGE was performed using 4-20% Tris-glycine gels (Invitrogen EC6025BOX) under reducing conditions at 140 volts for 1 hour then stained with Coomassie blue utilising Invitrogen's SeeBlue system.

3.4 Results and Discussion

3.4.1 HETP

Figure 3.1 shows the results for each column design and ligand.

It can be seen that there are no statistical differences in HETP between the different column designs.

However, resolution is not determined by HETP; with all other factors such as column geometry and extra-column volume being equal, it is defined by the total number of theoretical plates present within the column. Despite having the same internal volume all three column designs were of different lengths and this influences how many total theoretical plates are available for a separation.

Figure 3.2 demonstrates the substantial difference this makes. PRC columns are over 4 times as long as the Hi-Trap and this corresponds to far more theoretical plates in total, much like a capillary chromatography system (Bartle & Myers, 2001). With all other chromatographic parameters remaining equal, the PRC columns should theoretically display greater resolution in a chromatographic separation relative to the AcroSep or Hi-Trap column designs.

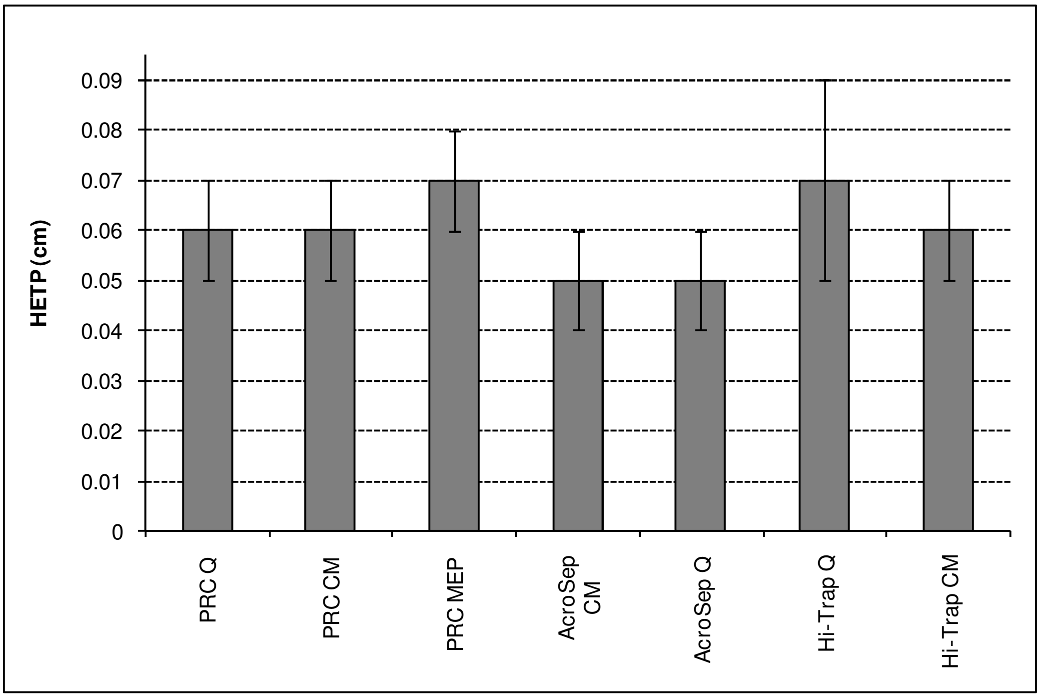


Figure 3.1: The HETP value for each column design and sorbent, as determined by the acetone pulse method as detailed in section 3.2.3. Error bars are calculated to 1 standard deviation from triplicate experiments with each column design.

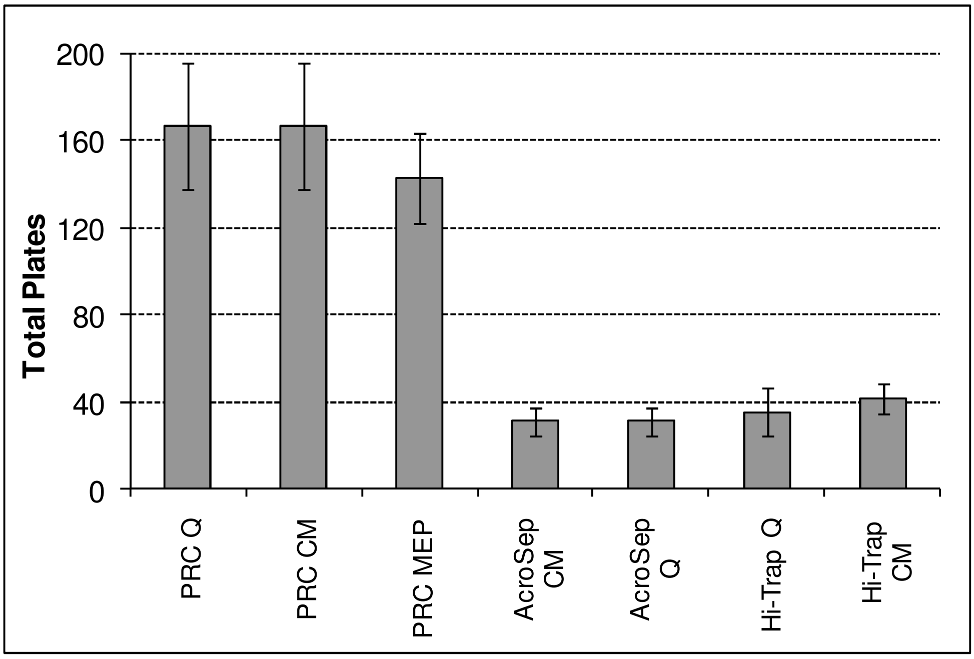


Figure 3.2: A bar graph showing total number of plates for each column design and ligand, ascertained by the multiplication of the HETP value presented in Figure 3.1 by the column length. Error bars are calculated to 1 standard deviation.

3.4.2 Dynamic Binding Capacity

The literature value for the dynamic binding capacity (DBC) of MEP HyperCel for human IgG, the molecule the ligand shows pseudo-affinity for, is 32mg/mL (Pall Life Sciences Product Note, USD 2629). The cost of procuring enough IgG to observe breakthrough, and hence gauge DBC, would be prohibitive and as such the dynamic binding capacity of MEP HyperCel was not tested.

Ion exchangers are characterised by high binding capacities. As seen in Figure 3.3, for the range of sorbents tested the DBC was between 20 and 60mg/mL for 5% breakthrough. The DBC was comparatively higher for the AcroSep and PRC columns compared to the Hi-Trap. The probable reason for this is that the hydrogel found within the ceramic sorbents packed into the PRC and AcroSep columns is characterised by superior mass transfer properties, as compared to Sepharose FF.

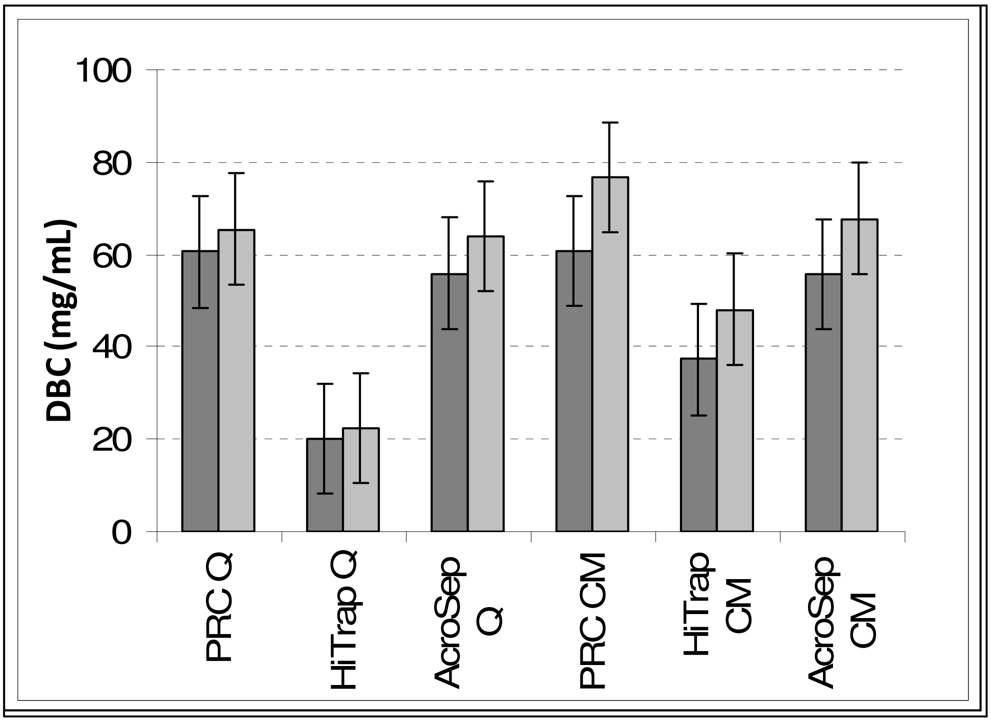


Figure 3.3: A bar graph showing the dynamic binding capacity values for each column design and ligand, for a suitable protein as defined in section 3.2.4. ■ *- 5% breakthrough,* ■ *- 50% breakthrough. Error is calculated to 1 standard deviation.*

The relatively quick 1mL/min loading flow rate gives a short residence time of 1 minute, which will favour those sorbents rapid mass transfer characteristics; if the residence time was extended then the apparent DBC of the Sepharose FF sorbents would likely to have increased substantially.

3.4.3 Test Separation

Figures 3.4 and 3.5 show a test separation of a valuable target molecule from a realistic feed, namely FAb fragments from a crude *E. coli* periplasmic lysate (section 2.3.2). As strong binding of the FAb fragments to the Q ligand would require significant pH adjustment of the feed liquor (see section 3.2.5) and hence only the CM ligand sorbents were tested. The MEP ligand was also not tested as it is only available in the PRC column design format; the separation of FAb fragments utilising this particular sorbent is investigated in Chapter 4.

As seen in Figure 3.5, with the relatively slow gradient over 10CV only a single peak was observed at approximately 20mS conductivity for all three column designs. What was notable however was that the peak height of the elution peak was much higher with the PRC format profile than with the other column designs, and eluted in a very similar volume. This indicates a higher chromatographic resolution for the PRC design. As the residence time for all the experiments was kept constant and the PRC and AcroSep designs shared the same ceramic sorbent, this can only be attributed to the additional number of theoretical plates found within the PRC designs, as demonstrated in section 3.3.1.

Qualitative analysis by reducing SDS-PAGE of the elution peak is shown in Figure 3.4. A significant concentrating effect of the FAb fragments was observed but selectivity was poor with several major contaminants being collected in the peak fraction for all three columns designs.

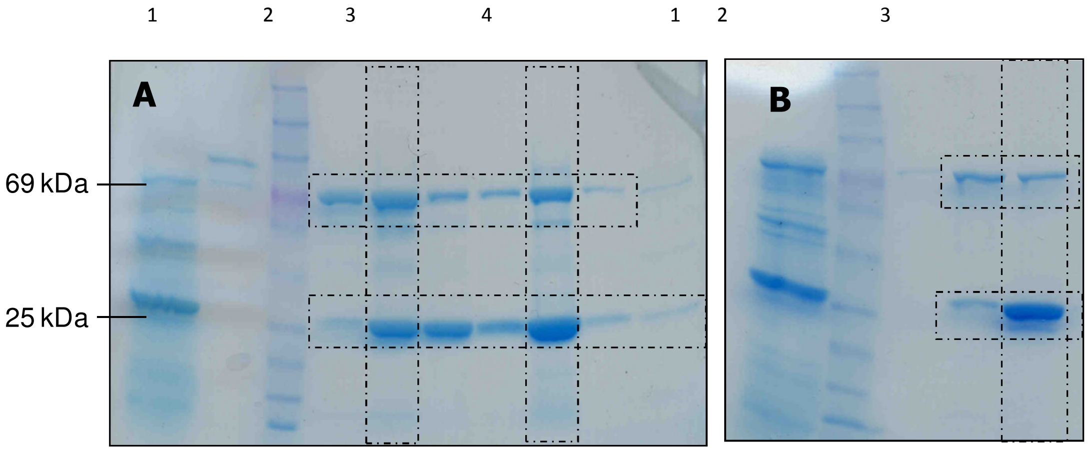


Figure 3.4: A reducing SDS-PAGE gels showing 1mL sample fractions taken from the first elution peak from a range of separations utilising a CM ligand and different column designs and sorbent backbones. Staining was with Coomassie blue and the marker was Benchmark pre-stained marker (Invitrogen). The peak fractions and FAb bands are boxed.

Gel A: 1. FAb feedstock 2. Marker 3. AcroSep elution peak 4. PRC elution peak; Gel B: 1. FAb feedstock 2. Marker 3. Hi-Trap elution peak.

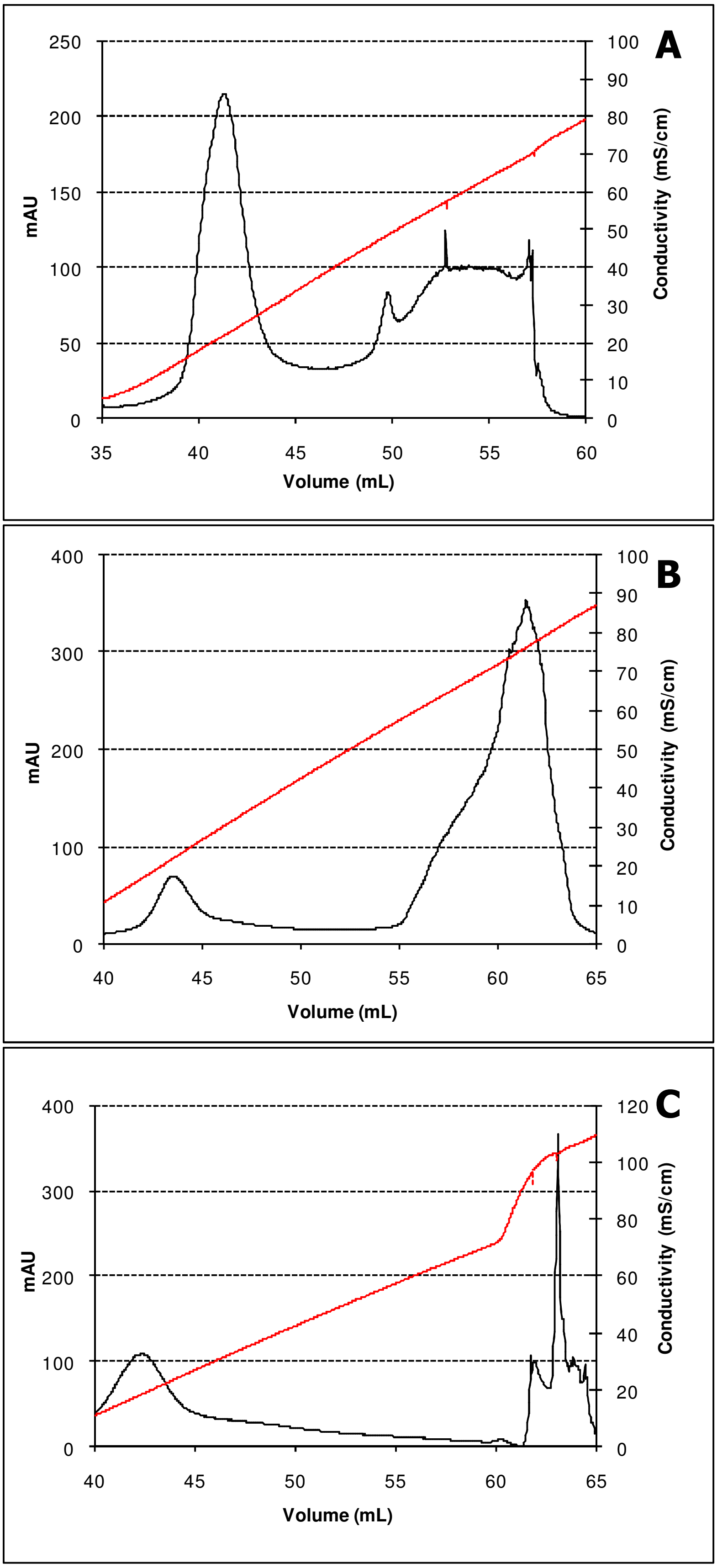


Figure 3.5: Three chromatograms showing the elution profile for the separation of FAb fragments from a crude periplasmic lysate using a gradient elution with a variety of column designs and sorbent backbones. The red dashed series is the conductivity.

A: CM Ceramic HyperD in PRC format, B: CM Ceramic HyperD in AcroSep format, C: CM Sepharose in Hi-Trap format.

3.5 Conclusions

The objective of this chapter was to investigate three different small (1mL internal volume) pre-packed column designs regarding common chromatographic parameters affecting operation and separation characteristics, namely HETP and dynamic binding capacity. A test separation of an *E. coli* periplasmic lysate was also performed to evaluate chromatographic separation of a complex and realistic feedstock.

It was found that as regards dynamic binding capacity there was little difference between the column designs.

There was little significant difference between the HETP measurements, but the PRC column design demonstrated approximately four times the total number of theoretical plates compared to the AcroSep and Hi-Trap designs, due to its long and narrow aspect ratio. The PRC design also displayed, albeit qualitatively, a sharper elution profile that the Hi-Trap and AcroSep column designs.

There were major differences in dynamic binding capacity, particularly with the Hi-Trap Q binding BSA. The highest dynamic binding capacities were observed with the PRC Ceramic HyperD, for both CM and Q ligands.

This data clearly recommends that the PRC design is the most suitable for future small scale experiments and this was therefore taken forward for further studies.

4 Separation of FAb Fragments by Mixed-Mode Chromatography

4.1 Abstract

FAb fragments have become important targets for biological therapeutic development. They are formed from immunoglobulin G proteins by the removal of the F_C region, leaving a stable 2 chain polypeptide of approximately 50KDa in size that contains the antigen binding epitope. Mixed mode chromatography ligands, including the very closely related HCIC ligands, are novel ligands that bind by a mild hydrophobic interaction and elute via a charge repulsion brought about by a pH change, providing separation on basis of hydrophobicity in low salt conditions. The MEP ligand shows pseudo-affinity binding behaviour with immunoglobulin G molecules.

Two mixed-mode sorbents, MEP HyperCel and PPA HyperCel (Pall Life Sciences, Portsmouth, UK) were investigated as to their chromatographic performance when separating FAb fragments contained within a complex *E. coli* periplasmic lysate feed liquor. The dynamic binding capacity, at 150cm/h loading linear velocity and pH 8.8 for purified FAb fragments at 50% breakthrough, was found to be 9.8mg/mL for PPA HyperCel and 5.0mg/mL for MEP HyperCel; this dropped to 2.1mg/mL for PPA HyperCel and 2.4mg/mL for MEP HyperCel at pH 7.4 with the complex lysate feedstock. This value was increased by the addition of 0.5M ammonium sulphate to the loading buffer and by raising the buffer pH to 8.8.

Chromatographic resolution was excellent with both sorbents, and FAb purities of >95% were observed. Two individual FAb containing peaks were identified by Protein G HPLC in both PPA and MEP chromatograms, with the second peak postulated as to being a different FAb species, potentially di-FAb or FAb fragments with an uncleaved leader sequence.

A CM Ceramic HyperD (Pall Life Sciences, Portsmouth, UK) ion exchanger sorbent was also investigated as a comparator and the resolution and purity obtained was significantly lower than those seen with PPA HyperCel and MEP HyperCel under the same linear velocity conditions.

4.2 Introduction

4.2.1 FAb Fragments

FAb and F(ab)$_2$ fragments have proven to be useful tools in targeted clinical therapies (Huhalov & Chester, 2004; Mountain & Adair, 1992), affinity based biosensors (Rogers et al, 2000) as well as in-vitro diagnostic tests (Borrebaeck et al, 2000) and have several advantages over using complete immunoglobulin G molecules. The F$_c$ region (absent in FAb fragment preparations) has associated immunological effector functions which can cause unexpected pharmacokinetic behaviour *in vivo*; its use for intravenous human therapies when glycosylated by appropriate mammalian or yeast hosts can have significant effects on blood clearance and immunogenicity. F(ab)$_2$ in particular can have increased affinity for the target epitope due to the lack of stearic hindrance from the F_c region which makes them appropriate for use in high sensitivity applications. Novel uses for engineered fragments include use as bespoke and specific affinity ligands (Kleymann et al, 1995).

4.2.2 Chromatographic Separation of FAb Fragments

The lack of the F_c region however can present bioprocess purification challenges. Previously immobilised metal ion affinity chromatography (IMAC) based methods have been shown to be successful in isolating FAb fragments (Skerra et al, 1994), but potential immunogenic effects of the His tags makes this unattractive in a regulated process environment. Protein A and G chromatography sorbents typically bind the F_c region with the greatest affinity, although the low affinity binding of protein A/G to the FAb and F_v regions can be exploited in analytical assays. This is in contrast to the newer Protein L, which binds strongly to the F_v region and is suitable, at least as regards specificity, for the purification of FAb fragments and scFv fragments (Sasso et al, 1991). Other issues also exist with using the Protein A/G/L class of affinity sorbents as a primary capture step, such as protease leaching, the need for very low pH elution and high cost media. Protein A sorbents also typically have a reduced lifetime due to the lack of resistance to alkali, though there has been progress in resolving this particular issue (Hahn et al, 1998). Ion exchangers are likewise not strong contenders as they require low ionic strength feeds to bind efficiently and most feeds for primary capture purification steps have a salt concentration that is far too high for effective binding, though some success in secondary purification has been reported using pH gradients rather than typical salt gradients (Mhatre et al, 1995). Issues of high salt content can be solved by expensive early diafiltration or by dilution of the feed, but the massive increase in process volume leads to greatly increased capital costs and difficult waste disposal issues (Queiroz et al, 2001). Hydrophobic interaction chromatography is also inefficient as this involves expensive disposal of large quantities of lyophilic salts, as well as difficulties such as protein precipitation and the essential inclusion of downstream diafiltration steps such high salt concentrations require (Roettger & Ladisch, 1989; Jungbauer et al, 2005).

Mixed-mode chromatography and the related hydrophobic charge induction chromatography (HCIC) ligands were designed to address many of the issues inherent in the binding of antibodies and their associated fragments. Originally mixed-mode functionality was proposed with two different ligands on the same bead (Belew et al, 2001), but this has since been replaced with a single ligand with multiple functional groups with different binding mechanisms (Burton et al, 1997; Lindner et al, 2005) . Mixed-mode ligands have been used previously for purification of target molecules as diverse as chemicals (Xu et al, 2007), antibodies (Burton et al, 2001), enzymes (Burton et al, 1997) and nucleic acids (McLaughlin et al, 1989). The ligands consist of both hydrophobic and hydrophilic groups, utilising aliphatic and aromatic amines on the same ligand which provides two modes of operation dependent on the working pH. HCIC sorbents bind principally by pseudo-affinity to IgG and its variants while modern mixed-mode sorbents principally bind by hydrophobic interaction, even in low salt conditions and at physiological pH. This simultaneous orthogonal mechanism of operation allows excellent separation and purification in a single step, even with complex feeds. However, the protein binding mechanisms are complex and can be difficult to predict, although recent investigations using quantitative structure-property relationships are starting to provide a more rational basis of separation and operation (Ghose et al, 2005; Yang et al, 2007). In both mixed-mode systems and HCIC, desorption is obtained by a reduction in the pH at which point the ligand becomes charged and bound molecules elute selectively by electrostatic repulsion. Key advantages of mixed mode chromatography are the flexible binding conditions with respect to feed pH and ionic strength, resistance to cleaning cycles and good selectivity for common bioprocess target molecules.

A range of mixed-mode and HCIC sorbents are commercially available for a range of applications. Pall Life Sciences (Portsmouth, UK) produce MEP HyperCel, PPA HyperCel and HEA HyperCel which have proven purification capabilities for immunoglobulins, ß-lactoglobulins and metallo-enzymes. (Lees et al, 2009; Brenac Brochier et al, 2008). GE Healthcare (GE Healthcare, Chalfont St Giles, U.K.) produce Capto MMC for similar separations, including purification of small proteins like human growth factor (Kaleas et al, 2010) and antibodies (Voitl et al, 2010), and there are over 18 different ligands under investigation or available for chromatography applications (Zhao et al, 2007).

4.2.3 Aims of This Chapter

In this chapter two sorbents, the HCIC MEP HyperCel and the mixed mode PPA HyperCel, were investigated with respect to the separation of FAb fragments from a complex *E. coli* lysate. This was compared and contrasted to a standard separation using a CM ion exchanger.

This chapter has two main objectives. The first objective was to investigate the effectiveness of a selected mixed mode and hydrophobic charge induction sorbent at separating an antibody fragment from a complex feed liquor, using a conventional carboxy-methyl ion exchange separation of the same feed as a reference.

The second is to produce a reference chromatographic separation of the assorted proteins found within the E. coli lysate, especially with reference to the FAb fragments which were considered the separation target. This reference separation will then be carried forward as an experimental system for the ultra scale-down methodology investigation as described in Chapter 5. The separation needs to be reproducible with more than one peak and with excellent purity (>90%) for the main FAb peak.

4.3 Materials and Methods

Unless otherwise stated all reagents and chemicals were sourced from Sigma-Aldrich (Dorset, UK).

4.3.1 FAb lysate feedstock production - Fermentation

The full method for the FAb lysate feedstock production may be found in Section 2.3.2. The single fermentation batch from which the feedstock for the study in this chapter is derived was referred to as “Batch 2”.

4.3.2 FAb lysate feedstock production – Periplasmic lysis

The full method for the liberation of FAb fragments from the periplasm is detailed in section 2.3.3.

4.3.3 Chromatographic separations – PRC pre-packed columns

The chromatography columns used were Pall Life Sciences Pre-Packed 1mL PRC columns packed with either MEP HyperCel, PPA HyperCel or CM Ceramic HyperD sorbents (Pall Life Sciences, Portsmouth, UK). All experiments were performed on an ÄKTA Basic chromatography system (GE Healthcare, Amersham, UK) utilising a absorbance 280nm UV flow cell.

4.3.3.1 MEP and PPA HyperCel Separation Procedure

The loading, washing and elution steps were performed at 150cm/h. Each pre-packed column was equilibrated with an equilibration buffer consisting of 3CV 50mM Tris-HCl, pH 8.8 or 50mM Tris-HCl/0.5M ammonium sulphate depending on the experimental conditions. The pH of the FAb fragment feed was adjusted with 0.5M NaOH to either pH 7.2 or pH 8.8 as experimental conditions dictated and then 1 CV was loaded onto the column. The column was then washed with 5CV of 50mM Tris-HCl equilibration buffer. Elution was performed via a linear pH gradient over 10CV, utilising an elution buffer consisting of 50mM sodium citrate, pH 2.2. 1mL fractions were taken throughout the elution gradient for the purposes of analysis, as described in section 4.3.5. Cleaning was performed with 5CV NaOH and the columns were stored in 20% ethanol.

4.3.3.2 CM Ceramic HyperD Separation Procedure

The loading, washing and elution steps were performed at 150cm/h. Each pre-packed column was equilibrated with an equilibration buffer consisting of 3CV 20mM pH 6 sodium phosphate. The pH of the FAb fragment feed was adjusted with 0.1M HCl to pH 6 and then 1 CV was loaded onto the column. The column was then washed with 5CV of 20mM sodium phosphate equilibration buffer. Elution was performed via a linear pH gradient over 10CV, utilising an elution buffer consisting of 20mM sodium phosphate pH6/1M NaCl. 1mL fractions were taken throughout the elution gradient for the purposes of analysis, as described in section 4.3.5. Cleaning was performed with 5CV NaOH and the columns were stored in 20% ethanol.

4.3.4 Generation of a pure FAb sample

Pure FAb was obtained by taking a 1mL fraction from the main peak as found in the separation of the lysate by MEP HyperCel.

4.3.5 Analysis

Protein concentrations were measured using the BCA Total Protein assay according to the method detailed in the instructions included within the kit. SDS-PAGE was performed using 4-20% tris-glycine gels (Invitrogen EC6025BOX) under reducing conditions at 140 volts for 1 hour then stained with Coomassie blue utilising Invitrogen's SeeBlue system. Protein G HPLC for analysis of FAb fragment concentration was performed on an Agilent 1200 HPLC using Protein G packed Sepharose Fast-Flow Hi-Trap columns (GE Healthcare) at a flow rate of 2mL/min. Elution was via pH gradient and controls consisted of pure FAb fragment purified using mixed-mode chromatography.

4.4 Results and Discussion

The objective of this chapter was to investigate and compare the chromatographic separation of a realistic feedstock (FAb fragment containing *E. coli* periplasmic lysate) on a range of sorbents, with the intention of purifying the FAb fragments from the other host cell proteins within the lysate. To this end a number of critical chromatographic performance indicators were investigated, namely the elution profile, purity, yield and dynamic binding capacity.

4.4.1 Dynamic Binding Capacity for FAb Fragments– MEP HyperCel

The apparent dynamic binding capacity (DBC) was investigated for both PPA HyperCel and MEP HyperCel sorbents in PRC pre-packed columns formats under a range of pH and conductivity conditions. FAb titre within the feed was 0.55mg/mL.

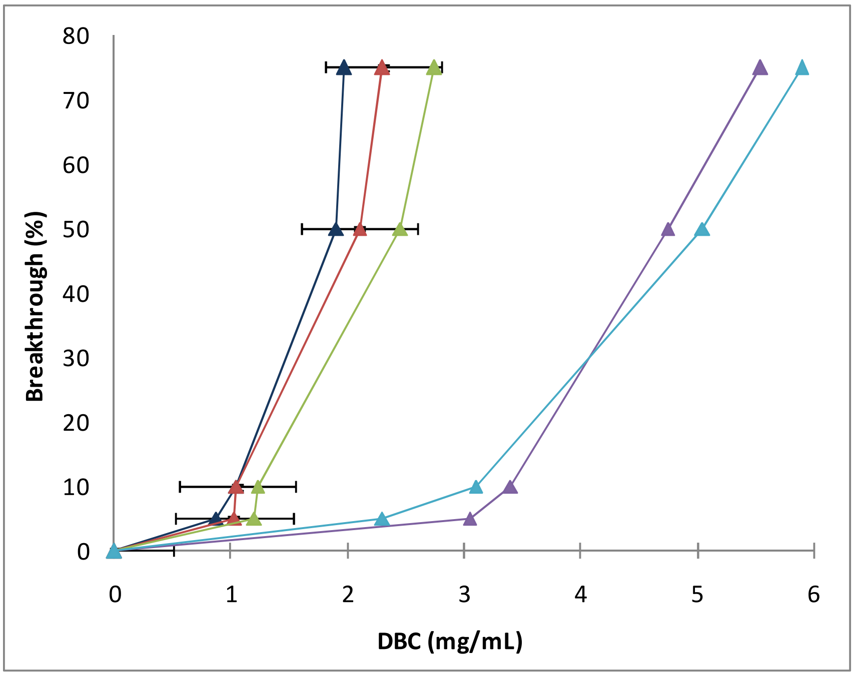


Figure 4.1: Dynamic binding capacities for increasing percentage breakthrough values, regarding FAb fragments bound to MEP HyperCel, loaded at 150cm/h. $+(NH_4)_2SO_4$ indicates feed and equilibration/loading buffers included 0.5M ammonium sulphate. Pure FAb was a >95% purified FAb fraction used as a feed. For clarity only the X error bars for the pH 8.8 (▲) series are shown, to 1 standard deviation. ▲ = pH 7.2, ▲ = pH 8.8, ▲ = pH 7.2 + $(NH_4)_2SO_4$, ▲ = pH 8.8 + $(NH_4)_2SO_4$, ▲ = pure FAb at pH 8.8.

It is clear that there are significant increases to dynamic binding capacities when FAb fragments are loaded at pH 8.8 (which approaches the pI of the FAb fragment, which is 9) rather than physiological pH.

It is understood that the primary binding mechanism of HCIC and mixed-mode sorbents are hydrophobic therefore bringing the loading pH closer to the pI will result in a less charged fragment and potentially stronger binding, as seen by a clear 15% increase in binding capacity between loading at physiological pH 7 and pH 8.8. This was further supported by the binding capacities observed when 0.5M ammonium sulphate was included in the loading buffers – a 34% increase in binding capacity was seen at pH 7.2 with a pronounced 160% increase seen at pH 8.8. A protein in a hydrophobic state is likely to be influenced more by salt concentration than one with a moderate net charge. Whether the binding capacity was increased only for the target FAb or for contaminants as well is discussed in section 4.4.4.

4.4.2 Dynamic Binding Capacity for FAb Fragments – PPA HyperCel

The addition of ammonium sulphate to the loading buffer for PPA HyperCel appeared to have an unpredictable and irreproducible effects on the binding of FAb fragments, most likely due to an increase in non-specific binding, and was therefore not investigated further.

The dynamic binding capacity results for PPA HyperCel broadly followed the results for MEP HyperCel. It was clear for both sorbents that a degree of non-specific binding was occurring as shown by the large increase in binding capacity when purified FAb fragments are loaded. Although PPA HyperCel binds more FAb fragments per mL of sorbent than MEP HyperCel more contaminant proteins were also bound, leading to roughly equivalent DBC using complex feeds. This supports the pseudo-affinity binding mechanism of MEP ligand that reduces non-specific binding, while the PPA ligand binds by a less selective mild hydrophobic interaction.

The DBC values for both sorbents are low compared to the quoted literature Figures; >20mg/mL for human IgG MEP HyperCel and >40mg/mL BSA for

PPA HyperCel (Pall Life Sciences, Technical Datasheet). There are various contributing factors however; the literature values were obtained using pure proteins at high concentrations (~5mg/mL) in ideal salt conditions and with target proteins that have physiological or near physiological pI. It also should be noted that FAb fragments are only 25KDa, much smaller than the 150KDa of an intact IgG molecule, and this renders a direct comparison by mass/mL difficult. The binding capacity data does however suggest that MEP HyperCel's pseudo-affinity is far more pronounced for the absent F_c region of IgG rather than the F_V/FAb fragments investigated here.

The DBC values for the sorbents and running conditions are summarised in Figure 4.3.

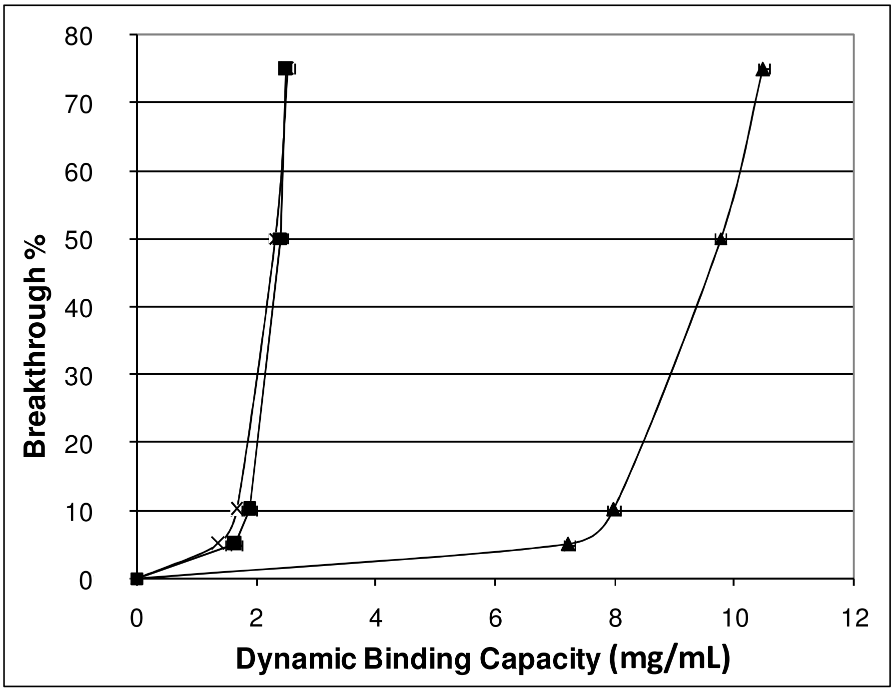


Figure 4.2: Dynamic binding capacities for various percentage breakthrough values, regarding FAb fragments bound to PPA HyperCel at 150cm/h. Pure FAb was a >95% purified FAb fraction used as a feed at 5 mg/mL. x = pH 7.4, ■ = pH 8.8, ▲ = Pure FAb from a 95% purified fraction at 5mg/mL.

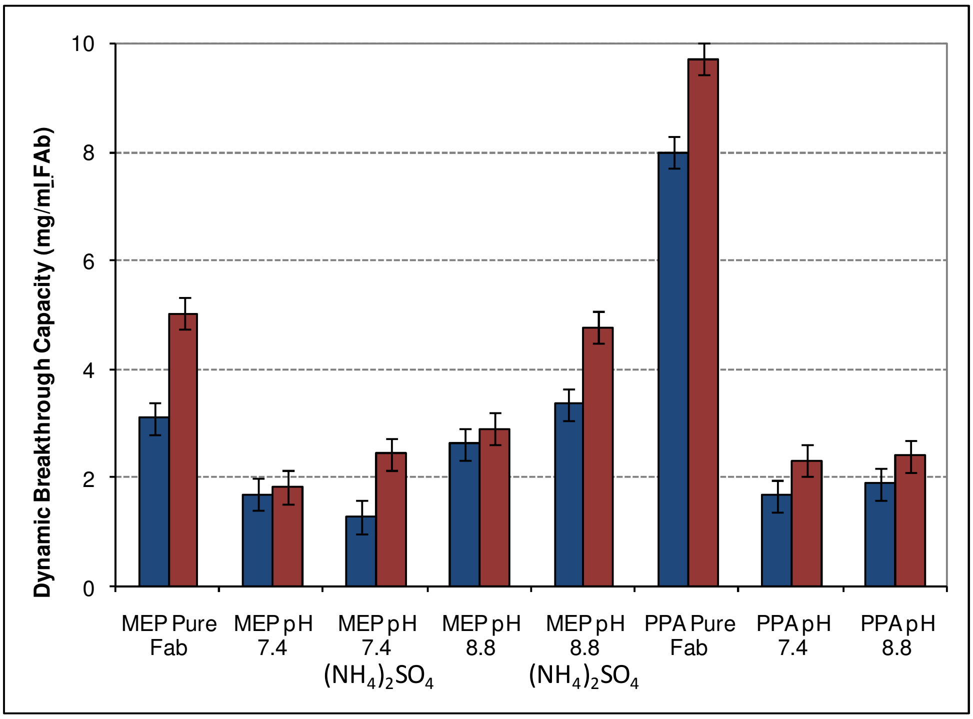


Figure 4.3: A comparison of dynamic binding capacities for FAb fragments on both PPA HyperCel and MEP HyperCel sorbents, loaded at 150cm/h under various pH and buffer conditions, from a crude periplasmic lysate. $(NH)_2SO_4$ *indicates 0.5M ammonium sulphate was added to the loading and equilibration buffer. Pure FAb indicates a purified preparation utilised as a feed rather than the complex lysate. Blue(■) bars are at 10% breakthrough and red (■) are at 50% breakthrough.*

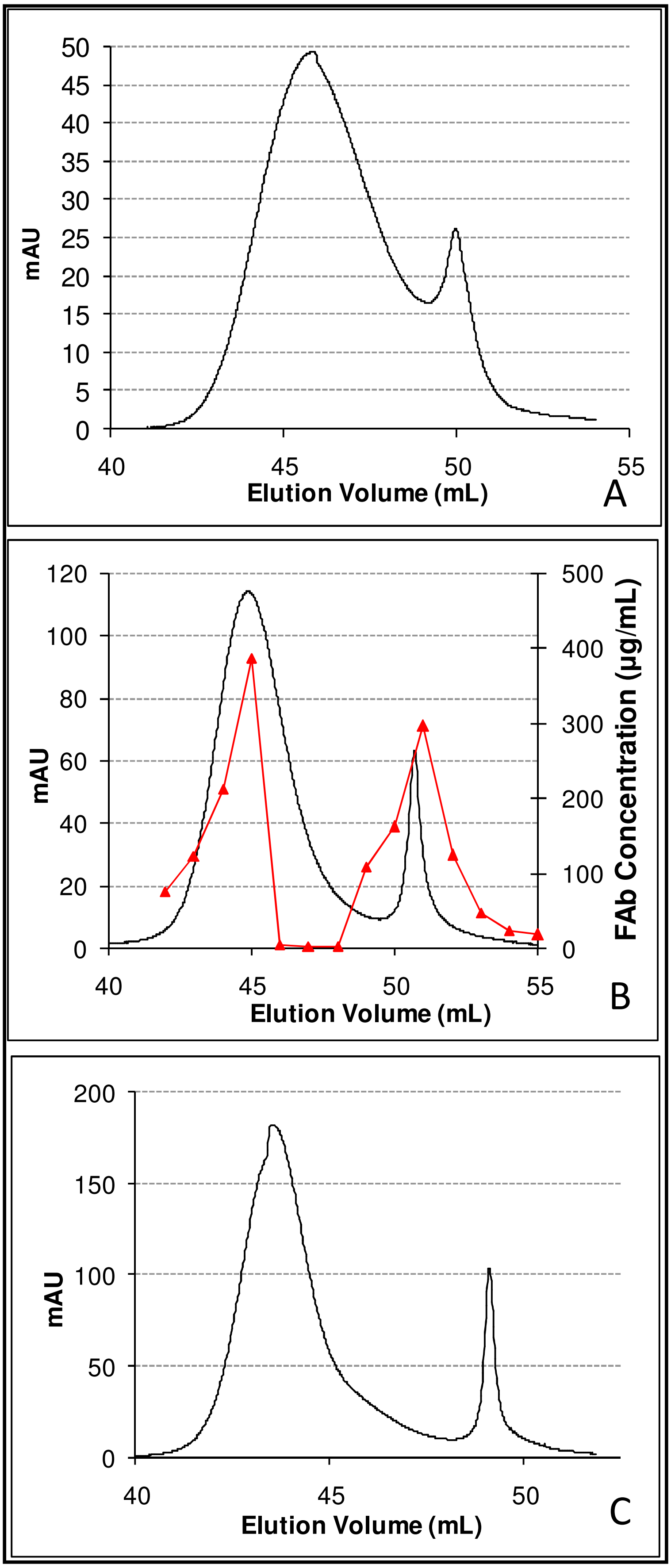


Figure 4.4: A set of elution profiles obtained by the separation of the proteins in FAb fragment containing E. coli *lysate. A: 300cm/h loading and elution linear velocities, B: 150cm/h loading and elution linear velocities, C: 50cm/h loading and elution linear velocities.*

The red series on the secondary axis of elution profile B is the FAb fragment concentration taken from 1mL fractions of the eluted material, as determined by Protein G HPLC.

The separation was performed using MEP HyperCel in PRC format with a 10CV pH gradient (pH 8.8-pH 2.2) at 150cm/h.

For all the separation analysis runs the PRC column was loaded with 0.55mg of FAb fragments per mL of sorbent as defined in section 4.3.3. The elution chromatograms are presented in Figure 4.4 for a range of loading and elution linear velocities.

It is clear from the elution profiles that separation over a gradient of 10 CV gives excellent resolution at acceptable flow rates. The recommended linear velocity for running separations using MEP HyperCel is 100cm/h although good separation is still shown at 150cm/h. The Protein G HPLC analysis (section 4.3.5) of the fractions show that both the peaks present here are in fact FAb fragments, as it is highly unlikely *E. coli* would produce any other antibody analogue that would bind to the Protein G ligand. Identification and purity is discussed in section 4.4.6.

The FAb fragment peak elutes at pH 5.5. This is atypical for MEP HyperCel, which tends to elute at pH 4.8. As this FAb fragment is a basic protein elution could be initiated by net charge changes in the protein itself rather than the ligand, at pH 5.5. The second peak elutes between pH 4.8 and pH 5.

4.4.3 Chromatographic Separation of FAb fragments using MEP HyperCel – Peak Purity and Yield

Figure 4.5 shows convincing qualitative evidence for high purity of FAb fragments in the first peak of the elution profiles obtained from the individual chromatograms. IgG FAb fragments are typically 50 KDa in size (Coleman & Mahler, 2003) consisting of two very similar chains bound with disulphide bonds. Under reducing conditions this produces a strong band at 25 KDa seen here. A very faint band is also visible at approximately 50 KDa and this is most likely this band is native FAb that has failed to reduce fully; FAb fragments are extremely stable even under reducing conditions (Röthlisberger et al, 2005).

The second peak, eluting at approximately 50mL, is more complex to define. Protein G HPLC analysis (section 4.3.5) identified this peak as consisting of FAb fragments and the 25KDa size of the main band potentially confirmed this, although there is a consistent faint contaminant band at approximately 22 KDa.

There are a number of potential identities for this second peak. The first is that it is formed primarily of FAb fragments with a leader sequence still attached; the production process directs the FAb fragments to the periplasm using a leader sequence which is then cleaved. Potentially the leader sequence could still be attached if the periplasm is lysed before cleavage is complete. The FAb fragment would still be 25KDa as the leader sequence would be reduced by the temperature and chemical reducing conditions prior to loading on the gel.

The second potential identity is that the feedstock comprises a mixture of FAb fragments and di-FAb – two FAb fragments still connected by a linker at the F_c terminus. Again di-FAb would be reduced to single FAb chains by the reducing conditions on the gel and hence appear at 25 KDa on the SDS-PAGE gel. The cleaning lane shows a small concentration of FAb fragments is retained on the column after the elution conditions have been applied and only eluted when the column is cleaned.

The chromatographic separation results are interesting as they demonstrate excellent separation of two very similar peptide entities using a chemical ligand. Such resolution is rarely seen in ion exchange, HIC or affinity systems, particularly in primary capture applications.

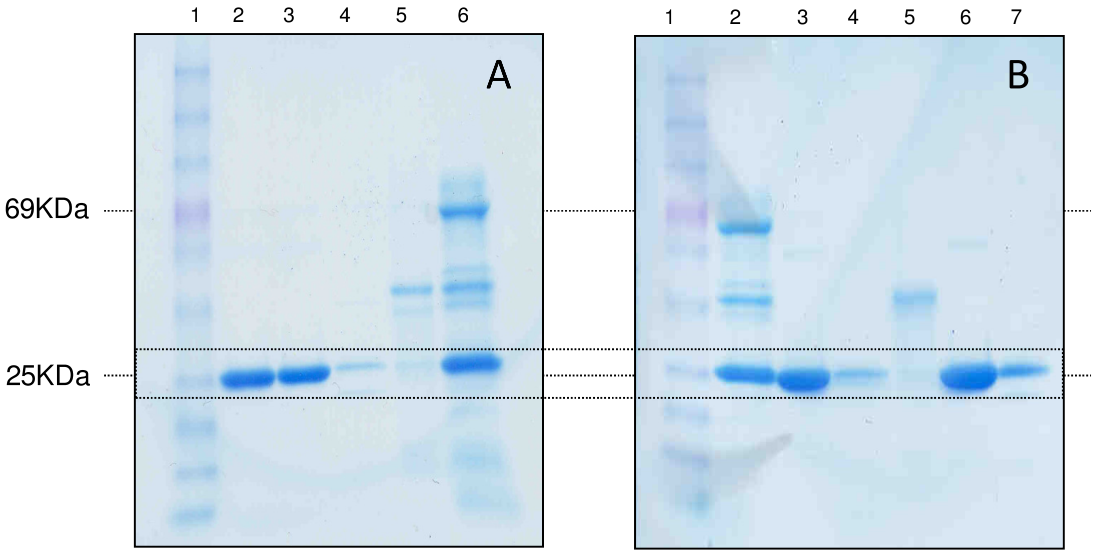


Figure 4.5: SDS-PAGE gels containing samples from fractions obtained from the elution profiles shown in Figure 4.4. Gel A presents the samples from the 300cm/h loading/elution linear velocities. Gel B presents the samples from the 150cm/h and 50cm/h loading/elution linear velocities respectively. FAb bands are boxed. Benchmark pre-stained marker (Invitrogen) was used in both gels and staining was performed with Coomassie blue.

Gel A lanes: 1. Marker 2. Pure FAb control 3. 1st peak 4. 2nd peak 5. Cleaning peak (300cm/h run) 6. FAb lysate feedstock.

Gel B lanes: 1. Marker 2. FAb lysate feedstock 3. 1st peak 150cm/h 4. 2nd peak 150cm/h 5. Cleaning peak (150cm/h run) 6. 1st peak 50 cm/h 7. 2nd peak 50cm/h.

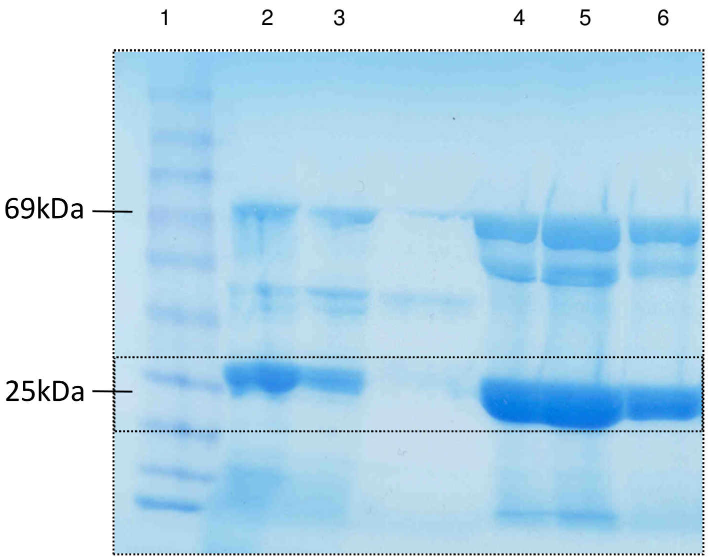


Figure 4.6: A reducing SDS-PAGE gel presenting samples taken from the elution stage of an MEP HyperCel separation of FAb fragments from an E. coli *periplasmic lysate. 0.5M ammonium sulphate was included in the loading and elution buffers. Gel lanes: 1. Marker. 2&3: FAb lysate feedstock. 4&5. 1st peak 6. 2nd peak.*

As Figure 4.6 demonstrates the increased binding capacity achievable with high salt conditions for MEP HyperCel, as seen in section 4.4.2, does however come at a cost. Operation at high salt concentrations reduces the selectivity seen previously and high concentrations of contaminants were observed co-eluting with the product FAb fragments. The Protein G HPLC assay could not give reproducible results due to the high ammonium sulphate concentration (Generon, Data Sheet Protein G Sepharose) but it is clear from the SDS-PAGE results that this separation is not viable at high salt levels.

As expected from the qualitative SDS-PAGE results in Figure 4.7, quantitative purity of the 1st peak is very high (>92%) at both pH conditions with no salt, as determined by BCA total protein analysis and Protein G HPLC (section 4.3.5). This is particularly significant for a primary capture step from a complex feed where acceptable selectivity is difficult to achieve in practice. It should be noted however that for industrial purposes it is likely both peaks would be taken and the di-FAb would be reduced in a further step.

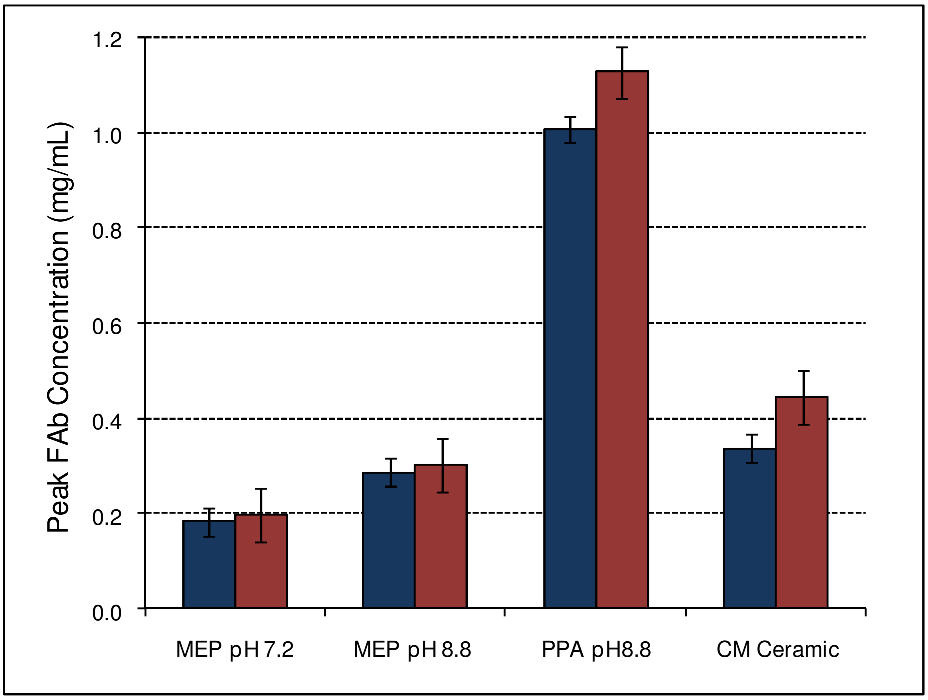


Figure 4.7: Samples from the highest concentration 1mL fraction from the 1st FAb peak of the elution profile, for all sorbents and relevant conditions tested, were subjected to Protein G HPLC and BCA Total Protein Assay analysis. Blue (■) bars show the concentration of FAb fragments as determined by Protein G HPLC and the red (■) bars show the total protein concentration in the same fraction as determined by BCA Total Protein assay. Error is shown to 1 standard deviation.

4.4.4 Chromatographic Separation of FAb fragments using PPA HyperCel – Elution Profile

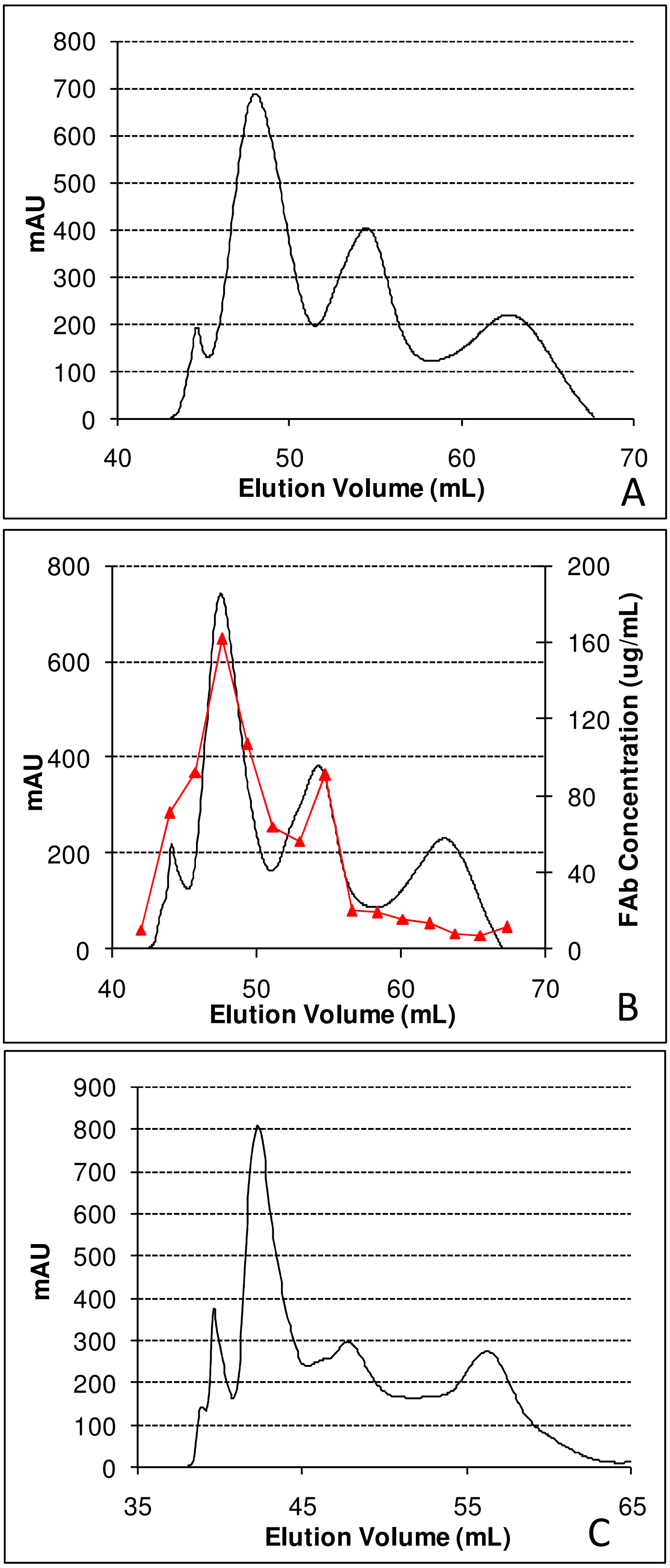


Figure 4.8: A set of elution profiles obtained by the separation of the proteins in FAb fragment containing E. coli *lysate. A: 300cm/h loading and elution linear velocities, B: 150cm/h loading and elution linear velocities, C – 50cm/h loading and elution linear velocities.*

The red series on the secondary axis of elution profile B is the FAb fragment concentration, as determined by Protein G HPLC, taken from 0.75mL fractions of the eluted material.

The separation was performed using PPA HyperCel in PRC format with a 10CV pH gradient (pH8.8-pH2.2).

Separation analysis runs were conducted with the PRC column loaded with 0.55mg of FAb fragments, as defined in section 4.3.3. The elution chromatograms are presented in Figure 4.8 for a range of flow rates.

The chromatograms in Figure 4.8 show a far more complex elution profile than seen with MEP HyperCel, with at least 6 distinguishable peaks visible in the chromatogram obtained with 50cm/h loading and elution linear velocities. This suggests a far less specific binding with the PPA ligand compared to the MEP ligand, which is consistent with the known binding mechanisms of the two ligands (Brenac Brochier et al, 2008). A pseudo-affinity ligand like MEP HyperCel, by definition, exhibits fairly specific binding; the hydrophobic mechanism of the PPA ligand will bind any protein or peptide with a suitable hydrophobic patch, which will be a significant proportion of the proteins found within the lysate feed.

The elution profile for FAb fragments was very similar to that seen for MEP HyperCel. The FAb eluted in two separate peaks with the higher concentration of FAb in the first peak to elute. The first FAb peak elutes at pH 5.6 and the second FAb peak eluted at approximately pH 5.

4.4.5 Chromatographic Separation of FAb fragments using PPA HyperCel – Peak Purity and Yield

Figure 4.9 shows an strong band at 25KDa in lanes 1 and 2, which was confirmed as FAb fragments by the Protein G HPLC data shown in Figure 4.8. Purity was very good for a primary capture step, as shown qualitatively in

Figure 4.8 and quantitatively in Figure 4.9 at approximately 89%, as ascertained by BCA assay and Protein G HPLC. This purity value was however below the value that was seen with MEP HyperCel. This is compensated by a much higher product concentration for the peak fraction, over 1mg/mL. The fact that PPA HyperCel and MEP HyperCel had similar DBC values for binding FAb fragments from a complex feed suggests elution is more complete for PPA HyperCel, while some FAb fragments are retained on the MEP HyperCel sorbent even after elution.

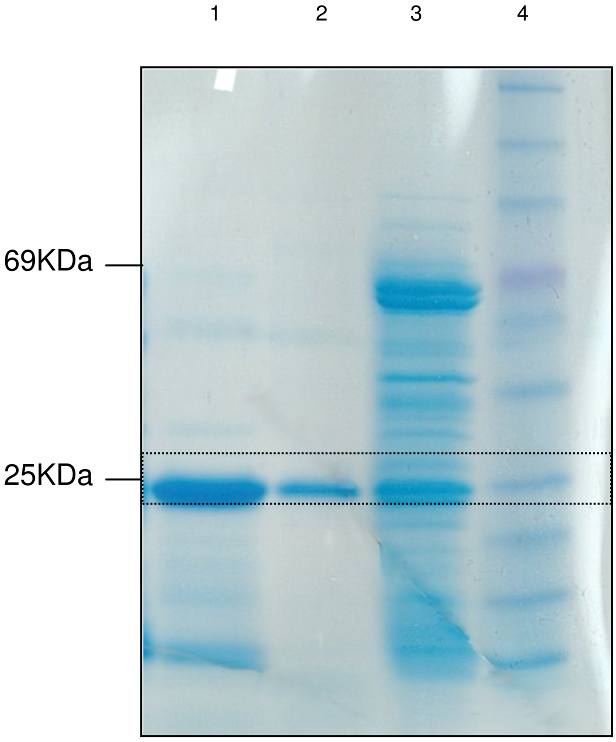


Figure 4.9: An reducing SDS-PAGE gel presenting the 1mL fractions obtained from the two most relevant peaks in the PPA HyperCel/FAb fragment lysate elution profile. Staining is with Coomassie blue and the marker is Benchmark pre-stained marker (Invitrogen). The FAb fragment bands are boxed. Gel lanes: 1: 1st FAb peak 2: 2nd FAb peak 3: FAb fragment feed 4: Marker.

4.4.6 Chromatographic Separation of FAb fragments using CM Ceramic HyperD – Elution Profile

A sample separation of the FAb containing feedstock was performed using a carboxy-methyl weak cationic ion exchanger, CM Ceramic HyperD. This was for comparison with the mixed-mode and HCIC separations detailed in sections 4.4.4 to 4.4.7.

As Figure 4.10 shows, little visible separation was achieved using the CM ion exchanger for the purification of FAb fragments. Elution of the FAb fragments proved very difficult, eventually requiring a combined pH and salt gradient. Elution was eventually achieved at a pH 10.2 and a conductivity of 22mS, with the FAb concentration being present in the latter section of the peak. Whether this challenging elution is due to the FAb being very basic and therefore tightly bound or due to an interaction with the ceramic sorbent backbones is unclear. The elution profile after 45mL of elution was essentially a cleaning peak due to the high pH.

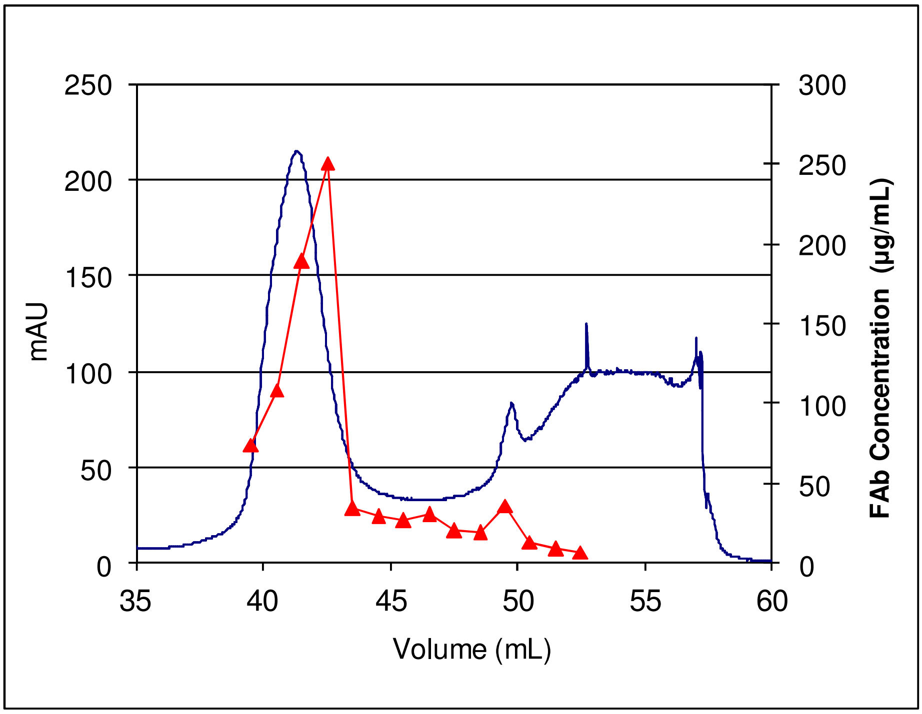


Figure 4.10: An elution chromatogram obtained by the separation of the proteins in FAb fragment containing E. coli lysate, at 150cm/h loading and elution conditions. The separation was performed using a CM Ceramic HyperD column in PRC format, eluting with a 10CV NaCl gradient (0-1M). Equilibration, loading and elution were performed at pH 7.4. ● – Elution profile at A280nm, ▲ - FAb concentration as determined by Protein G HPLC.

4.4.7 Chromatographic Separation of FAb fragments using CM Ceramic HyperD – Peak Purity and Yield

To keep the comparison to the mixed-mode sorbents consistent the feedstock was not pre-processed except for the necessary clarification steps for packed bed chromatography. Typically ion-exchangers require to be operated at a conductivity of <3mS. The clarified FAb feedstock had a conductivity of approximately 12-15mS. Within an industrial process setting such a feed would be either diafiltered or more likely diluted to reduce the conductivity levels so that effective binding could be achieved. This high ionic strength probably explains why the FAb concentration of the elution peak is so low (Figure 4.10), as compared to the norm for an ion-exchanger of this type. The purity is approximately 75%, revealing there are other proteins within the feed with a similar pI and which therefore elute in the same fraction.

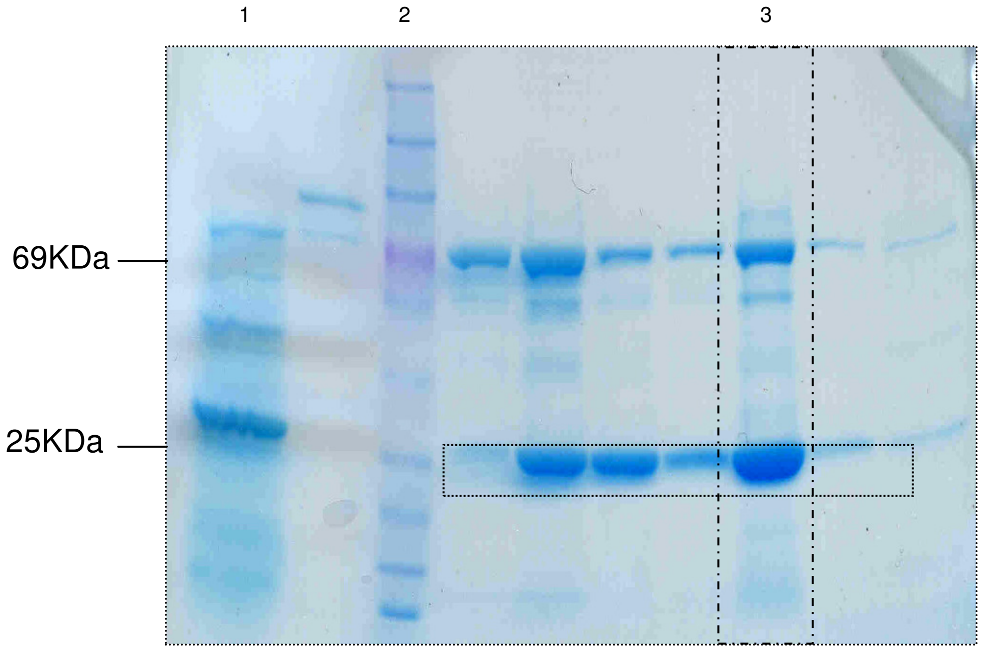


Figure 4.11: A reducing SDS-PAGE gel showing the protein composition of a 1mL fraction taken from the most concentrated part of the FAb peak. Staining is with Coomassie blue and the marker was Benchmark pre-stained marker (Invitrogen). Gel lanes are: 1. Feedstock 2: Marker 3: FAb peak fraction. FAb bands are boxed and the sample of interest are boxed.

Figure 4.11 shows several high concentration impurities, as well as a strong FAb band at 25KDa. The highest concentration contaminant at approximately 70KDa is potentially alkaline phosphatase; its periplasmic variant is liberated in the periplasmic lysis process in significant quantities.

4.5 Conclusions

The objective of this chapter was to investigate the effectiveness of the recently commercialised mixed mode and HCIC sorbents regarding the separation of FAb fragments from a periplasmic *E. coli* lysate, utilising a CM ion exchanger as a comparator.

In the case of purifying FAb fragments MEP HyperCel has been shown not to be as effective as some chromatographic alternatives with regard to a primary capture step, as the dynamic binding capacity is comparatively low. However, the selectivity and purity of the elution fraction is excellent, even when separating from a complex feedstock like the FAb fragment containing *E. coli* lysate examined here, which would make MEP HyperCel a good choice for secondary chromatographic purification.

It has been shown here that CM Ceramic HyperD can purify and concentrate FAb fragments to a degree and if the feed is pre-treated it could be postulated that binding capacities would rise to the ion-exchanger norm (>25mg FAb/mL sorbent). The high conductivity and high concentration of the eluted FAb fraction from such a separation would theoretically improve the binding to MEP HyperCel and would not need to be diafiltered before application. The degree of selectivity even with closely related species would also make MEP HyperCel an excellent choice for purifications that involve closely related contaminant species.

PPA HyperCel produced a far higher yield in its peak fraction than MEP HyperCel but at the expense of a small drop in purity. The mode of action of the PPA ligand is suited to a larger range of applications as it mainly binds by hydrophobicity rather than by pseudo-affinity like the MEP ligand. It is therefore logical that the binding capacity will be relatively higher in the favourable ionic conditions within the FAb lysate with the compromise being binding of non-target proteins with suitable hydrophobic patches.

HEA HyperCel, another mixed-mode cellulose based sorbent from Pall Life Sciences, was also investigated with respect to binding FAb fragments from the crude clarified lysate. It was found to be ineffective in binding and separation of FAb fragments (data not shown), which shows that two mixed-mode sorbents with very similar binding mechanisms (the HEA and PPA ligands) can demonstrate very different chromatographic behaviour and have a range of diverse chromatographic applications.

Both the PPA HyperCel and MEP HyperCel separations, although not optimised for an economic industrial process, give clear and reproducible results. For this reason both the separations are suitable targets for future experiments and investigations.

Of the three separations of FAb fragment containing lysate examined here (with the MEP, PPA and CM ligands), the MEP HyperCel experimental system is the most straight-forward to analyse. There are only two peaks that can be well resolved at slow linear velocities and the peak consists of a single species of FAb fragment. As such the MEP HyperCel experimental system, rather than the PPA or CM systems, is the most suitable to use for the development of the ultra scale-down methodology and will be adopted for future studies.

5 Formulation of an Ultra Scale-Down Methodology for Chromatography Scale-Up

5.1 Abstract

This chapter concerns the development of an ultra scale-down method and model to predict accurately laboratory scale (10-60mL internal column volume) chromatographic behaviour using very small (1mL internal column volume) pre-packed chromatographic columns. The experimental system was based upon a gradient separation of a crude *E. coli*-derived lysate utilising the mixed-mode sorbent MEP HyperCel, as defined in Chapter 4.

A methodology was formulated and based around an experimental regime created by application of factorial design. The derived chromatographic data was processed with the exponentially modified Gaussian equation to derive curve coefficients for each peak within the chromatograms. The curve coefficients were then transformed with a mathematically derived function to predict accurately the chromatogram produced at large scale when operated at a range of different packed bed heights and loading/elution linear velocities. The formulated methodology and transformation functions were very effective at accurately predicting the large scale chromatogram using only the small scale chromatographic data, with the majority of predicted

large scale chromatograms being less than 5% different to the large scale chromatogram generated experimentally, as determined by normalised root mean square analysis.

5.2 Introduction

5.2.1 Applications – Accelerated Process Development

The primary driver behind the development of an ultra scale-down system is to facilitate rapid process development. The most useful point of application is in the early stages of process development where feed material is of limited availability and financial resources are most likely to be minimal. This combination renders an extensive experimental programme at process scales infeasible. A scale-down model for a particular process operation, with the study focussed by the application of a design of experiments (DoE) regime, can reduce massively the amount of feedstock and resources required to obtain accurate data on which to base full scale process decisions (Godavarti et al, 2002). The limitation of conventional scale-down models is the extent of the scale down that is feasible before the physical size of a process is so small that geometric similarity to the large scale process is lost. For example, traditional scale down of chromatography requires the bed height to be maintained whilst the bed diameter is reduced and there is a practical limit as to how small a scale down column can be without introducing other variables, such as disproportionately large extra-volume to intra-column volume ratios which increase dispersion and reduce chromatographic resolution.

Ultra scale-down methods differ from conventional scale down in that the ultra scale-down device does not necessarily bear any geometric or physical resemblance to the original item of equipment it seeks to represent. There has been considerable success predicting large scale chromatographic behaviour using microwell based systems (Coffman et al, 2008) and in producing a performance mimic of tangential flow microfiltration using rotating disc shear devices (Ma et al, 2010). As these ultra scale-down methods are decoupled from the restrictions of physically mimicking the process scale device they can be very small indeed, often requiring only a few millilitres of feedstock.

The disadvantage of ultra scale-down devices is that a variety of physical effects, mostly hydrodynamic in nature, can produce non-linear performance differences that are non-trivial to translate between very small and large scale. A "bridging" method is therefore required which links the ultra scale-down data to the process scale data it is seeking to emulate. This bridging method can be theoretical in nature, such as application of computational fluid dynamics and mathematical modelling, but it can be notoriously difficult to apply simplistic models to real data obtained from experimentally separating complex feed streams. The second method is to evaluate apparent results, to analyse the changes in process operational behaviour in terms of chromatograms, fluxes and other readily obtainable and process relevant data rather than theoretical terms and coefficients. This has the advantage of providing industrially relevant results but lacks the capacity to provide any understanding of why the differences in purification behaviour between process scale and very small scale occur. Either way ultra scale-down methods are characterised by consisting of a methodology, as in how to use the device to obtain data and how to process this data, and potentially a mathematical model to bridge between the ultra scale-down data and process scale data.

The use of such a methodology and model is to accelerate process development in the early stages of process design, when feedstock is scarce and resources very finite. An understanding of how the purification will behave in certain unit operations, both upstream and downstream, will save considerable time and resources when it comes to process optimisation at pilot scale.

5.2.2 Applications - Quality by Design

Quality by design is a regulatory approved concept for pharmaceutical product and process development quality systems, as laid out in ICH Guidelines Q8R2-Q10. The concept is that a process, and to a lesser extent the pharmaceutical product itself, is designed with full understanding of its critical quality attributes (CQAs) and how these change with respect to changes in process parameters (Rathore & Winkle, 2009; Barker et al, 1994). Critical quality attributes are attributes which will directly impact upon product or process quality – in the case of a chromatography step this could be purification factor or impurity profile, or in an upstream process it could be product titre. Each unit operation however can have potentially dozens of CQAs depending on how complex the process is and how liberally the classification of CQA is applied to process parameters (Clausing & Simpson, 1990). This is in contrast to the conventional "quality by inspection" where process and product knowledge can be fairly minimal, and the product is released on the basis of its end process analysis and limited data from in-process controls.

The key limitation in applying quality by design is the sheer depth of process and product knowledge that is required to be able to justify "full understanding" of a process. Traditional experimental design is uni-variate in nature, with one variable or parameter being changed at a time and the results recorded. To understand fully a section of a biological production process (quality by design is nearly always applied in a modular fashion to each unit operation or even part of a unit operation) the analysis has to be multivariate in nature, with potentially very many separate yet inter-dependent parameters being studied. Computerised systems working with design of experiments principles can effectively reduce the time, effort and resources required to collect an adequate data set but at full scale this still requires a very large amount of feedstock and time. The eventual goal is to provide a “design space” for each CQA, an experimentally mapped area in which full understanding of how process parameters modify the attribute is gained. Design spaces have defined boundaries at which process understanding is judged to be limited and as long as the process parameters stay within those boundaries the CQA parameter can be predicted in a reproducible fashion.

An ultra scale-down methodology is therefore critical to applying an effective quality by design strategy, as under normal circumstances the extensive data sets required will be infeasible to achieve at large scales. Any ultra scale-down method will have to be extensively validated and justified to the regulatory authorities before its data sets can be used to qualify a process, but in principle the regulators are supportive of quality by design concepts. Currently any major changes to a manufacturing process or the product requires a variation submission to the regulatory authorities detailing the change and validation studies demonstrating that no change to product safety or efficacy has occurred. A fully validated design space will allow process changes within the space to be performed without regulatory approval, although changes outside the design space boundary will still be required to submit to the variation process.

5.2.3 Developed Ultra Scale-Down Methods

Ultra scale-down (USD) methods have been successfully implemented before for a range of unit operations. Typically, due to the potential complexity of utilising non-geometrically similar devices to scale down to such small volumes, these are limited to a single unit operation with a fixed feed. USD has been demonstrated well in centrifugation steps, particularly centrifugation of precipitates (Boulding et al, 2002; Boychyn et al, 2004), expanded bed chromatography (Fenneteau et al, 2003; Willoughby et al, 2004; Ghose & Chase 2000), ultra-filtration (Reynolds et al, 2005; Ma et al, 2009), micro-filtration (Ma et al, 2009) and conventional ion-exchange chromatography (Kaltenbrunner et al, 1997). The use of computer models for decision analysis during bio-process design has also been demonstrated, particularly with industrially relevant chromatographic separations such as those for antibodies (Chhatre et al, 2007; Chhatre et al, 2008). It has also been recognised that unit operation integration of ultra-scale down methods is critical for the ultimate aim of complete process prediction at very small scales and as such integrated ultra-scale down methods has been reported (Neal et al, 2003). As described in section 3.2.4 microwell based chromatography scale-down methods have also had some measure of success (Coffman et al, 2008; Rege et al, 2006) as have the emerging microfluidic technologies (Shapiro et al, 2009).

5.2.4 Modelling of Chromatographic Behaviour

There are many theoretical models for chromatography in existence, but the practical application is typically limited to simple separations, which are of limited industrial relevance. The theoretical basis of chromatographic separation, as demonstrated by the Van Deemter equation and general rate model, is discussed in sections 1.6.2 and 1.6.3.

The output of a chromatographic process is not a series of coefficients however, and a method of chromatographic peak analysis is required to interpret experimental and process data derived from chromatograms. Over

90 different empirical functions exist for the representation of chromatographic peaks (Di Marco et al, 2001), mostly with the objective of mathematically resolving individual peaks from a complex chromatogram. This process is referred to as deconvolution.

5.2.4.1 Deconvolution functions

Functions with the capability to resolve individual peaks from a complex, semi-resolved chromatogram are very important since they potentially enable the protein composition and yield at each point in the elution profile to be determined. These deconvolution functions have often been taken from other analytical methods, including mass spectrometry or other spectroscopic techniques, and then applied to chromatograms derived from packed bed elution chromatography. Another source of applicable functions is located within the field of statistics, as the similarities between chromatographic peaks and statistical peaks can be exploited. Most of the mathematical functions are used for deconvolution of poorly resolved peaks, in essence breaking a complex chromatogram into its constituent peaks. The three most common functions in use include the exponentially modified Gaussian function (EMG) (Naish et al, 1988; Foley et al, 1984; Foley et al, 1987), the Poisson function (Grimalt et al, 1982) and the log normal function (Olive et al, 1991). Each of these commonly used functions have their own merits. There is some debate with regard to the actual identity of the Poisson equation, as the name is attributed to different equations by different authors (Di Marco et al, 2001). The Log-Normal and EMG functions both give good results, but the EMG is considered to give a more consistent fit (Naish et al, 1988). The key disadvantage to the EMG equation is that the skew value must be >0, but as fronted peaks are not typically seen in elution chromatography this is not considered relevant in a chromatographic context. The skew value in the case of these equations is not equivalent to asymmetry but a value calculated using the third moment of the mean deviation of the peak data, as is typical in statistical analysis.

5.2.4.2 Chromatographic behaviour models

Few models describing how chromatograms change with respect to process conditions exist for complex and realistic feedstocks. Most are highly specific in nature, either relating to simple separations (such as affinity chromatography) or to particular feedstock/resin systems. A more comprehensive chromatography model (as in not just restricted to chromatogram deconvolution) has been investigated by identifying statistical moments within the chromatographic data that represent process relevant data, known collectively as chromatographic figures of merit (Papas et al, 1989). Some success has been demonstrated within the biochemical field using principal component analysis methods, on HPLC data (Pate et al, 1999), for database mining (Chandwani et al, 1997) and for the analysis of small scale data sets (Edwards-Parton et al, 2008). However, mostly the models are restricted to analytical chemistry applications which are relatively simplistic when compared to biological applications.

5.2.5 The Challenges of Application of Chromatographic Models

There is currently a gap in knowledge between the theoretical models, utilised successfully on simple separations that are mostly synthetic, and the complex biological separations so common in the bio-process industries which are often optimised and scaled up or down using experience and experimental studies. It has to be accepted that producing a theoretical model that engages first principles that can accurately describe chromatographic behaviour when applied to highly complex biological feeds is not currently possible. Therefore the challenge is to be innovative in creating ways of bridging the gap between the theoretical descriptions and the purely practical experimental systems, utilising a methodology and model that works using chromatographic data to provide a predictive basis for scale up. Industrial process engineers are not interested in models that describe eddy diffusion or resistance to mass transfer, only in how this translates to differences in purity and yield of a valuable target molecule separated from a realistic feedstock – with this in mind any model needs only describe the chromatogram itself, not the underlying physics of mass transfer. For this reason the challenge has to be approached with practicality firmly in mind.

5.2.6 Aims of This Chapter

- To formulate a methodology for data collection at small and large scales.
- To mathematically deconvolute the chromatographic data into data sets that are amenable to mathematical modelling.
- To develop a model that accurate predicts the changes in chromatogram between very small (1mL) scale to laboratory scale (20-60mL) systems. This acts as a basis for greater scale up using conventional scale up rules.
- To develop an experimental methodology that uses the model to generate design spaces for effective scale up in a way that would be useful in bio-process design and optimisation.

5.3 Materials and Methods

Unless otherwise stated all reagents and chemicals were sourced from Sigma-Aldrich (Dorset, UK).

5.3.1 FAb lysate feedstock production - Fermentation

The full method for the FAb lysate feedstock production may be found in Section 2.3.2. The single fermentation batch from which the feedstock for the study in this chapter is derived was referred to as “Batch 2”.

5.3.2 FAb lysate feedstock production – Periplasmic lysis

The full method for the liberation of the FAb fragments from the periplasm may be found in section 2.3.3.

5.3.3 Chromatographic separations – PRC pre-packed columns

The chromatography columns used were Pall Life Sciences Pre-Packed 1mL PRC columns packed with MEP HyperCel, PPA HyperCel or CM Ceramic HyperD sorbents (Pall Life Sciences, Portsmouth, UK). All experiments were performed on an ÄKTA Basic chromatography system (GE Healthcare, Amersham, UK) utilising a absorbance 280nm UV flow cell.

Each pre-packed column was equilibrated an equilibration buffer consisting of 3CV 50mM Tris-HCl, pH 8.8. The pH of the FAb fragment feed was adjusted with 0.5M NaOH to pH 8.8 and then 1 CV was loaded onto the column. The column was then washed with 5CV of 50mM Tris-HCl equilibration buffer. Elution was performed via a linear pH gradient over 10CV, utilising an elution buffer consisting of 50mM sodium citrate, pH 3. 1mL fractions were taken throughout the elution gradient for the purposes of analysis. Cleaning was performed with 5CV NaOH and the columns were stored in 20% ethanol.

The flow rate of loading and elution depended on the experimental variable being examined and ranged between 50cm/h and 300cm/h. Apparent bed height was increased by linking up to 3 PRC columns together in series. All columns were loaded with 0.55mg FAb fragments per mL of sorbent.

5.3.4 Packing – XK16 Columns

The XK16/20 and XK16/40 chromatography columns (GE Healthcare, Amersham, UK) were packed with the sorbent according to the manufacturer's instructions supplied with the sorbent. XK16 columns have a

1.6cm diameter. Packing buffer was 50mM Tris-HCl pH 8.8. Packing tolerances were 2000 plates per metre +/- 10% for the HyperCel sorbents with an asymmetry value < 2, as determined by acetone pulse.

5.3.5 Chromatographic separations – XK16

Separations were performed as described in section 5.3.3. Bed height was between 5cm-30cm depending on the experimental variables being studied, and was defined as the height of the bed when fully packed ± 1cm.

5.3.6 Matlab Analysis

Analysis of the UV chromatograms was performed using Matlab version 7.1 (The Mathworks, USA) utilising the Curve Fitting and Optimisation toolboxes. Code for the algorithms for chromatogram/peak deconvolution was developed and provided by S. Edwards-Parton (University College London, London, UK) and is detailed in Appendix 1. Full details of the Matlab analysis are given in section 5.4.

5.4 Experimental Premise

This section details the hypothesis of how large scale chromatographic behaviour will be predicted by data obtained from an ultra scale-down device. This is followed by how the analysis will be used to investigate this hypothesis utilising the experimental data.

5.4.1 Objective

The objective of the methodology developed in this chapter was to investigate how chromatographic peaks change with respect to linear velocity and packed bed height for non-geometrically similar columns. This will be investigated at PRC Pre-Packed column scale, henceforth referred to as small scale (1-3mL internal column volume) and at XK16/40 scale, henceforth referred to as laboratory scale (10-60mL internal column volume). This immediate experimental target was part of a larger objective, which was to design, test and validate an experimental and analytical methodology that allows the prediction of large scale chromatographic behaviour using small scale chromatographic data and very limited validation data obtained at larger chromatographic scales. The two parameters selected as variables for consideration were packed bed height (cm) and linear velocity of loading/elution (cm/h).

5.4.2 Analytical Method Concept

The uses of mathematical descriptions of peaks, whether chromatographic or statistical, are detailed in section 5.2.4. At the heart of the method is the assumption that a single peak can be broken down into four key parameters or curve coefficients. These are detailed below and shown in Figure 5.1:

Maximum peak height (*h*): Peak height at maximum, in units of milli-absorption units (mAU).

Width at half height (*w*) : Peak width at *h*/2, in units of elution volume (mL).

Peak skew (*s*) : The peak skew value as derived from the peak height width and location.

Peak location (*z*): The location of the peak maximum height (*h*) on the x-axis, expressed in units of elution volume (mL).

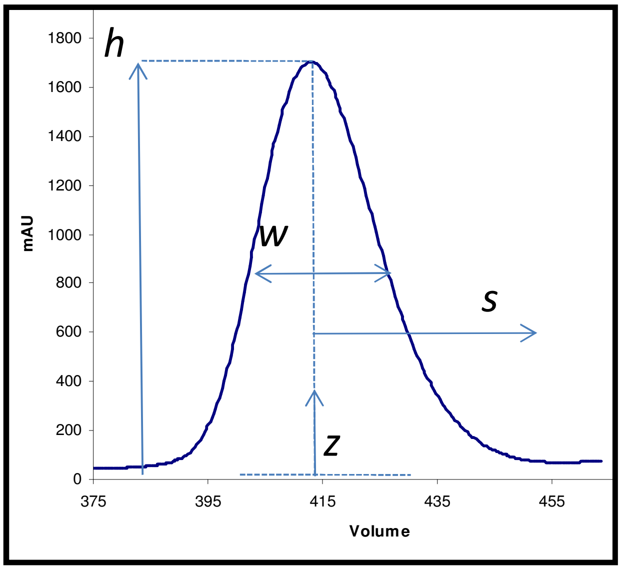


Figure 5.1: A graphical representation of the curve coefficients. h: peak height, w: peak width at half height, s: peak skew, z: peak location.

Utilising only these four curve coefficients a chromatographic peak can be produced using an empirical function, of which several exist. The function selected for use within this study was the Exponentially Modified Gaussian (EMG) due to its relative simplicity and proven good fit to experimentally derived peaks (Di Marco et al, 2001). The EMG equation is shown below in equation 5.1,

$$y = \frac{hw}{s}\sqrt{\frac{\pi}{2}}\exp\left(\frac{w^2}{2s^2} - \frac{x-z}{s}\right)\left\{1 - \operatorname{erf}\left[\frac{1}{\sqrt{2}}\frac{w}{s} - \frac{x-z}{w}\right]\right\} \quad (5.1)$$

where h is the peak height, s is the peak skew; w is the peak width at half height, z is the peak location and x is the elution volume in mL.

The only constraint of this function is that $s>0$. This constraint is however not an issue in this study as with the exception of overpacking scenarios individual chromatographic peaks are always trailing and as such s will always be positive.

It is very difficult to mimic accurately process scale chromatographic behaviour only utilising chromatographic data obtained at very small scales due to the large number of variables which affect the elution profile in different ways. The hypothesis to be investigated in this chapter is that it will be possible to find usable mathematical relationships between the individual curve coefficients (h, s, w, z) at the two scales of scrutiny. This concept is summarised in Figure 5.2.

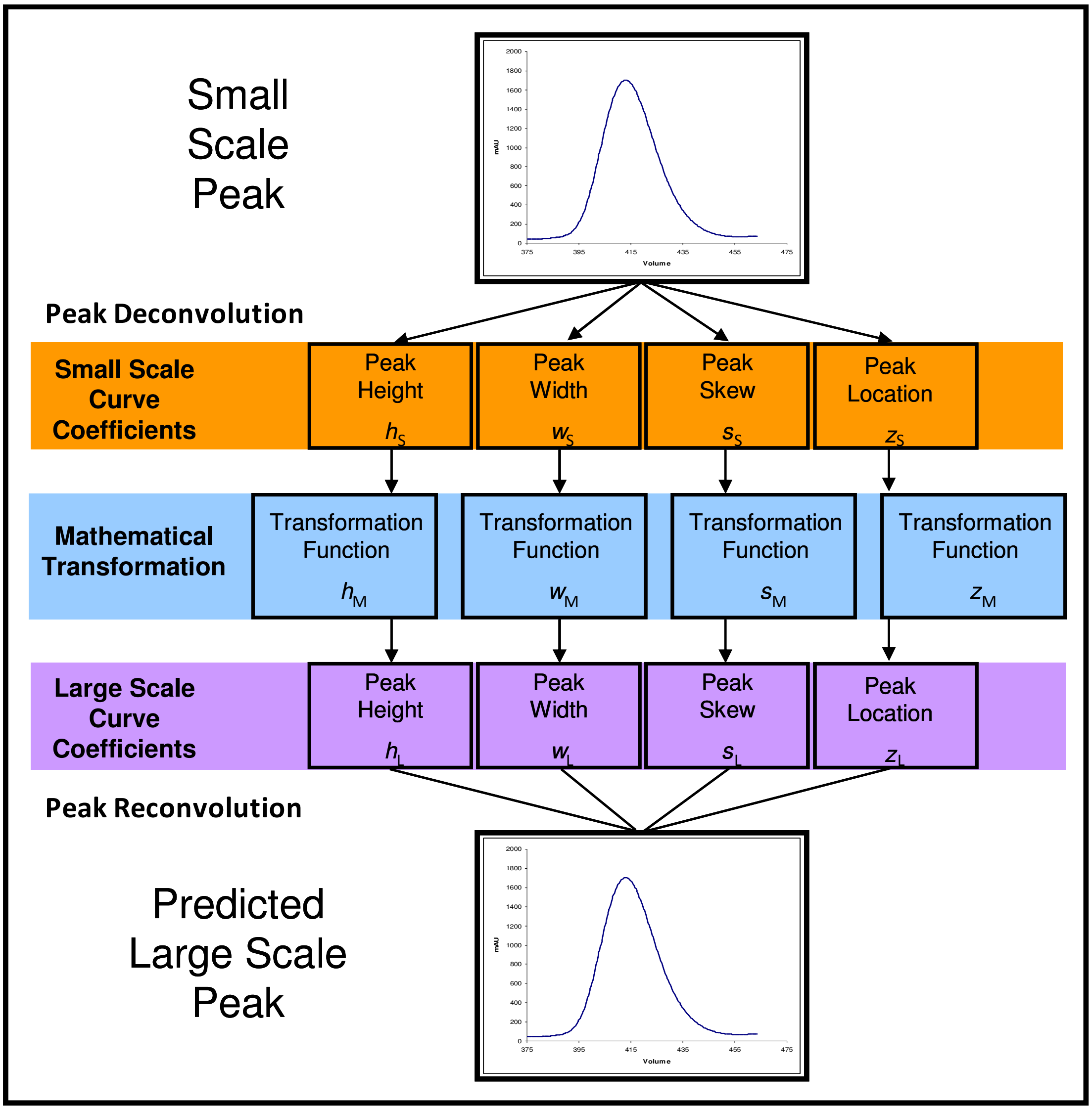


Figure 5.2: A flowchart summarising the proposed method for large scale chromatographic peak prediction using small scale data. Subscripts S, M and L denote the curve coefficients for the small scale peak, during transformation and for the large scale peak respectively.

The objective of the transformation function is to mathematically transform the small scale data until it accurately mimics the large scale data; this function will be critical to the success of the methodology. The transformation function will be a mathematical expression which describes the relationship between the small scale values (including the curve coefficient, the bed height and the linear velocity values) and the large scale equivalents. Once the small scale peak curve coefficients have been transformed into their large scale equivalents (the formation of which is described in section 5.5.8) the predictive large scale peak can be effectively rebuilt using the curve coefficients and the EMG equation.

5.4.3 Chromatogram Analysis and Algorithms

Single peak elution profiles with baseline resolution are unusual in elution chromatography except with the use of affinity sorbents or unrealistically slow elution gradients, particularly for primary capture unit operations. To address the multi-peak systems expected the Matlab curve fitting algorithms *EMGCurveFitLSQBig* and *PlotIt* (S. Edwards-Parton, University College London) were utilised, using the inbuilt Curve Fitting toolbox supplied with Matlab.

The EMGCurveFitLSQBig algorithm accepts input arguments from the user which specifies how many and the approximate location of the relevant peaks, at which point the sub-routine *Findnearest* will analyse the peak each side of the input and locate the exact highest point. The x value then becomes both *z* within the EMG equation and the start point for the rest of the algorithm for each peak. *EMGCurveFitLSQBig* then iteratively applies the *lsqcurvefit* function to the data, which deconvolutes the chromatogram into the respective number of peaks as inputted by the user. This curve fitting algorithm is a proprietary function of Matlab and works on the least squares principle; it is highly effective in practical terms but can be inexact with too few iterations or with poorly resolved peaks. The subroutine *EMG2* is then applied, which outputs the curve coefficients *h, w, s* and *z* for each peak.

The original algorithm was named *EMGCurveFitLSQ*. It was written for data generated on 1mL column volume Hi-Traps (GE Healthcare) and was found to be very inaccurate on data generated at laboratory scales above 10mL column volume. This was traced to the start point for peak width (*w),* which is a fixed value in the algorithm; within the number of iterations stated it was not reaching the width necessary for peaks generated on a much larger column. This value was increased by a magnitude of 10 and was found to be effective on both small scale and large scale data, so this modified code replaced the original algorithm and was renamed *EMGCurveFitLSQBig*.

The breakdown of a chromatogram by this algorithm is shown in Figure 5.3.

The *PlotIt* algorithm rebuilds a chromatogram from the curve coefficients using the EMG equation, as shown in Figure 5.4. Matlab code for these algorithms may be found in Appendix 1.

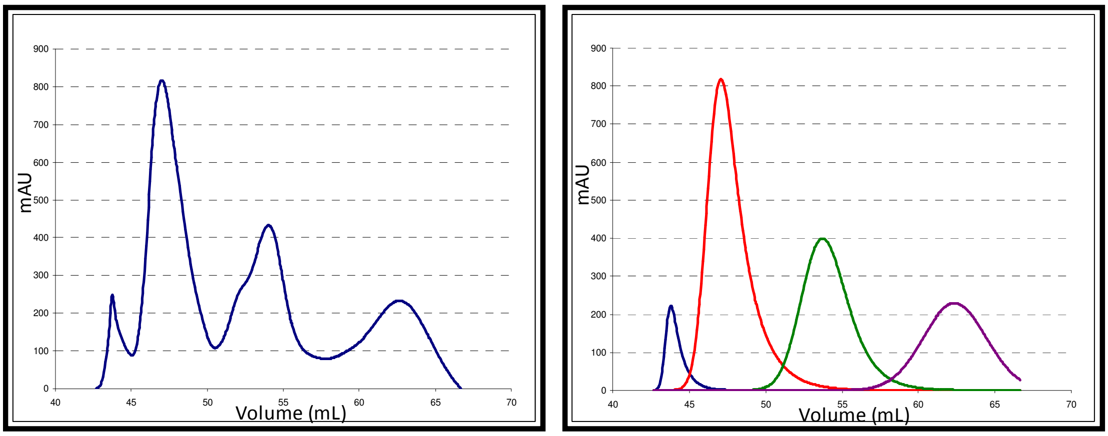


Figure 5.3: The left chromatogram is an elution profile of FAb fragments from a crude E. coli lysate separated on PPA HyperCel under a 10CV pH gradient (pH 8.8-2.2). The right chromatogram shows the individual peaks elucidated from this chromatogram by the EMGCurveFitLSQBig algorithm.

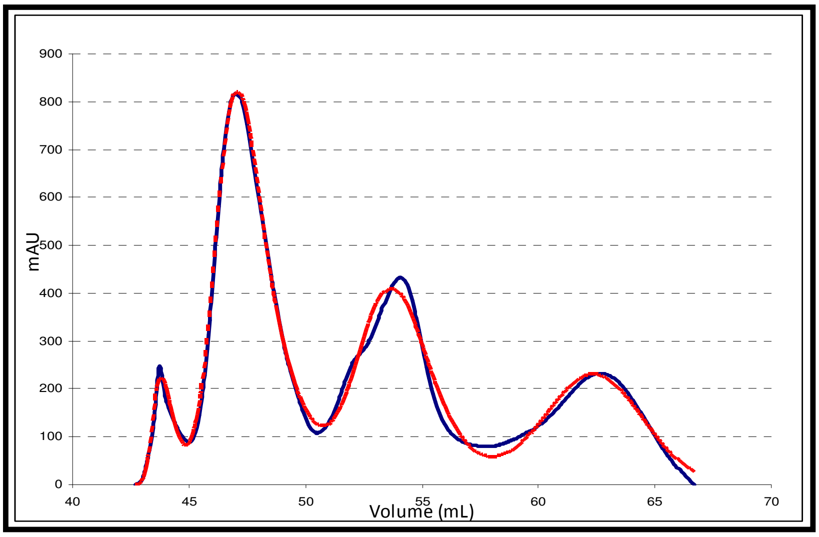


Figure 5.4: The • series in this chromatogram is the original experimentally derived elution profile. The • series is the pseudo-chromatogram produced by the 16 curve coefficients from the four separate peaks in Figure 5.3, produced using the PlotIt algorithm within Matlab.

5.4.4 Experimental Data Collection

The changes in the curve coefficients were investigated with regards to changing the height of the packed bed and linear velocity of loading/elution. Two data sets were required for comparison and contrast, one at small scale utilising PRC pre-packed columns and one at large or laboratory scale utilising XK16/40 columns packed with the appropriate resin.

For the small scale data set all data points within the design space were collected as minimal feedstock was required. For the large scale data set a factorial experimental design was employed to reduce the experimental duration and feedstock quantities required. The experimental grids identifying the experimental parameters are detailed in Figure 5.5.

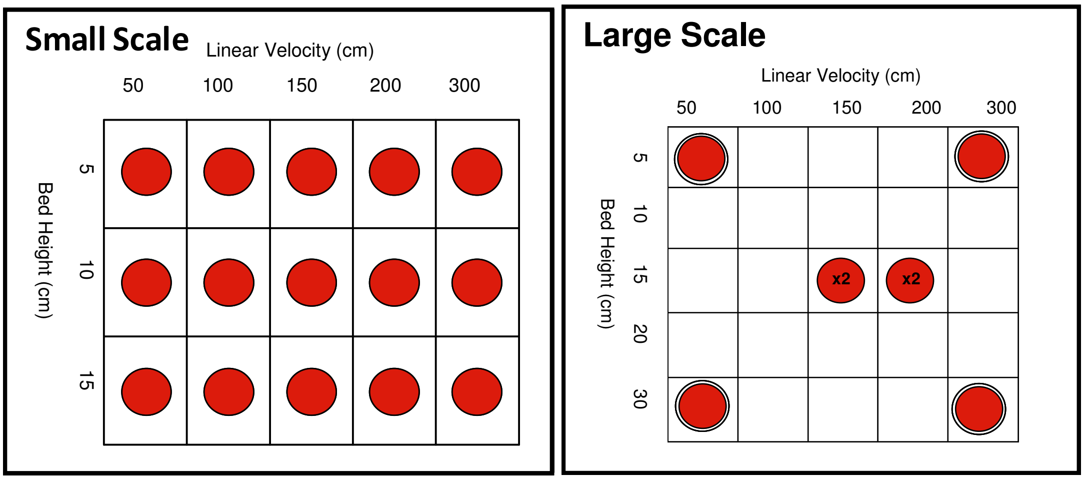


Figure 5.5: The experiments that were required to be performed to produce the chromatographic data sets for the variables bed height and linear velocity. The UV_{280nm} chromatogram was recorded for each set of conditions marked above.

The factorial design concept requires that the four corners of the design space be investigated together with a point central to the design space, which is typically repeated to ascertain a measure of experimental reproducibility. Having two separate points in the centre (as shown in Figure 5.5) is not strictly necessary; however, sufficient quantities of feedstock were available in this study and additional experimental data increases confidence in the accuracy of the trends seen within the design space.

Each chromatogram generated by the experiments proposed in Figure 5.5 is analysed separately as detailed in section 5.4.3. It is important to note that the experimental plan outlined in Figure 5.5 is designed to generate detailed data for the development of the ultra scale-down methodology, rather than being a proposed experimental method for the final methodology. The final methodology will only require a maximum of three experiments at large scale, to validate the small scale data set and to ensure the predicted large scale results are both accurate and consistent.

5.5 Results and Discussion

In this section chromatograms obtained from small scale and large scale experiments were analysed, as outlined in section 5.4 above. This includes the chromatograms themselves, the deconvolution of the peaks involved and the production of response surfaces to facilitate graphical analysis. The curve coefficients derived from this data were then correlated and transformation functions describing the relationship between the curve coefficients at large and small scale were defined.

5.5.1 Chromatograms

The chromatograms representing the edges of the design space for both the small and large scales are shown in Figure 5.6, the design space edges referring to the shortest bed height/fastest flow rate to the tallest bed height/slowest flow rate. The data obtained from the extreme edges of the design space will show the best chromatographic resolution expected and the worst, with all non-presented data falling between these two boundaries. As expected, increasing the flow rate and reducing the bed height had a detrimental effect to chromatographic resolution. Flow rate had relatively more impact than bed height. The two peaks seen in this chromatographic separation are labelled 1 and 2 in section D of Figure 5.6 and will be referred to as such throughout this chapter.

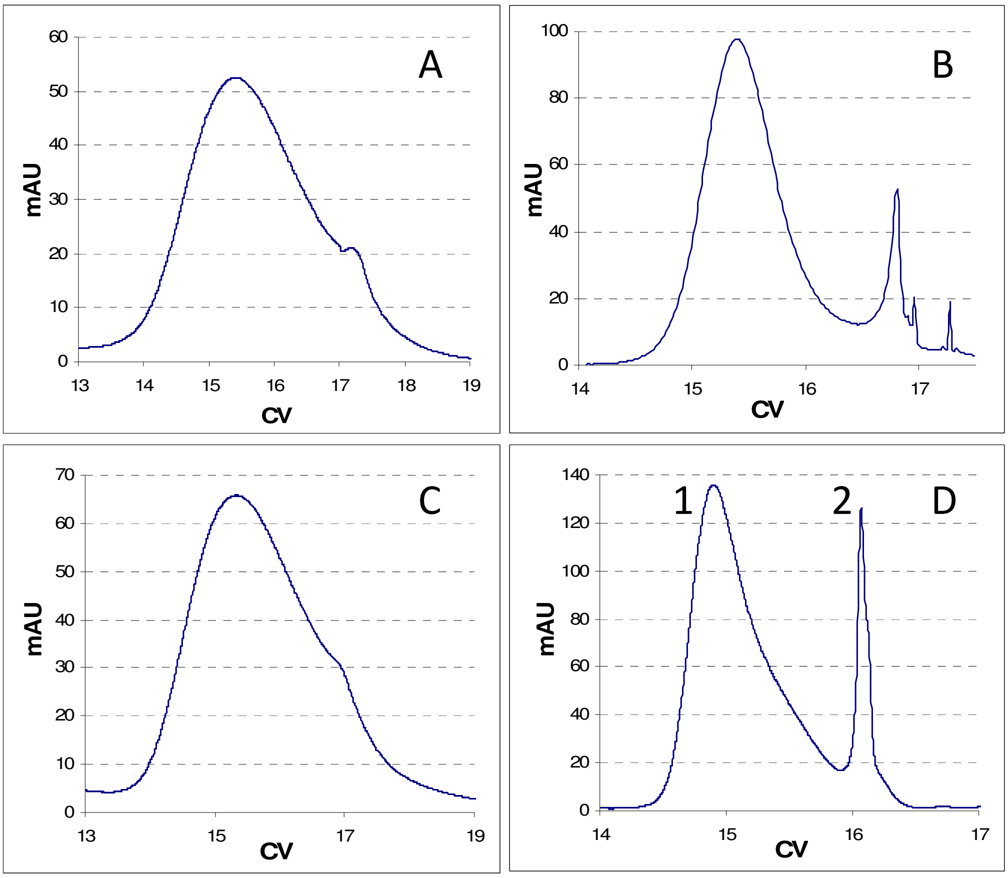


Figure 5.6: Four chromatograms demonstrating how the complete chromatogram changes between the edges of the factorial design space for both small scale and large scale systems. The chromatograms are an elution profile from the separation of a periplasmic E. coli lysate containing FAb fragments using the HCIC sorbent MEP HyperCel. A: Small scale 5cm bed height and 300cm/h linear velocity B: Small scale 15cm bed height and 50cm/h linear velocity, C: Large scale 5cm bed height and 300cm/h linear velocity, D: 30cm bed height and 50cm/h linear velocity. The stated linear velocity values refer to the linear velocity of both the loading and elution steps.

5.5.2 Chromatogram Analysis

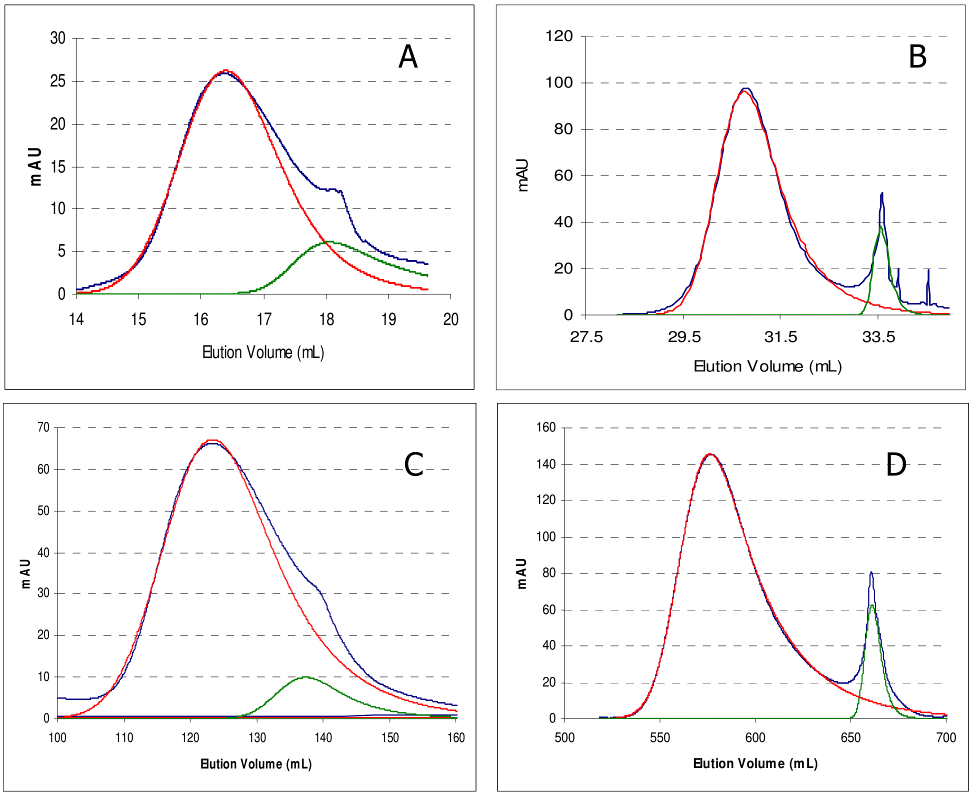


Figure 5.7: The four chromatograms shown in Figure 5.6 after processing through the EMGCurveFitLSQBig algorithm to produce a prediction of the individual peaks within the chromatogram. 🟦 series: original chromatogram, 🟥 series: peak one, 🟩 series: peak two.

Figure 5.7 demonstrates the effectiveness of the Matlab algorithm *EMGCurveFitLSQBig* in elucidating the individual peaks. There are several assumptions that are necessary to use this data reliably. It has to be assumed that there are no completely concealed peaks underneath other peaks, in that each peak consists of only a single species and if more than one species is present its peak can be located and inputted into the algorithm. The other assumption is that the operator can reliably and accurately input the peak location as an argument into the algorithm, although the curve fitting algorithm is robust enough to search for the peak maximum if the inputted location is in close proximity. Decisions also have to be made about what constitutes a peak, particularly with the smaller scale experiments in which the chromatograms can contain significant artefacts, though this can be solved by using the highest resolution chromatogram as a reference and always specifying the same number of peaks. This particular chromatographic separation system using MEP HyperCel and FAb fragment containing lysate has been well characterised in Chapter 4 and therefore two peaks were always specified for this study.

The curve fitting algorithm *EMGCurveFitLSQBig* processes artefacts by ignoring small peaks or short volume spikes such as those produced by small air bubbles or UV detector glitches, unless the peak is directly specified in the algorithm arguments. The algorithm outputs the series for each peak and also the peak curve coefficients *h,s,w,z*. The values for each experiment are shown in Table 5.1.

Scale	LV	BH	Peak One				Peak Two			
			h (mAU)	s	w (mL)	z (mL)	h (mAU)	s	w (mL)	z (mL)
Small	50	5	169.3	0.40	0.38	7.9	74.5	0.32	0.12	9.4
Small	100	5	109.6	0.57	0.44	8.1	25.7	0.20	0.12	9.9
Small	150	5	41.5	0.84	0.54	9.2	6.9	0.15	0.12	11.3
Small	200	5	50.7	1.07	0.51	8.5	8.2	0.20	0.12	10.4
Small	300	5	34.0	0.68	0.64	7.9	12.9	0.30	0.41	9.5
Small	50	10	146.4	0.75	0.46	14.3	51.8	0.15	0.12	17.4
Small	100	10	88.1	1.00	0.64	12.8	25.7	0.15	0.20	15.9
Small	150	10	99.6	1.13	0.75	12.8	25.2	0.20	0.17	16.4
Small	300	10	73.8	1.51	0.97	14.4	14.0	0.15	0.39	18.0
Small	50	15	126.1	0.86	0.6	25.8	28.8	0.15	0.25	30.6
Small	100	15	156.6	1.24	0.7	24.7	29.6	0.15	0.32	29.1
Small	150	15	88.4	1.43	0.9	21.0	19.0	0.15	0.27	26.0
Small	200	15	128.1	1.58	1.0	19.5	22.2	0.17	0.30	24.5
Small	300	15	66.7	1.88	1.2	21.9	15.0	0.15	0.30	27.3
Large	50	5	170.6	8.0	3.3	113.9	45.0	2.14	0.86	128.9
Large	300	5	97.1	8.8	6.2	118.0	15.7	6.17	3.49	134.0
Large	150	15	174.7	16.0	8.9	315.8	44.9	7.85	4.12	365.5
Large	200	15	199.9	18.0	9.3	279.9	28.1	1.03	5.21	334.8
Large	200	15	163.3	13.9	10.4	285.2	25.2	2.09	4.44	337.9
Large	50	30	315.8	27.2	7.5	553.9	157.0	2.63	1.73	632.9
Large	150	30	264.4	27.8	11.9	563.0	84.0	3.77	3.16	658.7
Large	300	30	149.7	42.8	19.5	599.7	49.7	4.92	5.98	695.1
Small-scale Error			8.9	0.057	0.014	0.39	3.8	0.035	0.057	0.35
Large-scale Error			15.9	2.8	0.778	3.8	2.0	0.75	0.54	2.1

Table 5.1: The curve coefficients generated for each peak, regarding the chromatograms gained at variable bed heights (cm) and linear velocities (cm/h). Error for each curve coefficient was calculated by triplicate experiments (shown in red) to 1 standard deviation.

5.5.3 Response Surface Production

To facilitate analysis of the data for each curve coefficient for both small and large scales the data was visualised using response surfaces. These response surfaces were graphical structures with the x-axis as bed height, the y-axis as linear velocity and the z-axis as the curve coefficient for either the small or large scale experimental data, be it height, skew, width or location, with the data interpolated across the surface. Representative examples of these response surfaces are shown in Figure 5.8. The complete sets can be found in Appendix 2.

In total there are 16 response surfaces for the MEP HyperCel/FAb fragment containing lysate data set, eight for the small scale and eight for the large scale.

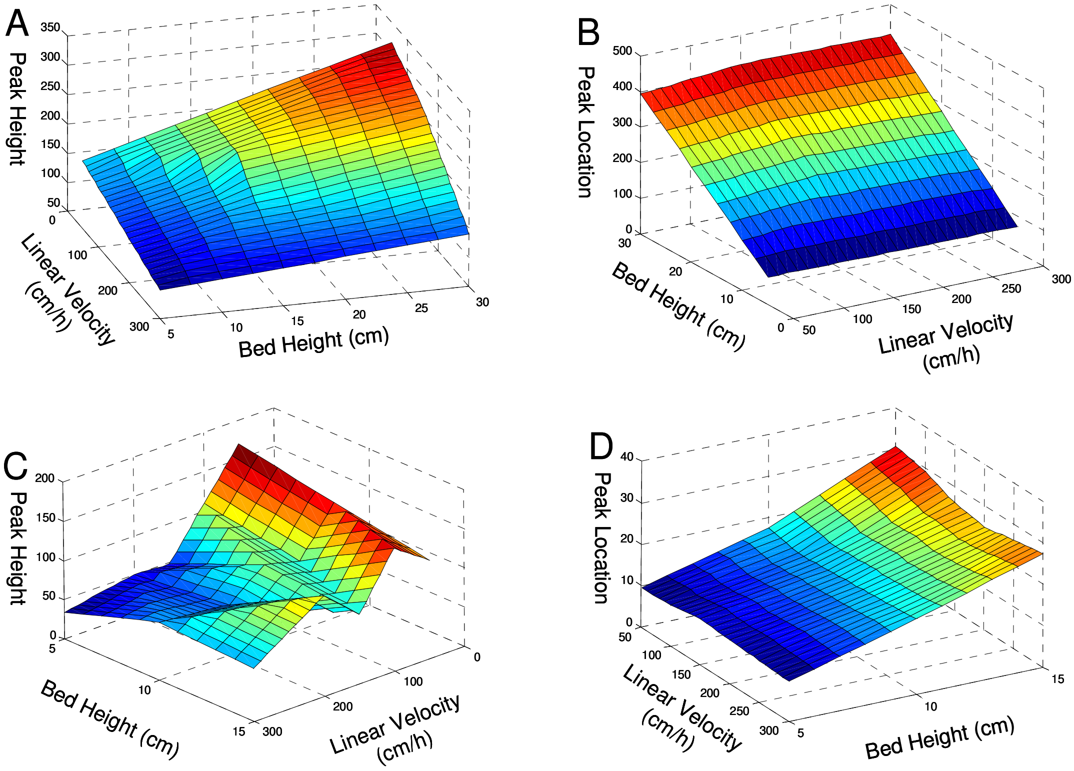


Figure 5.8: A selection of response surfaces to demonstrate how the curve coefficients changes with respect to bed height and linear velocity. A: Large scale peak one height, B: Large scale peak two location, C: Small scale peak one height, D: Small scale peak two location. All of the response surfaces for this study may be found in Appendix 2.

5.5.4 Peak Location

Peak location (*z*), as described by the EMG equation, is described in this work as the distance from the start of the sample loading to the maximum peak height, in terms of elution volume. Although this will very accurately describe a chromatographic peak the value itself is somewhat relative if the pre-elution stages of the chromatogram are included, such as loading and washing. For the purposes of the transformation function the peak location of peak one was taken as 0 (or equal to the large scale peak one location value) and the location of the second peak was calculated relative to this value. If the chromatogram included more than two peaks then each peak location would be described as its distance from the first peak. This was referred to as relative location in the transformation function calculations and was calculated back into an absolute peak location during the chromatogram reconstruction.

5.5.5 Response Surface Averaging

It was found that the response surfaces for the small scale curve coefficient data in particular had a high degree of error. For this reason all response surfaces generated were processed using a surface smoothing function that was designed in Excel; this averaged the results using 3 iterations of linear regression in 2 dimensions. This was performed by separating out the response surfaces into 9 separate series, presenting the 3D surface data as 9 individual 2D graphs. A linear regression was then performed on each and resulting data from the trend line incorporated into the response surface, with an average of the bed height and linear velocity value taken at each point on the surface. As seen in Figure 5.9 this renders experimental trends clearly visible on high error response surfaces while make very little difference to response surfaces that have clear, near linear slopes.

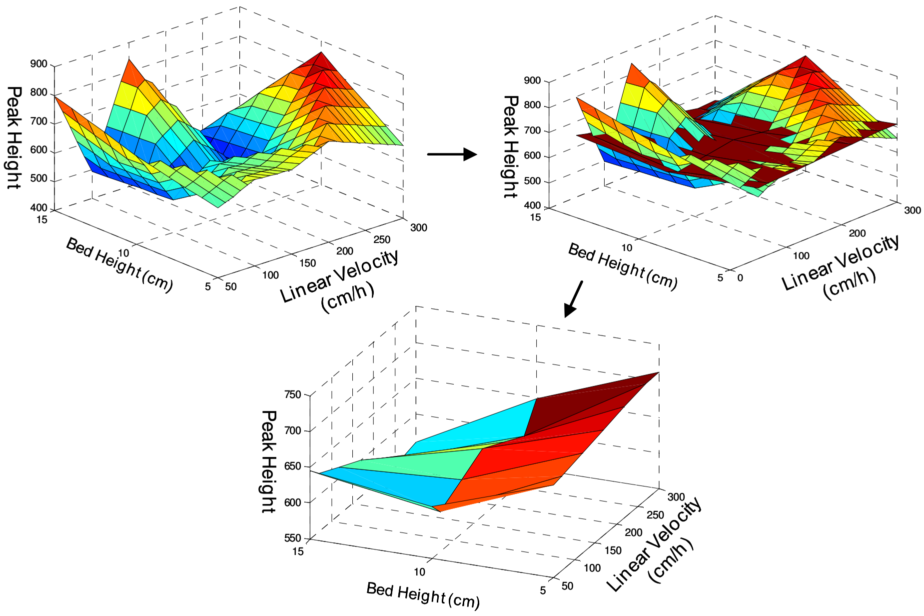


Figure 5.9: A series of response surfaces showing the original response surfaces, an intermediate with the averaged response surface on the same axes as the original and the final averaged response surface.

5.5.6 Response Surface Transformation – Axis Stretching

The objective of the transformation function is to mathematically transform the small scale response surface until it accurately mimics the large scale response surface. In the case of the majority of potential variables in a given process this can be done directly as the variables are in the same range; for example the range of linear velocities tested is 50cm/h to 300cm/h at both small and large scales and this would directly follow for all scale independent variables. However, in the experimental study detailed here the bed height is not scale independent and as such the small scale response surface will not be able to mimic the full large scale response surface without significant extrapolation.

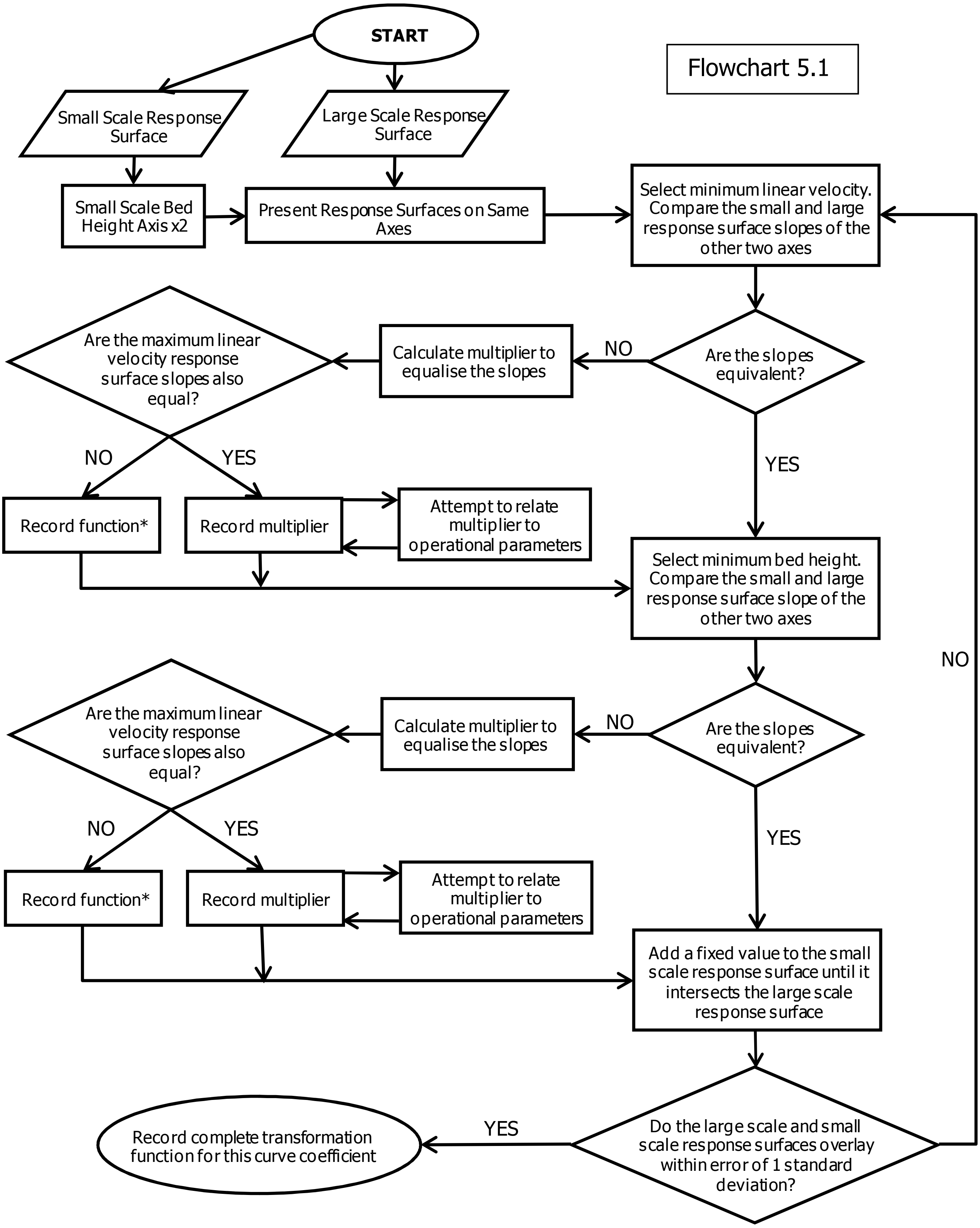
A solution to this was to include the extrapolation into the transformation function by elongating the bed height axis to match the large scale. This has the effect of making the surface slope shallower and then has to be corrected by the linear velocity and curve coefficient transformation functions. As the response surfaces for each curve coefficient are mostly linear this is very effective and allows direct prediction of the entire large scale design space. All small scale response surfaces had the bed height variable multiplied by a factor of 2 to achieve this.

5.5.7 Response Surface Transformation Function

It should be noted that the transformation functions produced within this thesis are still in development and were formulated as a way of testing the ultra scale-down methodology, to prove that a single transformation function could be utilised across the entire response surface. A great deal of time and resources would be required to fully investigate the correlations between the response surfaces at different scales and how these are potentially related to operational characteristics, which can then be included in any ultra scale-down model. The proposed work to fully investigate the transformation functions is outlined in Chapter 8, ”Future Work”. The method utilised for elucidating these basic functions is outlined in flowchart 5.1. The function* mentioned in the flowchart consists of relating operational parameters, such as the linear velocity or bed height, to ensure the response surface slopes at small and large scales are equal on all four sides.

As shown in Figure 5.10, the large and small scale response surfaces for a particular curve coefficient are very dissimilar. The transformation function development approach was to match the small and large scale surface slopes for bed height and linear velocity axes, then to multiply the entire small scale response surface upwards to match the large scale surface. The method for developing this is summarised in Figure 5.11.

Flowchart 5.1



Examples of the response surface fitting are demonstrated in Figure 5.11 and examples showing the fit within error are shown in Figure 5.12. Apparent fit is very good as seen by eye and all points for all response surfaces fall within the error of 1 standard deviation, as compared to the large scale response surface.

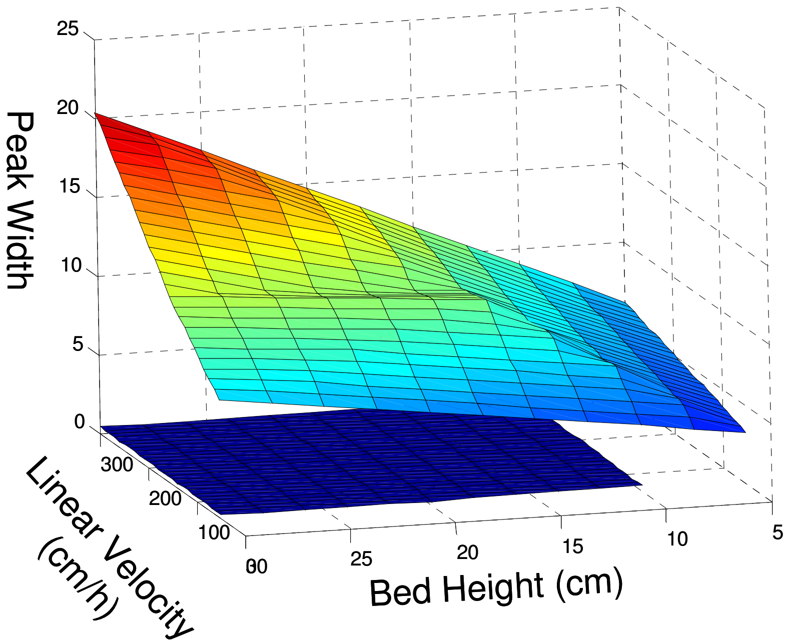


Figure 5.10: The response surfaces produced from the data for the peak width of peak one. The response surface on the bottom is the response surface for the small scale data, the one on the top is for the large scale data. The bed height axis has been stretched by a factor of 2 for the small scale response surface.

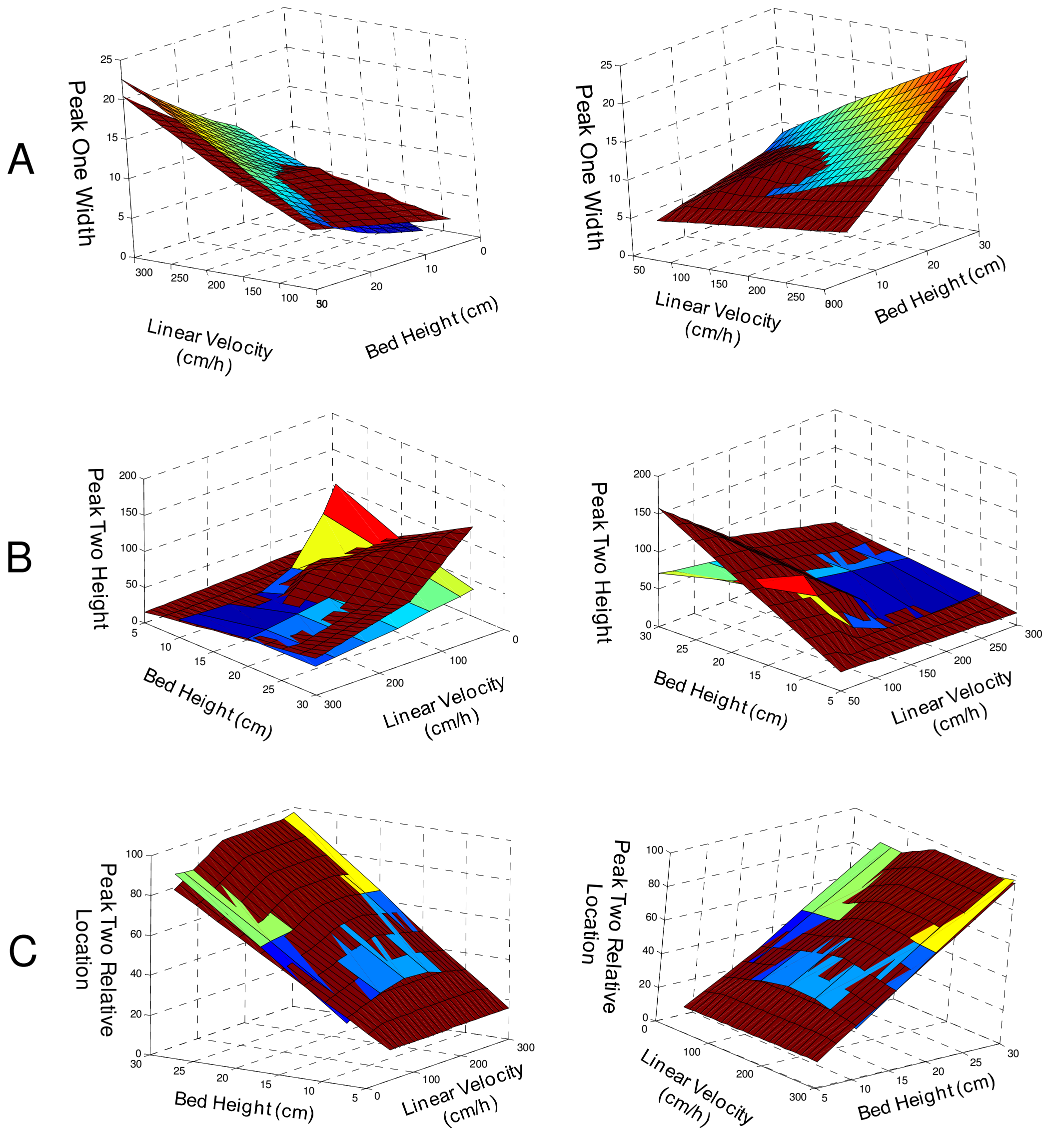


Figure 5.11: A selection of response surfaces showing the original response surface from the large scale experiments (maroon) overlaid the transformed small scale response surface (jet). As it is difficult to see a three dimensional image in two dimensions two aspects of the same response surface are displayed. A: Peak One Width, B: Peak Two Height, C: Peak Two Relative Location.

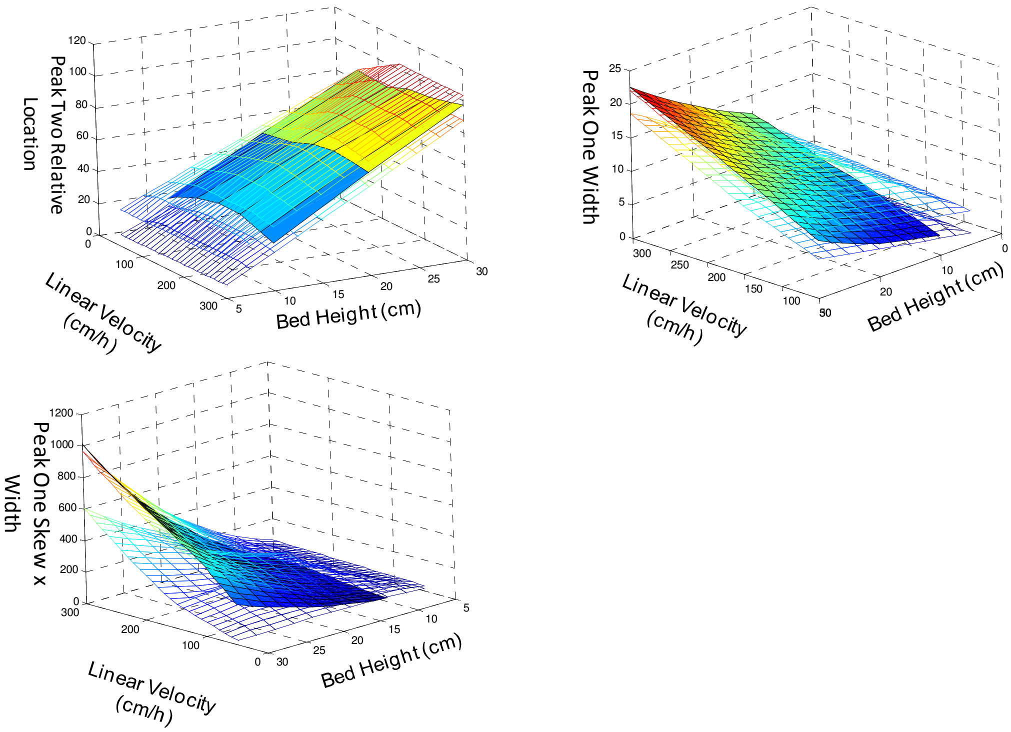


Figure 5.12: Three representative response surfaces showing the transformed small scale response surface (solid surface). The wire mesh surfaces above and below the small scale response surface represent the large scale surface to 1 standard deviation.

Figure 5.11 and Figure 5.12 demonstrate the good fit that is obtained with the small scale response surfaces once they have been transformed and for most surfaces all points are within 1 standard deviation of the large scale response surface standard, even for non-linear surfaces. It was found that the skew curve coefficient produced a more linear surface when multiplied by the peak width. As the skew curve coefficient is partially based on the peak width coefficient this calculation simply normalises the data.

There is still an issue with rogue points derived from the small scale experiments. Due to the very small scale of the experiments errors produce disproportionate effects in the response surface, whether the error derives from equipment or operator error. The response surface averaging function reduces the effect on the overall trends visible within a response surface to an extent. Response surfaces however provide very clear visual cues that a particular experiment in the data set has produced rogue points and the experiments can be repeated as necessary. The data obtained from the large scale experiments, in contrast, was very consistent.

5.5.8 Transformation Function

The transformation functions for each curve coefficient for this particular experimental system are listed in Table 5.2. Effort was made to relate linear velocity and bed height parameters within these functions rather than using an empirical approach, to better understand how the function relates to the hydrodynamics and potentially reducing system specificity of the functions.

Curve Coefficient	Peak One	Peak Two
Height	$2h$	$2h$
Skew	$30s^2wb$	$2sf$
Width	$w(0.7b)+\frac{f}{30}$	$w\left(\frac{f}{20}\right)+1.9$
Location	$9z$	$(25z)-30$

Table 5.2: The details of the transformation function for each curve coefficient for the MEP HyperCel separation of FAb fragments from crude lysate. h: peak height, s: peak skew, w: peak width, b: bed height (cm), f: linear velocity (cm/h), z: peak location. For each of the transformation functions the bed height axis was multiplied by 2.

5.5.9 Chromatogram Comparison

It has been shown that the individual curve coefficients can be accurately modelled from small scale to large scale for a two peak system. To review this the transformed curve coefficients were added as arguments to the *PlotIt* Matlab algorithm, which produces a predicted chromatogram via the EMG equation. Results are shown for a representative sample of experimental chromatograms in Figure 5.13.

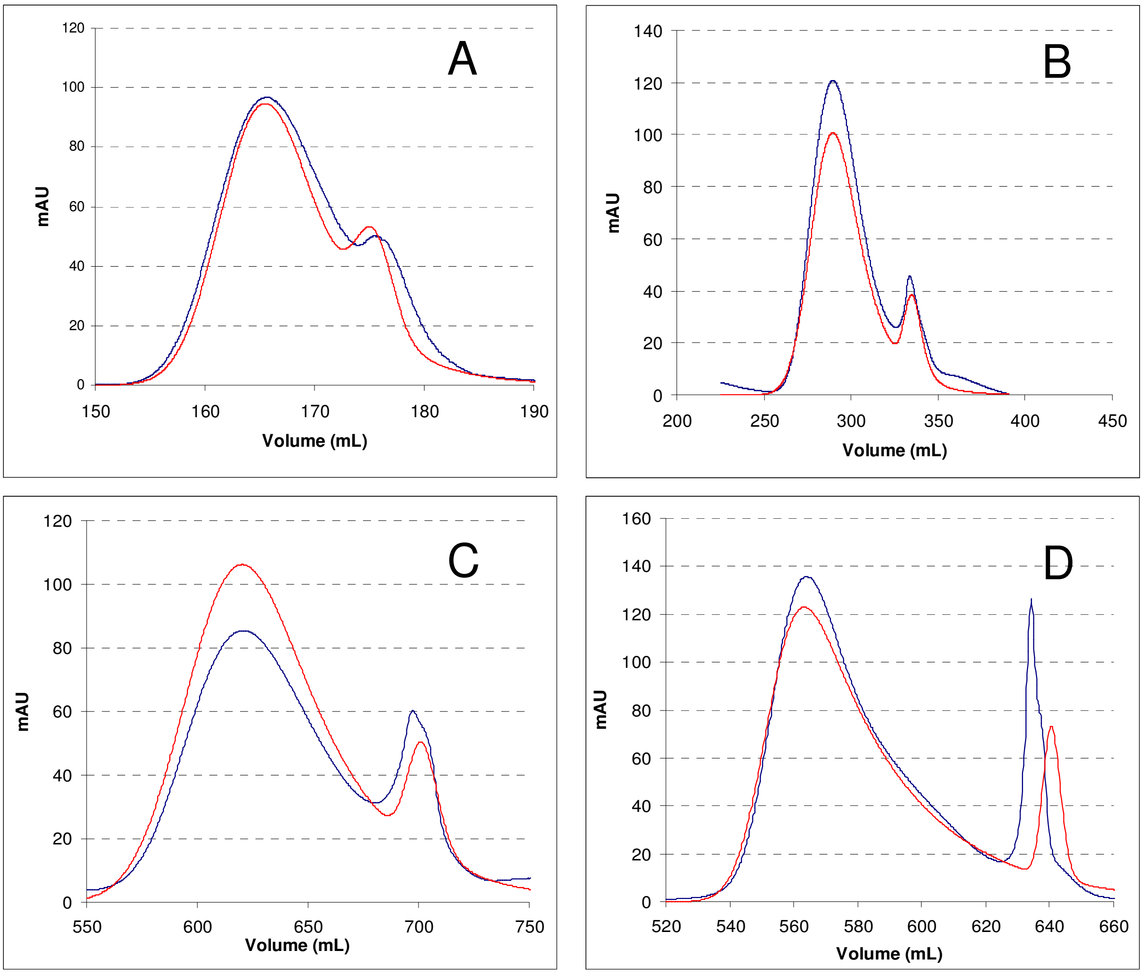


Figure 5.13: Four chromatograms comparing the chromatogram predicted by the modelled small scale data directly with the reference large scale chromatogram. Series 🟦: original large scale chromatogram, series 🟥: predicted chromatogram from transformed small scale data. A: 5cm bed height 50cm/h linear velocity, B: 15cm bed height 200cm/h linear velocity, C: 30cm bed height 300cm/h linear velocity, D: 30cm bed height 50cm/h linear velocity.

5.5.10 Normalised Root Mean Square

It is clear from the chromatograms in Figure 5.13 that, qualitatively, the predicted chromatogram is a very good fit. To quantify this, Table 5.3 shows the normalised root mean square values for the predicted chromatograms versus the reference large scale chromatograms.

Bed Height	**Linear Velocity**	**Normalised Root Mean Square**
5	50	3.5
5	300	11.1
15	150	3.8
15	200	3.8
30	50	3.0
30	300	4.0

Table 5.3: The normalised root mean square value for the compared chromatograms for each bed height and linear velocity parameter.

5% error was selected as a confidence limit to signify good fit, to provide a criterion of success for the chromatogram comparison. It is noticeable from that these values are very low, almost all beneath 5% error according to the root mean square. The exception is the fastest linear velocity and shortest bed height (5cm and 300cm/h) which has over 11% error. It was difficult to obtain a reproducible chromatogram at small scale with this bed height and linear velocity and this is demonstrated by a lack of a good predicted chromatogram fit. The experimental parameters in this case were testing the design space limits, with a very short bed and a very fast flow rate, so unpredictable chromatographic behaviour could be expected.

The normalised root mean square only relates to the entire chromatogram and averages the error thus. However, the same levels of inaccuracy in different parts of the chromatogram will have varying levels of impact upon the predicted purity and yield. A small level of inaccuracy at the point where the target peak intersects with the second peak will produce large inaccuracies in predicted purity values. A relatively larger level of inaccuracy at, for example, the start of the target peak or the trailing end of the second peak will have a correspondingly smaller effect on either purity or yield values. However, as a general method of determining the accuracy of the chromatogram prediction the normalised root mean square value is a good overview.

5.5.11 Purity and Yield

The values which a process engineer would find particularly valuable to obtain for a scale-down model are the purity and yield of a particular fraction cut. The yield/purity diagrams and original chromatogram are shown in Figure 5.14.

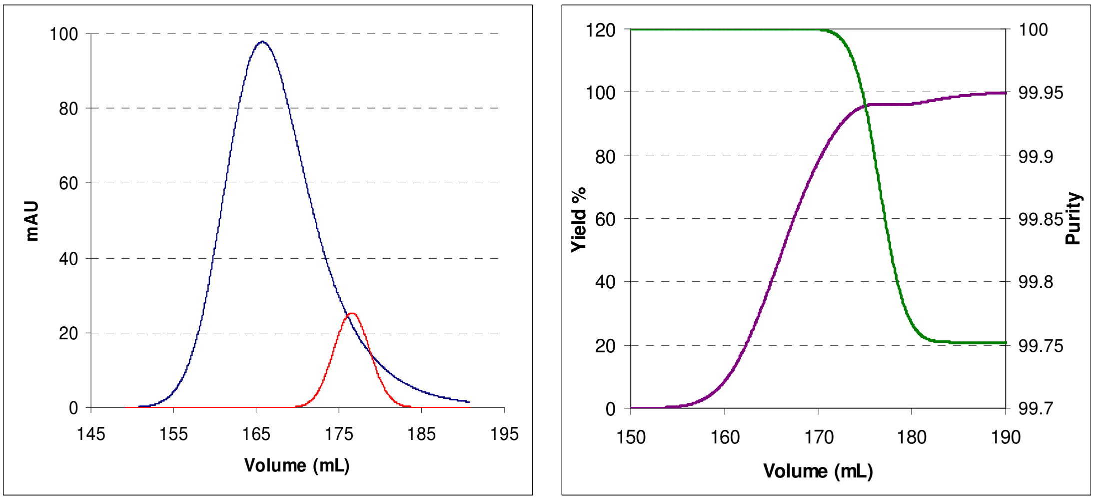


Figure 5.14: The left hand figure shows the large scale chromatogram for the parameters of 5cm bed height/50cm/h linear velocity, once the individual peaks have been elucidated by the algorithm EMGCurveFitLSQBig. The right hand figure shows the yield/purification diagram from this chromatogram. 🟦: peak one, 🟥: peak two, 🟩: purity, 🟪: yield.

For the experimental system examined here these values are complicated by the fact that even for the fastest linear velocities and shortest bed heights resolution was very good, leading to very high purities and high yields.

The data in table 5.4 shows a high level of accuracy in predicting expected purification factors using transformed small scale data, although differences are minimised by the high resolution of the separation.

Bed Height	Linear Velocity	Large Scale purification at 99.8% recovery	Predicted purification at 99.8% recovery
5	50	99.75	99.81
15	150	99.85	99.75
15	200	99.73	99.79
30	50	99.84	99.85
30	300	99.69	99.73

Table 5.4: A table showing the actual purification obtained for a set % recovery at large scale and the predicted purification factor generated from the transformed small scale data.

5.6 Conclusion

This chapter provides evidence that it is possible to scale mathematically chromatographic data obtained from very small, non-geometrically similar devices to laboratory scale chromatographic systems, even when separating complex and realistic feedstocks.

The set of transformation functions created as the basis of this method are not ideal however. The necessary empirical nature of some of the curve coefficient transformations means they are likely to be system specific. Ideally the transformation functions should relate only to system properties, such as bed height, linear velocity or column volume, but this did not prove possible. What this chapter does prove however is the validity of using the concept of scaling between very dissimilar scales by the deconvolution of chromatographic peaks into the curve coefficients and then using statistically derived equations to form an accurate transformation function.

A key weak point in this methodology is the accuracy of the curve fitting function *EMGCurveFitLSQBig*, as the curve coefficients it produces are critical to the transformation process. The fitting function has been shown to be very accurate in the system used within the chapter, but its capabilities when applied to a more complex chromatogram remains to be seen.

Analysis of the 5cm bed height experiments proved difficult. It is reasonable to postulate that constraints will apply to how large a design space the curve coefficient transformation methodology can apply to, a limit beyond which the chromatographic parameters become very difficult to predict. This would be accompanied by a combination of increased experimental error and hydrodynamic effects which only become apparent after a certain critical point. For example, the linear velocity parameter effect on chromatographic behaviour would be linear and produce an accurate model up until the critical velocity is reached and the bed compresses, adding a new factor to the equation. The edges of the design space would have to take this into account but also would be constrained by the feasibility in an industrial process; no industrial chromatography unit operation is going to run with a 5cm high bed or at linear velocities which compress the packed bed.

It would also be possible to extend this methodology to a range of other parameters involved with chromatography. Bed height and linear velocity of loading/elution were selected as they are hydrodynamic in nature and it was believed that they impact most on the chromatogram and facilitate analysis. The studies within this chapter utilised only two variables, bed height and linear velocity. However, the methodology developed could be equally applied to a more complex system via a factorial multivariate experimental model (applying DoE principles to multiple variables, rather than limiting to bed height and linear velocity). This would allow large number of parameters simultaneously including loading pH, packing quality, buffer conductivity and sorbent selection.

The ultra scale-down methodology presented within this chapter was developed for a single experimental system, using a single consistent feedstock on the same sorbent. To ensure the methodology was not system-specific the validation of the methodology needed to be performed, in which key parameters such as target protein concentration, chromatographic ligand, sorbent backbone and characteristics of the feed liquor were varied and the accuracy of the method validated. This validation study is detailed in Chapter 6.

6 Validation of Methodology and Model for the Ultra Scale-Down of Elution Chromatography

6.1 Abstract

This chapter concerns the validation of the ultra scale-down methodology by its application in 3 separate validation studies based upon the experimental methodology described in Chapter 5. These studies consisted of a change in feedstock target molecule titre, a change of chromatographic ligand and a multivariate change where the ligand, the feedstock and the sorbent backbone were all altered.

It was found that the ultra-scale down methodology, consisting of the collection of experimental data sets at both small and large scales then applying deconvolution algorithms to derive curve coefficients for mathematical analysis, was completely successful in all three validation studies.

The transformation functions that were developed to mathematically transform the small scale curve coefficients to their large scale equivalents were consistent and allowed chromatogram prediction with a high level of accuracy. However, it was found that these transformation functions were mostly unique to the experimental system except in the case of a reduction of target molecule titre, limiting the potential utility of the transformation function section of the methodology.

6.2 Introduction

6.2.1 Validation of Ultra Scale-Down Methodology

A functional methodology and model has been created for the prediction of large scale chromatographic behaviour using data generated by ultra scale-down devices. This was developed and tested using a single experimental system as detailed in Chapter 5.

This chapter seeks to validate the methodology and model by implementing the methodology and model developed previously using different experimental conditions selected to test the broad applicability of the approach. Three validation studies were performed:

- Validation study 1 - Reduction in product titre
- Validation study 2 - Change of chromatographic ligand
- Validation study 3 - Change of chromatographic ligand, sorbent backbone and feedstock

6.2.2 Process Optimisation

Production and purification processes for biotechnologically derived products go through changes and optimisations throughout a product lifecycle. These changes can occur for a multitude of reasons, such as regulatory requirement changes, improved process knowledge/optimisation and most commonly for scale up to produce sufficient quantities of material for late clinical development phases or for commercial supply. Due to the integrated nature of production and purification processes a change in the upstream or higher downstream conditions can have a cascading effect throughout the process, requiring re-evaluation and re-validation of subsequent unit operations with the new feed (Ladisch et al, 2001).

A key objective of upstream development is often stated as being to produce as high titres as possible of the desired product to keep costs of goods to a minimum. This can however have significant impact on the effectiveness of downstream operations; a raised product titre can lead to higher concentrations of host cell proteins (especially if the increased product titre was produced primarily via increased host cell biomass) and DNA/RNA contaminants. Even if the host cell protein and genetic material concentrations are kept constant a significant increase in feed viscosity will generally be observed, together with a higher incidence of aggregated products and insoluble product related impurities (Doran et al, 1995).

Downstream processes have to be robust in design and versatile in application to ensure good performance in the face of potential feed changes. Good application of Quality by Design (QbD), (Rathore et al, 2009; Clausing et al, 1990) principles are invaluable to this end as the extensive process knowledge, which is implicit in observing QbD principles, allows directed manipulation of process parameters to handle complex changes in process feeds whilst still maintaining product quality. Educated chromatography process decisions such as choice of ligand, sorbent backbone and operating conditions directly influence product critical quality attributes and as such are key to process effectiveness and economy.

6.2.3 PPA HyperCel

The phenylpropyl amine chromatography ligand (PPA) is considered, unlike the HCIC MEP ligand, to have a mechanism of action that is truly mixed mode. In the case of hydrophobic charged induction chromatography the binding mechanism is in part affinity and in part hydrophobic; the elution of the target protein, brought about by a drop in pH conditions, is accomplished by the charge induction of both the ligand and the bound protein. In the case of a true mixed mode ligand, such as PPA or hexyl-amine (HEA), binding is hydrophobic in nature and elution is typically brought about by a charge induction of the bound protein only. This occurs because the HEA and PPA molecular moieties have a PKa of approximately 8 and therefore distant from the elution pH, while the MEP moiety has a PKa of 4.8.

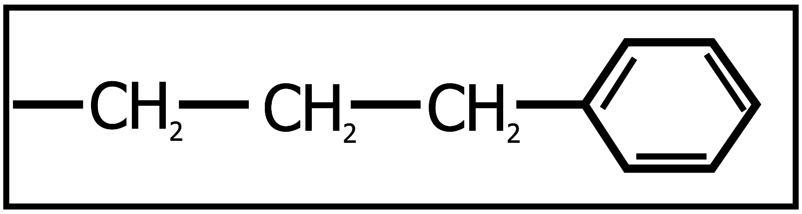


Figure 6.1: The PPA ligand

There is however significant cross-over in chromatographic mechanisms between HCIC and mixed-mode ligands and the exact nature of binding and elution can depend on process and feedstock conditions. All mixed mode sorbents (such as PPA and HEA) and HCIC sorbents have different affinities for different classes of proteins and so far the mechanism of binding has proven difficult to predict without experimentation (Brenac Brochier et al, 2008), although computerised structure-property relationship modelling techniques have had some success in this area (Yang et al, 2007). All mixed-mode and HCIC sorbents however can be characterised by a high binding affinity at physiological pH irrespective of salt conditions.

6.2.4 Cellulose Sorbent Backbone

Natural cellulose polysaccharides have been in use as a chromatographic structural backbone of sorbents for many years. Application was initially restricted as cellulose was only available in a micro-granular or fibrous form. Processes which allowed the formation of reproducible cellulose beads were introduced in the early 1970's and allowed cellulose to become established as packing material for chromatographic applications (Peska et al, 1976). Crystalline regions found within the structure of the beads produce a porous structure and by selecting for a macro-porous material high flow rates can be used without significant back pressure problems (Leonard et al, 1997). Modern cellulose backbones can withstand 2000cm/h without serious deformation or compression, unlike most modern agarose derivatives.

However, a consequence of this is that packing is complicated by the very high flow rates (around 1000cm/h) required to pack the beads adequately for high resolution elution chromatography and that the natural elasticity of cellulose can interfere with conventional flow packing processes.

The PPA and MEP sorbents used in this study each had a cellulose backbone marketed as HyperCel by Pall Life Sciences.

6.2.5 Ceramic Sorbent Backbone

The pressure drop developed in packed bed chromatography systems is due to a combination of different factors, including column dimensions, flow rate, feedstock viscosity, frit design and the physical properties of the sorbent.

A critical factor in the exhibited pressure drop is due to the backbone material of the sorbent and this is shown by the relative compressibility of different sorbents (Stickel et al, 2001). Most sorbents are inherently compressible due to the nature of the backbone material and the operational requirement for a porous cross linked structure with a large surface area and enhanced diffusivity characteristics. As the flow rate is increased a critical point will be reached where the sorbent suddenly compresses or crushes, leading to a substantial and rapid increase in pressure drop which is a particular issue with columns that utilise fast flow packing (Keener et al, 2004b) or mechanical packing methods (Keener et al, 2004a). For sorbents such as the original Sepharose this was rated at an operationally limiting 75cm/h, although the invention and subsequent marketing of the Sepharose Fast-Flow variants have significantly increased these values. The tendency for non-rigid sorbents to compress under the influence of fast flow rates can be countered somewhat by the support of the chromatography column walls which serve to stabilise the matrix structure of the beads within the column; however, industrial column designs have large diameters and sorbent compression is commonly seen when the beads packed into the centre of the column furthest from the walls compress and collapse. The column design and wall effects have significant impact on packing effectiveness and flow behaviour and have been studied extensively (Guiochon et al, 1997; Shalliker et al, 2000).

The immediate solution to the issue of packing material compressibility would be to utilise non-compressible materials as a sorbent backbone, such as a ceramic material. These have been shown to achieve a linear pressure/flow profile (Cortes et al, 1987). Incorporation of a macroporous structure into a ceramic backbone with the requisite diffusivity properties has however proved challenging, requiring an increased residence time/reduced linear velocity to obtain acceptable dynamic binding capacities and thus defeating the objective of a packing material with high compressible strength to withstand fast flow rates. Many ceramic materials have been tested, including mesoporous silica (Gallis et al, 1999) and titania and zirconia (Kawahara et al, 1989).

The Pall Life Sciences Ceramic HyperD range of sorbents utilise a giga-porous ceramic bead polymerised with a highly diffusive hydrogel which holds the ion exchange ligand. This design achieves the structural strength of a ceramic bead with the macroporous diffusive properties of a polysaccharide sorbent. With increasing product titres being generated by advances in fermentation technology feeds have become more viscous, which can apply further pressure drop to the chromatographic system. Feeds obtained by natural (rather than recombinant) means with high polysaccharide and lipid content are also viscous by nature, and this can include traditional feedstocks such as sweet whey and egg white.

6.2.6 Egg White Feedstocks

The protein composition of hen's egg white is shown in Table 6.1. Egg white, or albumin, consists of a variety of proteins (ranging in size from approximately 12.5 KDa for lysozyme to 77KDa for ovaltransferrin), saccharides and lipids. Many of the proteins found within egg white have industrial or research uses; lysozyme and ovalbumin are used as bacteriolytic enzymes and as a reference marker respectively in laboratories, while ovaltransferrin has uses in iron fortified food products.

Effective and efficient purification of these proteins from egg white has therefore been the subject of much research, particularly utilising ion-exchange chromatography (Ming et al, 1993; Li-Chan et al, 1986), affinity chromatography (Cuatrecasas et al, 1968; Awade et al, 1996) and precipitation (Tankrathok et al, 2009).

Protein	% Composition
Ovalbumin	54
Ovotransferrin	12
Ovomucoid	11
Ovaglobulins	8
Ovamucin	3.5
Lysozyme	3.4
Other proteins	8.1

Table 6.1: A table presenting the major proteins found in hen's egg albumin (Ternes et al, 2001).

Egg white, compared to other biological products, is relatively simple in its protein composition and very inexpensive, so has proven popular as a model feedstock for chromatography experimentation and modelling (Edwards-Parton et al, 2008; Levison et al, 1989). Egg white solutions are however relatively viscous due to the high protein and sugar content, and this may require dilution and pre-processing before use with packed bed chromatography. Dilution of feeds is easily performed at small scale, but this presents serious logistical difficulties at industrial scales where capital equipment and holding tanks are fixed and very expensive.

Despite the issues with viscosity egg white is an inexpensive feedstock that, although far more simple than a homogenate or lysate, is biologically complex enough to be relevant and contains a range of discrete and characterised proteins. This makes egg white an excellent model feedstock for chromatographic investigations.

6.3 Materials and Methods

Unless otherwise stated all reagents and chemicals were sourced from Sigma-Aldrich (Dorset, UK).

6.3.1 Validation Study 1

6.3.1.1 FAb Lysate Feedstock – Reduced Titre Study

The FAb containing *E. coli* lysate used within the study had a FAb fragment concentration of 0.27mg/mL. In all other respects materials and methods are as section 2.3.2. The single fermentation batch that produced the feedstock for this study in this chapter is referred to as "Batch 3", as described in Table 6.2 in section 6.5.1.

6.3.2 Validation Study 2

6.3.2.1 FAb lysate feedstock

The *E. coli* derived lysate containing FAb fragments used in this study was "Batch 4" as described in Table 6.2 in section 6.5.1. The method for its production is detailed in Section 2.3.2.

6.3.2.2 Chromatographic separations – PRC pre-packed columns MEP and PPA HyperCel

PPA HyperCel PRC pre-packed columns were obtained from Pall Life Sciences (Portsmouth, UK) and were packed and operated in an identical fashion to the MEP HyperCel PRC pre-packed columns used in Chapters 5 and 6. Full experimental details of the chromatographic separation are provided in section 5.3.3. In all other respects, except as explained in section 6.3.2.3, materials and methods were as described in section 5.3.3.

6.3.2.3 10cm Bed Height

A minor change to the experimental plan was implemented in this validation study. It was noticed and discussed in section 5.6 that it was very difficult to emulate the 5cm bed height results at laboratory (XK16) scale and that 5cm was outside of any feasible or reasonable design space for an industrial chromatography process. The small bed height axis was also stretched consistently by a factor of 2 during response surface transformation, restricting the design space it can model. For this reason the bed heights tested at large scale were 10cm, 20cm and 30cm rather than 5cm, 15cm and 30cm as seen in Chapter 5.

6.3.3 Validation Study 3

6.3.3.1 IgG4 Spiked Egg White Feedstock

Egg white was procured from standard large hen’s eggs (Sainsburys, London, UK), with the yolk manually separated. 100mL of the egg white was added to 900mL of 20mM sodium phosphate and gently swirled to mix. The feedstock was then spiked by the addition of bovine IgG4 (I5506, Sigma-Aldrich) to a final concentration of 0.1mg/mL. The feedstock was then centrifuged at 4000RPM in a Beckman JE-26 centrifuge for 30 minutes and then the supernatant was decanted into a clean Falcon tube. Storage was at -20°C and a single batch was used for all experiments within this study.

6.3.3.2 Chromatographic Separations – PRC pre-packed columns CM Ceramic HyperD

Each pre-packed column was equilibrated with an equilibration buffer consisting of 3CV 20mM pH 6 sodium phosphate. The pH of the FAb fragment feed was adjusted with 0.1M HCl to pH 6 and then 1 CV was loaded onto the column. The column was then washed with 5CV of 20mM sodium phosphate equilibration buffer. Elution was performed via a linear pH gradient over 10CV, utilising an elution buffer consisting of 20mM sodium phosphate pH6/1M NaCl. 1mL fractions were taken throughout the elution gradient for the purposes of analysis, as described in section 4.3.5. Cleaning was performed with 5CV NaOH and the columns were stored in 20% ethanol.

The flow rate of loading and elution depended on the experimental variable being examined and ranged between 50cm/h and 300cm/h. Apparent bed height was increased by linking up to 3 PRC columns together in series.

6.4 Validation Strategy

6.4.1 Validation Premise

The ultra scale-down methodology described in Chapter 5, together with the transformation function set required to translate the small scale curve coefficients into the large scale equivalents, has been shown to be effective (in section 5.5) for a chromatographic separation system incorporating the MEP HyperCel sorbent on a cellulose backbone separating FAb fragments from an *E. coli* periplasmic lysate under gradient conditions. However, a scale-up methodology that is entirely system-specific would be of very limited relevance to industrial chromatography process development therefore validation with other experimental systems is required.

These studies sought to validate two components of the methodology developed in this thesis:

- The ultra scale-down methodology: This study sought to validate the methodology of peak deconvolution via the EMG equation algorithm and resulting curve coefficient analysis over the bed height and linear velocity variables. To be considered successful this method will have to produce a predictable set of response surfaces for each curve coefficient and allow an accurate correlation between small scale and large scale curve coefficient values across the entire design space.

- The transformation function set – The purpose of this study was to establish how the specific transformation functions for each curve coefficient and each peak translate between experimental systems.

6.4.2 Validation Studies

Three separate validation studies are incorporated into this chapter, each altering various aspects of the experimental system.

- Reduction in product titre – upstream processes can produce a range of product titres due to the inherent variability of biological processes. Confirmation was required that large reductions in product titre or related changes in feedstock composition do not affect the effectiveness of the methodology or significantly alter the transformation functions for each curve coefficient. If a significant alteration in transformation function was observed with a change in product titre then batch to batch variability of the feedstock will render further validation of other experimental systems very difficult.

- Replacement of chromatographic ligand – The change of the MEP ligand for a PPA ligand will be used to determine how changes in ligand/protein chemistry alter elution profiles and also how these change when the bed height and linear velocity variables are altered.

The feedstock (FAb fragments within an *E. coli* periplasmic lysate), the sorbent backbone (cellulose), the ligand class (HCIC, mixed-mode) and the small scale device presentation (Pall PRC prepacked) remained fixed.

- A multi-variate change – The final validation study was to change multiple variables in a simultaneous fashion. The feedstock was changed to bovine IgG4 spiked egg white, the ligand to a carboxy-methyl (CM) ion-exchanger and the sorbent backbone to ceramic. Small scale device presentation remained unchanged with Pall PRC prepacked columns being used.

6.5 Validation Study Results and Discussion - Reduction in Product Titre

The results of the validation studies are presented in this section, with each study being individually presented, separated into analysis of the chromatographic separation, chromatographic peak analysis, curve coefficient data, the transformation function development and finally the chromatogram analysis. The feedstock equivalence data was pertinent for all three studies and is therefore presented first.

6.5.1 Feedstock Equivalence

Batch to batch variability is generally an issue when dealing with complex and realistic feedstocks such an *E. coli* lysate. The lysate from three separate fermentation batches has been utilised in the entire investigation, as summarised in Table 6.2.

To verify qualitative comparability a sample from each fermentation batch, purified and processed in an identical fashion, were run on a reducing SDS-PAGE shown in Figure 6.2 below.

Batch Number	**Production Bioreactor Scale**	**Titre (mg/mL)**	**Thesis Section**
1	20L	0.55	Chapter 4 – Separation of FAb fragments by mixed mode chromatography
2	20L	0.55	Chapter 5 – Ultra Scale-Down Method Production
3	75L	0.27	Chapter 6.5.2 – Validation of Ultra-Scale Down Method: Titre change
4	75L	0.19	Chapter 6.6 – Validation of Ultra Scale-Down Method: Ligand Change

Table 6.2: Details of the source of E. coli *lysate batches used in the studies within chapters 4-7.*

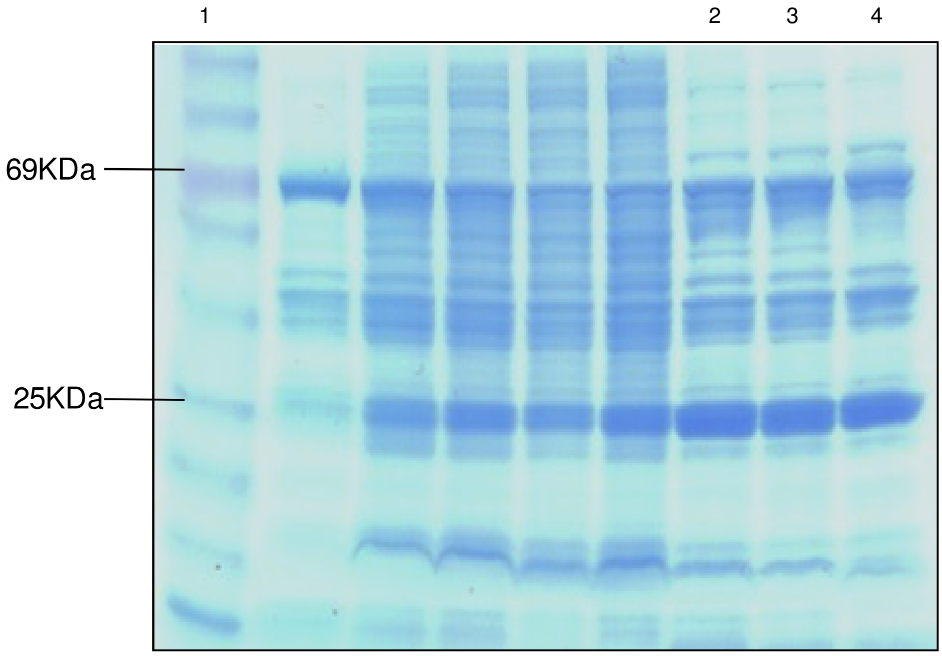


Figure 6.2: A reducing SDS-PAGE comparing samples from different fermentation batches of FAb fragment containing E. coli *lysate. Staining is with Coomassie blue and the marker was Benchmark pre-stained marker (Invitrogen). The gel lanes are: 1: Marker, 2: Batch 1, 3: Batch 2, 4: Batch 3.*

The sample were prepared to give the same concentration of FAb fragments on the gel. Each of the batches was qualitatively equivalent in terms of protein composition and all produced a clean, sharp Protein G HPLC peak when analysed, supporting batch comparability.

6.5.2 Chromatogram Analysis

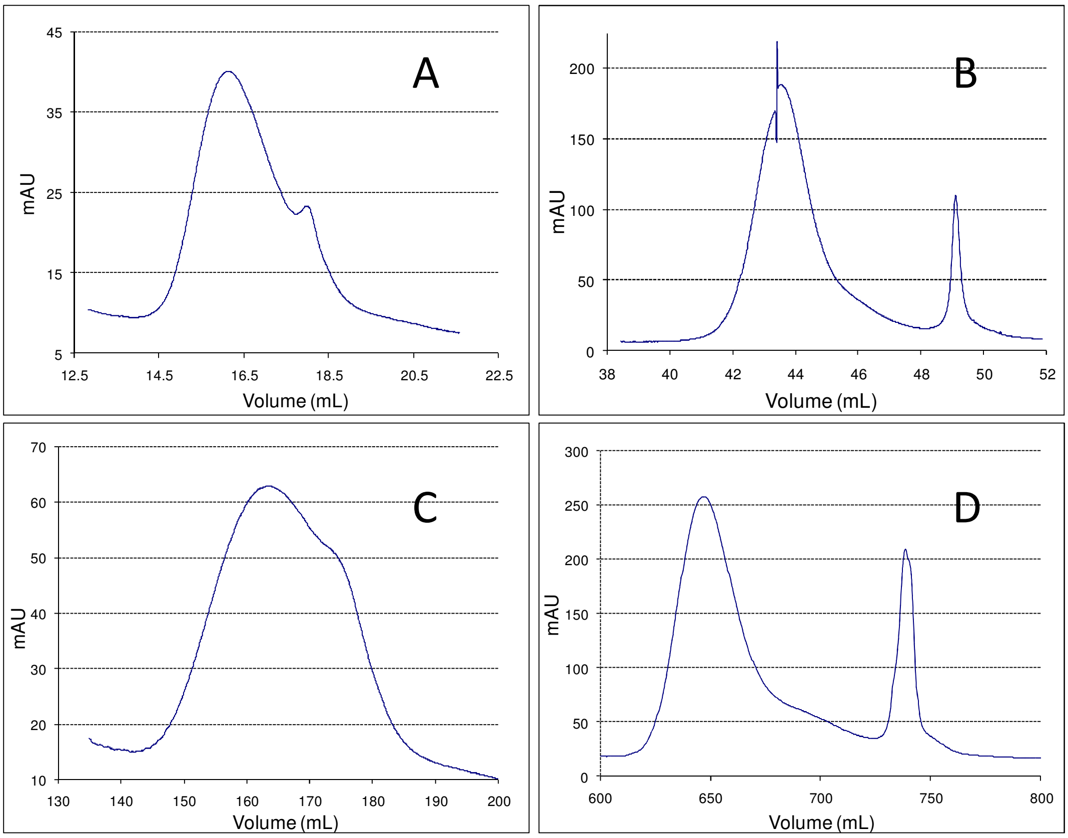


Figure 6.3: Four chromatograms demonstrating how the complete chromatogram changes between the edges of the factorial design space, with the variables being bed height and linear velocity of loading and elution. Chromatograms obtained with small scale and large scale systems are presented. A: Small scale 5cm bed height and 300cm/h linear velocity, B: Small scale 15cm bed height and 50cm/h linear velocity, C: Large scale 5cm bed height and 300cm/h linear velocity, D: 30cm bed height and 50cm/h linear velocity. Linear velocity values refer to both the loading and elution steps.

The above chromatograms were analysed and processed exactly as described in section 5.5.2. The elution profile for the feedstocks batch 2 and batch 3 are superficially similar in that a two peak pattern was observed.

6.5.3 Batch 2 and 3 Chromatogram Comparison

As a point of reference the chromatograms for the same device and conditions were compared using feedstock derived from batch 2 and batch 3. The results are shown in Figure 6.4.

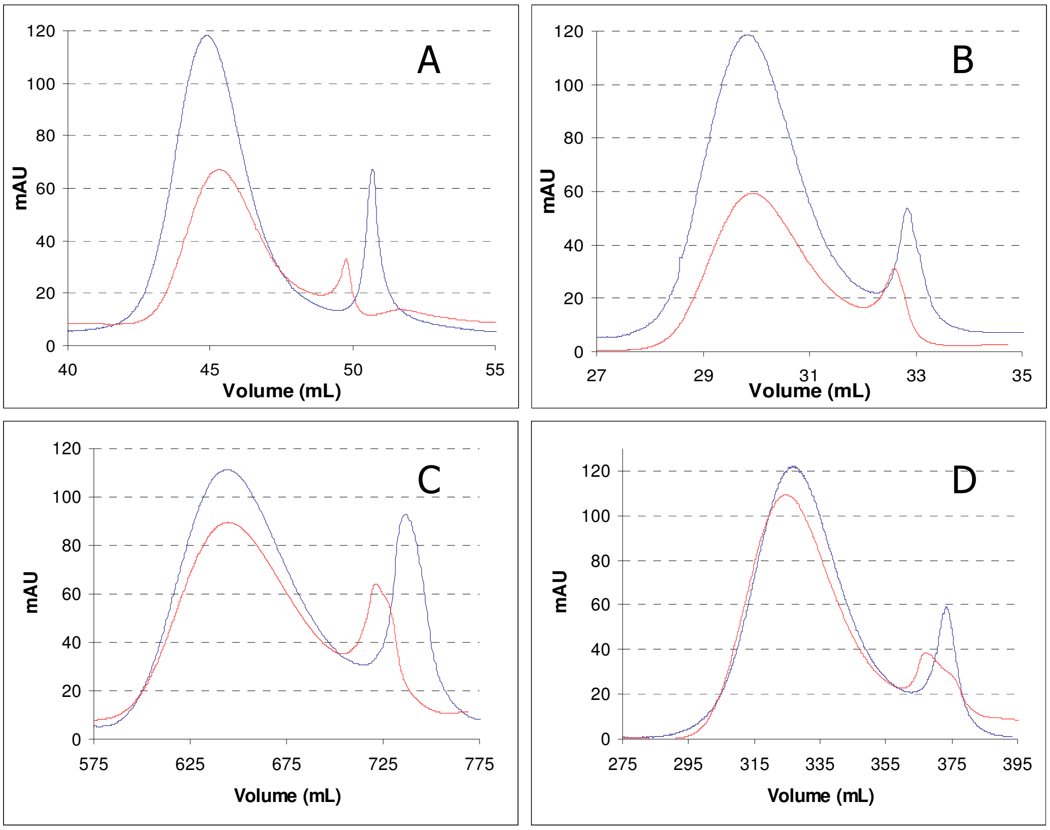


Figure 6.4: Eight representative chromatograms demonstrating how the elution profiles change when the feedstock titre was changed from 0.55mg/mL FAb fragment concentration to 0.27mg/mL, for the same bed height and linear velocity conditions. Series •: Batch 2 profile. Series •: Batch 3 profile. A: Small scale 15cm bed height and 150cm/h linear velocity B: Small scale 10cm bed height and 100cm/h linear velocity C: Large scale 30cm bed height and 300cm/h linear velocity D: 15cm bed height and 150cm/h linear velocity. Linear velocity values refer to both the loading and elution steps.

As can be observed the elution profile is different between batch 2 and batch 3 in scale, although the characteristic two peak chromatogram was still seen. It is interesting to note that the differences are far more pronounced between the small scale chromatograms than the large scale chromatograms .

6.5.4 Curve Coefficient Data

Table 6.3 below details the curve coefficients for all linear velocity and bed height conditions at both large and small scales. Response surfaces were produced from this data in the method described in sections 5.5.3-5.5.6, including the axis stretching and response surface averaging.

Scale	LV	BH	Peak One				Peak Two			
			h (mAU)	s	w (mL)	z (mL)	h (mAU)	s	w (mL)	z (mL)
Small	50	5	98.7	0.53	0.33	16.6	31.5	0.20	0.09	18.4
Small	**100**	**5**	**73.5**	**0.70**	**0.41**	**15.9**	**20.7**	**0.20**	**0.11**	**17.9**
Small	**150**	**5**	**48.2**	**1.05**	**0.51**	**17.8**	**11.0**	**0.28**	**0.18**	**19.8**
Small	**200**	**5**	**60.2**	**1.04**	**0.56**	**15.8**	**11.8**	**0.20**	**0.14**	**18.0**
Small	300	5	51.4	1.12	0.64	15.5	7.8	0.21	0.19	17.8
Small	**100**	**10**	**198**	**1.15**	**0.58**	**29.1**	**63.4**	**0.29**	**0.16**	**32.6**
Small	**150**	**10**	**116**	**1.11**	**0.78**	**35.4**	**27.3**	**0.20**	**0.17**	**39.3**
Small	**200**	**10**	**124**	**1.16**	**0.85**	**30.4**	**41.5**	**0.20**	**0.21**	**34.1**
Small	300	10	69.5	1.13	1.24	32.8	24.6	0.22	0.29	37.1
Small	**50**	**15**	**272**	**1.09**	**0.62**	**42.9**	**159**	**0.23**	**0.11**	**48.9**
Small	100	15	202	1.22	0.76	43.2	71.8	0.22	0.23	49.5
Small	**150**	**15**	**162**	**1.29**	**0.90**	**44.0**	**94.4**	**0.44**	**0.19**	**50.4**
Small	200	15	126	1.28	1.03	44.1	33.0	0.20	0.23	49.6
Small	**300**	**15**	**76.2**	**1.97**	**1.14**	**44.1**	**17.5**	**0.20**	**0.36**	**49.8**
Large	50	5	140	5.22	3.7	162	26.2	0.60	2.1	176
Large	**300**	**5**	**61.8**	**5.61**	**8.1**	**159**	**20.4**	**3.65**	**2.3**	**173**
Large	**150**	**15**	**174**	**13.8**	**9.5**	**343**	**50.0**	**1.23**	**3.2**	**397**
Large	**200**	**15**	**184**	**17.4**	**9.3**	**333**	**71.2**	**2.74**	**3.3**	**380**
Large	300	15	129	14.4	13.5	339	28.7	1.35	6.3	385
Large	**50**	**30**	**518**	**25.2**	**7.9**	**635**	**219**	**2.40**	**3.3**	**736**
Large	300	30	187	39.0	18.7	624	91.8	5.34	6.8	733
Small scale error			4.79	0.02	0.02	0.87	2.98	0.03	0.05	0.99
Large scale error			14.17	1.14	0.46	5.81	6.73	0.17	0.19	3.62

Table 6.3: The curve coefficients generated for each peak from the elution profile of a FAb fragment containing lysate separated using MEP HyperCel, regarding the chromatograms gained at variable bed heights (cm) and linear velocities (cm/h). Error for each curve coefficient was calculated by triplicate experiments (shown in red) to 1 standard deviation.

6.5.5 Transformation Function

The transformation function formulated in Chapter 5 was applied to the response surfaces obtained from the batch 3 separation data. The results are shown below in Table 6.4.

Curve Coefficient	Peak One	Peak Two
Height	$2h$	$2h$
Skew	$30s^2wb$	$30bsw$
Width	$w(0.7b)+\frac{f}{30}$	$w\left(\frac{f}{20}\right)+1.9$
Location	$9z$	$(25z)-30$

Table 6.4: A table detailing the transformation function for each curve coefficient for the system MEP HyperCel separating FAb fragments from batch 3 derived crude periplasmic lysate. In each case the bed height axis was multiplied by 2. h: peak height (mAU), s: peak skew, w: peak width (mL), b: bed height (cm), f: linear velocity (cm/h), z: peak location (mL).

For the justification of these functions see section 6.8. The transformation functions for each curve coefficient are identical between batches 2 and 3, with the exception of the skew of peak 2. The skew of the second peak is difficult to model in the MEP HyperCel gradient separation as a third peak does exist, but is usually very small and completely concealed by the second peak. In the large bed height systems with slow flow rates this third peak can resolve, albeit poorly, after the second peak. A user-inputted argument for the *EMGCurveFitLSQBig* algorithm for deconvolution of the chromatogram is the number and location of the peaks – if only two peaks are selected it will consider the third peak to be part of the second peak and skew the second peak accordingly. Therefore in this case the skew transformation function for peak can be considered an artefact.

6.5.6 Chromatogram Comparison

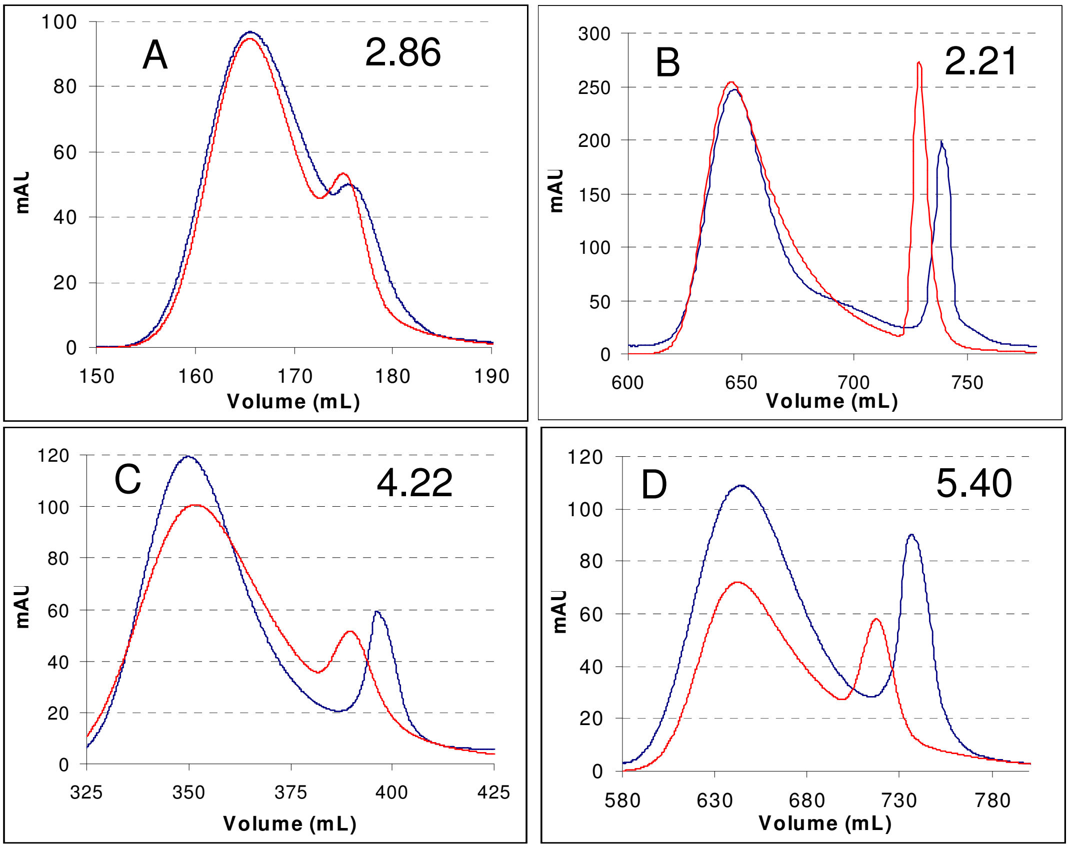


Figure 6.5: Eight chromatograms comparing the predicted chromatogram directly with the reference large scale chromatogram, for the separation of a periplasmic lysate using MEP HyperCel. Series 🟦: original large scale chromatogram, series 🟥: predicted chromatogram from transformed small scale data. A: 5cm bed height 50cm/h linear velocity, B: 30cm bed height 50cm/h linear velocity, C: 15cm bed height 150cm/h linear velocity, D: 30cm bed height 300cm/h linear velocity. The value in the top right of each chromatogram overlay is the normalised root mean square value between the two chromatograms presented on the same axes.

A representative sample of large scale chromatograms and predicted chromatograms from the small scale data is shown in Figure 6.5. Even for the most dissimilar result (30cm bed height, 300cm/hr) a normalised root mean square value of only 5.4% is observed, with all other results having a root mean square of under 5%.

6.6 Validation Study Results and Discussion – Change in Chromatographic Ligand

Once it had been determined in section 6.5 that a change in the feedstock product titre does not affect the outputs of the methodology and the resulting transformation function, a more challenging validation study could be performed. This involved implementing the same experimental system and the same feedstock, but utilising the mixed-mode PPA HyperCel instead of HCIC MEP HyperCel.

6.6.1 Chromatogram Analysis

Much higher resolution is observed at the lower linear velocities with the elution profiles obtained from the PPA HyperCel system, compared to the elution profiles obtained from the MEP HyperCel system as seen in sections 4.4.3 and 6.5.2. The number of visible peaks ranged from 3 (in chromatogram A in Figure 6.6) to 8 (in chromatogram B in Figure 6.6).

It is difficult for the *EMGCurveFitLSQBig* curve fitting deconvolution algorithm to accurately portray peaks that are very poorly resolved, therefore the peak limit is set at those visible in the least resolved chromatogram, A, with the analysed peaks numbered in Figure 6.6.

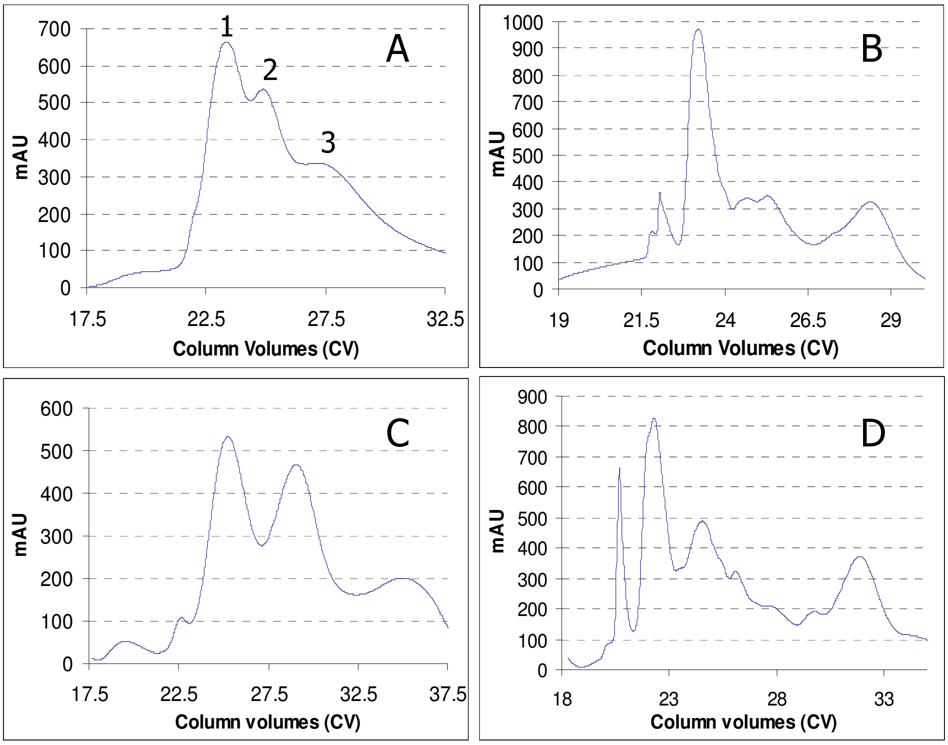


Figure 6.6: Four chromatograms demonstrating how the complete chromatogram changes between the opposing edges of the factorial design space. The elution profiles were generated by the separation of FAb fragments from a periplasmic E. coli *lysate using PPA HyperCel. The variables were bed height and the linear velocity of both loading and elution steps. Relevant peaks are numbered for future reference in chromatogram A. A: Small scale 5cm bed height and 300cm/h linear velocity, B: Small scale 15cm bed height and 50cm/h linear velocity, C: Large scale 5cm bed height and 300cm/h linear velocity, D: 30cm bed height and 50cm/h linear velocity.*

It is demonstrated in Figure 6.7 that the FAb fragments are concentrated in peak 1 and peak 2 in a similar fashion to the MEP HyperCel separations. This was further supported by the SDS-PAGE shown in Figure 6.8 which also demonstrates good qualitative purity of the separation for these conditions.

These results provide a good justification for restricting the deconvolution and mathematical analysis to the three peaks that are most relevant to the purification of the FAb fragments from the *E. coli* lysate.

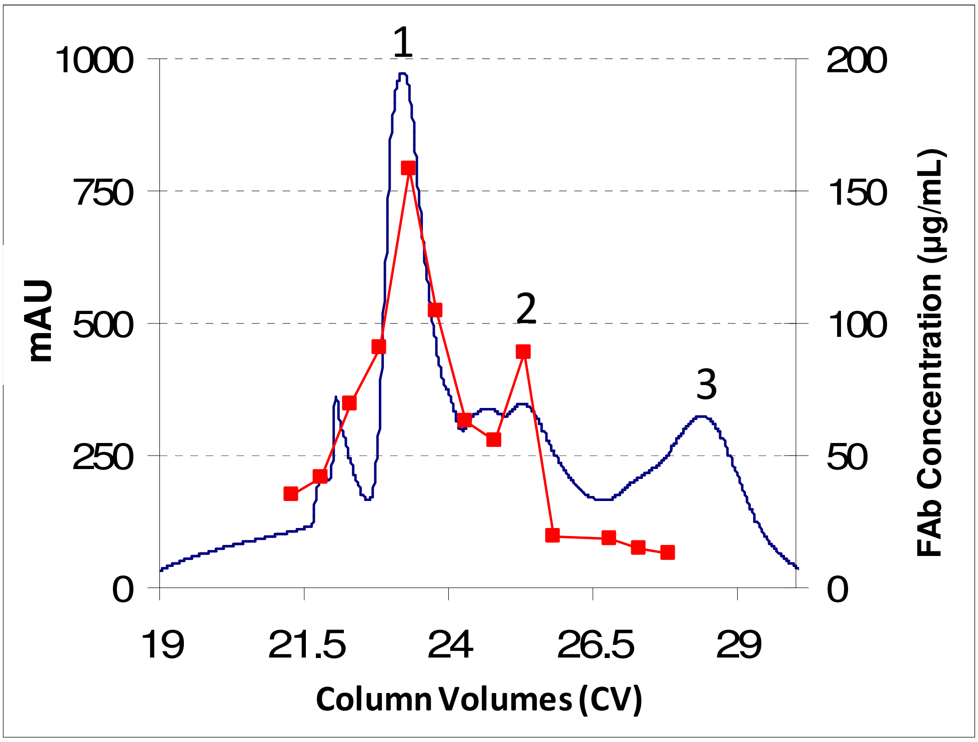


Figure 6.7: The elution profile of the separation of FAb fragment containing E. coli lysate using a pre-packed PRC PPA HyperCel chromatography column. Elution was performed by linear pH gradient (8.8-2.2) over 10CV. Series 🟦: chromatogram 15cm bed height, 50cm/h linear velocity at PRC scale, series 🟥: - (second axis) Protein G HPLC data showing FAb concentration for each fraction. The numbered peaks are the ones taken forward for mathematical analysis.

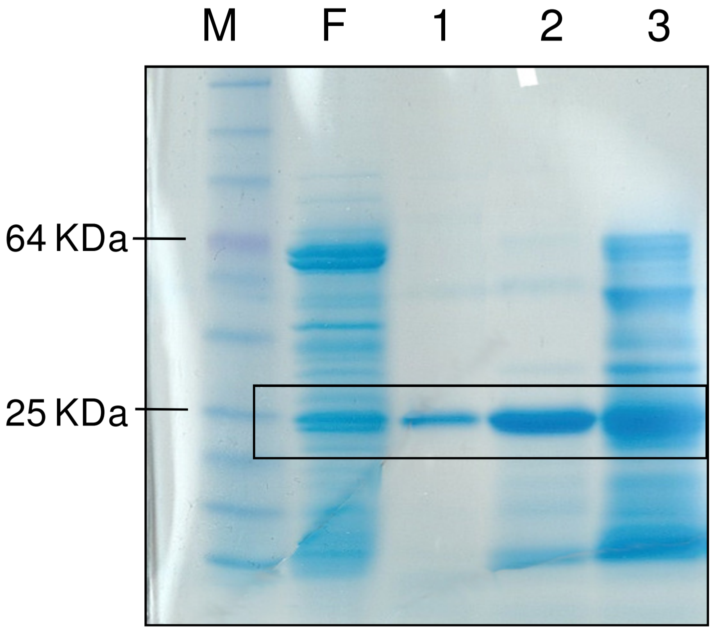


Figure 6.8: A reducing SDS-PAGE showing the protein composition of peak fractions from the separation of FAb fragment containing E. coli *lysate by PPA HyperCel in PRC pre-packed format. Staining was by Coomassie blue and the marker was Benchmark pre-stained marker (Invitrogen). Conditions were 15cm bed height and 50cm/h linear velocity loading/elution. M: Marker, F: Feedstock, 1: Peak 2, 2: Peak 1, 3: Peak 3.*

6.6.2 Curve Coefficient Data

Table 6.5 below details the curve coefficients for all linear velocity and bed height conditions at both large and small scales. Response surfaces were produced from this data in the method described in sections 5.5.3-5.5.6, including the axis stretching and response surface averaging.

Scale	LV	BH	Peak One				Peak Two				Peak Three			
			h (mAU)	s	w(mL)	Z(ml)	h (mAU)	s	w(mL)	Z(ml)	h (mAU)	s	w(mL)	Z(ml)
Small	50	5	680	0.45	1.06	36.5	706	1.45	0.79	39.6	527	4.55	1.38	43.9
Small	100	5	659	0.26	0.57	23.1	643	0.61	0.36	24.7	488	2.31	0.88	26.5
Small	150	5	1893	1.81	0.50	41.6	796	6.63	1.08	46.1	198	0.53	1.71	55.7
Small	200	5	841	0.30	0.54	23.3	700	0.74	0.40	24.8	225	0.26	1.11	27.7
Small	300	5	1077	0.47	0.68	37.6	512	0.40	0.98	40.8	470	2.57	1.70	43.8
Small	100	10	1249	1.32	0.78	46.3	496	1.15	1.26	52.9	232	0.41	2.07	61.9
Small	150	10	769	0.29	0.56	22.6	693	0.61	0.43	24.3	225	0.35	1.07	26.8
Small	200	10	1173	0.97	0.62	36.3	609	1.87	0.94	39.6	273	0.42	1.41	46.3
Small	300	10	763	0.72	2.10	47.0	843	2.55	1.22	52.7	414	2.23	2.40	61.3
Small	50	15	785	0.40	0.64	23.1	751	0.96	0.42	24.7	541	2.62	0.85	26.7
Small	100	15	702	0.55	1.11	36.6	504	1.97	1.00	40.1	212	0.49	1.86	46.8
Small	150	15	771	0.69	1.48	46.9	473	1.95	1.40	51.9	206	0.59	2.13	61.6
Small	200	15	952	0.99	0.61	22.7	654	1.30	0.40	24.6	473	2.71	0.88	26.7
Small	300	15	723	0.73	1.53	36.5	860	1.74	0.83	40.4	521	3.75	1.66	45.1
Large	50	5	992	1.64	1.16	47.0	549	1.74	1.19	53.1	217	0.43	2.41	62.2
Large	300	10	1013	7.57	5.24	256	560	2.00	9.79	294	341	1.9	11.4	346
Large	150	10	622	8.59	9.26	246	636	14.2	9.04	282	173	3.9	17.7	346
Large	150	20	786	13.2	9.66	556	559	22.4	18.0	618	264	7.9	26.3	734
Large	150	30	1484	21.9	8.46	656	467	15.8	15.4	726	279	12.1	19.9	945
Large	200	30	1092	28.2	14.1	696	658	35.1	24.0	785	343	29.3	36.2	940
Small scale error			23.8	0.03	0.03	0.68	9.21	0.11	0.03	0.80	10.1	0.10	0.04	1.0
Large scale error			24.4	0.82	0.36	21.9	5.26	1.05	0.66	24.3	5.61	0.89	0.97	30.3

Table 6.5: Peak data for the chromatograms gained at variable bed heights (cm) and linear velocities (cm/h) for Batch 4 FAb fragment containing periplasmic lysate feedstock material, separated using PPA HyperCel chromatography sorbent. Error was calculated by triplicate experiments (shown in red) to one standard deviation.

6.6.3 Transformation Function

The transformation function formulated in Chapter 5 was applied to the response surfaces obtained from the Batch 3 separation data. The results are shown below in Table 6.6.

Curve Coefficient	Peak One	Peak Two	Peak Three
Height	No change	$0.9h$	$h+275$
Skew	$21swb$	$21swb$	$25swb$
Width	$11w$	$18w$	$w\left(\frac{f}{30}\right)+b+3$
Location	$12z$	$16z$	$\frac{(18z)+(250b)}{f}$

Table 6.6: A table detailing the transformation function for each curve coefficient for the system PPA HyperCel separating FAb fragments from batch 4 derived crude lysate. In each case the bed height axis was multiplied by 2. h: peak height, s: peak skew, w: peak width, b: bed height (cm), f: linear velocity (cm/h), z: peak location.

It is clear from the transformation functions in Table 6.6 that there are major differences between the PPA HyperCel system and the MEP HyperCel system. In the case of the curve coefficients width and location the transformation functions were simpler for the PPA HyperCel model, involving a single numeric multiplier. In the case of the MEP HyperCel experimental system references to the linear velocity and bed height parameters were necessary to achieve acceptable accuracy within the function for these curve coefficients. Peak three required a relatively complex transformation function and did not give acceptably accurate prediction of the large scale peak; however it should be noted that the PPA HyperCel elution profile did contain a significant trail on the final peak which would have contributed to complex and potentially misleading outputs of the deconvolution of the third peak.

6.4.6 Chromatogram Comparison

A representative sample of large scale chromatograms and predicted chromatograms from the small scale data is shown in Figure 6.9. The difficulties with elucidating a suitable transformation function for peak three is readily apparent in Figure 6.9, with the third peak being visibly very dissimilar between the predicted chromatogram and the experimentally derived one. However, the normalised root mean square values for the first two peaks in which the FAb fragments elute is lower than 2.5%, signifying a very accurate prediction.

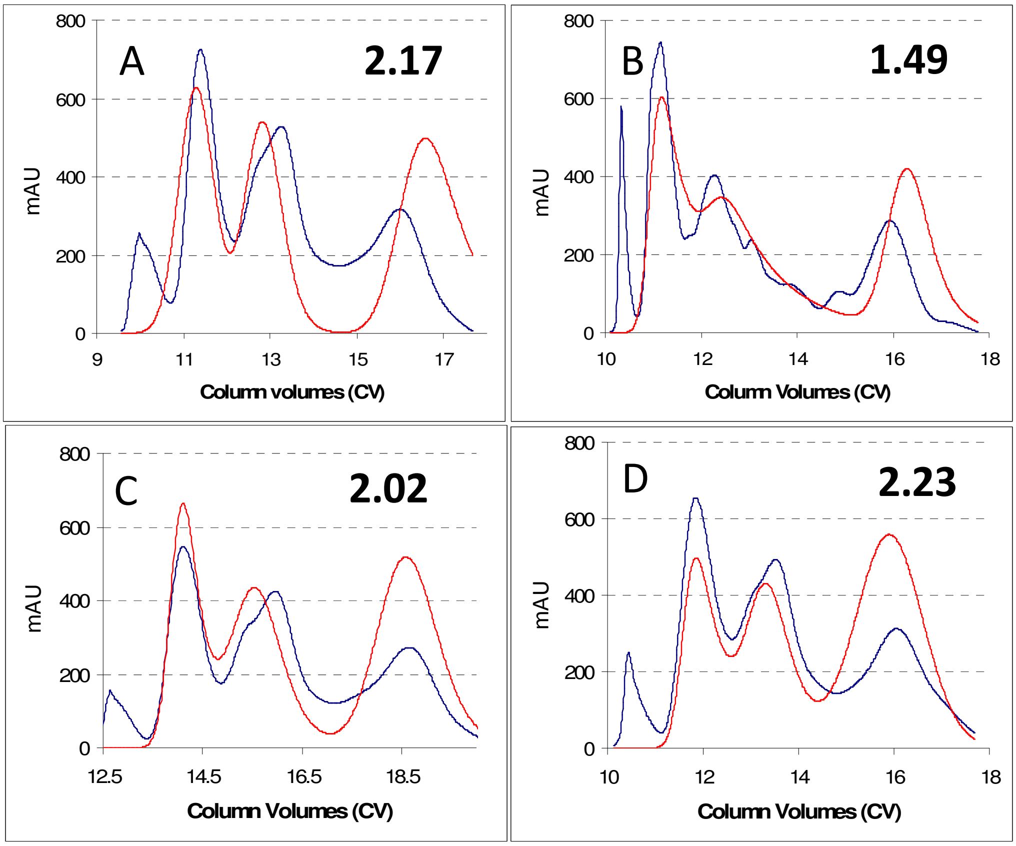


Figure 6.9: Eight chromatograms comparing the predicted chromatogram directly with the reference large scale chromatogram, for the separation of a periplasmic lysate using PPA HyperCel. Series 🟦: original large scale chromatogram, series 🟥: predicted chromatogram from transformed small scale data. A: 5cm bed height 50cm/h linear velocity, B: 30cm bed height 50cm/h linear velocity, C: 15cm bed height 150cm/h linear velocity, D: 30cm bed height 300cm/h linear velocity. The value in the top right of each chromatogram overlay is the normalised root mean square value between the two chromatograms presented on the same axes.

6.7 Validation Study Results and Discussion – Change of Ligand, Sorbent Backbone and Feedstock

This validation study investigated the effects of a multi-variate change in experimental parameters and evaluated whether the ultra scale-down methodology, which was developed on a different chromatographic platform, would be effective under these conditions. This study involved a change of sorbent to a CM ion exchange ligand with a ceramic backbone and a change of feedstock to an IgG spiked egg white. The target molecule class (an immunoglobulin/immunoglobulin fragment) and the column presentation (Pall PRC) remained the same as in the previous two studies.

6.7.1 Chromatogram analysis

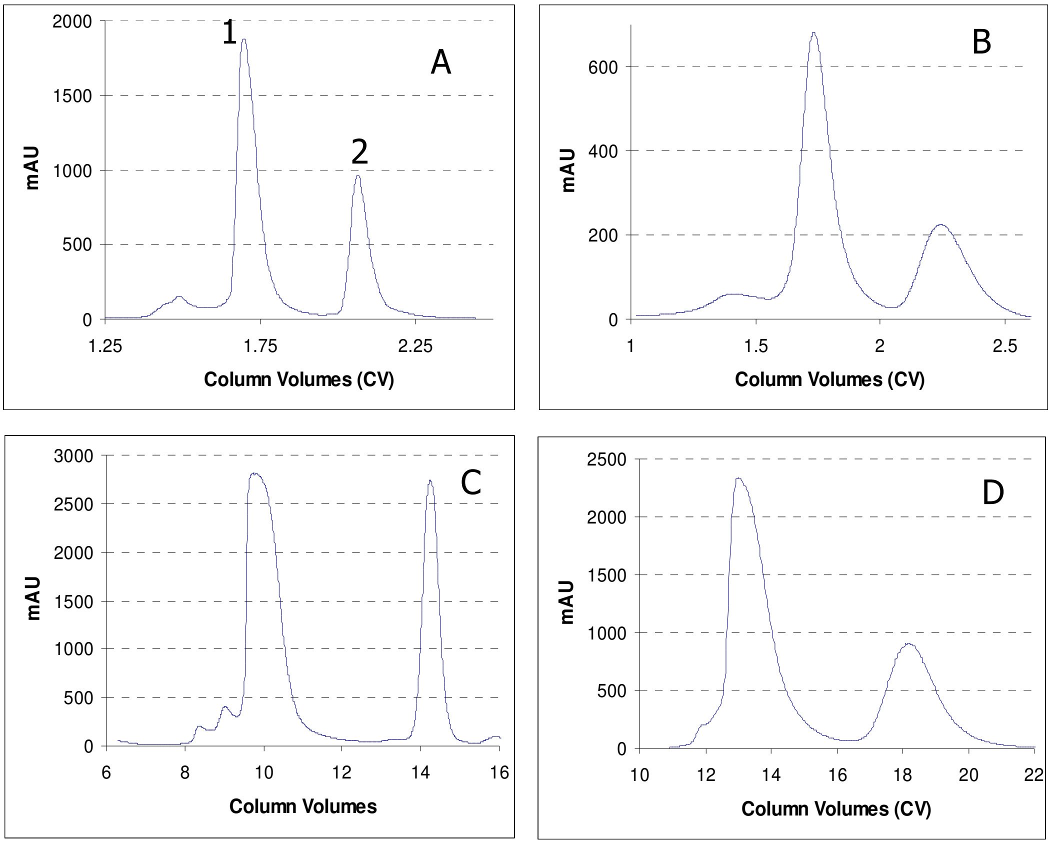


Figure 6.10: Four chromatograms demonstrating how the complete chromatogram changes between the opposing edges of the factorial design space. The elution profiles were generated by the separation of IgG spiked egg white using CM Ceramic HyperD. The variables were bed height and the linear velocity of both loading and elution steps. Relevant peaks are numbered for future reference in chromatogram A. A: Small scale 15cm bed height and 50cm/h linear velocity, B: Small scale 5cm bed height and 300cm/h linear velocity, C: Large scale 30cm bed height and 50cm/h linear velocity, D: Large scale 10cm bed height and 300cm/h linear velocity.

The separation of bovine IgG4 from hen's egg white produced a clear elution profile with 2 main peaks that were well resolved from each other. Peak 1 is shown (by Protein G HPLC) in Figure 6.11 to contain the majority of the IgG, although the antibody does elute over 0.5 CV prior to peak 1. Although Figure 6.12 demonstrates that this elution profile is not a particularly effective separation the two main peaks are well defined and change in shape and size as the bed height/linear velocity variables are altered. As such the separation was deemed suitable for this validation study.

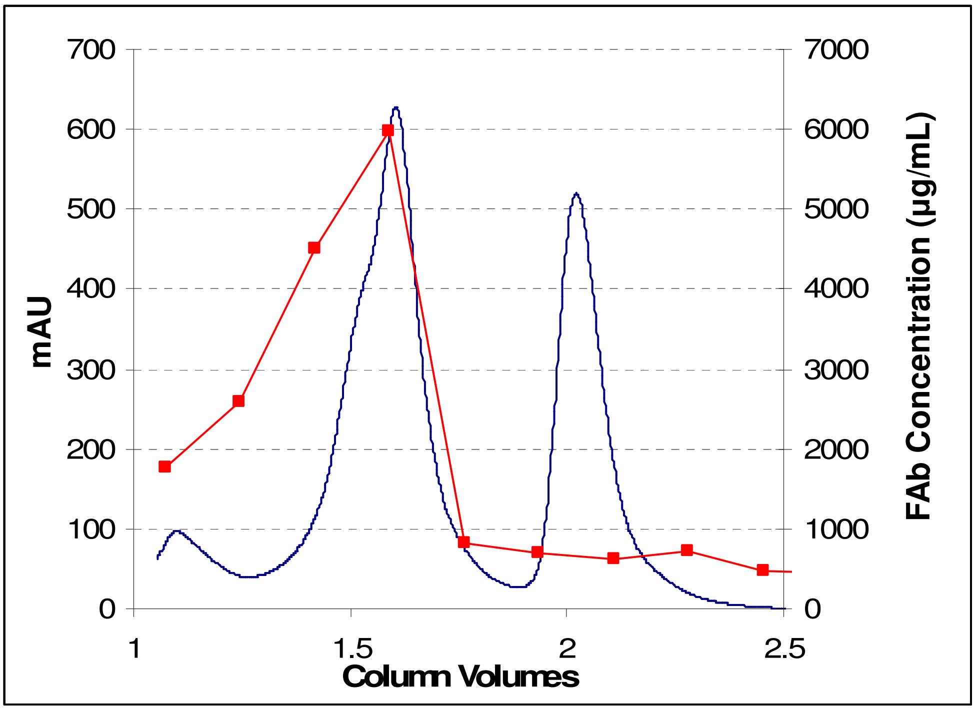


Figure 6.11: The elution profile of the separation of FAb fragment containing E. coli *lysate on a pre-packed PRC CM Ceramic HyperD chromatography column. Elution was performed by linear NaCl gradient (0-1M) over 10CV. Series 🟦: chromatogram 15cm bed height, 50cm/h linear velocity at PRC scale, series 🟥: - (second axis) Protein G HPLC data showing FAb concentration for each fraction.*

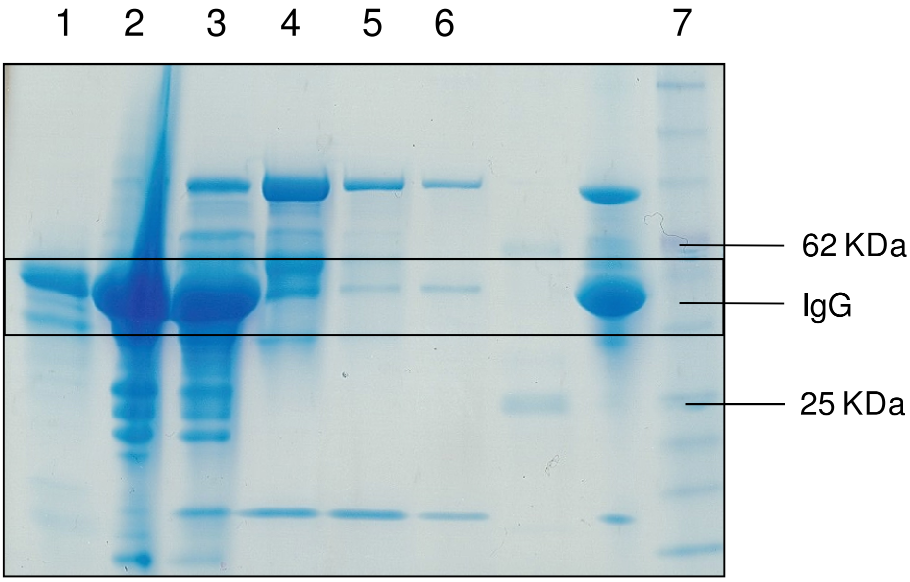


Figure 6.12: A reducing SDS-PAGE showing the protein composition of the peak fractions from the separation of IgG spiked egg white by CM Ceramic HyperD in PRC pre-packed format. Conditions were 15cm bed height and 50cm/h linear velocity loading/elution. Staining was with Coomassie blue. Key: 1: Peak eluted at 1.1 CV, 2: First half of peak 1, 3: Second half of peak 1, 4: Peak 2, 5/6: NaOH cleaning peaks, 7: Marker.

6.7.2 Curve Coefficient Data

Table 6.7 below details the curve coefficients for all the linear velocity and bed height conditions investigated at both large and small scales. Response surfaces were produced from this data by the method developed in sections 5.5.3-5.5.6, including the axis stretching and response surface averaging.

Scale	LV	BH	Peak One				Peak Two			
			h (mAU)	s	w (mL)	z (mL)	h (mAU)	s	w (mL)	z (mL)
Small	5	50	1220	0.40	0.12	8.81	629	0.40	0.12	11.1
Small	**5**	**100**	**788**	**0.40**	**0.28**	**8.20**	**482**	**0.40**	**0.27**	**10.7**
Small	5	150	1118	0.40	0.15	8.45	344	0.40	0.25	10.8
Small	**5**	**200**	**755**	**0.41**	**0.22**	**9.40**	**226**	**0.45**	**0.34**	**11.7**
Small	5	300	1127	0.40	0.21	9.42	552	0.40	0.32	11.8
Small	**10**	**50**	**931**	**0.40**	**0.51**	**20.8**	**1743**	**0.40**	**0.22**	**24.9**
Small	**10**	**150**	**607**	**0.40**	**0.80**	**15.4**	**865**	**0.65**	**0.33**	**19.8**
Small	**10**	**300**	**1718**	**0.65**	**0.31**	**16.2**	**740**	**0.64**	**0.43**	**20.7**
Small	15	50	3507	0.61	0.27	25.2	1611	0.55	0.28	30.6
Small	**15**	**100**	**3030**	**0.77**	**0.31**	**23.4**	**1320**	**0.54**	**0.37**	**29.3**
Small	15	150	3349	0.81	0.35	34.1	1438	0.65	0.46	40.2
Small	**15**	**200**	**2862**	**0.81**	**0.34**	**23.6**	**1063**	**0.64**	**0.48**	**29.6**
Small	15	300	2677	0.79	0.31	22.9	903	0.69	0.53	28.9
Large	**10**	**50**	**5455**	**5.82**	**2.56**	**131**	**2830**	**3.49**	**2.24**	**176**
Large	10	300	4600	7.53	3.05	128	1303	7.29	5.26	177
Large	**20**	**150**	**5822**	**11.2**	**4.07**	**198**	**2606**	**6.41**	**4.75**	**280**
Large	30	50	4644	11.9	7.43	290	3270	3.55	5.07	424
Large	**30**	**300**	**6750**	**18.4**	**5.52**	**276**	**2358**	**9.79**	**7.81**	**406**
Small scale error			**83.8**	**0.01**	**0.01**	**0.64**	**37.5**	**0.01**	**0.01**	**0.76**
Large Scale Error			**178**	**0.97**	**0.40**	**15.4**	**147**	**0.53**	**0.40**	**23.9**

Table 6.7: The curve coefficients generated for each peak from the elution profile of IgG spiked egg white separated using CM Ceramic HyperD, regarding the chromatograms gained at variable bed heights (cm) and linear velocities (cm/h). Error for each curve coefficient was calculated by triplicate experiments (shown in red) to 1 standard deviation.

6.7.3 Transformation Function

Curve Coefficient	Peak One	Peak Two
Height	$h + 4000$	$2h + 600$
Skew	$25swb$	$(300sw) - 20$
Width	$w\left(\frac{b}{3}\right) * 3 + 1$	$12.5w + \left(\frac{b}{10}\right)$
Location	$10z$	$z(1.1b) + 37.5$

Table 6.8: A table detailing the transformation function for each curve coefficient for the system CM Ceramic HyperD separating bovine IgG spiked egg white. In each case the bed height axis was multiplied by 2. h: peak height, s: peak skew, w: peak width, b: bed height (cm), f: linear velocity (cm/h), z: peak location (mL).

The transformation functions for the curve coefficients shown in Table 6.8 are again very different to both the MEP HyperCel and PPA HyperCel experimental systems. However, as shown in Figure 6.13, the above transformation functions produce a very accurate prediction of large scale chromatographic behaviour.

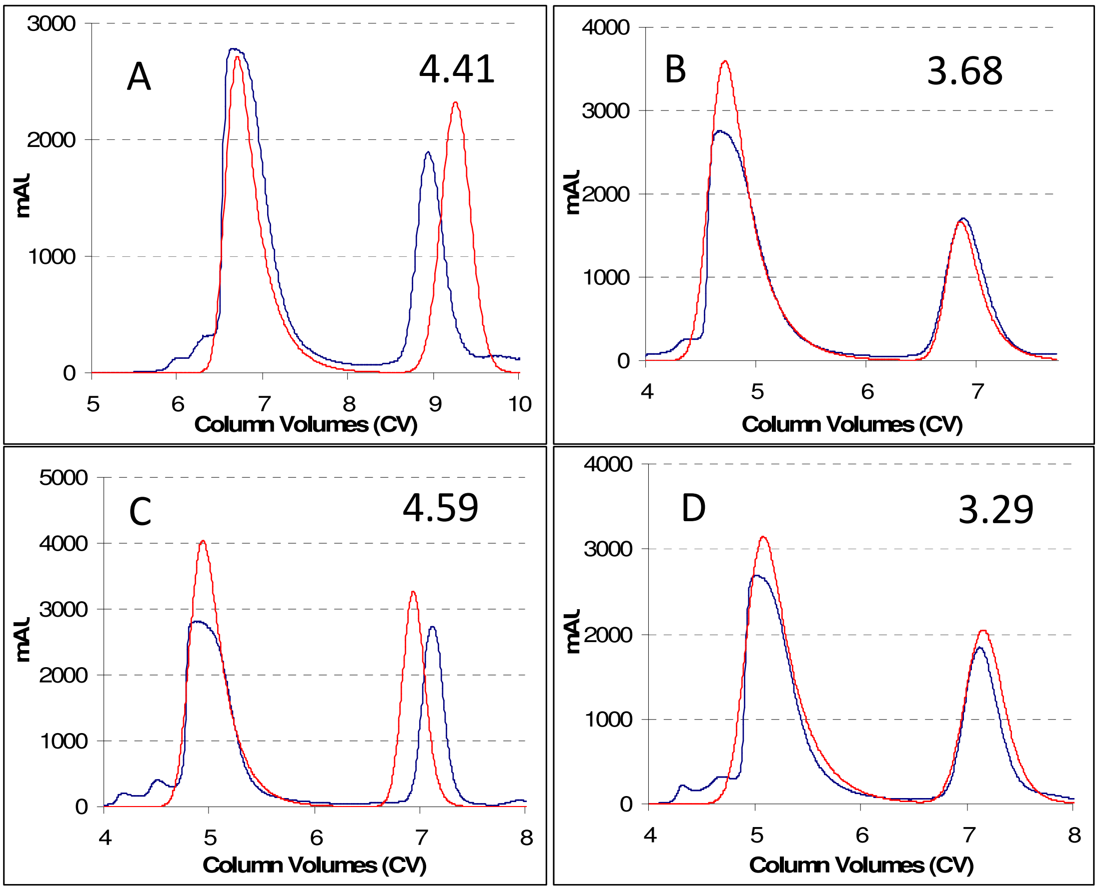


Figure 6.13: Eight chromatograms comparing the predicted chromatogram directly with the reference large scale chromatogram, for the separation of bovine IgG spiked egg white using CM Ceramic HyperD. Series 🟦: original large scale chromatogram, series 🟥: predicted chromatogram from transformed small scale data. A: 5cm bed height 50cm/h linear velocity, B: 30cm bed height 50cm/h linear velocity, C: 15cm bed height 150cm/h linear velocity, D: 30cm bed height 300cm/h linear velocity. The value in the top right of each chromatogram overlay is the normalised root mean square value between the two chromatograms presented on the same axes.

6.8 Conclusions

The objective of this chapter was to challenge and validate the ultra scale-down methodology and transformation functions as developed in Chapter 5 by its application to different experimental systems. The different experimental systems consisted of a reduction in product titre, a change of the chromatographic ligand to PPA and a multi-variate change that altered the chromatographic ligand, the feedstock and the sorbent backbone material.

6.8.1 Validation Study – Reduction in Product Titre

It is apparent from the results that reducing the product concentration by 50% in the feed has very little effect on how the curve coefficients change between small and large scales, allowing interchangeable use of the transformation function formed in Chapter 5. This can be explained as to the nature of the process change; it has been postulated that the major changes in elution profile are seen through hydrodynamic and geometric effects, such as changing linear velocity conditions and bed height. Other than potential changes to the viscosity of the feed product titre, the change is purely chemical in nature and as the total quantities of FAb loaded in both data sets was identical the difference between the chromatograms at small and large scale was kept constant. However, it was essential to extensively test the effects a reduction in product titre would have on the outputs of the scale-down methodology, otherwise batch to batch variability of feedstock would render any validation study very difficult to analyse.

6.8.2 Validation Study – Change of Chromatographic Ligand

It was not possible to apply directly the transformation function obtained from the MEP HyperCel separation study to the PPA HyperCel study. This implies that the functions are not totally generic for different chromatographic ligands, although some of the curve coefficients do exhibit limited correlations between experimental systems (see Section 6.8.5). However, the new set of transformation functions formed for the PPA HyperCel experimental system proved to give excellent prediction capability between large and small scales for the two critical peaks that contain the target FAb fragments.

Changing the chromatographic ligand produces a complex chromatogram, which is only fully resolved at the largest bed heights and slowest linear velocities. The presence of unresolved chromatographic peaks under the three main peaks analysed in the study would have significant effects on the curve coefficients and therefore the transformation functions required for the translation of small scale data into large scale predictive data. A potential solution to this problem would be for the deconvolution algorithm to take the number of peaks seen in the highest resolution chromatogram (8 in this case, as seen in the 50cmh/h linear velocity/30cm bed height elution profile) and extrapolate the peak positions to the less resolved chromatograms. This would however be an extrapolated figure and as such would introduce an uncertainty into the algorithm outputs that would be difficult to validate.

6.8.3 Validation Study – Change of Ligand, Sorbent Backbone and Feedstock

The objective of this study was to investigate whether the ultra scale-down methodology and set of curve coefficient transformation functions described in Chapter 5 could be applied to a very dissimilar chromatography separation platform, where this change involves multiple responses. In this case the changes included a change of sorbent backbone , a change in ligand class and a change in feedstock. The translation was partially successful in that the ultra-scale down methodology was capable of producing a set of transformation functions that gave a good prediction of large scale elution profiles from a small scale data set. However, the transformation functions that were produced with the MEP HyperCel/FAb fragment containing lysate experimental system proved to be highly unique to the system, with a new set of transformation functions required to produce an accurate chromatographic prediction.

6.8.4 The Validation of the Ultra-Scale Down Methodology

The validation of the ultra scale-down method methodology as described in Chapter 5 can be considered successful. The same methodology and algorithm set as utilised in Chapter 5 successfully deconvoluted the elution profiles seen in the MEP HyperCel, PPA HyperCel and the CM Ceramic HyperD experimental systems and allowed the production of transformation functions for the translation of small scale curve coefficients into their large scale equivalents with a high degree of accuracy.

6.8.5 Correlation of Transformation Functions

It was deemed necessary for the transformation functions to utilise the hydrodynamic experimental parameters (the bed height and the linear velocity) if at all possible. The response surfaces graphically represented how the bed height, linear velocity and curve coefficient values were related and as such the transformation function would need to include these arguments. This however was not necessary when the transformation function was a simple multiplier, as seen in the function for peak one location in all three studies.

There was a significant difference in the transformation functions for peak one, peak two and (in the case of the PPA HyperCel experimental system) peak three. It was observed that the transformation functions were simpler and generally contained far fewer terms for peak one compared to peaks two and three, and there could be a number of reasons for this. Peak one in all cases was the largest and best resolved peak, and as such higher accuracy of the deconvolution function could be expected. Also, peaks two and three are in someway dependent on the data for peak one, depending on the resolution of the chromatogram and how severely peak one is skewed down the elution profile. This would mean that the relationship between the small and large scale curve coefficients would be more complicated for peaks two and three and the transformation functions would need to reflect that.

The transformation functions were developed for accuracy rather than simplicity, as a key objective was to mimic a large scale chromatogram with small scale data as closely as possible. Many of the numerical terms in the transformation functions are for "fine tuning" to ensure this accuracy and could be dropped if a slightly less accurate chromatographic mimic was acceptable.

It is clear that the transformation functions are not entirely generic and there are significant differences observed between experimental systems. However, there are some correlations that can be drawn, particularly with the target molecule elution peak (peak one).

As far as correlations between small and large scale data are concerned, skew can be predicted on the basis of the skew curve coefficient itself, the width curve coefficient and, unexpectedly, the bed height. The linear velocity is critical for determining how skewed a chromatographic peak will be, but it is constant as far as scaling from small to large scales are concerned.   The only difference between the experimental systems is the multiplier that is required.

Peak location showed strong connections between experimental systems, at least for the first peak. This was logical as the elution volume is dependent almost entirely on column volume. It was however very difficult to provide linking correlations for peaks two and three.

Peak height also seemed to have little obvious correlation between the experimental systems, ranging from a simple multiplier, a fixed increase or even no change between small scale and large scale. Peak width was similar to peak height, in that the transformation function required a complex association of numbers, the bed height parameter and the flow rate parameter, and was therefore dissimilar between experimental systems.

This analysis was not designed to be exhaustive and there would be numerous ways in which the available parameters could be manipulated to produce a more generic set of transformation functions that had additional applicability between experimental systems. This analysis would be facilitated by additional data sets from other experimental systems and additional mathematical analysis on the existing data sets, and as such would be suitable for future work.

7 Business Case

7.1 Product Development and Intellectual Property

The ultra scale-down methodology developed in this thesis represents a proven method but not a tangible product. As such there are a number of options concerning its potential commercialisation. The commercialisation could expose the methodology and algorithms to potential reverse-engineering, and as such protection of the associated intellectual property will be critical for any long term marketing strategy.

As discussed in future work, the scale-down methodology developed in this thesis is not complete and can instead be considered a beta prototype. As the transformation functions are still broadly system specific this limits potential commercial uptake – further work to produce more generic transformation functions that are applicable across a range of chromatography systems and feedstocks would be required to ensure the methodology is suitable for a customer's needs.

7.2 Software Development

The methodology and transformation function could not be marketed in its current form. Matlab is currently required for all the algorithms that are run, not including the response surface averaging function which was written in MS Excel. Other than the requirement for specific and expensive software, the interface is not user friendly or self-explanatory. As such there would be a requirement for a polished and user-accessible front end to facilitate the use of the methodology and also to conceal proprietary algorithms that are key to the effectiveness of the ultra scale-down methodology.

It can however be expected that a potential client will have MS Excel, as the MS Office suite is ubiquitous amongst many computing platforms. Mathworks, the developers of the Matlab application, also produce a plug-in for Excel that will allow code in M file format to run natively within the Excel system. With some development this could be integrated into an Excel macro to provide a user friendly experience that will generate the required outputs.

Ideally the only user input the Excel macro would need is the chromatographic data itself, identification of the peak locations and the operating parameters such as linear velocity, bed height and sorbent ligand class, manipulated via a GUI.

7.3 Pre-Packed PRC Chromatography Columns and Integrated System Solutions

It is unlikely that the marketing of the ultra scale-down methodology software package on its own would be successful as it would be difficult to integrate it with the client’s existing equipment. However, the methodology was developed utilising a Pall Life Sciences product, the PRC pre-packed chromatography columns, which are commercially available.

It would therefore be logical to market the USD methodology software and the PRC chromatography columns together as an integrated unit. This would also restrict the quality testing required of the software to Pall Life Science products only, rather than having to validate the software with competing chromatography products.

The cost of goods for the USD methodology software would be difficult to ascertain from a development perspective, but actual production and maintenance would be inexpensive. Some training would be required of technical support departments to handle technical enquiries with the software and there would need to be some support of the licensing model in the long term, but compared to the manufacture of a quality controlled precision product the ongoing costs are negligible.

It is envisaged then that the USD methodology software would be used as an intellectual loss leader for the PRC pre-packed chromatography columns themselves. Once the software system is in use with a client and satisfactory results are gained it would be very difficult to use a competitor's chromatography products without the risk of losing the effectiveness of the USD methodology software.

From a very long term perspective the PRC columns and the USD methodology software could be part of an integrated systems approach to downstream processing of biologics, where the software and it's associated products are implemented as a "service" from Pall Life Sciences regarding the scale up of all unit operations. Being able to offer an integrated system and product solution to minimise time and resources required for downstream scale up, rather than a combination of various non-integrated products, would be extremely attractive to clients involved heavily in downstream process development.

8 Future work

This thesis has a structured narrative but many of the earlier chapters can be considered to be self-contained. As such future work will be considered on a chapter by chapter basis, followed by a section for future work on the thesis as a whole.

8.1 Chapter 2 – Feedstock Selection

A number of other potential feedstocks could be utilised as model feeds. The other possibilities considered were:

- CHO cell line supernatant: CHO cell lines have become very common hosts in bioprocessing as the cell lines are well characterised and can perform complex post-translational modifications to a recombinant protein product, including complex glycosylation.
- *E. coli* homogenate: The periplasmic lysate utilised as a feedstock in Chapters 5 and 6 has considerably fewer contaminants than a typical *E. coli* homogenate, making for a less complex separation.

8.2 Chapter 3 – Small Scale Device Selection

There are a large number of manufacturers of very small chromatography columns, with a range of sorbents and columns designs. These are mostly, in essence, very similar to the columns investigated within Chapter 3 so further work would be unlikely to be needed if different commercial designs were to be used.

8.3 Chapter 4 – Separation of FAb Fragments by Mixed-Mode Chromatography

This chapter is essentially a methods chapter, similar to an industrial investigation to optimise a downstream chromatography process for a new product. The chapter however only investigated the separation to a level of detail sufficient to provide a robust model system for use in the ultra scale-down methodology development in Chapter 5 and 6, rather than a fully optimised process.

The investigation in Chapter 4 concluded that PPA HyperCel would be the most suitable mixed-mode sorbent for the separation of FAb fragments from a crude *E. coli* periplasmic lysate. Further pertinent work would be in applying Quality by Design principles to further optimise this separation as regards linear velocities, loading buffer pH and composition and modifications of the feed, such as the addition of differing concentrations of ammonium sulphate. The ultra scale-down methodology developed in Chapter 5 could be well used to this effect.

An investigation into a different feedstock would also be very relevant, for example the aforementioned CHO cell supernatant containing an intact IgG, rather than an antibody fragment.

8.4 Chapter 5 - Formulation of an Ultra Scale-Down Methodology for Chromatography Scale-Up

There would be two sections to the future work to this chapter.

8.4.1 Experimental Ultra Scale-Down Methodology

The concept of the ultra scale-down methodology was shown to be sound in the form described in Chapter 5. Future work would therefore concern expansion of this concept to other parameters rather than just bed height and linear velocity. This could involve operational decisions, such as buffer conditions, gradient length, feed processing and ligand selection. The eventual aim would be to input all critical quality attributes (CQAs) into the methodology and allow the entire chromatographic development and optimisation process to be initiated at very small scales, with progression to large scales only for validation.

8.4.2 Transformation Functions

As described in Chapter 6, the transformation functions are not entirely generic and contain components unique to each experimental system. Further mathematical analysis could potentially identify important correlations within this data and harmonise more of the components within the transformation functions for each experimental system. This most likely would require more validation data sets, as mentioned in the next section.

8.5 Validation of Methodology and Model for the Ultra Scale-Down of Elution Chromatography

There are currently four different data sets to which the ultra scale-down methodology has been applied. The addition of extra data sets would further validate the ultra scale-down methodology itself. This would also provide more data for the optimisation of the transformation functions.

The extra validation studies would have to focus on uni-variate studies to fully understand how each chromatographic parameter affects the scale up within the methodology. Now that the methodology has been validated with a complex feed, a synthetic feed of four analytical grade proteins (for example the major proteins in egg white, but pre-purified and purchased from Sigma) would allow clearer analysis of the relationships, as the experimental error from the non-protein contaminants would be reduced.

9 Appendices

9.1 Appendix 1 - MatLab Algorithms Code (all code provided by S. Edwards-Parton, UCL, London, UK)

9.1.1 Deconvolution Algorithm (EMGCurveFitLSQBig)

```
function [Peaks, VerticalData]=EMGCurveFitLSQBIG(XAxis, Curve,
NoPeaks, KnownPos);
%function [Peaks, VerticalData]=EMGCurveFit(XAxis, Curve, NoPeaks,
KnownPos);
%known pos specifies which curves if any you have a known location
for, if %you know 3 peak location of 1 4 10 then known pos should be
[1 4 10], %program will then ask you to make guesses on the rest

flag_reduce=0;

%if max(XAxis)>100
%    XAxis=XAxis/100;
%    flag_reduce=1;
%end

%how many do we know
[NoKnownPeak IGNORE]=size(KnownPos);

Maximums=[];

for i=1:NoKnownPeak
    Maximums(i)=KnownPos(i);
end

if NoKnownPeak<NoPeaks
for i=1:NoPeaks-NoKnownPeak
    %get user to select remaining peaks
    figure
    plot(XAxis, Curve, 'r', XAxis(Maximums), Curve(Maximums), 'bo');
    [Points]=ginput(1);
    [Distance, Points]=findnearest(Points, [XAxis Curve]);
    Maximums=[Maximums find(XAxis==Points(1))];
    close
end
end
LowerWZ=XAxis(2);
UpperWZ=max(XAxis)*1;

PeakMove=max(XAxis);

for i=1:NoPeaks
    h=Curve(Maximums(i));%'h
```

```
width(1)=(XAxis(max(find(Curve(1:Maximums(i))<Curve(Maximums(i))*0.6)
))-XAxis(Maximums(i))); %low curve width

width(2)=(XAxis(Maximums(i)+min(find(Curve(Maximums(i):length(XAxis))
<Curve(Maximums(i))*0.6)))-XAxis(Maximums(i))); %low point of curve

    StartPoint((i-1)*4+1)=h*1.5;
    StartPoint((i-1)*4+2)=2;  %s
%     StartPoint((i-
1)*4+2)=max([2*min(abs(width))*abs(width(2)/width(1)) 1]); %s

    StartPoint((i-1)*4+3)=min(abs(width)); %w
    StartPoint((i-1)*4+4)=XAxis(Maximums(i))-min(abs(width)); % %z

    Lower((i-1)*4+1)=Curve(Maximums(i))*0.5;
    Lower((i-1)*4+2)=0.2*StartPoint((i-1)*4+2); %s
    Lower((i-1)*4+3)=min(abs(width))*0.01; %w
    Lower((i-1)*4+4)=XAxis(Maximums(i))-max(abs(width))*2; %z



    Upper((i-1)*4+1)=Curve(Maximums(i))*4;
    Upper((i-1)*4+2)=100; %s
    Upper((i-1)*4+3)=min(abs(width))*2; %w
    Upper((i-1)*4+4)=XAxis(Maximums(i))+max(abs(width))*2; %z

end
StartPoint
Lower
Upper

Options=Optimset;
Options.MaxFunEvals=48000;
Options.TolX=1e-17;
Options.TolFun=1e-17;
 x = lsqcurvefit(@EMG2, StartPoint, XAxis , Curve, Lower, Upper,
Options)

VerticalData=x;

for i=1:NoPeaks

    Peaks(:,i)=EMG2(x( ((i-1)*4+1) : ((i-1)*4+4) ), XAxis);
    if flag_reduce==1
        VerticalData((i-1)*4+3)=VerticalData((i-1)*4+3)*10;
    end

end
```

9.1.2 Peak Locator (Findnearest)

```
function [Distance, Points]=FindNearest(Points, Data)
%function [Distance, Points]=FindNearest(Points, Data)
%finds the nearest actual point from Data to Points

%find x and y points on curve that match x and y points from click
%selection

%VariableDeclare=   {'VariableDeclare' 'ans' ... %Standard info
%                    'Points' ... %input of rough points to find
nearest from
                                 %and output variable giving nearest
points based on Data
%                    'Data' ... %list of all known points
%                    'XNearest' ... %Nearest points on X axis
%                    'YNearest' ... %Nearest points based on Y axis
%                    'YPoint', 'XPoint' ... %Corresponding complete
point data for XNearest and YNearest
%                    'YDist', 'XDist' ... %Distance of these points
from Points variable
%                    }

[XNearest]=Nearest(Data(:,1)', Points(1));
[YNearest]=Nearest(Data(:,2)', Points(2));

%set two pairing pointy and point x that are based on nearest x and y
vals
%(could be same points but easier to assume there not rather than
check and
%then process)
YPoint=Data(find(Data(:,2)==YNearest),:);
XPoint=Data(find(Data(:,1)==XNearest),:);

%Find distance to original point for each of these
YDist=(((YPoint(1)-Points(1))^2)+((YPoint(2)-Points(2))^2))^(0.5);
XDist=(((XPoint(1)-Points(1))^2)+((XPoint(2)-Points(2))^2))^(0.5);

%take points with smallest distance
if YDist<=XDist
    Points=YPoint;
    Distance=YDist;
else
    Points=XPoint;
    Distance=XDist;
end

%VariableCheck(who,VariableDeclare);
```

9.1.3 Nearest (sub-procedure of Findnearest)

```
function NearestValue = Nearest(Numbers,PinPoints)
%
% NEAREST Find nearest values to the pin points provided.
%
%   Example: If a = 0:0.01:1; b = sin(2*a*pi); plot(a,b)
%
%            then nearest(b,[-0.5 0 0.25 2]) is [-0.4818 0 0.2487
1.0000].
%
%   See also FIND.

%   Copyright (c) 2000- by Heekwan Lee (heekwan.lee@reading.ac.uk)
%   $Revision: 1.1 $  $Date: 2000/04/16 16:17:18 $


if nargin < 2
   NearestValue = [];
   return
end


Numbers = real(Numbers);
Pins = length(PinPoints);

Max = max(Numbers);
Min = min(Numbers);

for iPin = 1:Pins

   if PinPoints(iPin) > Max
      NearestValue(iPin) = Max;
   elseif PinPoints(iPin) < Min
      NearestValue(iPin) = Min;
   else
      nul = abs(Numbers - PinPoints(iPin));
      iValue = find(nul==min(nul));
      NearestValue(iPin) = Numbers(iValue);
   end

end
```

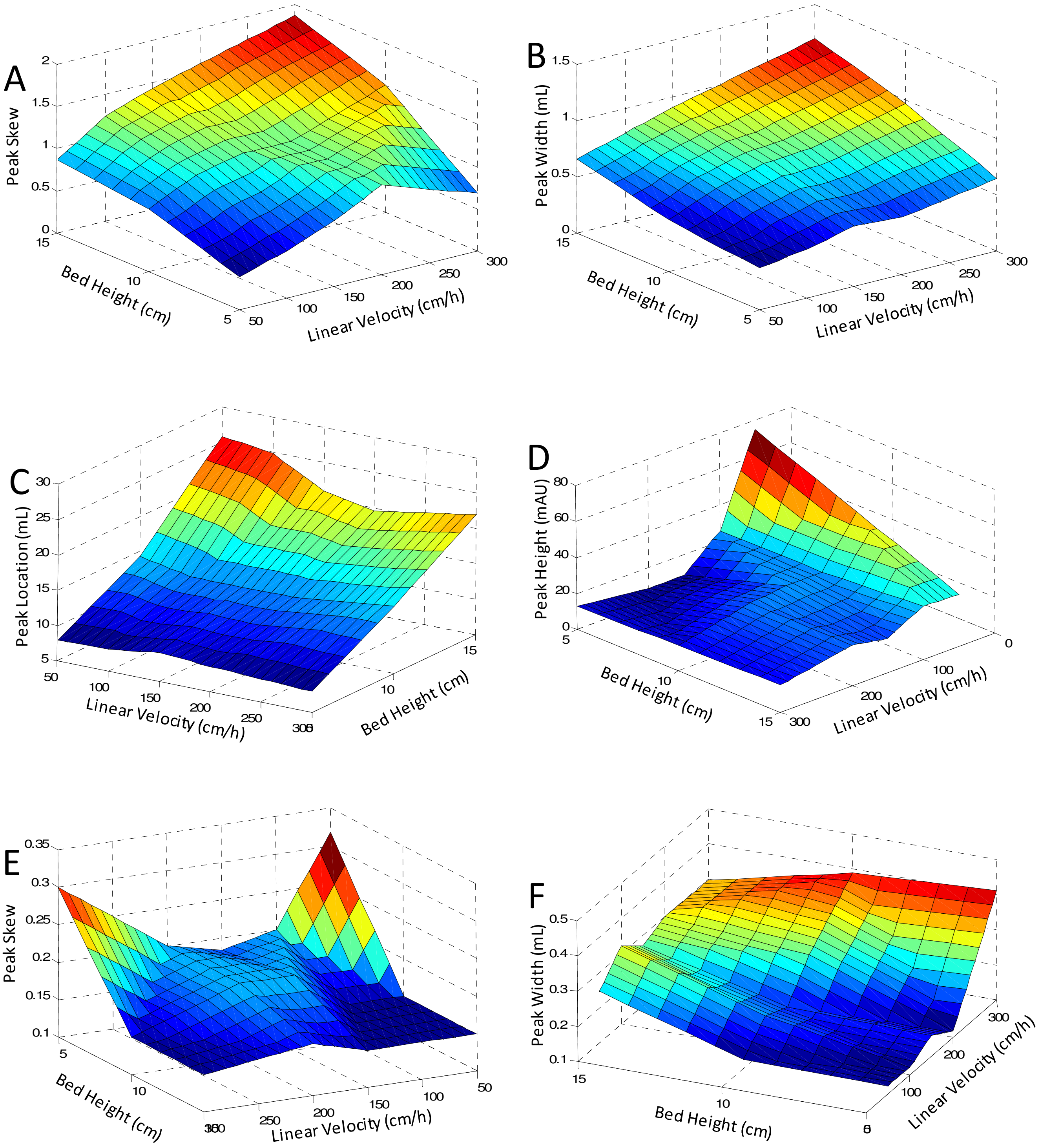
9.1.4 PlotIt (Plots chromatographic peaks from curve coefficients)

```
function Y=PlotIt(XAxis, CurveData, Peaks, Type);
%function Y=PlotIt(XAxis, CurveData, Peaks, Type);

for i=1:Peaks
    StartP=(i-1)*4;
    switch upper(Type)

    case 'EMG'
        Y(:,i)=EMG2(CurveData(StartP+1:StartP+4), XAxis);
    case 'GAUSSIAN'
        Y(:,i)=Gaussian(XAxis, CurveData(StartP+1),
CurveData(StartP+2), CurveData(StartP+3));
    case 'LOGNORMAL'
        Y(:,i)=LogNormal(XAxis, CurveData(StartP+1),
CurveData(StartP+2), CurveData(StartP+3), CurveData(StartP+4));
    otherwise
        Y=[];
        disp('No curve mathc function found');
    end

end
```

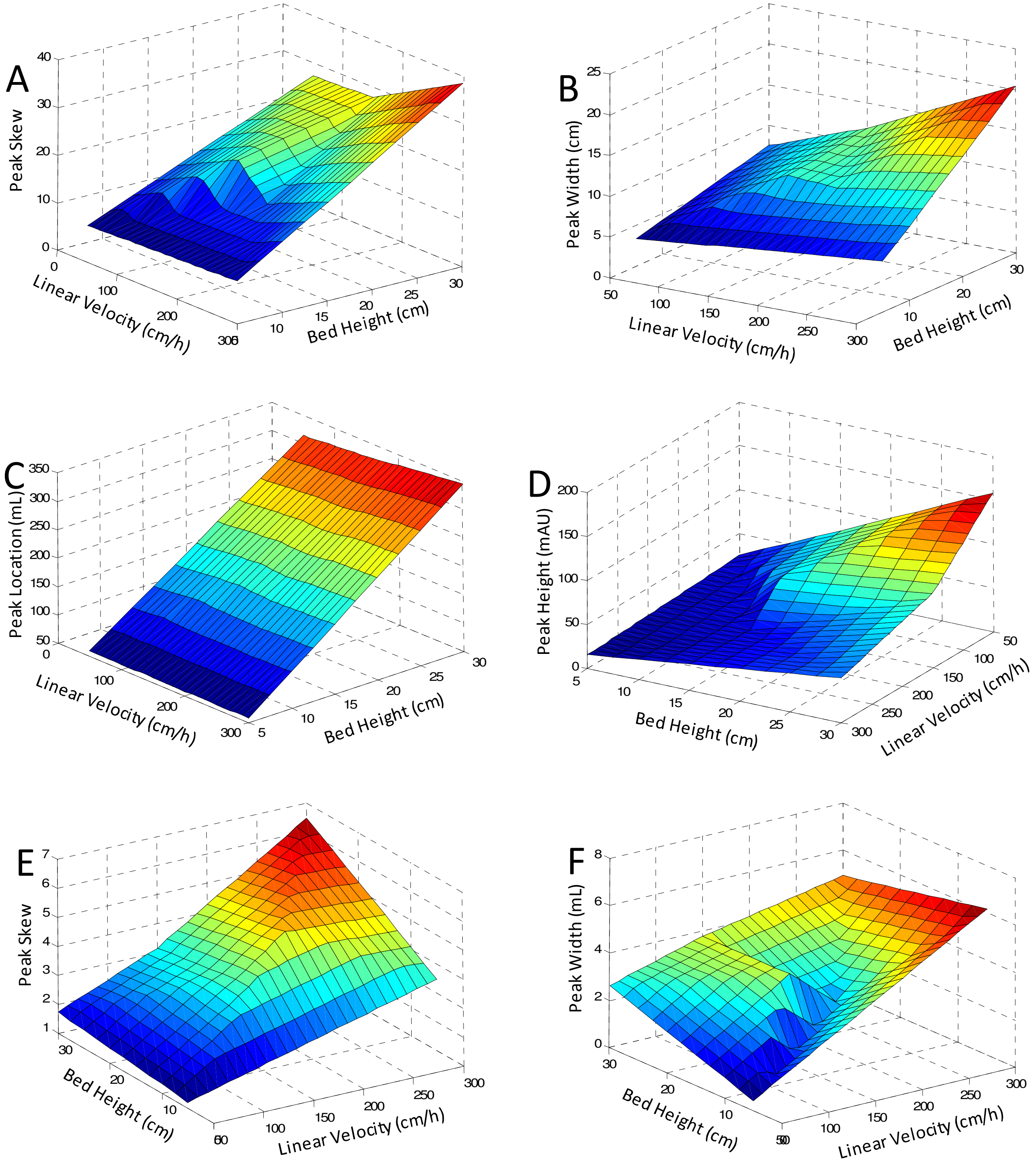
9.1.5 EMG (Incorporates the EMG equation)

```
function [Y]=EMG(x, xdata); %x is coefficients xdata is the xaxis
%[Y]=EMG(x, xdata);
%x is coefficients xdata is the xaxis

[rows cols]=size(x);

if mod(rows,4)~=0
    error('GAH wrong number of coefficients')
end

nocurves=rows/4;


if (s<0.01) && (s > (-0.01))
    if s>0
        s=0.01;
    else
        s=-0.01;
    end
end

Section1=((h.*w)./abs(s)).*((pi/2)^0.5);

Section2=((0.5*(w.^2))./(s.^2))   -   ((X-z)./s);

Section3=(2^(-0.5));

Section4=((w./abs(s))   -   (((X-z)./w).*(s./abs(s))));

Y=Section1.*exp(Section2).*(1-erf(Section3.*(Section4)));
```

9.2 Appendix 2 – Response Surfaces



Response surfaces for bed height and linear velocity against a range of curve coefficients at small (PRC) scale. A – Peak one skew, B – Peak one width, C – Peak one location, D – Peak two height, E – Peak two skew, F – Peak two width



Response surfaces for bed height and linear velocity against a range of curve coefficients at large (XK16) scale. A – Peak one skew, B – Peak one width, C – Peak one location, D – Peak two height, E – Peak two skew, F – Peak two width

10 Bibliography

Awade, A.C. 1996. On hen egg fractionation: applications of liquid chromatography to the isolation and the purification of hen egg white and egg yolk proteins. *Zeitschrift für Lebensmitteluntersuchung und -Forschung A* 202, no. 1 (January 1): 1-14. doi:10.1007/BF01229676.

Anspach, F. B, D. Curbelo, R. Hartmann, G. Garke, and W. D Deckwer. 1999. Expanded-bed chromatography in primary protein purification. *Journal of Chromatography A* 865, no. 1: 129–144.

Baines, D. 2001. Analysis of purity. Protein purification techniques: a practical approach: 27.

Baneyx, François. 1999. Recombinant protein expression in Escherichia coli. *Current Opinion in Biotechnology* 10, no. 5 (October 1): 411-421. doi:10.1016/S0958-1669(99)00003-8.

Barker, T. B. 1994. *Quality by experimental design*. Marcel Dekker.

Bartle, K. D, and P. Myers. 2001. Theory of capillary electrochromatography. *Journal of Chromatography A* 916, no. 1: 3–23.

Beisler, Amy T., Kathleen E. Schaefer, and Stephen G. Weber. 2003. Simple method for the quantitative examination of extra column band broadening in microchromatographic systems. *Journal of Chromatography A* 986, no. 2 (February 7): 247-251. doi:10.1016/S0021-9673(02)02018-6.

Belew M., Johansson, B-L., Maloisel, J-L., 2001. Method for mixed mode adsorption, US Patent 7008542

Bensch, M., P. Schulze Wierling, E. Von Lieres, and J. Hubbuch. 2005. High throughput screening of chromatographic phases for rapid process development. *Chemical Engineering & Technology* 28, no. 11.

Bergander, Tryggve, Kristina Nilsson-Välimaa, Katarina Öberg, and Karol M. Lacki. 2008. High-Throughput Process Development: Determination of Dynamic Binding Capacity Using Microtiter Filter Plates Filled with Chromatography Resin. *Biotechnology Progress* 24, no. 3: 632-639. doi:10.1021/bp0704687.

Berthod, A. 1989. On the Use of the Knox Equation. I. the Fit Problem. *Journal of Liquid Chromatography & Related Technologies* 12, no. 7: 1169–1185.

Bhattacharjee, S., C. Bhattacharjee, and S. Datta. 2006. Studies on the fractionation of beta-lactoglobulin from casein whey using ultrafiltration and ion-exchange membrane chromatography. *Journal of Membrane Science* 275, no. 1: 141–150.

Borrebaeck, Carl A. K. 2000. Antibodies in diagnostics – from immunoassays to protein chips. *Immunology Today* 21, no. 8 (August): 379-382. doi:10.1016/S0167-5699(00)01683-2.

Boulding, N., S. S. S. Yim, E. Keshavarz-Moore, P. Ayazi Shamlou, and M. Berry. 2002. Ultra scaledown to predict filtering centrifugation of secreted antibody fragments from fungal broth. *Biotechnology and Bioengineering* 79, no. 4: 381-388. doi:10.1002/bit.10306.

Boychyn, M., S. S. S. Yim, M. Bulmer, J. More, D. G. Bracewell, and M. Hoare. 2004. Performance prediction of industrial centrifuges using scale-down models. *Bioprocess and Biosystems Engineering* 26, no. 6 (December 1): 385-391. doi:10.1007/s00449-003-0328-y.

Brenac Brochier, V., A. Schapman, P. Santambien, and L. Britsch. 2008. Fast purification process optimization using mixed-mode chromatography sorbents in pre-packed mini-columns. *Journal of Chromatography A* 1177, no. 2: 226–233.

Broyles, B. Scott, R. Andrew Shalliker, and Georges Guiochon. 1999. Visualization of sample introduction in liquid chromatography columns: The effect of the frit diameter. *Journal of Chromatography A* 855, no. 2 (September 10): 367-382. doi:10.1016/S0021-9673(99)00737-2.

Burton, S. C., and D. R. K. Harding. 1998. Hydrophobic charge induction chromatography: salt independent protein adsorption and facile elution with aqueous buffers. *Journal of Chromatography A* 814, no. 1 (July 24): 71-81. doi:10.1016/S0021-9673(98)00436-1.

Burton, S. C., N. W. Haggarty, and D. R. K. Harding. 1997. One step purification of chymosin by mixed mode chromatography. *Biotechnology and Bioengineering* 56, no. 1 (10): 45-55. doi:10.1002/(SICI)1097-0290(19971005)56:1<45::AID-BIT5>3.0.CO;2-V.

Burton, Simon C., and David R. K. Harding. 2001. Salt-independent adsorption chromatography: new broad-spectrum affinity methods for protein capture. *Journal of Biochemical and Biophysical Methods* 49, no. 1 (October): 275-287. doi:10.1016/S0165-022X(01)00204-4.

Burton, S., Haggarty, N., Harding, D., Becker, N., Bulthius, B., Steele, L.; Chromatographic resins and methods for using same; 1997, US Patent 5945520

Chandwani, R. M., M. E. Pate, J. Savery, N. F. Thornhill, B. Horstmann, M. Hoare, and N. J. Titchener-Hooker. 1997. The use of principal component analysis as a database mining tool for the exploratory diagnosis of chromatographic processes. *Bioprocess and Biosystems Engineering* 17, no. 4: 229–234.

Chen, T. F, H. Yu, and D. F Barofsky. 1992. Centrifugal size-exclusion chromatographic method for rapid desalting and filtering of carbohydrate samples prior to fast atom bombardment mass spectrometry. *Analytical Chemistry* 64, no. 17: 2014–2019.

Chhatre, Sunil, and Nigel J. Titchener-Hooker. 2009. Review: Microscale methods for high-throughput chromatography development in the pharmaceutical industry. *Journal of Chemical Technology & Biotechnology* 84, no. 7: 927-940. doi:10.1002/jctb.2125.

Chhatre, Sunil, Pranavan Thillaivinayagalingam, Richard Francis, Nigel J. Titchener-Hooker, Anthony R. Newcombe, and Eli Keshavarz-Moore. 2007. Decision-Support Software for the Industrial-Scale Chromatographic Purification of Antibodies. *Biotechnology Progress* 23, no. 4: 888-894. doi:10.1021/bp070062u.

Chhatre, Sunil, Richard Francis, Anthony R Newcombe, Yuhong Zhou, Nigel Titchener-Hooker, Josh King, and Eli Keshavarz-Moore. 2008. Global sensitivity analysis for the determination of parameter importance in the chromatographic purification of polyclonal antibodies. *Journal of Chemical Technology & Biotechnology* 83 (February): 201-208. doi:10.1002/jctb.1801.

Clausing, D., and B. H Simpson. 1990. Quality by design. *Quality Progress* 23, no. 1: 41–4.

Coffman, Jonathan L., Jack F. Kramarczyk, and Brian D. Kelley. 2008. High-throughput screening of chromatographic separations: I. Method development and column modeling. *Biotechnology and Bioengineering* 100, no. 4: 605-618. doi:10.1002/bit.21904.

Coleman, Lynette, and Stephen M Mahler. 2003. Purification of Fab fragments from a monoclonal antibody papain digest by Gradiflow electrophoresis. *Protein Expression and Purification* 32, no. 2 (December): 246-251. doi:10.1016/j.pep.2003.07.005.

Cortes, H. J., C. D. Pfeiffer, B. E. Richter, and T. S. Stevens. 1987. Porous ceramic bed supports for fused silica packed capillary columns used in liquid chromatography. *Journal of High Resolution Chromatography* 10, no. 8: 446-448. doi:10.1002/jhrc.1240100805.

Creasey, A.; Pipette tips; Patent application: 9/899,027

Cuatrecasas, P., and M. Wilchek. 1968. Single-step purification of avidine from egg white by affinity chromatography on biocytin-Sepharose columns. *Biochemical and biophysical research communications* 33, no. 2: 235.

Di Marco, Valerio B., and G. Giorgio Bombi. 2001. Mathematical functions for the representation of chromatographic peaks. *Journal of Chromatography A* 931, no. 1 (October 5): 1-30. doi:10.1016/S0021-9673(01)01136-0.

Doran, P. M, and A. Ison. 1995. *Bioprocess engineering principles*. Academic Press London.

Edwards-Parton, S., N. F. Thornhill, D. G. Bracewell, J. M. Liddell, and N. J. Titchener-Hooker. 2008. Principal Component Score Modeling for the Rapid Description of Chromatographic Separations. *Biotechnology Progress* 24, no. 1: 202-208. doi:10.1021/bp070240j.

Fenneteau, Frédérique, Hafida Aomari, Parminder Chahal, and Robert Legros. 2003. Modeling of scale-down effects on the hydrodynamics of expanded bed adsorption columns. *Biotechnology and Bioengineering* 81, no. 7: 790-799. doi:10.1002/bit.10522.

Foley, J. P. 1987. Equations for chromatographic peak modeling and calculation of peak area. *Analytical chemistry* 59, no. 15: 1984–1987.

Foley, J. P., and J. G. Dorsey. 1984. A review of the Exponentially Modified Gaussian(EMG) function: evaluation and subsequent calculation of universal data. *Journal of chromatographic science* 22, no. 1: 40–46.

Foucault, A. P. 2001. Enantioseparations in counter-current chromatography and centrifugal partition chromatography. *Journal of Chromatography A* 906, no. 1 (January 12): 365-378. doi:10.1016/S0021-9673(00)00499-4.

Gallis, K. W., J. T. Araujo, K. J. Duff, J. G. Moore, and C. C. Landry. 1999. The use of mesoporous silica in liquid chromatography. *Advanced Materials* 11, no. 17: 1452–1455.

Geciova, Jana, Dean Bury, and Paul Jelen. 2002. Methods for disruption of microbial cells for potential use in the dairy industry--a review. *International Dairy Journal* 12, no. 6: 541-553. doi:10.1016/S0958-6946(02)00038-9.

Ghose, Sanchayita, Brian Hubbard, and Steven M. Cramer. 2005. Protein Interactions in Hydrophobic Charge Induction Chromatography (HCIC). *Biotechnology Progress* 21, no. 2: 498-508. doi:10.1021/bp049712+.

Ghose, Siddhartha, and Howard Chase. 2000. Expanded bed chromatography of proteins in small diameter columns. I. Scale down and validation. *Bioseparation* 9, no. 1 (January 1): 21-28. doi:10.1023/A:1008193312969.

Godvarti, R., Petrone, Robinson, R. Wright, and B. D Kelley. Scale-down models for purification processes: Approaches and applications. *Guidelines, Current Practices, and Industrial Case Studies* 2002; 69.

Goetz, H., M. Kuschel, T. Wulff, C. Sauber, C. Miller, S. Fisher, and C. Woodward. 2004. Comparison of selected analytical techniques for protein sizing, quantitation and molecular weight determination. *Journal of Biochemical and Biophysical Methods* 60, no. 3 (September 30): 281-293. doi:10.1016/j.jbbm.2004.01.007.

Golay, Marcel J.E. 1981. The height equivalent to a theoretical plate of retentionless rectangular tubes. *Journal of Chromatography A* 216 (October 30): 1-8. doi:10.1016/S0021-9673(00)82330-4.
Gottesman, S. 1996. PROTEASES AND THEIR TARGETS IN ESCHERICHIA COLI 1. *Annual review of genetics* 30, no. 1: 465–506.

Grimalt, J., H. Iturriaga, and X. Tomas. 1982. The resolution of chromatograms with overlapping peaks by means of different statistical functions. *Analytica Chimica Acta* 139: 155–166.

Grimes, B. A., S. Lüdtke, K. K. Unger, and A. I. Liapis. 2002. Novel general expressions that describe the behavior of the height equivalent of a theoretical plate in chromatographic systems involving electrically-driven and pressure-driven flows. *Journal of Chromatography A* 979, no. 1 (December 6): 447-466. doi:10.1016/S0021-9673(02)01470-X.

Grob, R. L, and E. F Barry. 2004. *Modern practice of gas chromatography*. Wiley-IEEE. 978-0471229834

Guiochon, Georges, Tivadar Farkas, Hong Guan-Sajonz, Joon-Ho Koh, Matilal Sarker, Brett J. Stanley, and Tong Yun. 1997. Consolidation of particle beds and packing of chromatographic columns. *Journal of Chromatography A* 762, no. 1 (February 21): 83-88. doi:10.1016/S0021-9673(96)00642-5.

Guiochon, G., Felinger, A. 2004. Comparison of the Kinetic Models of Linear Chromatography. *Chromatographia.* 60 Supp. 1, S175-S180, doi: 10.1365/s10337-004-0288-7

Hahn, R., P. M. Schulz, C. Schaupp, and A. Jungbauer. 1998. Bovine whey fractionation based on cation-exchange chromatography. *Journal of Chromatography A* 795, no. 2 (February 6): 277-287. doi:10.1016/S0021-9673(97)01030-3.

Hannig, G., and S. C Makrides. 1998. Strategies for optimizing heterologous protein expression in Escherichia coli. *Trends in Biotechnology* 16, no. 2: 54–60.

Hefti, M. H, F. J Milder, S. Boeren, J. Vervoort, and W. J.H van Berkel. 2003. A His-tag based immobilization method for the preparation and reconstitution of apoflavoproteins. *BBA-General Subjects* 1619, no. 2: 139–143.

Heller, Renu A., Mark Schena, Andrew Chai, Dari Shalon, Tod Bedilion, James Gilmore, David E. Woolley, and Ronald W. Davis. 1997. Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. *Proceedings of the National Academy of Sciences of the United States of America* 94, no. 6 (March 18): 2150-2155. doi:VL - 94.

Huhalov, A., and K. A. Chester. 2004. Engineered single chain antibody fragments for radioimmunotherapy. *QJ Nucl Med Mol Imaging* 48, no. 4: 279–88.

Ingraham, R. H, S. Y.M Lau, A. K Taneja, and R. S Hodges. 1985. Denaturation and the effects of temperature on hydrophobic-interaction and reversed-phase high-performance liquid chromatography of proteins:: Bio-gel tsk-phenyl-5-pw column. *Journal of Chromatography A* 327: 77–92

International Conference of Harmonisation – Guidelines: Quality 8-10

Issaq, Haleem J., Timothy D. Veenstra, Thomas P. Conrads, and Donna Felschow. 2002. The SELDI-TOF MS Approach to Proteomics: Protein Profiling and Biomarker Identification. *Biochemical and Biophysical Research Communications* 292, no. 3 (April 5): 587-592. doi:10.1006/bbrc.2002.6678.

Joseph, John R, Andrew Sinclair, Titchener-, Nigel J Hooker, and Yuhong Zhou. 2006. A framework for assessing the solutions in chromatographic process design and operation for large scale manufacture. *Journal of Chemical Technology & Biotechnology* 81 (June): 1009-1020. doi:10.1002/jctb.1508.

Jungbauer, Alois. 2005. Chromatographic media for bioseparation. *Journal of Chromatography A* 1065, no. 1 (February 11): 3-12. doi:10.1016/j.chroma.2004.08.162.

Kaleas, Kimberly A., Charles H. Schmelzer, and Shelly A. Pizarro. 2010. Industrial case study: Evaluation of a mixed-mode resin for selective capture of a human growth factor recombinantly expressed in E. coli. *Journal of Chromatography A* 1217, no. 2 (January 8): 235-242. doi:10.1016/j.chroma.2009.07.023.

Kaltenbrunner, Oliver, Jungbauer, Alois and Shuichi Yamamoto. 1997. Prediction of the preparative chromatography performance with a very small column. *Journal of Chromatography A* 760, no. 1 (January 24): 41-53. doi:10.1016/S0021-9673(96)00689-9.

Kawahara, M., Nakamura, H., and Nakajima, T. 1989. Titania and zirconia as new ceramic column packing materials for high performance liquid chromatography. *Analytical sciences* 5: 485.

Keener, R. N, J. E Maneval, and E. J Fernandez. 2004a. Toward a robust model of packing and scale-up for chromatographic beds. 2. Flow packing. *Biotechnology progress* 20, no. 4.

Keener, R. N, J. E Maneval, and E. J Fernandez. 2004b. Toward a robust model of packing and scale-up for chromatographic beds. 1. Mechanical compression. *Biotechnology progress* 20, no. 4.

Kleymann, Gerald, Christian Ostermeier, Bernd Ludwig, Arne Skerra, and Hartmut Michel. 1995. Engineered Fv Fragments as a Tool for the One-Step Purification of Integral Multisubunit Membrane Protein Complexes. *Nat Biotech* 13, no. 2 (February): 155-160. doi:10.1038/nbt0295-155.

Ladisch, M. R. 2001. Bioseparations engineering: principles, practice, and economics. Wiley New York.

Langer, T., C. Lu, H. Echols, J. Flanagan, M. K Hayer, and F. U Hartl. 1992. Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. *Nature* 356: 683–689.

Lees, A., A. Topping, A. Razzaq, K. Reiter, and A. L Acosta. 2009. Purifying a Recalcitrant Therapeutic Recombinant Protein with a Mixed-Mode Chromatography Sorbent. *BioProcess International* 7, no. 2: 42–48.

Leonard, M. 1997. New packing materials for protein chromatography. *Journal of Chromatography B: Biomedical Sciences and Applications* 699, no. 1: 3–27.

Lesley, S. A. 2001. High-throughput proteomics: protein expression and purification in the postgenomic world. *Protein Expression and Purification* 22, no. 2: 159–164.

Levison, P., D. Toome, S. Badger, B. Brook, and D. Carcary. 1989. Influence of mobile phase composition on the adsorption of hen egg-white proteins to anion-exchange cellulose. *Chromatographia* 28, no. 3: 170-178. doi:10.1007/BF02319642.

Levison, Peter R. 2003. Large-scale ion-exchange column chromatography of proteins: Comparison of different formats. *Journal of Chromatography B* 790, no. 1 (June 25): 17-33. doi:10.1016/S1570-0232(03)00087-4.

Li-Chan, E., S. Nakai, J. Sim, D. B. Bragg, and K. V. Lo. 1986. Lysozyme separation from egg white by cation exchange column chromatography. *Journal of Food Science* 51, no. 4: 1032–1036.

Lillehoj, E. P., and V. Malik. n.d. S. 1989. Protein purification. *Adv. Biochem. Eng./Biotechnol* 40: 19–71.

Lindner W., Maemmerhofer M., Mixed-modal anion-exchanged type separation material; 2005, US Patent 7648636

Lindsay, S., and J. Barnes. 1992. *High performance liquid chromatography*. Wiley.

Low, Duncan, Rhona O'Leary, and Narahari S. Pujar. 2007. Future of antibody purification. *Journal of Chromatography B* 848, no. 1 (March 15): 48-63. doi:10.1016/j.jchromb.2006.10.033.

Lowe, C. R. 2001. Combinatorial approaches to affinity chromatography. *Current Opinion in Chemical Biology* 5, no. 3: 248–256.

Ma, Guijun, Jean Aucamp, Spyridon Gerontas, Ranna Eardley-Patel, Alan Craig, Mike Hoare, and Yuhong Zhou. 2009. Mimic of a large-scale diafiltration process by using ultra scale-down rotating disc filter. *Biotechnology Progress* 9999, no. 9999: NA. doi:10.1002/btpr.327.

Makrides, S. C. 1996. Strategies for achieving high-level expression of genes in Escherichia coli. *Microbiology and Molecular Biology Reviews* 60, no. 3: 512.

Marston, A., and K. Hostettmann. 2006. Developments in the application of counter-current chromatography to plant analysis. *Journal of Chromatography A* 1112, no. 1 (April 21): 181-194. doi:10.1016/j.chroma.2005.10.018.

McLaughlin, L. W. 1989. Mixed-mode chromatography of nucleic acids. *Chemical Reviews* 89, no. 2: 309–319.

Mhatre, R., W. Nashabeh, D. Schmalzing, X. Yao, M. Fuchs, D. Whitney, and F. Regnier. 1995. Purification of antibody Fab fragments by cation-exchange chromatography and pH gradient elution. *Journal of Chromatography A* 707, no. 2 (July 21): 225-231. doi:10.1016/0021-9673(95)00319-I.

Middelberg, Anton P. J., Brian K. O'Neill, I. David L. Bogle, and Mark A. Snoswell. 1991. A novel technique for the measurement of disruption in high-pressure homogenization: Studies on *E. coli* containing recombinant inclusion bodies. *Biotechnology and Bioengineering* 38, no. 4: 363-370. doi:10.1002/bit.260380406.

Ming, F., Howell, J., Acosta, F., and Hubble, J. 1993. Study on separation of conalbumin and lysozyme from high concentration fresh egg white at high flow rates by a novel ion-exchanger. *Biotechnology and Bioengineering* 42, no. 9: 1086-1090. doi:10.1002/bit.260420910.

Missiakas, D., and Raina, S.. 1997. Protein folding in the bacterial periplasm. *Journal of bacteriology* 179, no. 8: 2465.

Mori, S., and H. G Barth. 1999. *Size exclusion chromatography*. Springer Verlag.

Mountain, A, and Adair, JR. 1992. Engineering antibodies for therapy. *Biotechnology & Genetic Engineering Reviews* 10: 1-142.

Naish, P. J, and Hartwell, S. 1988. Exponentially Modified Gaussian functions—A good model for chromatographic peaks in isocratic HPLC? *Chromatographia* 26, no. 1: 285–296.

Neal, G., J. Christie, E. Keshavarz-Moore, and P. Ayazi Shamlou. 2003. Ultra scale-down approach for the prediction of full-scale recovery of ovine polycolonal immunoglobulins used in the manufacture of snake venom-specific Fab fragment. *Biotechnology and Bioengineering* 81, no. 2: 149-157. doi:10.1002/bit.10454.

Olive, J., and J. O. Grimalt. 1991. Log-normal derived equations for the characterization of on-line acquired chromatographic peaks. *Journal of chromatographic science* 29, no. 2: 70–77.

O'Neil, Paul, and Ian Sellick. 2004. Innovations that improve process column chromatography. *Filtration & Separation* 41, no. 2 (March): 24-27. doi:10.1016/S0015-1882(04)00070-9.

Papas, A. N. 1989. Chromatographic data systems: a critical review. *Critical Reviews in Analytical Chemistry* 20, no. 6: 359–404.

Pate, M. E., M. K. Turner, N. F. Thornhill, and N. J. Titchener-Hooker. 1999. The use of principal component analysis for the modelling of high performance liquid chromatography. *Bioprocess and Biosystems Engineering* 21, no. 3: 261–272.

Pedersen, L, J Mollerup, E Hansen, and A Jungbauer. 2003. Whey proteins as a model system for chromatographic separation of proteins. *Journal of Chromatography B* 790, no. 1 (6): 161-173. doi:10.1016/S1570-0232(03)00127-2.

Peska, J., J. Stamberg, J. Hradil, and M. Ilavskı. 1976. Cellulose in bead form:: Properties related to chromatographic uses. *Journal of Chromatography A* 125, no. 3: 455–469.

Pollock, Daniel P., Joseph P. Kutzko, Eszter Birck-Wilson, Jennifer L. Williams, Yann Echelard, and Harry M. Meade. 1999. Transgenic milk as a method for the production of recombinant antibodies. *Journal of Immunological Methods* 231, no. 1 (December 10): 147-157. doi:10.1016/S0022-1759(99)00151-9.

Queiroz, J. A., C. T. Tomaz, and J. M. S. Cabral. 2001. Hydrophobic interaction chromatography of proteins. *Journal of biotechnology* 87, no. 2: 143–159.

Rathore, A. S, and H. Winkle. 2009. Quality by design for biopharmaceuticals. *Nature Biotechnology* 27, no. 1: 26–34.

Rege, K., M. Pepsin, B. Falcon, L. Steele, and M. Heng. 2006. High-throughput process development for recombinant protein purification. *Biotechnology and bioengineering* 93, no. 4.

Regnier, Fred E., Bing He, Shen Lin, and John Busse. 1999. Chromatography and electrophoresis on chips: critical elements of future integrated, microfluidic analytical systems for life science. *Trends in Biotechnology* 17, no. 3 (March 1): 101-106. doi:10.1016/S0167-7799(98)01294-3.

Reynolds, T., M. Boychyn, T. Sanderson, M. Bulmer, J. More, and M. Hoare. 2005. Scale-down of continuous filtration for rapid bioprocess design: Recovery and dewatering of protein precipitate suspensions. *Biotechnology and Bioengineering* 83, no. 4 (February 3): 454-464. doi:10.1002/bit.10687.

Roettger, B. F, and M. R Ladisch. 1989. Hydrophobic interaction chromatography. *Biotechnology advances* 7, no. 1: 15–29.

Rogers, K. R. 2000. Principles of affinity-based biosensors. *Molecular biotechnology* 14, no. 2: 109–129.

Roque, A. C.A, M. Â Taipa, and C. R Lowe. 2005. An artificial protein L for the purification of immunoglobulins and Fab fragments by affinity chromatography. *Journal of Chromatography A* 1064, no. 2: 157–167.

Röthlisberger, Daniela, Annemarie Honegger, and Andreas Plückthun. 2005. Domain Interactions in the Fab Fragment: A Comparative Evaluation of the Single-chain Fv and Fab Format Engineered with Variable Domains of Different Stability. *Journal of Molecular Biology* 347, no. 4 (April 8): 773-789. doi:10.1016/j.jmb.2005.01.053.

Sasso, EH, GJ Silverman, and M Mannik. 1991. Human IgA and IgG F(ab')2 that bind to staphylococcal protein A belong to the VHIII subgroup. *J Immunol* 147, no. 6 (September 15): 1877-1883.

Schein, C. H. 1989. Production of soluble recombinant proteins in bacteria. *Nature Biotechnology* 7, no. 11: 1141–1149.

Schena, M., R. A Heller, T. P Theriault, K. Konrad, E. Lachenmeier, and R. W Davis. 1998. Microarrays: biotechnology's discovery platform for functional genomics. *Trends in Biotechnology* 16, no. 7: 301–306.

Schwartz, W., D. Judd, M. Wysocki, L. Guerrier, E. Birck-Wilson, and E. Boschetti. 2001. Comparison of hydrophobic charge induction chromatography with affinity chromatography on protein A for harvest and purification of antibodies. *Journal of Chromatography A* 908, no. 1: 251–263.

Scopes, R K. 1987. Dye-ligands and multifunctional adsorbents: an empirical approach to affinity chromatography. *Analytical Biochemistry* 165, no. 2 (September): 235-246.

Shalliker, R. Andrew, B. Scott Broyles, and Georges Guiochon. 2000. Physical evidence of two wall effects in liquid chromatography. *Journal of Chromatography A* 888, no. 1 (August 4): 1-12. doi:10.1016/S0021-9673(00)00517-3.

Shapiro, Michael S., Steve J. Haswell, Gary J. Lye, and Daniel G. Bracewell. 2009. Design and characterization of a microfluidic packed bed system for protein breakthrough and dynamic binding capacity determination. *Biotechnology Progress* 25, no. 1: 277-285. doi:10.1002/btpr.99.

Sherma, J., and B. Fried. 2003. *Handbook of thin-layer chromatography*. CRC.

Shih, Y. P, W. M Kung, J. C Chen, C. H Yeh, A. H.J Wang, and T. F Wang. 2002. High-throughput screening of soluble recombinant proteins. *Protein Science: A Publication of the Protein Society* 11, no. 7: 1714.

Skerra, A. 1994. A general vector, pASK84, for cloning, bacterial production, and single-step purification of antibody Fab fragments. *Gene* 141, no. 1: 79–84.

Sofer, G. K, and L. Hagel. 1997. Handbook of process chromatography: a guide to optimization, scale up, and validation. Academic press London.

Sorensen, H. P., and K. K Mortensen. 2005. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microb Cell Fact* 4, no. 1: 1.

Stickel, J.J., and A. Fotopoulos. 2001. Pressure-Flow Relationships for Packed Beds of Compressible Chromatography Media at Laboratory and Production Scale. *Biotechnology Progress* 17, no. 4 (8): 744-751. doi:10.1021/bp010060o.

Tallarek, U., K. Albert, E. Bayer, and G. Guiochon. 1996. Measurement of transverse and axial apparent dispersion coefficients in packed beds. *AIChE Journal* 42, no. 11: 3041–3054.

Tankrathok, Anupong, Sakda Daduang, Rina Patramanon, Tomohiro Araki, and Sompong Thammasirirak. 2009. Purification Process for the Preparation and Characterizations of Hen Egg White Ovalbumin, Lysozyme, Ovotransferrin, and Ovomucoid. *Preparative Biochemistry and Biotechnology* 39, no. 4: 380. doi:10.1080/10826060903209646.

Ternes, W. 2001. Egg proteins. Chemical & Functional Properties of Food Proteins.

Tswett, M. 1906. Adsorption analysis and chromatographic methods. *Application to the chemistry of chlorophyll. Ber Dtsch Botan Ges* 24: 384–393.

Usher, K. M, C. R Simmons, and J. G Dorsey. 2008. Modeling chromatographic dispersion: A comparison of popular equations. *Journal of Chromatography A* 1200, no. 2: 122–128.

van Deemter, J. J., F. J. Zuiderweg, and A. Klinkenberg. 1956. Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography. *Chemical Engineering Science* 5, no. 6 (September): 271-289. doi:10.1016/0009-2509(56)80003-1.

van der Wal, F. J, G. Koningstein, C. M. Ten Hagen, B. Oudega, and J. Luirink. 1998. Optimization of bacteriocin release protein (BRP)-mediated protein release by Escherichia coli: random mutagenesis of the pCloDF13-derived BRP gene to uncouple lethality and quasi-lysis from protein release. *Applied and Environmental Microbiology* 64, no. 2: 392–398.

van Hout, M. W.J, R. A. de Zeeuw, and G. J. de Jong. 1999. Coupling device for desorption of drugs from solid-phase extraction-pipette tips and on-line gas chromatographic analysis. *Journal of Chromatography A* 858, no. 1: 117–122.

Vanhoutte, K., W. Van Dongen, I. Hoes, F. Lemiere, E. L. Esmans, H. Van Onckelen, E. Van den Eeckhout, R. E. J. Van Soest, and A. J. Hudson. 1997. Development of a nanoscale liquid chromatography/electrospray mass spectrometry methodology for the detection and identification of DNA adducts. *Anal. Chem* 69, no. 16: 3161–3168.

Venn, Richard F., James Merson, Susan Cole, and Paul Macrae. 2005. 96-Well solid-phase extraction: a brief history of its development. *Journal of Chromatography B* 817, no. 1 (March 5): 77-80. doi:10.1016/j.jchromb.2004.08.029.

Voitl, Agnes, Thomas Müller-Späth, and Massimo Morbidelli. 2010. Application of mixed mode resins for the purification of antibodies. *Journal of Chromatography A* 1217, no. 37 (September 10): 5753-5760. doi:10.1016/j.chroma.2010.06.047.

Wan, E. W.M, and F. Baneyx. 1998. TolAIII Co-overexpression Facilitates the Recovery of Periplasmic Recombinant Proteins into the Growth Medium ofEscherichia coli. *Protein expression and purification* 14, no. 1: 13–22.

Wenger, Marc D, Peter Dephillips, Colleen E Price, and Daniel G Bracewell. 2007. An automated microscale chromatographic purification of virus-like particles as a strategy for process development. *Biotechnology and Applied Biochemistry* 47, no. 2 (June): 131-139. doi:10.1042/BA20060240.

Wierling, P. S, R. Bogumil, E. Knieps-Grunhagen, and J. Hubbuch. 2007. High-throughput screening of packed-bed chromatography coupled with SELDI-TOF MS analysis: monoclonal antibodies versus host cell protein Pierre Schulze Wierling and Ralf Bogumil contributed equally to this work. *Biotechnology and Bioengineering* 98, no. 2.

Willoughby, Nik, Phillip Martin, and Nigel Titchener-Hooker. 2004. Extreme scale-down of expanded bed adsorption: Purification of an antibody fragment directly from recombinant E. coli culture. *Biotechnology and Bioengineering* 87, no. 5 (2004): 641-647. doi:10.1002/bit.20173.

Xu, J., T. Tan, and J. C Janson. 2007. One-step purification of epigallocatechin gallate from crude green tea extracts by mixed-mode adsorption chromatography on highly cross-linked agarose media. *Journal of Chromatography A* 1169, no. 1: 235–238.

Yang, T., C. M Breneman, and S. M Cramer. 2007. Investigation of multi-modal high-salt binding ion-exchange chromatography using quantitative structure–property relationship modeling. *Journal of Chromatography A* 1175, no. 1: 96–105.

Yuan, Q. S., A. Rosenfeld, T. W. Root, D. J. Klingenberg, and E. N. Lightfoot. 1999. Flow distribution in chromatographic columns. *Journal of Chromatography A* 831, no. 2 (January 29): 149-165. doi:10.1016/S0021-9673(98)00924-8.

Zhao, Guofeng, Xiao-Yan Dong, and Yan Sun. n.d. Ligands for mixed-mode protein chromatography: Principles, characteristics and design. *Journal of Biotechnology* 144; 1: 3-11 doi:10.1016/j.jbiotec.2009.04.009.